Data Loss Prevention and a Point of Sales Breach

Nicholas Kollasch

Data Loss Prevention and a Point of Sales Breach

GIAC (GSEC) Gold Certification

Author: Nicholas Kollasch, nicholas.kollasch@gmail.com

Advisor: Richard Carbone

Accepted: July 12th 2014

Abstract

Target could have used a data loss prevention solution to mitigate the success of its infamous data breach. However, organizations typically deploy data loss prevention with simple policies and rules that detect 15- or 16-digit number strings that might represent a credit card number; this strategy, would not be effective in the case of the Target attack due to the attackers packaging the "loot" with Base64 encoding directly on the point of sales systems. Therefore, a security practitioner requires alternative detection measures to detect this type of anomalous activity. Data loss prevention can support an organization's ability to implement the Critical Security Controls, thereby providing the capability to detect such a sophisticated attack during the key stage of the Kill Chain model: Actions on Objective. Data loss prevention, when implemented with robust rules that reflect current attack tactics, techniques, and procedures, can reduce the likelihood of success by making it a bit more difficult to extract the valuable data.

Introduction

Data breach headlines seem to appear quite frequently in our daily news sources, and retail breaches involving credit card numbers are no exception. The specifics of the breach might vary to a certain degree, and the number of credit card numbers might be tens of thousands, hundreds of thousands, or even millions. However, the overarching trend indicates that these types of headlines will continue at the current rate without any notable shifts in our protection measures.

One of the most widely publicized attacks involving the exfiltration of credit card numbers involved Target, which reported a breach of approximately 40 million customer credit and debit card numbers in December 2013; the attackers also managed to acquire the personally identifiable information (PII) of another 70 million customers shortly thereafter in January 2014 (Jarvis & Milletary, 2014). Jarvis and Milletary (2014), who authored an analysis of the Target malware for Dell SecureWorks, noted that the malware was one of the many variants of RAM "scrapers" or "skimmers" at the disposal of attackers. The sheer volume of account information and PII has certainly contributed to the amount of exposure and analysis of this attack.

Although the tactics used to execute the attack might not be considered sophisticated by themselves, taken in their collective nature, an argument can be made that this attack represented an extremely organized and dynamic effort to acquire valuable information. In fact, sentiments by security practitioners support both sides of this perspective illustrated by the comments by McAfee's Director of Threat Intelligence Operations, Jim Walter: "the malware utilized is absolutely unsophisticated and uninteresting" (Smith, 2014). Conversely, it has been reported that the Director of the Department of Homeland Security's (DHS) National Cybersecurity and Communications Integration Center (NCCIC), Lawrence Zelvin, assessed the malware as "incredibly sophisticated" (Committee on Commerce, Science, and Transportation, 2014, p. 3). This level of sophistication in the malware itself is also supported by Neiman Marcus CIO, Michael Kingston. In his testimony to the Committee on Commerce, Science, and Transportation (2014), Kingston stated that their forensic investigation, supported by the Secret Service, revealed that the malware used in their attack was "exceedingly

sophisticated." Both Target and Neiman Marcus were attacked by variants of the BlackPOS malware (Jarvis & Milletary, 2014; Newman, 2014). Ultimately, it must be understood that an attack often involves many different threads that can only wreak havoc when effectively woven together.

Leveraging Lockheed Martin's Kill Chain framework, the Target breach will be analyzed in this paper to identify specific areas where a targeted security solution may have reduced the likelihood of the attack to succeed. Data loss prevention (DLP) solutions do not represent a new capability that organizations are just beginning to implement. The Radacati Group, a technology market research group, assesses that the "content-aware" DLP market continues to grow at a rate of 20% annually (Radacati & Buckley, 2013). The DLP market growth appears to parallel the apparent increase in headlines relating to data breaches.

Many organizations use DLP to detect number sequences that look like credit card numbers. However, organizations need to realize that DLP's capabilities can and must go far beyond the simple detection of "structured data." DLP can also support existing Defense in Depth efforts by detecting anomalous behavior in the environment – a requirement to detect a breach like the Target attack. Best practices dictate that organizations layer their security efforts to provide Defense in Depth. DLP falls in line with these efforts, as most organizations will want to detect simple "structured data" when it is being transmitted in an unauthorized manner. However, with encrypted or encoded data elements, content-aware DLP solutions will possess little or no ability to detect structured number sequences (Symantec, 2014; Websense 2014; Fiscus, 2011). Therefore, organizations must respond appropriately and implement additional mechanisms to detect unauthorized transmissions of this type.

The SANS Institute developed the Critical Security Controls (Controls) to aid organizations with the mitigation of risk for security incidents, such as data breaches. The Controls permit organizations to identify, prioritize, and mitigate vulnerabilities in their defenses with an "offense must inform defense" approach, regardless of the security framework in use (SANS Institute, 2015). The Controls will aid the discussion of the

Target breach and provide a foundation to illustrate where the DLP strategy may need to shift based on the success of the Target breach.

