Dissect the Phish to Hunt Infections

GIAC (GCED) Gold Certification

Author: Seth Polley, readingroom@nightvisionsecurity.org
Advisor: Christopher Walker, CISSP, GSEC, GCED, GWEB, GCWN, GCUX, GCISO
Accepted: February 2, 2017

Abstract

Internal defense is a perilous problem facing many organizations today. The sole reliance on external defenses is all too common, leaving the internal organization largely unprotected. The times when internal defense is actually considered, how many think beyond the fallible antivirus (AV) or immature data loss prevention (DLP) solutions? Considering the rise of phishing emails and other social engineering campaigns, there is a significantly increased risk that an organization’s current external and internal defenses will fail to prevent compromises. How would a cyber security team detect an attacker establishing a foothold within the center of the organization or undetectable malware being downloaded internally if a user were to fall for a phishing attempt?

Defense in Depth (DiD), when terminated at the host AV level, will eventually fail a company. There are several methods to extend the depth of your defense, to combat host-based security failures, and gain awareness of compromises in progress. An infrastructure of hosts that are Microsoft-centered can be difficult to monitor, troubleshoot, and/or investigate when these security concern arise. Command-line process auditing can bridge these gaps, becoming an incredibly powerful tool when combined with a Security Information & Event Management (SIEM) product. The M-Trends 2016 Report, based on Mandiant’s experience responding to breaches, states that “The median number of days an organization was compromised in 2015 before the organization discovered the breach (or was notified about the breach) was 146.” (Mandiant Consulting, 2016). What steps will you take to decrease your detection times? Will you detect these compromises or will you be notified by another party first (if at all)?

By using a case study method and analyzing processes developed by a world-class Security Operations Center (SOC), we will evaluate and compare tools that can be used to audit Windows hosts, analyze phishing emails at depth utilizing a multifaceted approach, and search for user-initiated compromises that a security stack can fail to identify. By identifying key Indicators of Compromise (IoCs), a company can begin detecting malicious activity and begin remediation in a timely manner, often within hours of the malware first being executed. Start hunting and stop threats before they have an overwhelming impact!
1. Introduction

Phishing is a technique that involves tricking one or more users into divulging sensitive information (such as usernames and passwords), clicking a link, or executing malware; this is done by masquerading as a confidential, legitimate, or trustworthy source. Some common types of phishing are deceptive phishing, CEO fraud, and malware-based phishing. The most common type of phishing, deceptive phishing, involves the impersonation of a legitimate company in an attempt to steal people’s personal information or login credentials, generally incorporating threats or urgency to scare users through emotional triggers. Messages are crafted instructing users to verify account information, suggesting fictitious or undesirable account changes, or hastening the acceptance of new services before the window of opportunity closes. CEO fraud tends to involve impersonation of key executives in an attempt to authorize fraudulent wire transfers to an attacker’s bank of choosing or disclosure of confidential information (leading to fraudulent tax returns and identity theft). Though we will touch on the first two types of phishing through the course of this case study, the greatest amount of time will be spent discussing malware-based phishing. These scams deceive the recipients into executing malicious software on the device(s), introducing the malware through an email attachment, a linked file download hosted on a compromised or malicious domain, or by exploiting security vulnerabilities through code or command injection.

Organizations tend to incorporate industry standards for perimeter defense, such as the implementation of firewalls, IDS/IPS, network antivirus, proxies, and/or other means of segmentation. These standards focus on detecting and/or preventing inbound attacks, but neglect or opt not to secure against outbound connections from the ‘impenetrable’ core. Attackers will exploit the easiest vulnerabilities and weaknesses of the targets, which all too often is the unsecured human element – the trusted users with administrative (or otherwise elevated) permissions. These social engineering techniques materialize most frequently (supported by many research studies) in the form of email phishing campaigns. One study, conducted by the Anti-Phishing Working Group (APWG), has reported that the number of phishing websites it detected jumped an alarming 250% between October 2015 and March 2016 (Anti-Phishing Working Group (APWG), 2016). By identifying attack methodologies and understanding the trends, a company can take steps to mitigate these risks and actively hunt for phishing campaigns.
2. Tools and Configurations

Before we dive into analysis of these phishing campaigns however, discussion will begin with some prerequisites that should first be implemented within an organization. There are many tools available to cyber security teams, but it can often be difficult to select the best one for the job. A review of the top tools for email management, logging and monitoring client devices, and Security Information & Event Management (SIEM) will be included. The review will be concentrated on Windows platforms, as it is the preferred Operating System for many corporate client and server devices, beginning at the perimeter and continuing towards the internal tools.

2.1 Secure Email Gateways

Secure email gateways are solutions that monitor inbound and outbound enterprise email for undesirable content and attempts to prevent these messages from being delivered to the intended recipients. The two leading providers of messaging security appliances, Cisco Email Security Appliance (or ESA, formerly known as IronPort) and Proofpoint Enterprise Protection focus on trying to prevent unwanted phishing campaigns, malware attachments, and spam messages. Configurations are sometimes utilized for basic Data Loss Prevention (DLP) capabilities through detection and/or prevention rules which monitor for sensitive data (corporate credit cards, Social Security Numbers, etc.) being sent through the inherently insecure email channels.

