THE ROLE OF AUTOMATION IN EFFECTIVE PHISHING DEFENSE

2019

COFENSE.COM

One of the fundamental 'Uncomfortable Truths' about phishing defence is that no matter how good your perimeter controls, malicious emails are still being delivered to user inboxes. In recent research[1], Cofense^{TM} observed that 90% of malicious emails were found in environments using Secure Email Gateways. The specific reasons for this high rate of failures are broad, but are rooted in the fact that the evolution of threat actor tactics and techniques far outpaces the ability of technology to keep up. This results in a capability gap that threat actors are only too ready to exploit.

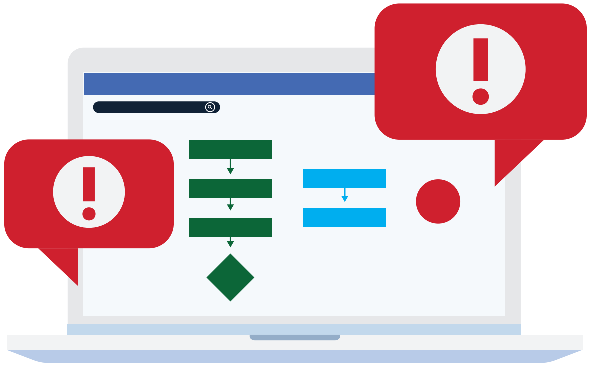
Mature information security organisations recognise that users are not the problem. Users are vital to improving organisational security posture, and specifically, the ability to better defend against phishing attacks. These same organisations know that effective phishing defence is more than raising user awareness to drive down clicks on malicious payloads. Users must be empowered to become an active part of defence – identifying, reporting and providing visibility of suspicious emails that technology has failed to stop.

When users have a simple one-click mechanism to report suspicious emails, they rise to the challenge.
In large organisations, they can quickly provide a stream of data, much of it useful to Security Operations teams. However, all this reporting creates another challenge—sifting through and analyzing a high volume of emails.

From a threat visibility perspective, over-reporting is always better than under-reporting. But before phishing threats become visible, Security Operations teams must prioritise, analyse, and understand a steady flow of reported emails – in essence turn them into actionable intelligence. As a result, Security Operations teams look to orchestration as a means to consume, analyse, categorise and respond to these reports, without analyst input. The mindset behind this approach is understandable, as even users in the best conditioned organisations still report significant volumes of non-malicious emails. With Security Operations teams already resource constrained, cutting through this noise to 'find bad quickly' can seem an overwhelming task.

THE ROLE OF SECURITY ORCHESTRATION, AUTOMATION & RESPONSE (SOAR) TOOLS

As the threat landscape evolves, organizations deploy increasing numbers of point solutions aimed at addressing specific evolving threats. However, as the number of tools proliferates, so does the volume of alerts and data that needs to be consumed and processed. As a result, organizations deploy SOAR tools to integrate these disparate tools and address a wide range of Security Operations use cases.



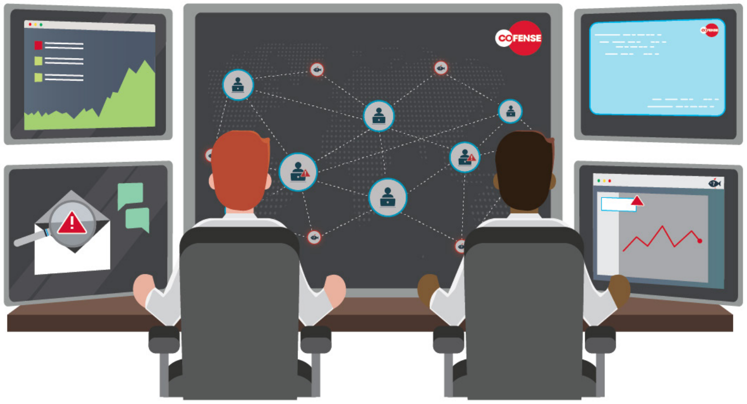
These use cases typically fall into one of three high-level categories:

1. **Data Enrichment/Enhancement** – collection of data from disparate sources and aggregated into a single alert or incident
2. **Data Determination** – making decisions on whether something is good or bad
3. **Data Actioning** – doing something with data such as IOCs to streamline response

In the context of phishing defense, SOAR tools add the greatest value in the third category where identified Indicators of Compromise (IOCs) are passed to the SOAR tool to be actioned, for example, blocking malicious sites at a proxy or updating Secure Email Gateway policies.