The lessons learned from this attack will span many different technologies, processes, and procedures. However, organizations must have a starting point and need to utilize implemented tools to their fullest potential in order to reap their benefits. Because of the highly dynamic and evolutionary nature of the risk landscape, security practitioners must avoid a complete shift of focus but rather create extra layers of protection to account for additional types of vulnerabilities.

1. Target Breach and the Kill Chain

The Target data breach has been thoroughly analyzed by various security professionals. One of the most comprehensive analyses of the attack, aggregating sources across the industry to provide security practitioners with enough information to react, was contributed by the U.S. Senate's Committee on Commerce, Science, and Transportation (2014). While the report should not to be viewed as the absolute answer to advanced attacks, the report provided the industry with a starting point to mitigate the additional risk of this attack methodology, which was being identified in other organizations (Schwartz, 2014).

The authors of the Senate Committee report used Lockheed Martin's model to describe the lifecycle of an Advanced Persistent Threat (APT). The Intrusion Kill Chain model promotes an intelligence-driven approach to combat advanced threats (Hutchins, Cloppert, & Amin, 2011). The term *Advanced Persistent Threat*, widely attributed to analysts in the United States Air Force in 2006, appropriately describes the sophistication of some of these attacks, as the threat agents constantly alter strategies and tactics, techniques, and procedures (TTPs) to ensure the attack succeeds (Binde, McRee & O'Connor, 2011). The individual attack TTPs may not be considered advanced on their own.

The developers of the Kill Chain model designed the model to describe the attack lifecycle and inherently split the intrusion into phases that provide a security practitioner multiple opportunities to reduce the likelihood of success. Hutchins et al. (2011) designed

seven phases for the Kill Chain. Aggregating multiple media reports and analyses by security practitioners, the Senate Committee was able to put together a comprehensive timeline of events (Committee on Commerce, Science, and Transportation, 2014), which has already been described by numerous papers.

1.1. Reconnaissance

Definition: Research, identify, and select target.

It was reported that the attackers identified a Target HVAC vendor and were able to map out much of the network approximately two months prior to the attack (Krebs, 2014). Brian Krebs (2014), an independent security researcher, also explains that simple open source research into Target's vendors and supply chain would have yielded an exorbitant amount of data using metadata analysis of accessible files; the attackers would have little problem mapping out the Target network and identifying the specific targets.

1.2. Weaponization

Definition: Creating a "weapon" – typically, remote access mechanism via malware and using an exploit to deliver the malware (e.g., compromised PDF or Microsoft Office files).

Krebs (2014) also cited sources close to the breach investigation that the initial "weapon" in this attack included a password-stealing bot called Citadel. With this malware, the attackers would have been able to retrieve passwords using several mechanisms in the Target environment (Segura, 2012; Sherstobitoff, 2013).

1.3. Delivery

Definition: Sending the weapon to the target (e.g., transmitting the "infected" PDF or Office file to the target via an email, web page, or USB device).

Both Krebs (2014) and the U.S. Senate Committee on Commerce, Science, and Transportation (2014) assessed that the Target HVAC vendor Fazio fell victim to an email malware attack, which may have used either phishing or spear phishing tactics. Krebs (2014) goes on to note that this attack occurred approximately two months prior to the attackers actually exfiltrating data from the Target network.

1.4. Exploitation

Definition: "Detonation" of the "weapon" – typically exploiting vulnerabilities in applications or operating systems (OS) or the users themselves.

Notably, "Exploitation" at this phase of the Kill Chain centers on the exploit used to steal the credentials. As was noted in the "Weaponization" phase, some reports indicate that the Citadel Trojan was used to siphon the credentials of the vendor, Fazio (Krebs, 2014). Citadel, which typically has been used to steal passwords from targets in the Financial Services sector using a technique called "man-in-the-browser," can also retrieve other types of passwords, such as those for password management applications (Constantin, 2014). Citadel operates in part by using a web injection attack, prompting users to enter credentials, which are then captured by the malware; the malware can capture credentials using a variety of methods including key strokes, screen capture, or video capture (Segura, 2012).

However, the Senate Committee's report focused more on the "Exploitation" of the subsequent RAM-scraping malware, a variant of BlackPOS, rather than the prerequisite piece of malware permitting the attackers to gain an initial foothold into the environment (Committee on Commerce, Science, and Transportation, 2014; Jarvis & Milletary, 2014). Without the initial password malware or "Exploitation," the attackers would not have had the ability to install additional malware, so this phase should account for all prerequisites. The POS RAM-scraping malware, which exploits the vulnerable credit card numbers in memory, will be discussed further down the Kill Chain in Actions on Objective.

1.5. Installation

Definition: Installing the remote access mechanism or "backdoor" to permit persistent access to the target.

