Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: 16 July 2022

K. Paine
Splunk Inc.
O. Whitehouse
NCC Group
J. Sellwood
Twilio
A. Shaw
UK National Cyber Security Centre

12 January 2022

Indicators of Compromise (IoCs) and Their Role in Attack Defence

draft-paine-smart-indicators-of-compromise-04

Abstract

Cyber defenders frequently rely on Indicators of Compromise (IoCs) to identify, trace, and block malicious activity in networks or on endpoints. This draft reviews the fundamentals, opportunities, operational limitations, and best practices of IoC use. It highlights the need for IoCs to be detectable in implementations of Internet protocols, tools, and technologies - both for the IoCs' initial discovery and their use in detection - and provides a foundation for new approaches to operational challenges in network security.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This draft describes the various types of Indicator of Compromise (IoC) and how they are used effectively in attack defence (often called cyber defence). It introduces concepts such as the Pyramid of Pain [PoP] and the IoC lifecycle to highlight how IoCs may be used to provide a broad range of defences. This draft provides best practice for implementers of controls based on IoCs, as well as potential operational limitations. Two case studies which demonstrate the usefulness of IoCs for detecting and defending against real world attacks are included. One case study involves an intrusion set (a collection of indicators for a specific attack) known as APT33 and the other an attack tool called Cobalt Strike. This document is not a comprehensive report of APT33 or Cobalt Strike and is intended to be read alongside publicly published reports (referred to as open source material among intelligence practitioners) on these threats (for example, [Symantec] and [NCCGroup], respectively).

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Terminology

Attack defence: the activity of providing cyber security to an environment through the prevention, detection and response to attempted and successful cyber intrusions. Successful defence is achieved through the blocking, monitoring and response to adversarial activity at a network, endpoint or application levels.

Command and control (C2) server: an attacker-controlled server used to communicate with, send commands to and receive data from compromised machines. Communication between a C2 server and compromised hosts is called command and control traffic.

Domain Generation Algorithm (DGA): used in malware strains to generate domain names periodically. Adversaries may use DGAs to dynamically identify a destination for C2 traffic, rather than relying on a list of static IP addresses or domains that can be blocked more easily.

Kill chain: a model for conceptually breaking down a cyber intrusion to allow defenders to think about, discuss, plan for, and implement controls to defend discrete phases of an attacker's activity [KillChain].

Tactics, Techniques, and Procedures (TTPs): the way an adversary undertakes activities in the kill chain - the choices made, methods followed, tools and infrastructure used, protocols employed, and commands executed. If they are distinct enough, aspects of an attacker's TTPs can form specific Indicators of Compromise (IoCs), as if they were a fingerprint.

3. IoC Fundamentals

3.1. IoC Types and the Pyramid of Pain

Indicators of Compromise (IoCs) are observable artefacts relating to an attacker or their activities, such as their tactics, techniques, procedures, and associated tooling and infrastructure. These indicators can be observed at network or endpoint (host) levels and can, with varying degrees of confidence, help network defenders (blue teams) to pro-actively block malicious traffic or code execution, determine a cyber intrusion occurred, or associate discovered activity to a known intrusion set and thereby potentially identify additional avenues for investigation. Examples of protocol-related IoCs can include:

* IPv4 and IPv6 addresses in network traffic.

* DNS domain names in network traffic, resolver caches or logs.

* TLS Server Name Indication values in network traffic.

* Code signing certificates in binaries or TLS certificate information (such as SHA256 hashes) in network traffic.

* Cryptographic hashes (e.g. MD5, SHA1 or SHA256) of malicious binaries or scripts when calculated from network traffic or file system artefacts.

* Attack tools (such as Mimikatz [Mimikatz]) and their code structure and execution characteristics.

* Attack techniques, such as Kerberos golden tickets [GoldenTicket] which can be observed in network traffic or system artefacts.

The common types of IoC form a ’Pyramid of Pain’ [PoP] that informs prevention, detection, and mitigation strategies. Each IoC type’s place in the pyramid represents how much ’pain’ a typical adversary experiences as part of changing the activity that produces that artefact. The greater pain an adversary experiences (towards the top) the less likely they are to change those aspects of their activity and the longer the IoC is likely to reflect the attacker’s intrusion set - i.e., the less fragile those IoCs will be from a defender’s perspective. The layers of the PoP commonly range from hashes up to TTPs, with the pain ranging from simply recompiling code to creating a whole new attack strategy. Other types of IoC do exist and could be included in an extended version of the PoP should that assist the defender to understand and discuss intrusion sets most relevant to them.

```
                             /\
                            /  \                             MORE PAIN
                           /    \                           LESS FRAGILE
                          /      \                          LESS PRECISE
                         /  TTPs  \
                        /          \                            / \
                       ==============                            |
                      /              \                           |
                     /      Tools     \                          |
                    /                  \                         |
                   ======================                        |
                  /                      \                       |
                 / network/host artefacts \                      |
                /                          \                     |
               ==============================                    |
              /                              \                   |
             /          domain names          \                  |
            /                                  \                 |
           ======================================                |
          /                                      \               |
         /              IP addresses              \              |
        /                                          \            \ /
       ==============================================
      /                                              \       LESS PAIN
     /                   Hash values                  \     MORE FRAGILE
    /                                                  \    MORE PRECISE
   ======================================================
```