However, due to the fast evolution of the phishing threat landscape and rapidly changing associated threat actor tactics and techniques, SOAR tools are less effective in the second category – giving a reliable determination of whether something is malicious or not.

The Cofense Phishing Defense Center™ (PDC) receives and analyses emails reported by some 2 million users globally. 1 in 7 of the emails reported into the PDC are found to contain malicious content, a significant volume of bad.



In 2018, it included the following highlights:

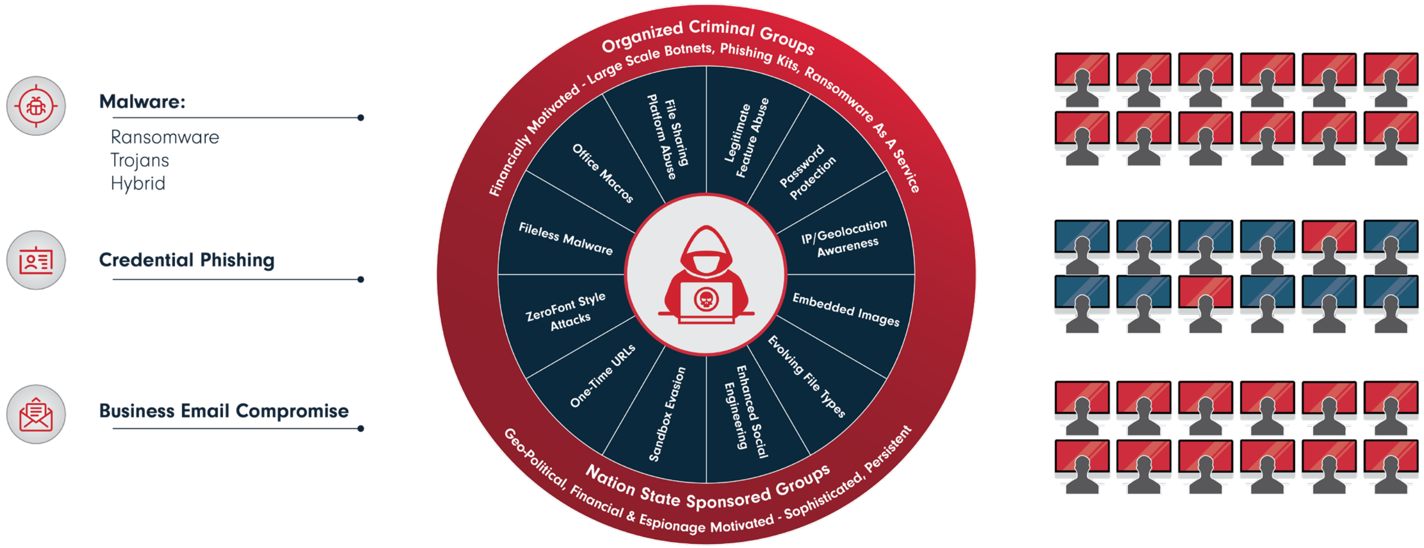
- 55,404 Credential Harvesting attacks
- 27,501 campaigns delivering malicious attachments, including abuse of filesharing services
- 4,152 Business Email Compromise attacks

These are big numbers, and it should be remembered that every one of the emails received by the PDC has bypassed some form of automated analysis by a secure email gateway or other inline threat scanning tool. Therefore, Security Operations teams must consider the fact that if automated controls have failed to identify a threat prior to delivery to the recipient, can technology alone be relied upon to reliably analyse threats reported by end users after delivery to the inbox?

THE RISK OF FALSE NEGATIVE ANALYSIS RESULTS

False negative results from automated phishing threat analysis are a significant issue. These false negatives occur when automated analysis identifies a reported email as benign, but it is later observed to be a threat. In the period between the initial false negative result and the true threat being understood, the reporting end user typically receives an email that the reported email is safe and so takes an action leading to compromise, such as exposing corporate credentials.