Whether you take the Gartner “Magic Quadrant” research methodology at face value or simply as ‘best of the worst’ solutions, the Cisco and Proofpoint technologies can be regarded as having the greatest competitive positioning for secure email gateways (Firstbrook & Wynne, 2015). A company may be ingesting logs containing email metadata into their SIEM, but for sake of impartial comparison here (as SIEM logs can vary in appearance and extracted fields), the core features for email searching as they apply to hunting phishing emails and correlating campaigns with different senders and/or subjects will be reviewed.

As a security professional that has been tasked with email analysis in both products, Proofpoint does provide a greater edge than Cisco within the Graphical User Interface (GUI). As Proofpoint notes in their data sheet on the “Smart Search”, they provide an intuitive interface in which an
analyst can quickly trace senders, recipients, subjects, attachments, and more, getting the results displayed with easy-to-understand dispositions (Proofpoint, Inc., 2012):

**Figure 1 – Proofpoint Smart Search**

**Figure 2 – Proofpoint Smart Search Time filter**

All noted fields and more (totaling 30 columns) are exported to Excel and can be filtered/searched:

**Figure 3 – Proofpoint Smart Search Results**

The ESA interface has similar search functions, but is a little more difficult to navigate and the exported results are more limited, as seen below:
From experience, each product does have the shortfall of habitually allowing a handful of messages through the filters when queuing buffers are reached, resulting in unwanted messages being dumped into your network. The gateways may block $x$ number of messages, then allow a few identical messages to be delivered, before continuing to block the remaining $x$ number of emails in the campaign. The Proofpoint Targeted Attack Protection (TAP) Attachment Defense Service and the Cisco Advanced Malware Protection (AMP) can be a powerful complement to a security stack by delivering dynamic malware analysis and sandboxing to provide protection against all malicious attachments, thus helping to reduce the threats delivered to users.

### 2.2 Monitored Mailbox
Cyber security teams should have a monitored mailbox where users can report questionable emails. Some teams designate the primary SOC mailbox for all the cyber-related security concerns that arise, while others may create a secondary mailbox just for fraud, phishing, and spam messages. Common designations for these mailboxes are user-friendly names like “malicious”, “phishing”, “spam”, and/or “suspicious”. Users should be trained and encouraged to report the emails in question to the SOC team by forwarding it as an attachment, to make full header details available for scrutiny. In Outlook, this can be accomplished by right-clicking one or more messages, selecting “Actions” or “More Actions”, and then “Forward as Attachment” from the menu. Alternatively, the keyboard shortcut “Ctrl + Alt + F” can be used. To simplify these actions, organizations may choose to implement Outlook plugins that allow the user to simply click once or twice and properly report the email as an attachment to a pre-configured mailbox. Some tools that provide this functionality, often in conjunction with the capability of managed phishing drills, are PhishAlarm, PhishMe, PhishReporter, ThreatSim, etc.

2.3 Audit Process Logging

An auditing policy specifies the categories of security-related events that will be audited by the system. When performing Digital Forensics and Incident Response (DFIR), these logs become invaluable and are the core foundation to some of the hunting techniques discussed later in the paper. Windows auditing allows not only the monitoring of Process Creation and Process Termination, but the tracking of Process Command Line events too. These later become useful when trying to determine where malware was initiated from and what, if any, malicious scripts were executed. The two screenshots below illustrate the basic logging capabilities – Figure 6 showing in the “Process Command Line” that “ping google.com” was executed when the new process “PING.EXE” was
launched and Figure 7 showing “mmc.exe” initializing the “eventvwr.msc”:

Figure 6 – Process Command Line for ping

Figure 7 – Process Command Line for eventvwr.msc
As Audit Process logging is the feature that allows for detailed examination of host events and provides the ability to track user activity, you will want to enable these policies through Group Policy Object (GPO) for the entire organization. The key policies that need to be enabled can be viewed through the Local Group Policy Editor (gpedit.msc); the minimum settings are noted below:


- OR -


- AND -

- Local Computer Policy > Administrative Templates > System > Audit Process Creation > Include command line in process creation events > Enabled

A resource that may be useful when configuring the auditing in your environment:


2.4 Operating System (OS) Logging and Monitoring Agents

Are you looking to further increase your monitoring and logging capabilities? By incorporating OS logging and monitoring agents, you can begin to capture greater detail such as Alternate Data Stream (ADS) execution, certificate monitoring, enhanced network connection logging, the process image file hashes, changes to file creation time, performance data, or even integration of custom logging sources. Three solutions are discussed, starting with some free, readily available tools, and then moving into others that are more specialized with feature rich options.

2.4.1 Windows Sysinternals’ System Monitor (Sysmon)
Windows Sysinternals’ System Monitor (Sysmon) is a free tool provided by Microsoft to monitor and log advanced system activity to the Windows event logs. Greater levels of detail can be obtained (configurable) about process creations for current and parent processes (as discussed earlier), network connections (optional), process image file hashes (MD5, SHA1, SHA256, and/or IMPHASH), changes to file creation time, and even captures events occurring early in the boot processes which may help identify actions made by kernel-mode malware. These logs can be collected using Windows Event Collection (Microsoft, 2016) or SIEM agents (as we will discuss shortly). The data can subsequently be analyzed at the central compilation point (read SIEM) and analyzed to identify malicious activity, system anomaly detection, or other statistical trends that can trace intruder activity across a network.