Figure 1

On the lowest (and least painful) level are hashes of malicious files. These are easy for a defender to gather and can be deployed to firewalls or endpoint protection to block malicious downloads or prevent code execution. While IoCs aren't the only way for defenders to do this kind of blocking, they are a quick, convenient, and unintrusive method. Hashes are precise detections for individual files based on their binary content. To subvert this defence, however, an adversary need only recompile code, or otherwise modify the file content with some trivial changes, to modify the hash value.

The next two levels are IP addresses and domain names. Interactions with these may be blocked, with varying false positive rates (misidentifying non-malicious traffic as malicious, see Section 5), and often cause more pain to an adversary to subvert than file hashes. The adversary may have to change IP ranges, find a new provider, and change their code (e.g., if the IP address is hard-coded, rather than resolved). Domain names are more specific than IP addresses (as multiple domain names may be associated with a single IP address) and are more painful for an adversary to change.

Network and endpoint artefacts, such as a malware's beaconing pattern on the network or the modified timestamps of files touched on an endpoint, are harder still to change as they relate specifically to the attack taking place and, in some cases, may not be under the direct control of the attacker. However, more sophisticated attackers use TTPs or tooling that provide flexibility at this level (such as Cobalt Strike's malleable command and control [COBALT]) or a means by which some artefacts can be masked (see [Timestomp]).

Tools and TTPs form the top two levels of the pyramid; these levels describe a threat actor's methodology - the way they perform the attack. The tools level refers specifically to the software (and less frequently hardware) used to conduct the attack, whereas the TTPs level picks up on all the other aspects of the attack strategy. IoCs at these levels are more complicated and complex - for example they can include the details of how an attacker deploys malicious code to perform reconnaissance of a victim's network, that pivots laterally to a valuable endpoint, and then downloads a ransomware payload. TTPs and tools take intensive effort to diagnose on the part of the defender, but they are fundamental to the attacker and campaign and hence incredibly painful for the adversary to change.

The variation in discoverability of IoCs is indicated by the numbers of IoCs in the open threat intelligence community Alienvault [ALIENVAULT]. As of June 2021, Alienvault contained:

* Groups (i.e., combinations of TTPs): 441

* Malware families (i.e., tools): ~24,000

* URL: 1,976,224

* Domain names: 34,959,787

* IPv4 addresses: 4,305,036

* SHA256 hash values: 4,767,891

The number of domain names appears out of sync with the other counts, which reduce on the way up the PoP. This discrepancy warrants further research; however, a contributing factor may be the fact that threat actors use domain names to masquerade as legitimate organisations and so have added incentive for creating new domain names as they are identified and confiscated.

3.2. IoC Lifecycle

To be of use to defenders, IoCs must first be discovered, assessed, shared, and deployed. When a logged activity is identified and correlated to an IoC this detection triggers a reaction by the defender which may include an investigation, potentially leading to more IoCs being discovered, assessed, shared, and deployed. This cycle continues until such time that the IoC is determined to no longer be relevant, at which point it is removed from the control space.

3.2.1. Discovery

IoCs are often discovered initially through manual investigation or automated analysis. They can be discovered in a range of sources, including in networks and at endpoints. They must either be extracted from logs monitoring protocol runs, code execution or system operations (in the case of hashes, IP addresses, domain names, and network or endpoint artefacts), or be determined through analysis of attack activity or tooling. In some cases, discovery may be a reactive process, where IoCs from past or current attacks are identified from the traces left behind. However, discovery may also result from proactive hunting for potential future IoCs extrapolated from knowledge of past events (such as from identifying attacker infrastructure by monitoring domain name registration patterns).

Crucially, for an IoC to be discovered, the indicator must be extractable from the internet protocol, tool, or technology it is associated with. Identifying a particular protocol run related to an attack is of limited benefit if indicators cannot be extracted and subsequently associated with a later related run of the same, or a different, protocol. If it is not possible to tell the source or destination of malicious attack traffic, it will not be possible to identify and block subsequent attack traffic either.

3.2.2. Assessment

Defenders may treat different IoCs differently, depending on the IoCs' quality and the defender's needs and capabilities. Defenders may, for example, place differing trust in IoCs depending on their source, freshness, confidence level, or the associated threat. These decisions rely on associated contextual information recovered at the point of discovery or provided when the IoC was shared.