The diagram below provides a snapshot of the current phishing threat landscape as observed by the Cofense ResearchTM, Cofense IntelligenceTM, and Phishing Defense teams. The types of attacks observed (malware, credential phishing and Business Email Compromise) are typically well understood. However, the toolbox of tactics and techniques that threat actors have at their disposal to maximise email delivery and payload execution are not so well understood. Each of the tactics highlighted in the diagram below have been observed to successfully bypass 'anti-phishing' controls and automated analysis.


Malware:
Ransomware
Trojans
Hybrid
Credential Phishing
Business Email Compromise
Organized Criminal Groups
Financially Motivated - Large Scale Botnets, Phishing Kits, Ransomware As A Service
File Sharing Platform Abuse
Legitimate Feature Abuse
Password Protection
IP/Geolocation Awareness
Embedded Images
Evolving File Types
Enhanced Social Engineering
Sandbox Evasion
One-Time URLs
ZeroFont Style Attacks
Fileless Malware
Office Macros
Nation State Sponsored Groups
Geo-Political, Financial & Espionage Motivated - Sophisticated, Persistent

To understand some of these tactics further, following are key examples. They show how automated analysis can be rendered ineffective.

FILE SHARING PLATFORM ABUSE

Threat actors recognise that organisations continue to invest in technical controls to identify, analyse, and remove malware threats attached to emails. They also recognise that organisations make use of file sharing platforms such as Dropbox and Google, amongst others, and these are not routinely blocked. Users are also typically comfortable in downloading content from these 'trusted' locations. As a result, the Cofense Phishing Defense Center regularly observes threat actors using these file sharing platforms to host and spread malicious content, including a 'legitimate' link in the phishing email to the content. This approach also makes it difficult for automated URL analysis tools to determine whether the link is malicious, particularly if user input of provided credentials is required. Worse still, as services such as Dropbox, OneDrive, or Sharepoint.com are used for legitimate business purposes, automated approaches commonly do not flag these services as malicious.

MAN-IN-THE-INBOX ATTACKS & ZOMBIE PHISH

Out of nowhere, someone responds to an email conversation that wrapped up months ago. It's a real conversation that actually happened. Maybe it's about a meeting, a job opportunity, or a reply to that problem you had over a year ago; this email is highly relevant to you. But something is off, the topic of the email is months out of date and now there is a weird error message. What is this devious tactic of reviving an email conversation long dead? It's the Zombie Phish. The Cofense Phishing Defense Center has recently defended against an extensive Zombie Phishing campaign against multiple clients. In such campaigns, fraudsters hijack a compromised email account and, using that account's inbox, reply to long dead conversations with a phishing link or malicious attachment. Due to the subject of the email being directly relevant to the victim, a curious click is highly likely to occur.

These types of attacks are dangerous as they can involve internal-to-internal communication or communication between trusted third parties. When combined with other techniques, such as malicious content being hosted in cloud-sharing services, inline controls and automated analysis can be neutralised.

URL EVASION TECHNIQUES

As organisations invest further in controls to defend against malicious URLs, phishing threat actors iterate their campaigns with tactics and techniques to defeat analysis or hinder subsequent analysis. The Cofense Intelligence, Research and Phishing Defense Center teams have observed use of the following tactics and techniques to circumvent controls. In many cases, these tactics and techniques can render URL pre-fetching and domain reputation services ineffective or results from threat analysis tools unreliable and inaccurate.

One-Time URLs

To disrupt efforts to understand and subsequently contain compromise following phishing attacks, threat actors increasingly employ one-time URLs. These URLs are typically configured to serve malicious content on the first click only and redirect all subsequent calls to benign content.

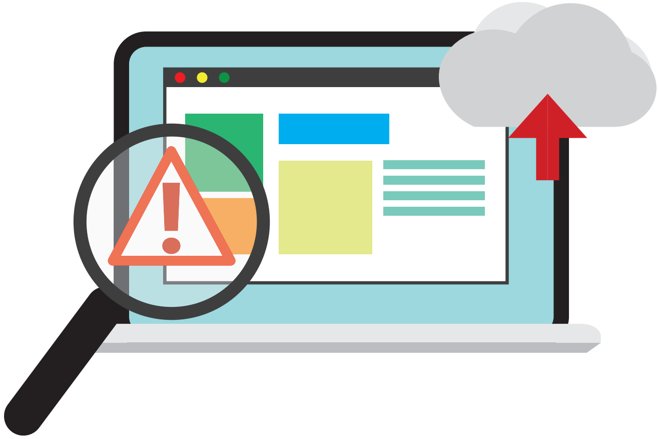
Post-Delivery Host Registration & Time-Delay URLs

With many URL protection solutions utilizing domain reputation services, it is not uncommon for phishing campaigns to delay the registration of hosts serving malicious content until after the campaign has been sent, rendering domain reputation services irrelevant.