![Figure 8 – Sysmon Event Log](https://technet.microsoft.com/en-us/sysinternals/sysmon)


### 2.4.2 Intersect Alliance’s Snare Agent

Consider Intersect Alliance’s Snare Agent (formerly BackLog) for log monitoring that elevates the event logging and syslog capabilities above that of Sysmon’s offerings. Though Snare is a powerful standalone SIEM, it can be used to compliment any other 3rd party SIEM (SIEMs will be discussed soon). Unlike Sysmon, which is geared solely for Windows platforms
(Windows 7+ and Windows Server 2012+), Snare Agents are available for Windows (XP+ and Windows Server 2003+), Linux, Solaris, and OSX. As Intersect Alliance states, “Snare Enterprise Agent for Windows is the industry standard for capturing and filtering audit and event log data, in a supported package, and with an enterprise-level feature set including highly reliable delivery, encryption, and custom event sources.” (InterSect Alliance International Pty Ltd, 2016). The agent provides front end filtering, interacts with the Windows Event log subsystem as a standalone auditing tool, or can be used to facilitate real-time transfer of event log information to a remote SIEM, Snare Server, or Syslog server. The two types of agents that Intersect Alliance issues are Enterprise and OpenSource Agents. The OpenSource Agents allow basic audit and event log collection as a stable solution, but the Enterprise Agent should be considered if a company needs to address audit or regulatory complains requirements. Some features that the Enterprise Agent provides are caching and confirmed delivery of log messages over TCP (in case of network disruptions), custom event logs, encryption protocols (TLS/SSL or 3DES), and a supported security platform (InterSect Alliance International Pty Ltd, 2016).

Snare Enterprise Agent – [https://www.intersectalliance.com/try-snare-eval-now/]
Snare Open Source Downloads – [https://www.intersectalliance.com/open-source-downloads/]

2.4.3 Kansas City Plant (KCP)’s Windows Logging Service (WLS)

The third and arguably the most feature-rich enhanced operating system logging and monitoring tool is Kansas City Plant (KCP)’s Windows Logging Service (WLS). As the name suggests, this is a Windows-only agent, but one that provides incredibly detailed logs, phenomenal support, and a continually growing list of additional features available. Some powerful features that are part of WLS are obfuscation detection (ADS execution), certificate monitoring, cryptographic hashes (MD5, SHA1, SSDeep (fuzzy hash), and more), file tailing, metadata collection, performance data, registry monitor, etc.

The example below is of an incredibly detailed Windows log showing a virus claiming to be a Windows Calculator application file being executed by Internet Explorer from a Temp folder (McCord & Green, 2010):
Dissect the Phish to Hunt the Infection - 11

Seth Polley - readingroom@nightvisionsecurity.org

Figure 9 – WLS Event Log

Sysmon and WLS can be coupled with a SIEM’s client-side event forwarder to aggregate the logs and index the events. The Snare Agent can be used to independently forward events to a chosen SIEM. The pairing of these logging and monitoring tools with the Splunk Universal Forwarders will be discussed later.

2.5 Security Information & Event Management (SIEM)

Security Information and Event Management (SIEM) seeks to provide a holistic view of an organization’s information security through the implementation of products and services that aggregate logs, normalize the data, and support real-time analysis of security alerts. Centralized collection and storage of security-related event logs allows for chronological timetabling of events and faster analysis of diverse device logs, making event correlation and trend analysis easier to perform. A SIEM works through the deployment of collection agents or event forwarding to gather
logging from a plethora of devices – end-user workstations (desktops, laptops), network equipment (proxies, routers/switches, VPNs), network security tools (antivirus, firewalls, intrusion detection/prevention systems), and even server platforms (Linux, Windows, Virtualized Servers). Not only is data aggregated from across a network, by forwarding the logs off-box, you prevent the deletion of all logs by attackers seeking to cover their tracks and you gain the ability to trace activities between systems. These are crucial to Digital Forensics and Incident Response (DFIR).

Vendors sell SIEM products as appliances, software, and even now as managed services. Leading vendors in this space are HPE (ArcSight), IBM (QRadar), LogRhythm, and Splunk. Open Source solutions that exist are AlienVault’s Open Source SIEM (OSSIM), ElasticSearch, Logstash, and Kibana (ELK), Enterprise Log Search and Archive (ELSA), LOGalyze, Security Onion (SO), and Snare. For the purposes of providing examples within this paper, Splunk will be the SIEM of choice. Don’t have Splunk? Get started with Splunk Free ([https://www.splunk.com/en_us/download/splunk-light.html]). “The Free license lets you index up to 500 MB per day and will never expire” (Splunk Inc., 2016).