As was noted by the Senate Committee's and Dell SecureWorks' reports, the attackers used Citadel to acquire credentials; these credentials permitted the attackers to move laterally within the environment, and the credentials were elevated enough to install tools as services that require administrator privileges (Committee on Commerce, Science, and Transportation, 2014; Jarvis & Milletary, 2014).

1.6. Command and Control (C2)

Definition: Control of the compromised host by a command server outside the target environment, permitting "hands on the keyboard" access.

Jarvis and Milletary (2014) noted that the attackers had to establish C2 nodes internally in order to circumvent the segmentation required in PCI environments; another server with access to the Internet and internal dump servers were used to exfiltrate data via FTP at periodic intervals.

1.7. Actions on Objective

Definition: Attacker carries out the intended mission on the target, which could simply be monitoring, data exfiltration, data destruction, or moving laterally within the environment.

The specific objective of the Target attackers was to acquire cardholder data. Thus, Actions on Objective would encompass a number of actions. These actions included the installation of the RAM scraper on the point-of-sale (POS) systems, packaging the credit card numbers, transmitting the encoded credit card numbers to the internal dump servers, using the C2 node to retrieve the encoded data from the dump servers, and sending the complete package to external FTP servers (Jarvis & Milletary, 2014).

2. Capabilities of DLP

While the Committee did outline methods that Target could have used to stifle the attack at each phase in the Kill Chain, focus will be placed on phases and Controls where DLP solutions would have contributed to the mitigation of the attack. DLP solutions operate as a content and context-based security tool that can aid organizations in the identification of intentional or unintentional data loss events or broken business processes.

Traditional DLP can leverage a variety of detection mechanisms across multiple channels. Regular expressions, keywords, dictionaries, and validators such as Luhn checks can be used to detect specific data elements in email messages, uploads to

websites, transmissions to USB devices, and other types of communications (Symantec, 2014). Unlike traditional security tools including firewalls and IDS/IPS, DLP can provide the detection of confidential data elements when combined with a multitude of other factors, such as sender, recipient, protocol, or file type. DLP can also provide protection at different levels of the protocol stack. Detection can occur at the network or transport level, where email gateways and web proxies reside or at the application level on the specific host or workstation where the communication is originating.

2.1. Credit Card Number Detection

Detection of a credit card number can take several forms. Usually, detection centers on the use of regular expressions that detect a 15 or 16-digit number string, with or without common credit card number special characters or spaces. Gartner has repeatedly ranked Symantec's DLP solution as one of the leaders in their Magic Quadrant for content-aware DLP (Law, 2013). Symantec DLP uses regular expressions in association with data normalizers and validators to limit false positives. The regular expressions often involve number sequences that begin with and align to valid bank identification numbers (BINs), in association with various special character delimiters like spaces, dashes, or periods (Symantec, 2014).

Normalizers ensure that the detected data elements represent the specific form expected, such as numbers, digits, or a combination of both; validators enhance detection by providing supporting components to tune or verify the detected content (Symantec, 2014). One of the most common forms of validators is the Luhn check, which uses an algorithm to aid in the verification of a potential credit card number (Freeformatter.com, 2015). Other validators include the use of keywords in conjunction with number detection, a number delimiter that ensures the targeted credit card string does not reside in a larger number string, and exclusion of known test or invalid combinations of numbers (Symantec, 2014).

Organizations can also detect credit card numbers using what is known as an exact data match (EDM) profile. An EDM is essentially an export of actual confidential data elements from an existing system, such as a database, which DLP then uses to detect content (Symantec, 2014). For example, Target could have imported a hashed form of a

credit card number database into their DLP solution. DLP would then search for matches of these known credit card numbers instead of "guessing" with the use of regular expressions with normalizers and validators. Nevertheless, an EDM would not represent the best type of detection for an organization as large as Target, which processes countless numbers of credit card transactions per day. Although Symantec (2014) DLP ensures that the contents of an EDM is secure, many organizations are not willing to accept the additional exposure risk. Moreover, as Target processes new credit card numbers every day, the EDM would quickly become "out-of-date" and would require constant updating to be effective.

2.2. DLP Network Detection

It is unclear if Target had a DLP solution in place prior to the attack, and it is even more difficult to assess what types of detection policies and rules were in place. Assuming that a DLP solution was deployed, it would be typical that Target had specific policies to detect credit card numbers transmitted outside the environment. These policies would be able to detect a 15 or 16-digit number in an email or web transmission such as FTP using standard out-of-the-box rules for most DLP solutions. However, the attackers used Base64 encoding and retrieved the encoded credit card numbers directly from the POS servers (Jarvis & Milletary, 2014). Therefore, the traditional method for detecting valid 15 or 16-digit credit card numbers would be fruitless at the network level; the files would simply be gibberish to any network detection DLP component sitting at the perimeter.