Email gateway solutions are increasingly delivering URL pre-fetching capabilities to analyse URL content prior to delivery to the recipient. To defeat these solutions, threat actors employ the use of time-delay URLs to only serve malicious content 'x' minutes after the campaign starts. The intent with time-delay URLs is to serve benign content to analysis tools, then once cleared, serve malicious content to the intended recipient.

IP Range Filters & Traffic Misdirection

Many of the technical controls utilized by organisations are services hosted by major cloud providers such as Amazon Web Services, Microsoft Azure, or Google Cloud. Knowing this, threat actors increasingly include IP range filters to reject connections from these providers or employ IP traffic misdirection to serve benign content when they suspect analysis is taking place. To address this, the Cofense Phishing Defense Center employs a comprehensive network infrastructure across multiple providers to neutralise the effects of IP range filters and traffic misdirection, ensuring appropriate analysis and capture of IOCs.



IP/GEOLOCATION AWARE THREATS

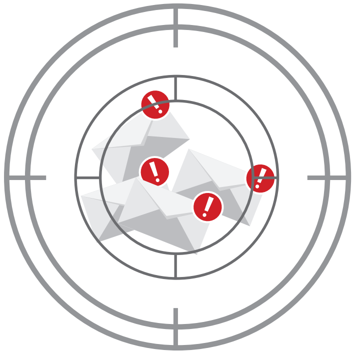
It is becoming increasingly common for threat actors to deliver phishing emails where the payload behaves differently, depending upon the location of the recipient. For example, a user in one country could be served benign content, whereas a user in a different country could be served a malicious payload. Or, different malware is retrieved and delivered depending on the determined location of the target. The use of IP/Geolocation awareness is another tactic employed to prevent analysis by security tools or researchers. For example, if a threat actor knows or suspects that specific cloud-based security tools might be in use, benign content is served to the known security vendor IP addresses. However, once deemed to be safe and delivered to the recipient, the malicious payload is delivered upon execution.

The Cofense Phishing Defense Center observed a campaign where reported emails used the proven theme of ‘Attached Invoice.’ Upon analysis, the attachment appeared benign – no malicious behaviour was observed. However, it had all the hallmarks of a phish and the analysts could see more reports arriving – all from Brazil. With this in mind, they put on their metaphorical Brazilian hat and gave their analysis workstation a Brazilian IP address.

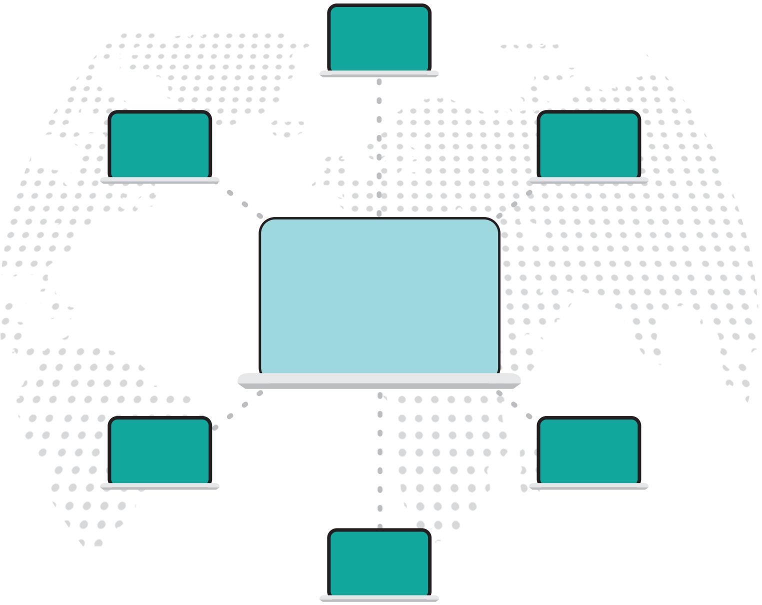
This time, upon execution, the analyst observed different behaviour with the attachment. A connection was made to payload infrastructure and a malicious script was downloaded. The script didn't execute, but deeper analysis identified further location validation checks. After configuring the analysis workstation with a Brazilian locale and keyboard layout, the sample was executed again, and, voila, IOCs were captured. The net result? Automated analysis would have had a hard time identifying this threat, as this customer's perimeter controls clearly did.