Once a product is chosen, how are you going to get the client logs into your SIEM? To accomplish this, “Splunk forwarders consume data and send it to an indexer. Forwarders require minimal resources and have little impact on performance, so they can usually reside on the machines were the data originates.” (Splunk Inc., 2016). There are two types of Splunk forwarders (Splunk Inc., 2016):

- Universal forwarder – contains only the components that are necessary to forward data.
- Heavy forwarder – a full Splunk Enterprise instance that can index, search, and change data as well as forward it.

Now that the client-side logging has been increased for DFIR capabilities, the Splunk Universal Forwarder is utilized to send the data into the SIEM. The logs will be aggregated to allow for quick searching across all client machines. Note: Splunk supports both Windows and *nix forwarders. Should the Snare Agent have been used, it could provide similar forwarding functionality.
With the client-side logging addressed, take a moment to consider web traffic. At a minimum, basic client web traffic metadata should be logged in order to capture any URL (Uniform Resource Locator) hits or redirections to malicious websites. Some fields of interest in DFIR will be: HTTP Method, HTTP Status Codes, Hostname, Filename/Path, IP (Source and Destination), Protocol, Referrer, URI/URL, User Agent, and Username.

3. Phishing Analysis

Though there are many types of phishing techniques (bluejacking, phishing, smishing, vishing, etc.), one of the most common and commonly abused is email phishing (often put simply as just ‘phishing’). Given the prevalence of these attacks and their higher success rates, this discussion will focus solely on the email aspects.

Email has become a fact of life now and is the backbone of communication for most organizations. Despite its popularity, most would not likely be able to identify the four main elements of an email address (functionally, three key portions). When considering the email address [lumbergh@initech.com] (brackets added to prevent hyperlinking), the first is the mailbox, the second a delimiter character, the third is the domain of the mail server, another delimiter, and finally the Top-Level Domain (TLD). Though various networks and systems have different formats for the mailboxes, generally nicknames or usernames, these must be unique within each individual domain for proper message routing. The domain usually identifies the organization that owns the mail server (or at least manages it). The domains can be further subdivided, but this will not be discussed here. The last part of the email address is the Top-Level Domain (or the portion immediately following the “dot” symbol, which is primarily classified into two categories – generic TLDs and country-specific TLDs. A query of Google’s index shows the top TLDs as .com (Commercial), .org (Noncommercial), .edu (Education), and .gov (U.S. Government), though .net (Network services) is commonly used too.

3.1 Email Headers

The elements of the email address will come into play shortly, but first we will discuss email headers and the importance of them as it relates to cyber security. When an email is sent, it consists of three integral components – the envelope, the headers, and the body of the message. The envelope is part of the internal routing process and a component that the user will never see, so it
will not be discussed here. The body of an email is the boring filler almost everyone focuses on first. Now the header, arguably the most interesting part of an email, should be reviewed as part of your standard response process.

Why bother reviewing the email headers? Not only are email headers present on every email you receive over the Internet, they can provide you with valuable insight into the true origins of the message. If you suspect that an email is a phishing attempt or spoofed, you will want to inspect the headers to view the routing information in order to determine who the actual sender is and the return path of the email.

There are many methods to view email headers and these will be contingent on Operating System (OS), email client, and/or tools being utilized. As Windows operating systems and Outlook clients are frequently utilized in corporate environments, they will be discussed here solely. As the Microsoft Office Support article points out (Microsoft, 2016), there are even different methods across Outlook versions, but they can be viewed in Outlook 2010, 2013, and 2016 with the following steps:

1. In an open email message, click the File tab.
2. On the Info tab, click Properties.
   Header information appears in the Internet headers box.

Some headers, such as the “From”, “To”, and “Date” headers, are mandatory. Others that are optional, but commonly used, are “Return-Path”, “X-Mailer”, “CC”, “Subject”, and “Body”. Tools like MX Toolbox ([http://mxtoolbox.com/EmailHeaders.aspx]) can help parse the blob of text and make the email headers more human readable by parsing them according to the RFC 822 standard for message format. A normal header should contain some of the following primary parts:

Seth Polley - readingroom@nightvisionsecurity.org

© 2019 The SANS Institute
Author retains full rights.
Figure 10 – Email Header Example

Spoofed email headers will likely contain different Reply-To addresses, as in the following example:

Figure 11 – Spoofed Email Header

Though, some emails just don’t require much header review to get started on your investigation…

Figure 12 – Phishing Example

After confirming that the message is a phishing (or otherwise malicious) attempt, you will utilize your secure email gateway for further investigation to determine whether this was an isolated attempt or ascertain who else may have received messages in this campaign. As the sender was identified through the headers above, initiate a broad search on the sending domain (ESA – Envelope Sender Contains; Proofpoint – default search functions as ‘contains’, not ‘equals’). Should the
attempts have come from a common email service, such as Google (Gmail) or Microsoft (Hotmail > Live > Outlook), a search of these domains is likely to generate extremely high volumes of false positives (unless you work within an environment with strict policies against inbound/outbound personal email, as certain defense or financial groups might) and you may be better off including the mailbox with the domain.