2.3. DLP Endpoint Detection

Most organizations tend to avoid the installation of applications on critical production systems due to the risk of interference with the critical business functions. The retail industry is no exception. Endpoint detection for DLP necessitates the installation of a lightweight DLP application, similar to an antivirus application, on the host. This DLP endpoint agent then detects communications and actions involving confidential data directly on the host and sends this data to an intermediary server that manages all of the DLP agents. Types of activity that the DLP agent can detect include transfers to USB, emails, web posts, cut/copy/paste, printing, and association with specific applications like

encryption and packaging software, cloud storage applications, and others (Symantec, 2014). The detection occurs before any "downstream" tools add encryption or other protection measures.

In the case of the Target breach and with traditional detection of common number strings, detection by a DLP agent would only have occurred if the attackers decided to encode the data at a location other than the POS system. In other words, even if Target had a DLP endpoint agent installed on their POS systems, standard detection of valid credit card numbers would still be negated by the Base64 encoding.

2.4. DLP People and Process

A technology can only really provide actual benefits if capable people and defined processes are in place to support the technology, and vice versa. This concept is particularly important with DLP solutions because the technology integrates into so many different areas of an organization's IT infrastructure and corporate policy structure. For example, DLP integrates into components of email, networking, desktop, databases, and storage while aiding in the support of corporate acceptable use policies (AUPs), data classification policies, and incident response workflows.

A detected communication with suspected confidential data will generate a DLP event that has a lifecycle similar to events or incidents detected by other, traditional security tools like firewalls, IDS/IPS, or SIEM solutions. This DLP event lifecycle has several high-level components, and terminology may vary according to the organization and is examined in more detail in the following subsections:

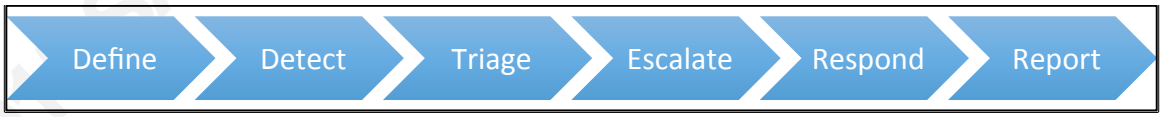

Figure 1: DLP Event Lifecycle. This figure depicts the stages through which a DLP event goes from creation of policy to metrics depicting closure statuses.

2.4.1. Define

In the case of Target, definition could simply be credit card numbers or magnetic stripe data, the encoded Track 1 and Track 2 data that includes cardholder PII and the card number itself (PCI Security Standards Council, 2015; Magtek, 2011). Define will

also involve the development and tuning of DLP policies or rules to ensure accurate detection.

2.4.2. Detect

Detection simply means generating accurate DLP events that could indicate data loss of the data elements previously defined.

2.4.3. Triage

Until the accuracy rate reaches a point when automatic responses can be implemented, detected DLP events must be validated, which is generally a manual process performed by an analyst. Even when a DLP policy is highly accurate, an analyst may still need to validate certain events to distinguish between personal use events and those deemed more nefarious. Triage adds context to the content detected by the DLP solution. For example, DLP can detect both a user shopping at an e-commerce site and an unencrypted FTP transfer to an external site. The former will not raise any red flags while the latter might provoke an investigation to find out more details. Triage also facilitates the detection accuracy of the DLP policies and rules in place because the analyst should be making recommendations for improving the logic to eliminate false positives and content that does not provide value (e.g., shopping at e-commerce sites).

2.4.4. Escalate

Once a valid DLP event is identified, appropriate resources must be notified. If the same FTP scenario described above in the Triage section existed in the Target environment, then Target's Incident Response Team and other relevant resources should be immediately notified.

2.4.5. Respond

It has been reported that Target's internal team did not respond to alerts from their FireEye security solution regarding the installation of the malware used in the attack (Riley, Elgin, Lawrence, & Matlack, 2013). Detection without a response is worthless, and DLP solutions act as no exception to this philosophy. The response should include analysis of other DLP activity for the sender and receiver, correlations, and supporting activity from other security technologies.

2.4.6. Resolve

Ensuring follow-through with the incident response workflow all the way to closure is fundamental for effective DLP program management. Depending on the type of DLP incident detected, a business process could be altered, an ongoing transmission could be terminated, or an employee could be disciplined. These situations all act as potential resolutions for detected DLP incidents.

2.4.7. Report

This final component to the DLP event lifecycle closes the loop and provides management with relevant metrics regarding the status of the DLP solution and associated incidents.

3. Critical Security Controls and DLP

DLP has been suggested as a mechanism that may have reduced the attackers' likelihood of success in the Target breach, illustrated in Jarvis and Milletary's (2014) Dell SecureWorks' report, but no details have been offered as to *how* DLP could have provided protection. Remember, the traditional DLP approach of detecting common number strings would not work with the malware used in Target's breach. The Critical Security Controls were designed with an "offense must inform defense" approach, which essentially leverages actual attack methodologies (offensive) to implement associated protective measures (defensive) (SANS Institute, 2014; Kellermann, 2012). These protective measures, or the Controls, must be continuously validated and their outputs monitored in order to reap their true benefits.