In addition to Geolocation-aware malware threats, the Cofense Phishing Defense Center has observed increasing volumes of Geolocation-aware credential phishing campaigns. In these campaigns, the credential harvesting pages will only accept input from clients in specific target locations.

MALWARE-SPECIFIC AUTOMATED ANALYSIS CHALLENGES – GEODO/EMOTET



Identification of a threat that has been reported by one or more end users is one step in the process of analysis, but merely identifying and then quarantining the threat is not enough. It is essential that Security Operations teams can fully understand the threat, including all Indicators of Compromise. This information is necessary to enable threat hunting, to identify users, or endpoints, that may have become compromised prior to the threat being reported and identified.



Typically delivered via a malicious Word document, Geodo is a banking trojan sharing some similarities with Dridex, Cridex, and other derivatives of the same codebase. An interesting feature of this malware is that it is also a worm, able to spread through e-mail. When the trojan is executed, it establishes a connection with its Command & Control (C&C) server in order to obtain the e-mail addresses and e-mail bodies and will start sending out messages, further spreading the malware.

The core functionality of the Geodo trojan lies in its ability to collect sensitive information from infected machines and their users. This is facilitated by sophisticated browser-based information-stealing capabilities, including form grabs, HTTPS man-in-the-middle attacks, and extensive injection of hostile content from third-party locations. Sandbox-based approaches to analysis of Geodo malware typically results in failure to capture all IOCs. Geodo always has 5 payload URLs, and at least 10 C&C URLs (often over 50 C2 URLs). If the first payload URL is available at the time of analysis, the sandbox will typically only provide that IOC. Similarly, if the first C2 server is available, that is the IOC reported by the sandbox. This means, on average, a sandbox in a default configuration will only provide 2 IOCs for malware that actually contains many more. As a result, the Cofense Phishing Defense Center ensures that Geodo/Emotet, and other malware strains like it are analysed manually. By intercepting network traffic at the appropriate time, it is possible to capture all IOCs for customers to consume for threat hunting purposes.

VALIDATION OF HUMAN INTERACTION PRIOR TO PAYLOAD DETONATION

Circumvention of automated analysis is a game of cat and mouse, engaged in by attacker and defender, with some techniques being surprisingly low tech. One such example is the use of CAPTCHA technology. In a twist that utilizes technology conceived to protect us against us, threat actors are including CAPTCHA verification in their campaigns to verify human interaction prior to payload detonation. By doing this, technical controls can be rendered ineffective as the content appears otherwise benign unless the CAPTCHA verification is successfully completed.

Another example of validation of human interaction is to redirect to malicious content or render malicious content only if user movement is detected. Using core HTML functionality such as on mouseover events, threat actors aim to circumvent automated analysis tools which perform no user input or struggle to reliably mimic real user interaction.

USE OF QR CODES TO CIRCUMVENT CORPORATE CONTROLS

An emerging credential phishing tactic observed by the Cofense Phishing Defense CenterTM involves the use of embedded QR code images in phishing emails, rather than the link itself. By doing this threat actors aim to achieve two primary objectives:

1. Ensure delivery of the email to the intended recipient by removing actual URLs from the email that can be analysed by automated security tools.
2. Aim to get the recipient to use a personal smart phone to scan the QR code, and access the malicious credential harvesting link outside of any corporate security controls.

When emails utilizing this tactic are analyzed, they appear to be simple HTML emails containing just text and images, and would typically be considered very low risk - significantly increasing the chances of a false negative analysis result.