A search of the sending domain “@bene3.cloudapp.net” (from a different phishing campaign than pictured above) reveals only one result, the message reported to us by the user containing the Subject “Detran Informa (907463700) - Multa /152B/ (61740)” - as seen below:

Sender: www-data@bene3.cloudapp.net
Recipient: lumbergh@initech.com
Subject: Detran Informa (907463700) - Multa /152B/ (61740)

Reviewing the Subject, the two words “Detran Informa” seem to be unique, whereas other portions such as the numeric strings can and often are randomized by malicious senders. Next, performing a cross-search on that smaller portion of the Subject reveals three messages – noting two new Senders that are different than the one reported to us:

Sender: www-data@live.com
Recipient: smykowski@initech.com
Subject: Detran Informa (907463700) - Multa /152B/ - [ 648437045980 ]

Sender: www-data@bene3.cloudapp.net
Recipient: lumbergh@initech.com
Subject: Detran Informa (907463700) - Multa /152B/ (61740)

Sender: www-data@bene20.cloudapp.net
Recipient: waddams@initech.com
Subject: Detran Informa (907463700) - Multa /152B/ (4962)

As suspected, the Subjects do vary; the tailing numbers having been randomized. The initial search shows that the phishing email was not an isolated event and this was only one email in a
larger campaign. It can be ascertained that the malicious sender uses different email domains, but
the source mailbox appears to be consistently “www-data”. There are three new data points that can
be cross-searched – incoming messages from mailboxes “www-data”, the domain “@live.com” (if
this was unlikely to produce a high volume of false positives in your environment), and the domain
“@bene20.cloudapp.net”. Based on the first indicator, a search is performed to identify all the
Sender addresses containing “www-data”. In addition to the three messages noted above, there are
three new messages identified:

Sender: www-data@live.com
Recipient: gibbons@initech.com
Subject: Urgente Formulario De Retirada. (69912)

Sender: www-data@vxokxs.cloudapp.net
Recipient: bolton@initech.com
Subject: URGENTE - Entrega n?o Efetuada. (87829)

Sender: www-data@fsdrew.cloudapp.net
Recipient: nagheenanajar@initech.com
Subject: URGENTE - Entrega n?o Efetuada. (98416)

New indicators to review may be portions of the Subjects likely to be unique, such as
“Urgente Formulario” and “URGENTE - Entrega”, and the two new domains
“@vxokxs.cloudapp.net” and “@fsdrew.cloudapp.net”. Continue pivoting down the rabbit hole
until no new messages or senders have turned up. Should there be any attachments present, the same
methodology can be applied to the attachment name(s). If an attachment named “Urgente
Formulario.doc” was found, begin with a search on attachment name containing “Urgente
Formulario”. Were any other attachment types found (.xls, .zip, etc.)?

Subsequent pivoting may be performed on “client-ip” (discussed more from a threat
intelligence perspective later) and the “X-Mailer”, if desired. These can provide great points of
analysis too. Which client-ips might you expect to see? Those belonging to Google and Microsoft
may not raise concern and/or may generate a high volume of false positives, but what if the client-ip
is directly associated with a malware domain? An analyst should investigate these further to see what other Senders/Subjects have traversed the company’s email appliances. There are many potential X-Mailers to be found when analyzing emails too, but which of them (if any) are concerning to you? Some may be expected; others may surprise you with malicious indications:

- client-ip=52.29.64.96
- client-ip=65.55.116.46
- client-ip=104.47.41.224
- client-ip=195.110.35.120
- client-ip=209.85.214.179

- X-Mailer: Apple Mail (2.753.1)
- X-Mailer: Barok... email.passwords.sender.trojan---by: spyder
- X-Mailer: iPad Mail (7B405)
- X-Mailer: Microsoft Office Outlook, Build 12.0.4210
- X-Mailer: PHPMailer 5.1 (phpmailer.sourceforge.net)
- X-Mailer: PyMailGUI 2.1 (Python)
- X-Mailer: YahooMailWebService/0.7.260.1

Having identified all of the email messages possible by pivoting through the unique Senders and Subjects, now is the time for the initial remediation. Should a jump box (a computer for remote administration that is specially secured) be available to you, utilize the Exchange Management Shell in conjunction with the Search-Mailbox cmdlet ([https://technet.microsoft.com/en-us/library/dd298173(v=exchg.160).aspx] and [https://technet.microsoft.com/en-us/library/jj983804(v=exchg.150).aspx]) to search and destroy. TechNet states that “Exchange Search indexes many item properties, including sender, recipients, message body, and attachments for email messages.” (Microsoft, 2014). Properties that you may find particularly useful are “Attachment”, “Body”, “From”, “Received”, and “Subject”:
Whether you have direct access to Exchange to sanitize the emails from the users’ mailboxes, need to engage the Messaging team for deletion of any emails, or have to contact each user individually – it is advisable to initiate the removal of these emails to ensure they are not an active threat waiting for an unwary user to open and execute the malicious content.