Currently, 20 Critical Security Controls have been identified and defined, all listed in order of prioritization or "must do first" order to ensure an organization can acquire the "low hanging fruit" as soon as possible (SANS Institute, 2014). A comprehensive DLP program can align with the definitions and objectives of several notable Controls. These Controls include Limitation and Control of Network Ports, Protocols, and Services (11), Boundary Defense (13), and Data Protection (17).

3.1. Limitation & Control of Network Ports, Protocols, & Services

This Control represents the eleventh Control in the list. DLP's ability in this realm is often overlooked because content-aware DLP solutions are designed to detect specific data elements in various forms of communications and storage locations. Many organizations typically only view DLP from this traditional perspective. However, DLP can also detect anomalous network activity much like a host-based and/or a network-based IPS detecting inbound activity. It is important to note that DLP's ability to aid in the control of this type of traffic does not insinuate that it can replace specific solutions designed for this function. Defense in Depth is the key, and DLP simply adds another layer to existing host and network-based security solution functions.

For example, Target could have implemented network-based DLP policies that prevented the transmission of FTP traffic, regardless of content. This type of DLP policy is common in many organizations, given the vulnerable nature of the protocol, and would act as a "backup" to existing firewall or IPS policies that prevent the same type of traffic. Additionally, a DLP policy created for a DLP endpoint agent residing on the POS systems could have limited the ability of applications to communicate on specific ports, using whitelists. In other words, the Target attackers mounted network shares from the POS systems to the internal dump servers; an endpoint DLP policy could have prevented this communication on ports 445 and/or 137-139 (Jarvis & Milletary, 2014).

DLP detection polices can be configured to only permit certain combinations of communications using both sender and recipient information. Thus, organizations can use the sender and recipient combination as a whitelist, in conjunction with content-aware detection mechanisms like file type, keywords, or specific numbers. Although the policy set is much more difficult to maintain due to the complexity, this type of layered defense should be fundamental to organizations with highly valuable data like credit card numbers and other PII elements. This methodology also utilizes the "offense must inform defense" approach of the Controls.

3.2. Boundary Defense

Boundary Defense is typically controlled using traditional security solutions like firewalls, IPS, and proxies. The capabilities of these traditional security tools are

expanding quite rapidly to better detect anomalous activity using a combination of signatures, heuristics and deep packet inspection. Supporting these capabilities, organizations can leverage DLP to act as a content-aware layer to Boundary Defense. Just as in the previous Control, Limitation and Control of Network Ports, Protocols, and Services, DLP can be configured to whitelist specific sender/recipient combinations at the perimeter. FTP transfers again act as a good example to illustrate notable DLP activity due to the vulnerable nature of the protocol. Target could have blocked any undefined FTP transfers from the network out to the Internet with DLP.

Another positive capability of DLP is that organizations do not necessarily need to define specific content to trigger the DLP policy. Rather, organizations only need to identify if FTP should be used to transfer *any* type of data from the network. Because the Target attackers maintained a persistent presence within its network and were able to move laterally almost at will, layering DLP into existing boundary defense measures may have reduced the attackers' likelihood of success.

3.3. Data Protection

The Data Protection Control represents the most applicable Control for DLP solutions when one thinks of traditional DLP capabilities. However, content-aware DLP solutions, including Symantec, would have little ability to detect these credit card numbers moving within and external to the Target environment with traditional content-based detection policies. This important fact represents a significant "offense must inform defense" tactical shift to how an organization leverages a DLP solution. That is, because attackers packaged the credit card numbers with Base64 encoding, DLP would need to detect other attributes of the communication(s), such as the sender/recipient combinations mentioned in the descriptions of the previous Controls.

It is quite important that the DLP strategy for implemented detection policies should not ignore the ability to detect sequences of numbers that may represent a credit card number. Organizations still need these content-aware and content-specific policies in place as a layer of overall Defense in Depth. Other RAM scraping variants may not encode or encrypt the content before moving the data internally or externally, and

organizations would still be interested in the detection of unencrypted or un-encoded communications involving credit card numbers.

4. Enterprise DLP Use Cases

Vendors for enterprise content-aware DLP solutions include a few "major players:" Symantec, Websense, RSA (end of life at the end of 2017), McAfee, and Digital Guardian (formally Verdasys) (DLP Experts, 2014; Ouellet, 2013; EMC, 2015). Ouellet (2013) in his analysis of the 2013 Magic Quadrant for content-aware DLP, provides Gartner's definition of content-aware DLP as "those that can perform content inspection of data at rest or in motion, and can execute responses." Each of the above five solutions fall into this definition and were in the Magic Quadrant for 2013 (Ouellet, 2013).