CHALLENGES WITH IDENTIFICATION OF CAMPAIGN SCOPE AND IMPACT

When a malicious email has been identified, seconds count. When threats are identified by Security Operations teams, an early question in the incident response process is 'who else has received this threat?' A task that seems trivial can be more complex than expected. Greater automation in email threat hunting is beneficial and organisations often leverage scripts and orchestration tools to hunt more efficiently. However, email threat hunting activities can be hampered by several critical factors:

Poor Search Performance

Native search mechanisms in Exchange and Office365 are not optimized for fast threat hunting. These capabilities exist for less time dependent compliance search tasks. Native capabilities in these platforms utilize Powershell scripts or other Exchange Web Services-based methods. To avoid the overhead of these search requests impacting mail delivery, they are commonly throttled, resulting in a wide variance in time to receive results – often into hours for even modest sized environments. Mechanisms that utilize tracking logs are dependent upon the required logs being available for searching. With no SLA in place to make these logs available within a specified time, there can be unknown delays in log availability, thus impacting ability to effectively hunt for email-based threats.

Given the ability for email searches to access potentially sensitive information, they require elevated rights within the mail environment, which mail admins can be reluctant to provide to multiple SOC Analysts. Occasionally these rights will be provided for use within a script - where they are often hard-coded – resulting in privileged credentials being widely shared and retained by analysts when they leave. As a result, SOC teams are often reliant on already burdened mail administrators to perform search and destroy tasks. Cross team collaboration and process delays further slow response and increase threat exposure time.

Limited Search Scope

Threat hunting within the mail environment is typically restricted to limited search criteria – most commonly sender and subject. When under attack from a morphing phishing attack, the actual threat payload is more important than the sender and subject. Therefore, it is necessary to quickly identify all emails in a campaign that deliver the same threat payload, particularly when all senders or subject lines in the campaign may not be known.

Meeting Compliance Obligations

Corporate email is a platform that commonly contains much sensitive data, such as intellectual property, customer data, information related to mergers and acquisitions, and employee PII. Access to this data must be strictly controlled, especially in heavily regulated industries. To meet compliance obligations, organizations enforce the principle of least privilege – giving any user the minimum rights needed to do their job – and require strict separation of duty across critical or high-risk tasks and processes.

With privileged rights needed to perform native search tasks in Exchange and Office365 – and with shared credentials in scripts presenting significant compliance issues relating to accountability, integrity, and password management – mail administration teams are reluctant to give up these keys to the kingdom to Security Operations teams. Therefore, it is necessary to ensure that email search and quarantine tasks are fully logged, enabling the necessary auditing of who searched for what, when.

THE COFENSE APPROACH TO PHISHING DEFENCE

First, it is important to acknowledge that automation has its place within phishing defence. However, it is important to recognise that automation should be appropriate and should be able to deliver consistent, repeatable, and reliable benefits.

Today's SOAR tools offer great benefits in eliminating human effort and reducing human error on many repeatable tasks. From a phishing defence perspective, the data actioning use case described above is particularly relevant, enabling threat mitigation actions to be undertaken more efficiently with identified IOCs, for example:

- Blocking malicious URLs at a proxy
- Blocking senders or malicious hashes at the email gateway
- Running SIEM searches of proxy logs to identify users who have connected to malicious URLs prior to blocking
- Changing user passwords or increasing Active Directory (AD) monitoring on potentially compromised accounts
- Querying Endpoint Detection & Response tools for malicious artefacts and other IOCs

Automating Reported Email Analysis

However, due to the reasons described above, traditional SOAR and automation tools are less effective at making an accurate and reliable determination of whether a reported email is malicious or not. To address this gap, Cofense TriageTM is a phishing incident response solution to enable Security Operations teams to consume large volumes of user-reported emails and turn them into actionable intelligence. To complement Cofense Triage, the Cofense Phishing Defense Center provides specialist managed services for reported email analysis and response.

Cofense Triage provides advanced capabilities to enable Security Operations teams to cut through the noise of reported emails to find bad quickly. Automation is employed to lighten the load of Security Operations teams as they consume reported emails and prioritise, appropriately analyse, and mitigate threats – in essence, shifting the analyst's workload from 'analysis' to 'action.' One major obstacle to traditional approaches is the inability of analysts to understand threats holistically, rather than as a long list of individual emails being reported by users in an abuse inbox. To address this, Cofense Triage delivers comprehensive clustering capability to group reported emails by payload fingerprint. This allows analysis at a threat campaign level, rather than an individual email level.