### 3.2 Command-and-Control (C&C or C2)

Once the immediate threat of the phishing campaign is mitigated, a closer analysis of the environment for potential compromises should be performed. One or more URLs may have been contained in the emails, directing the user to compromised or malicious web servers for credential harvesting or execution of web-based scripts. Search your SIEM for web traffic to the C&C – has any user clicked the email link? If you detect any traffic to the C&C in question, past or present, determine the intent of the website. Whether you evaluate the traffic passively through packet capture, actively from a sandboxed environment using an anonymized network like Tor, or some other means, evaluate not only the face value (f.e. fake Dropbox website) but the source code too – ensuring no web-based scripts are present either and it doesn’t POST a different domain/IP. If the site is simply a credential harvesting web site, evaluate the traffic to determine if only GET requests were made, or if the user POSTed credentials to the site (often in cleartext). Should the user have fallen victim and submitted credentials, initiate a password reset for their account, notify them to
change the password at any other site they have reused it on, and help educate them against future attacks.

An attachment may contain a link to C&C or, perhaps, something more sinister. Whether you can detonate the malware in a sandboxed environment or need to Reverse Engineer (RE) it, determine the callouts – any domains or IP addresses it attempts a connection to, note these. If you are new to RE, visit Appendix A to view some great tools that can get you started.

What about attachment hashes? Search your SIEM for any callouts (above) or execution of files containing the same hashes. These are both indications of a compromised host that need to be rebuilt. If any code execution is identified, send the computers for rebuild and/or have the network access immediately disconnected.

Should you find that a user has clicked a link from a phishing email, consider searching Exchange for the Body text in the user’s mailbox or performing a web traffic search against the user’s account. There are times when a phisher will repeatedly use a phishing domain in otherwise unrelated email campaigns. Utilize the Body property to determine any emails containing the domain (see search examples below). Should the search not identify any emails (Figure 14), perhaps it did not originate within your email system. Users can be notorious for forwarding questionable emails to their personal email or those of family members (especially in the case of fake DHL, FedEx, and UPS packages – ‘did you use my work email for this?’). Check for web traffic from the time that the user clicked the phishing link going back maybe five or ten minutes. Is there a high volume of traffic to an online email provider (such as Gmail or Hotmail/Outlook)? Bingo! Figure 15 shows the user accessing an online email provider before clicking the phishing link.

![Exchange Sanitization Queries](image)

Figure 14 – Exchange Sanitization Queries
Dissect the Phish to Hunt the Infection - 21

Seth Polley - readingroom@nightvisionsecurity.org

Figure 15 – Splunk Web Traffic Events

<table>
<thead>
<tr>
<th><em>time</em></th>
<th><em>url</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>2016-12-19 05:02:03</td>
<td><a href="http://messaging.bigpond.com/">http://messaging.bigpond.com/</a></td>
</tr>
<tr>
<td>2016-12-19 05:02:36</td>
<td>signon.bigpond.com:443</td>
</tr>
<tr>
<td>2016-12-19 05:02:00</td>
<td>messaging.bigpond.com:8080/</td>
</tr>
<tr>
<td>2016-12-19 05:02:03</td>
<td>messaging.bigpond.com:8080/</td>
</tr>
<tr>
<td>2016-12-19 05:02:02</td>
<td>messaging.bigpond.com:8080/</td>
</tr>
<tr>
<td>2016-12-19 05:02:02</td>
<td>messaging.bigpond.com:8080/</td>
</tr>
<tr>
<td>2016-12-19 05:02:02</td>
<td>messaging.bigpond.com:8080/</td>
</tr>
</tbody>
</table>

If your policies and tools permit, create blocks for these C&C domains and/or IP addresses at your firewall or proxy. Many phishers and malware authors reuse the same C&C for later campaigns. Why take the chance that another campaign comes in reusing the C&C and another user falls victim? Take the initiative and block them from being able to access the domain(s) or IP address(es) in question. Is there doubt as to the legitimacy of the domain (i.e. a customer or supplier that may have been compromised)? Perform 30, 60, or 90-day searches to ensure you don’t have daily legitimate traffic.

3.3 Domain Intel Analysis

When analyzing phishing emails, domain intelligence can be helpful in identifying other potential attack vectors a threat actor may utilize. Consider the sending domain of the email, the client-ip from which it originated, and even the C&C itself. When reviewing these domains and IP addresses, are there other malicious or suspicious indicators identified? Have others reported different C&C or file hashes associated with those you identified? Go at least one level deep, searching your SIEM for these indicators, and evaluate the advantages to blocking them (when possible). The emails may signify a large phishing campaign affecting the entire organization or a targeted spear phishing attack on your C-level executives. Why take the chance by ignoring other intelligence available to you – search the IOCs and block them! Note: If you are unsure where to begin with domain analysis, see Appendix B for a list of commercial and freely available resources.

© 2019 The SANS Institute
Author retains full rights.
4. Hunting the Infection

The direction of analysis to this point has been largely focused on phishing analysis and pivoting to find all the related Indicators of Compromise (IoCs). Now that you have been gathering information on the threats, what is the next step? When analyzing the C&C, we performed cursory searches to determine if the users of this current campaign had hit any of the domains/IPs and file hashes in question. The next step in maturity for any SOC is to move away from the sole format of reactive response and towards the process of proactively (and iteratively) searching through your environment’s logs to identify any previous or future undetected attacks or compromises. Begin hunting to mitigate risks and actively monitor for any new or advanced threats establishing a foothold in your networks. Though a basic understanding of Regular Expressions (RegEx) is preferred and will be of great benefit to you in the long run, some of the examples utilized can likely be applied ‘as is’ to your own initial searches. Take a look at some of the typical attack methods below to get started.