In addition, Coles (2015), in a blog post for Skyhigh Networks, reveals that Symantec has almost 50 percent of the DLP market share while Websense has nine percent; McAfee has the second largest market share with 18 percent, but its growth has been somewhat stagnated by recent acquisitions and lack of product updates (DLP Experts, 2014). Therefore, discussions around enterprise level DLP technologies tend to start with these vendors.

As of versions 12.5 and 7.8, respectively, neither Symantec nor Websense possess the capability to decode Base64 encoded data (Symantec, 2014; Websense, 2012). However, Websense can detect data that has been encoded with Base64, regardless of original content (Websense, 2012), and Symantec (2014) can detect data that has been encrypted with PGP, GPG or S/MIME, or protected using other password mechanisms. These capabilities, however, are not real-time Base64 decoding, which would be fundamental for a timely and effective DLP solution.

If an enterprise DLP solution had the capability to decode Base64, every piece of data going through the DLP system would need to be decoded, creating a tremendous amount of overhead. As Kevin Fiscus (2011) mentions in his white paper for SANS, custom regular expressions can be developed to detect Base64 encoded social security numbers, but as the format changes, so does the output. The regular expressions would

need to account for every potential format, every potential preceding character(s), and every potential trailing character(s). Therefore, other detection mechanisms for this type of anomalous activity would be preferred.

To illustrate these mechanisms, several use cases are listed in the following sections. These use cases depict nefarious activity as opposed to scenarios involving accidental leakage or broken business processes often detected by DLP solutions. Thus, the use cases represent activity within the last stage of the Kill Chain lifecycle, Actions on Objective, and can provide organizations with a foundational policy set developed from real world attacks and within current capabilities.

4.1. Use Case 1: Clear Text FTP Transfer

This first use case will illustrate the traditional approach to DLP policy and detection. Any enterprise DLP solution will possess the capability to detect a plain text or clear text credit card number being sent via FTP to an external recipient (Ouellet, 2013; Symantec, 2014; Websense, 2014). The DLP policy will leverage regular expressions to detect valid credit card numbers in various formats. The credit card number may be one continuous string or split by dashes, periods, or spaces, but will most likely utilize a Luhn check to increase validity; the DLP policy may also require validators, such as keywords and/or phrases: Visa, MasterCard, account, account number, credit card number, among many other terms (Symantec, 2014). These validators can aid in tuning the policy but may increase the likelihood of false negatives. Furthermore, exceptions for common test credit card numbers may also be implemented to tune the policy (Symantec, 2014). Unfortunately, neither Symantec (2014) nor Websense (2014) would be able to detect the transfer if the attacker used SFTP instead of FTP.

4.2. Use Case 2: PGP Encrypted File FTP Transfer

This second use case expands on the first. Instead of the credit card numbers sent in clear text, the attacker uses PGP to encrypt the file first. The attacker then transfers the file to an external FTP server. The policy that detected Use Case 1 would not be able to inspect this encrypted file – the file would simply look like gibberish to the DLP solution. Thus, the DLP policies would have to be expanded to include logic for authorized versus unauthorized senders and recipients, and perhaps authorized file types. For example,

organizations that deal with large amounts of credit card numbers should identify if there are any known, authorized business processes that permit FTP as a transmission mechanism. If organizations cannot identify any authorized processes using FTP, then DLP should be configured to block FTP, regardless of content.

If, however, an authorized FTP process is identified, then DLP should only permit that specific combination of sender to recipient along with the specific content or file type. That is, if the internal server 10.1.1.1 is authorized to send encrypted files to the external server acme.ftpserver.com because the connection uses a site-to-site VPN, then the DLP policy would need to whitelist this activity and block everything else (Symantec, 2014). These detection policies would reside at the "network level" using DLP servers designed to inspect various web traffic like FTP, HTTP, and HTTPS near the perimeter or web proxy (Symantec, 2014; Websense, 2014).

4.3. Use Case 3: Encoded Internally, Encrypted Externally

The third use case takes elements from Use Case 1 and Use Case 2 and features some components from the Target breach. The attackers collect the credit card numbers on internal POS systems, encode the numbers in simple text files, and use file shares to send the encoded numbers to an internal dump server. Then, the dump server aggregates the credit card numbers, encrypts the encoded credit card numbers with GPG, and sends the encrypted file via FTP to an external server.

If the same policy to block all unknown/unauthorized FTP transmission used in Use Case 2 is still in place, this transfer could be detected without any additional mechanisms. However, if the organization has not blocked any FTP transmissions, the DLP policy set can be expanded to attempt to detect the unauthorized internal file transfers. This expanded policy would bring the detection down to the "host level" rather than the "network level;" each POS system would need a DLP endpoint agent installed. The organization would also need to identify if any legitimate service or process on the POS systems should use the standard file share ports of 445 or 137 to 139; if legitimate uses cannot be identified, the endpoint DLP policies can block attempts to transfer any file using these standard ports (Symantec, 2014).