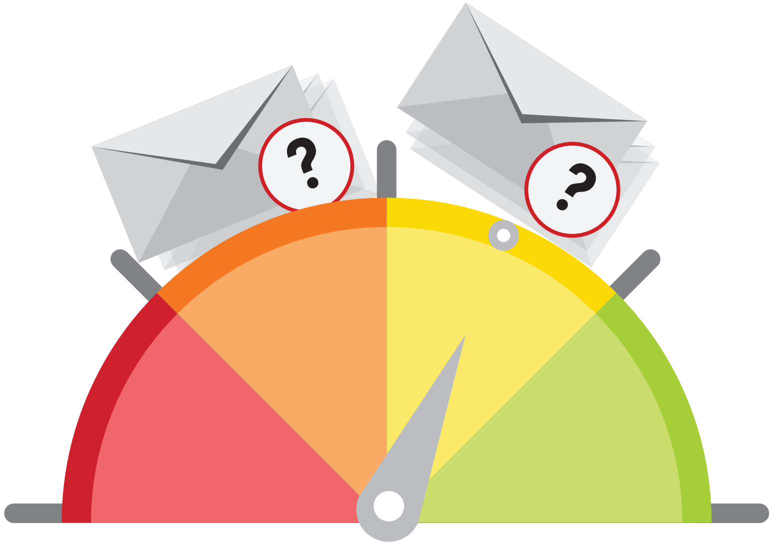
Once reported emails have been clustered, analysts require guidance classifying the type of email received. This is important to enable Security Operations teams to quickly identify known bad or filter out false-positives such as legitimate marketing emails that may have been previously opted in. The Triage Noise Reduction Engine assists this classification effort by using capabilities such as malware blacklists, content analysis, and heuristics to classify reported emails. When used in conjunction with email gateway x-header analysis, fast decisions can be made to automatically process non-malicious emails and respond to users accordingly.

To enable the analyst to undertake more efficient threat analysis, Cofense Triage contains an extensive library of rules to identify 'Indicators of Phishing.' The evaluation of reported emails against these rules provides insights into the key attributes of reported emails and their actual or potential risk. Cofense Triage rules are native, or YARA-based, and can apply to all elements of reported emails such as content, headers, and attachments. By identifying high-risk attributes, such as those related to new or evolved tactics and techniques, threats can be identified, even if other threat analysis tools observe behaviour as benign. The Triage rules library is updated regularly and is underpinned by the Cofense Intelligence, Cofense Research, and Cofense Phishing Defense Center teams, ensuring that Triage rules are able to identify the most current threat actor tactics, techniques, and campaigns. In addition, the ability for Triage operators to share rules they have created enables an intelligence sharing network that all

Triage operators are able to benefit from. Rule customization by Triage operators also enables further identification of known good emails, such as those from trusted third-parties or service providers to enable automatic processing, allowing analysts to focus their attention on the mails presenting the highest risk.

Integrations with third-party tools such as sandboxes and services like VirusTotal enable Cofense Triage to make automated submissions of files, hashes and URLs to these services for analysis. The results are then returned to Triage as referential data for analysts to use in building a picture of the threat as needed. Cofense Triage also includes a phishing-specific threat intelligence feed from Cofense Intelligence, allowing rapid identification of known threats

To focus analyst attention on reported emails with the greatest risk, Cofense Triage flags each potentially malicious or suspicious report and cluster based on results received from configured third-party threat analysis solutions and Cofense's own Triage Rules and Noise Reduction capability. In addition, Cofense Triage recognizes that users who report suspicious emails are not created equal. Some can consistently identify



and report suspicious emails with a high degree of accuracy, whereas others will report significant volumes of false positives. Cofense Triage automatically awards reputation points to users who consistently demonstrate that they can identify and report malicious emails. The presence of high reporter reputation scores on a report or cluster is an important mechanism to prioritize threats, even if those threats are zero-day and do not register on existing threat analysis tools and feeds. Cofense Triage operators can also designate appropriate users as VIPs. Typically, VIPs in Cofense Triage are users who an organization would expect are at a greater risk of phishing attacks, or those users where compromise would have a significant impact – such as Executives, Finance, or any user with a privileged level of access to systems, data, or business processes.