4.1 Phishing Links and Repeated URL Formats

The problem:
Phishers will randomize credential harvesting links sent through email, making it even more difficult to implement blocks when email Sender and Subject correlations don’t detect all of the messages from a particular phishing campaign. As you search and sanitize, if you notice URL patterns begin to emerge, consider implementing Splunk RegEx searches to help fill in the gaps in your detections.

The resolution:
Use RegEx to parse the URLs based on the patterns you have identified. The following RegEx looks for two blocks of randomized letters and/or numbers of 5-15 characters in length following the TLD and being submitted to a PHP page with the ID of your user’s mailbox and domain:

```
daysago=30 index=web_proxy \ .php\?id
| regex url=".\/*.\*\(a-zA-Z0-9\)\(5,15\)/\(a-zA-Z0-9\)(5,15)\ .php\ ?id=.*\(a-zA-Z0-9\)(1,15)\@[a-zA-Z0-9\]0-9\(1,10\)\.[a-zA-Z](2,3)\$"
| table time user url
```

Figure 16 – Splunk Search with RegEx

The detections:
All the events look harmless, right? Nope! These are each credential harvesting domains:
4.2 Randomized Numeric Attachment Names

The problem:
As with the credential harvesting links, malware authors will randomize file names and their lengths to avoid detection. When attachments are sent from multiple Senders and contain many Subjects, don’t just rely on your pivoting to catch all of the messages in a campaign – find the emerging patterns and hunt them!

The solution:
More RegEx! Search Splunk for the email metadata on attachments. The following RegEx searches for DOC and XLS attachments containing only numeric names of varying lengths:

The detections:
Seem legitimate? Randomized senders, subjects, and attachment names were used to help avoid detection; each document contains a malicious macro:
4.3 PowerShell Downloaders

The problem:
After reverse engineering only a handful of malware scripts, you may already notice a pattern of PowerShell invocations attempting to download second stage malware (namely Dridex). Though the presence of cmdlets like Copy* and New-Object* alone are not necessarily indicators of maliciousness, these are a great asset for attackers and items worth keeping your eye on.

The solution:
Did you guess more RegEx? Not this time… The search looks for Windows scripting processes, commonly abused cmdlets, and the utilized CommandLine parameters:
The detections:

Those files look safe, don’t they? These are scripted attempts to download second stage malware:

---

**Figure 20 – Splunk Search for Scripting Methods**

---

**Figure 21 – Splunk Search Results**

---

### 4.4 Carriers of Malicious Scripts

The problem:

Attackers commonly use Office Word (.doc/.docm), Excel (.xls/.xlsm), and PowerPoint (.ppt/.pptm) or Adobe products (.pdf) to launch malicious scripts from inside your defenses. The common Windows script handlers are “cscript”, “wscript”, and “powershell”. You generally don’t want to see Word launching PowerShell from Temp folders. Or when Excel calls “cscript
‘C:\Users\<username>\Desktop\Databases\_Public Loto Permit Excel\reg_setting.vbs’, you might want to take a second look and see what the user just executed.

The solution:
Proactively search for any carrier files an attacker may email or provide a link to download. When you see script handlers executed from CommandLine for abnormal CreatorProcessNames, investigate the source:

![New Search](image)

Figure 22 – Splunk Search for Malicious Carriers

The detections:
Some examples of detections you might notice include VBS & EXE CommandLine execution from Microsoft Office Excel & Word:

![Figure 23 – Splunk Search Results](image)

5. Conclusion
Phishing is a highly successful technique and one that will only continue to grow in popularity. Attackers generally follow the path of least resistance and email has proven itself to be just that threat vector. “91% of targeted attacks commence with using email as a point of entry. Further, … 78% of targeted email attacks utilize malware that has been embedded within an attachment. Given these points, clearly attackers perceive email to be a path of least resistance to evade existing security defences and to breach your network.” (Corson, 2014). Start preparing your organization to defend against these attacks by educating your users as to how they can detect these attacks. Consider denoting external emails by appending a warning banner to the Subject stating this is from an “**External Email**” or to the Body stating “**Warning: External Email - Think before you click!**”. Incentivize users, not with corporate concerns like bad publicity, downtime, or loss of revenue, but with
personal motivators that will strike home. Make them aware of risks such as identity theft from divulging sensitive information, loss of personal data like family pictures from ransomware infections, and even develop a 3-5x strike policy in which corrective action takes place at work for consecutive phishing drill failures. Users are susceptible to phishing campaigns and drills today, but what happens when the quality of the attacks develop to such a degree that it prevents the majority of your user base from detecting them? As with many techniques, phishing will continue to develop and become more sophisticated as time progresses. By implementing phishing awareness programs now, you can begin training your employees to identify phishing attacks and avoid their pitfalls. The benefit of training will not only be seen in monetary and time savings to the Information Technology (IT) departments and the organization as a whole, but in the mentality of the end users who are now able to assist in the fight against the cyber-attacks being leveraged against your organization.