Ideally, organizations should have DLP policies implemented to account for each of the three uses cases above. As with other security technologies, protection measures should support a Defense in Depth strategy, and these three use cases reflect the "offense must inform defense" approach of the SANS Critical Security Controls. Both Symantec (2014) and Websense (2014), as leaders in the DLP market according to a recent blog post by Coles (2015) for Skyhigh Networks, are capable of such a tiered detection strategy, but Symantec will require some effort around the prioritization of triage due to the likelihood of duplicate incidents. Ultimately, DLP policy sets do not tend to have a "one size fits all" scheme. What works for one organization may not work for the next, but having a starting point is extremely useful and effective.

5. Conclusion

The trend of attackers targeting companies in the United States is not likely to subside any time soon, and the cost of data breaches in the United States continues to exceed that of the global market. The Ponemon Institute (2015) has released their annual United States and Global data breach reports, concluding that data breaches in the U.S. cost, on average, approximately 71 percent more than the global average – $6.5 million in the U.S. versus $3.79 million globally.

The Kill Chain model also provides an organization with plenty of opportunities to respond to attacks occurring within its networks. Specifically, DLP solutions implemented in support of the Critical Security Controls will permit the organization to reduce the actual data exfiltration during the Actions on Objective phase of the Kill Chain. Regardless of type of detection, alerts from the DLP solution along with IPS, SIEM and other technologies, must provoke a response to properly mitigate the threat.

Credit card companies and financial institutions in the U.S. are also responding to these type of attacks by moving forward with EMV (Europay, MasterCard, and Visa) chip technology. Although the U.S. will only be taking advantage of the "anti-cloning" protective measures rather than the stronger form of authentication (PIN versus signature), the chip may provide retailers and financial institutions with a greater capability to protect the credit card number directly on the POS system via tokenization (Johnston, 2014; Huq, 2015).

Nonetheless, the importance of altering detection capabilities based on current attacks cannot be highlighted enough by the Target breach. This "offense must inform defense" approach ensures the organization has proper capabilities in place to protect critical assets and in turn, the bottom line. While its important that the existing mechanism of detecting "clear text" credit card numbers remains in place, new types of DLP detection policies and rules should be layered on top of existing policies in response to what is occurring in the wild and what is happening in the larger effort for credit card protection.

DLP technology is not a magic pill to protect organizations from data loss. A comprehensive program must surround the technology, and the program must remain fluid to react to the dynamic threat landscape. Organizations should realize that this threat landscape warrants various forms of detection, and content-aware DLP possesses a tremendous amount of detection capability. However, the implemented DLP capabilities must align with the strategy and approach of the Critical Security Controls, which may require a less traditional approach to DLP detection.

References

Binde, B. E., McRee, R., & O'Connor, T. J. (2011, May 22). *Assessing Outbound Traffic to Uncover Advanced Persistent Threat.* SANS Technology Institute. Retrieved from http://www.sans.edu/student-files/projects/JWP-Binde-McRee-OConnor.pdf

Coles, C. (2015). What are the top data loss prevention tools?. Retrieved from https://www.skyhighnetworks.com/cloud-security-blog/what-are-the-top-data-loss-prevention-tools/

Committee on Commerce, Science, and Transportation. (2014, March 26). *A "Kill Chain" Analysis of the 2013 Target Data Breach.* United States Senate. Retrieved from http://www.commerce.senate.gov/public/?a=Files.Serve&File_id=24d3c229-4f2f-405d-b8db-a3a67f183883

Constantin, L. (2014, November 20). Attackers use Citadel malware to steal password management apps. Retrieved from http://www.computerworld.com/article/2849848/attackers-use-citadel-malware-to-target-password-management-apps.html

DLP Experts. (2014). *DLP Leading Vendors Review.* Available from http://dlpexperts.com/dlpreview

EMC. (2015). Product versions life cycle: RSA data loss prevention suite. Retrieved from http://www.emc.com/support/rsa/eops/dlp.htm

Fiscus, K. (2011, April 13). Base64 can get you pwned!. Retrieved from http://www.sans.org/reading-room/whitepapers/auditing/base64-pwned-33759

Freeformatter.com. (2015). Credit card number generator & validator. Retrieved from http://www.freeformatter.com/credit-card-number-generator-validator.html

Huq, N. (2015, March 11). Defending against PoS RAM scrapers. Retrieved from http://blog.trendmicro.com/trendlabs-security-intelligence/defending-against-pos-ram-scrapers/

Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2011). *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains* [White paper]. Lockheed Martin. Available from

http://cyber.lockheedmartin.com/intelligence-driven-defense-and-the-cyber-kill-chain