Once threats are appropriately analysed, clusters can be processed manually via the analyst, or automated through the creation of recipes. Processing of reported emails enables:

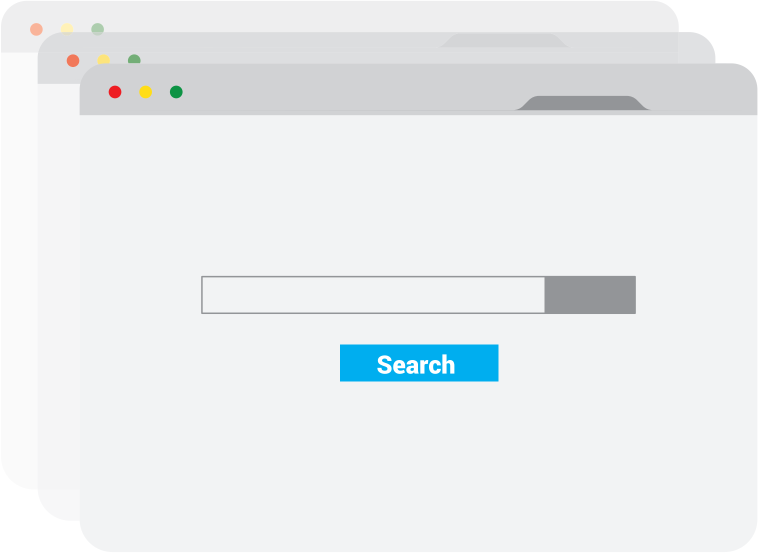
- Categorisation of the threat
- Responses to end users as appropriate
- Quarantine of threats through Cofense Vision™ (see next section)
- Sharing of IOCs with upstream and downstream teams for further threat mitigation actions

Fast and Compliant Email Threat Hunting and Quarantine

When a phishing threat is identified, seconds count. The faster Security Operations teams can take decisive action to identify all traces of the attack and quarantine the threat, the greater impact they can have on reducing exposure and protecting the organization from compromise or breach.

To address the growing threat of morphing attacks, Cofense Vision enables a wider range of search parameters than just sender and subject. The ability to hunt for threat payload, such as attachment name, attachment hash, mime type, domain, etc., increases the effectiveness of the search action.

Searches can be performed via the Cofense Vision User Interface, the API, or through integration with Cofense Triage. The integration with Cofense Triage enables an analyst to search for all instances of emails matching a specific cluster, including those not reported by users, directly from the Triage User Interface. Once identified, threats can be quarantined with a further click within the Triage UI.



To support strict compliance requirements, all search and quarantine actions are logged for auditing within Cofense Vision and within Cofense Triage when utilized, enabling organisations to demonstrate they are meeting specific compliance obligations.

Whilst Triage and Vision can easily be used and managed by internal Security Operations teams, many organisations choose to use the services of the Cofense Phishing Defense Center. Operating 24x7 and powered by Cofense Triage, the Cofense Phishing Defense Center is staffed by experienced phishing threat analysts to handle all elements of analysis of reported emails.

Supported by the Cofense Research and Intelligence teams, the Phishing Defense Center is able to utilise as needed an array of proprietary, open source, and commercial threat analysis tools. Cofense Phishing Defense Center customers participate in a collaborative phishing defense network, enabling a global perspective of threats. This approach also gives PDC analysts the most up to date understanding of evolving phishing threat actor tactics and the techniques needed for effective analysis to capture all IOCs, even when automated approaches fail. Once threats have been identified, the PDC passes actionable intelligence to customer teams.

By utilising the Cofense Phishing Defense Center, organisations can focus their often resource-constrained Security Operations teams on activities that mitigate attacks and proactively protect the business, rather than spending time on phishing email analysis.

For further information on Cofense Triage and Phishing Defense Services, please see:

https://cofense.com/product-services/triage/
https://cofense.com/product-services/vision/
https://cofense.com/product-services/phishing-defense-services/

[1] Cofense 2019 Phishing Threat & Malware Review (https://cofense.com/phishing-threat-malware-review-2019/)

ABOUT COFENSE