Malicious or criminal attacks account for 47% of data breach incidents and the average total cost of a successful phishing attack that leads to a data breach is $3.79 million. The average organizational cost of a data breach varies by country, but the US sample comes in the highest at $6.53 million (Ponemon Institute, 2015). A single click by just one person opening a malicious link contained within a phishing email can put your entire company at risk. Begin taking steps to incorporate alternative detection methods and reduce the cost of compromise for your organization. Develop and mature your cyber security detection capabilities by implementing hunting tactics as a standard process. No longer rely solely on your perimeter defenses or fallible antivirus and DLP solutions. Take the initiative to track down the phishing emails, pivoting through the IoCs and intelligence available to you, reverse engineer the malware, identify the patterns, and begin alerting on them before a significant compromise brings your organization to its knees.

The hunting examples given above were not intended to be an exhaustive approach to hunting infections from phishing campaigns, but rather a means to get you thinking about attacks that are unique to your organization. The threat landscapes will vary by industry and each approach should be tailored to the specific attacks being leveraged against your organization. Visit Appendix C for presentations and more resources to continue developing your hunting methods.
References


Appendix A - Malware Analysis Tools

Are you new to malware RE? Get started with these free tools:

- Buster Sandbox Analyzer [http://bsa.isoftware.nl/]
- Cuckoo Sandbox/Malwr [https://cuckoosandbox.org/] and [https://malwr.com/]
- IRMA (Incident Response Malware Analysis) [http://irma.quarkslab.com/]
- OfficeMalScanner [http://www.reconstructor.org/code.html]
- REMnux (Reverse-Engineering Malware Linux) [https://remnux.org/]
- Revelo [http://www.kahusecurity.com/tools/]
- Sandboxie [https://www.sandboxie.com/]
- VirusTotal [https://www.virustotal.com/]
- Virtual Environments (QEMU, Qubes OS, VirtualBox, Windows Virtual PC / Hyper-V)
- Zero Wine [https://sourceforge.net/projects/zerowine/]

Awesome Malware Analysis – [https://github.com/rshipp/awesome-malware-analysis]
DFIR with Windows Logging Service (WLS) –
Analyzing Suspicious PDF Files With PDF Stream Dumper – [https://zeltser.com/pdf-stream-dumper-malicious-file-analysis/]
Appendix B - Domain Analysis Tools

Commercial:
- Umbrella, formerly OpenDNS [investigate.umbrella.com]
- DomainTools [whois.domaintools.com]
- PassiveTotal [passivetotal.org]
- Mnemonic [mnemonic.no]

Free:
- ThreatMiner [threatminer.org]
- Threatcrowd [threatcrowd.org]
- Robtex [robtex.com]
- DNSDumpster [dnsdumpster.com]
- Censys.io [censys.io]
- IP Intel [ipintel.io]


IoC Automation Scripts – [https://github.com/BechtelCIRT/cassava], [https://github.com/BechtelCIRT/extract_iocs], and [https://github.com/BechtelCIRT/pivoteer]
Appendix C – Splunk Hunting Presentations

2016:


SplunkQueries – [https://github.com/BechtelCIRT/SplunkQueries]

2015:

SplunkLive! Santa Clara – [http://www.slideshare.net/Splunk/bechtel-customer-presentation],
[http://conf.splunk.com/session/2015/recordings/2015-splunk-117.mp4], and

2010:

# Upcoming Training

<table>
<thead>
<tr>
<th>Event Name</th>
<th>Location</th>
<th>Dates</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community SANS Columbia SEC501 @ UKI</td>
<td>Columbia, MD</td>
<td>Mar 16, 2020 - Mar 21, 2020</td>
<td>Community SANS</td>
</tr>
<tr>
<td>SANS Security West 2020</td>
<td>San Diego, CA</td>
<td>May 06, 2020 - May 13, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS San Antonio 2020</td>
<td>San Antonio, TX</td>
<td>May 17, 2020 - May 22, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS Cyber Defence Canberra 2020</td>
<td>Canberra, Australia</td>
<td>Jun 29, 2020 - Jul 11, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS London September 2020</td>
<td>London, United Kingdom</td>
<td>Sep 07, 2020 - Sep 12, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS Baltimore Fall 2020</td>
<td>Baltimore, MD</td>
<td>Sep 08, 2020 - Sep 13, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS Network Security 2020</td>
<td>Las Vegas, NV</td>
<td>Sep 20, 2020 - Sep 27, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS Paris September 2020</td>
<td>Paris, France</td>
<td>Sep 21, 2020 - Sep 26, 2020</td>
<td>Live Event</td>
</tr>
<tr>
<td>SANS OnDemand</td>
<td>Online</td>
<td>Anytime</td>
<td>Self Paced</td>
</tr>
<tr>
<td>SANS SelfStudy</td>
<td>Books &amp; MP3s Only</td>
<td>Anytime</td>
<td>Self Paced</td>
</tr>
</tbody>
</table>