Jarvis, K., & Milletary, J. (2014, January 24). *Inside a Targeted Point-of-Sale Data Breach.* Retrieved from Dell SecureWorks website: https://portal.secureworks.com/intel/mva?Task=ShowThreat&ThreatId=773

Johnston, S. (2014, October 28). Coming next fall: More chip and PIN cards in the U.S. Retrieved from http://money.usnews.com/money/personal-finance/articles/2014/10/28/coming-next-fall-more-chip-and-pin-cards-in-the-us

Kellermann, T. (2012, July 20). Offense must inform defense: The importance of continuous monitoring. Retrieved from http://blog.trendmicro.com/the_importance_of_continuous_monitoring/

Krebs, B. (2014, February 14). Email attack on vendor set up breach at Target. Krebs on Security. Retrieved from http://krebsonsecurity.com/2014/02/email-attack-on-vendor-set-up-breach-at-target/

Law, T. (2013, December 19). Symantec is an 8-time Leader in the Gartner Magic Quadrant for Data Loss Prevention (DLP). Retrieved from http://www.symantec.com/connect/blogs/symantec-8-time-leader-gartner-magic-quadrant-data-loss-prevention-dlp

Magtek. (2011). Magnetic stripe card standards. Retrieved from http://www.magtek.com/documentation/public/99800004-1.08.pdf

Mandiant. (2013). *APT1: Exposing one of China's Cyber Espionage Units*. Retrieved from http://intelreport.mandiant.com/Mandiant_APT1_Report.pdf

Newman, L. H. (2014, January 20). Target, Neiman Marcus credit card number hacks were caused by a 17-year-old Russian. Retrieved from http://www.slate.com/blogs/future_tense/2014/01/20/target_neiman_marcus_credit_card_number_hacks_were_caused_by_a_17_year_old.html

Ouellet, E. (2013, December 12). Magic Quadrant for Content-Aware Data Loss Prevention. Gartner. Available from http://pages.digitalguardian.com/Y0d4304PX00CH003rQO1AW0

PCI Security Standards Council. (2015). Glossary. Retrieved from https://www.pcisecuritystandards.org/security_standards/glossary.php

Ponemon Institute. (2015, May). *2015 Cost of Data Breach Study: United States.* Available from http://public.dhe.ibm.com/common/ssi/ecm/se/en/sew03055usen/SEW03055USEN.PDF

Radicati, S., & Buckley, T. (2013, March). *Content-Aware Data Loss Prevention Market, 2013-2017*. Retrieved from The Radicati Group, Inc. website: http://www.radicati.com/wp/wp-content/uploads/2013/03/Content-Aware-Data-Loss-Prevention-Market-2013-2017-Executive-Summary.pdf

Riley, M., Elgin, B., Lawrence, D., & Matlack, C. (2014, March 13). Missed alarms and 40 million stolen credit card numbers: How Target blew it. Retrieved from http://www.bloomberg.com/bw/articles/2014-03-13/target-missed-alarms-in-epic-hack-of-credit-card-data

SANS Institute. (2015, June 1). Critical Security Controls for Effective Cyber Defense. Retrieved from http://www.sans.org/critical-security-controls/

Segura, J. (2012, November 5). Citadel: a cyber-criminal's ultimate weapon?. Retrieved from https://blog.malwarebytes.org/intelligence/2012/11/citadel-a-cyber-criminals-ultimate-weapon/

Schwartz, M. J. (2014, January 21). Target, Neiman Marcus malware creators identified. Retrieved from http://www.darkreading.com/attacks-and-breaches/target-neiman-marcus-malware-creators-identified/d/d-id/1113510

Sherstobitoff, R. (2013). Inside the world of of the Citadel Trojan. McAfee Labs. Retrieved from http://www.mcafee.com/us/resources/white-papers/wp-citadel-trojan.pdf

Smith, C. (2014, March 13). It turns out Target could have easily prevented its massive security breach. Retrieved from http://bgr.com/2014/03/13/target-data-hack-how-it-happened/

Subcommittee on Commerce, Manufacturing, and Trade. (2015, February 5). *Protecting Consumer Information: Can Data Breaches Be Prevented?*. United States House of Representatives, Committee on Energy and Commerce. Retrieved from http://energycommerce.house.gov/hearing/protecting-consumer-information-can-data-breaches-be-prevented

Symantec. (2014). *Symantec Data Loss Prevention Administration Guide*. (Version 12.5). Moutain View, CA: Symantec Corporation. Available from http://www.symantec.com

Websense. (2012, August 21). Release notes for Data Security v7.7.2. Retrieved from http://www.websense.com/content/support/library/data/v772/release_notes/Data%20Security%20v7.7.2%20Release%20Notes.pdf

Websense. (2014). *Deployment Guide.* (Version 7.8.x). Retrieved from http://www.websense.com/content/support/library/data/v78/deploy/deploy_dss.pdf