An IoC without context is not much use for network defence. On the other hand, an IoC delivered with context (for example the threat actor it relates to, its role in an attack, the last time it was seen in use, its expected lifetime, or other related IoCs) allows a network defender to make an informed choice on how to use it to protect their network - for example, whether to simply log it, actively monitor it, or out-right block it.

3.2.3. Sharing

Once discovered and assessed, IoCs are most helpful when then shared at scale so many individuals and organisations can defend themselves. An IoC may be shared individually (with appropriate context) in an unstructured manner or may be packaged alongside many other IoCs in a standardised format, such as Structured Threat Information Expression [STIX], for distribution via a structured feed, such as one implementing Trusted Automated Exchange of Intelligence Information [TAXII], or through a Malware Information Sharing Platform [MISP].

While some security companies and some membership-based groups (often dubbed Information Sharing and Analysis Centres (ISACs)) provide paid intel feeds containing IoCs, there are various free IoC sources available from individual security researchers up through small trust groups to national governmental cyber security organisations and international Computer Emergency Response Teams (CERTs). Whomever they are, sharers commonly indicate the extent to which receivers may further distribute IoCs using the Traffic Light Protocol [TLP]. At its simplest, this indicates that the receiver may share with anyone (TLP WHITE), share within the defined sharing community (TLP GREEN), share within their organisation (TLP AMBER), or not share with anyone outside the original specific IoC exchange (TLP RED).

3.2.4. Deployment

For IoCs to provide defence-in-depth (see Section 6.1), which is one of their key strengths, and so cope with different points of failure, they should be deployed in controls monitoring networks and endpoints through solutions that have sufficient privilege to act on them. Wherever IoCs exist they need to be made available to security controls and associated apparatus to ensure they can be deployed quickly and widely. While IoCs may be manually assessed after discovery or receipt, significant advantage may be gained by automatically ingesting, processing, assessing, and deploying IoCs from logs or intel feeds to the appropriate security controls.

3.2.5. Detection

Security controls with deployed IoCs monitor their relevant control space and trigger a generic or specific reaction upon detection of the IoC in monitored logs.

3.2.6. Reaction

The reaction to an IoC’s detection may differ depending on factors such as the capabilities and configuration of the control it is deployed in, the assessment of the IoC, and the properties of the log source in which it was detected. For example, a connection to a known botnet C2 server may indicate a problem but does not guarantee it, particularly if the server is a compromised host still performing some other legitimate functions. Common reactions include event logging, triggering alerts, and blocking or terminating the source of the activity.

3.2.7. End of Life

How long an IoC remains useful varies and is dependent on factors including initial confidence level, fragility, and precision of the IoC (discussed further in Section 5). In some cases, IoCs may be automatically ’aged’ based on their initial characteristics and so will reach end of life at a predetermined time. In other cases, IoCs may become invalidated due to a shift in the threat actor’s TTPs (e.g., resulting from a new development or their discovery) or due to remediation action taken by a defender. End of life may also come about due to an activity unrelated to attack or defence, such as when a third-party service used by the attacker changes or goes offline. Whatever the cause, IoCs should be removed from detection at the end of their life to reduce the likelihood of false positives.

4. Using IoCs Effectively

4.1. Opportunities

IoCs offer a variety of opportunities to cyber defenders as part of a modern defence-in-depth strategy. No matter the size of an organisation, IoCs can provide an effective, scalable, and efficient defence mechanism against classes of attack from the latest threats or specific intrusion sets which may have struck in the past.

4.1.1. IoCs underpin and enable multiple layers of the modern defence-in-depth strategy

Firewalls, Intrusion Detection Systems (IDS), and Intrusion Prevention Systems (IPS) all employ IoCs to identify and mitigate threats across networks. Anti-Virus (AV) and Endpoint Detection and Response (EDR) products deploy IoCs via catalogues or libraries to all supported client endpoints. Security Incident Event Management (SIEM) platforms compare IoCs against aggregated logs from various sources - network, endpoint, and application. Of course, IoCs do not address all attack defence challenges - but they form a vital tier of any organisation's layered defence. Some types of IoC may be present across all those controls while others may be deployed only in certain layers. Further, IoCs relevant to a specific kill chain may only reflect activity performed during a certain phase and so need to be combined with other IoCs or mechanisms for complete coverage of the kill chain as part of an intrusion set.

As an example, open source malware can be deployed by many different actors, each using their own TTPs and infrastructure. However, if the actors use the same executable, the hash remains the same and this IoC can be deployed in endpoint protection to block execution regardless of individual actor, infrastructure, or other TTPs. Should this defence fail in a specific case, for example if an actor recompiles the executable binary producing a unique hash, other defences can prevent them progressing further through their attack - for instance, by blocking known malicious domain name look-ups and thereby preventing the malware calling out to its C2 infrastructure.

Alternatively, another malicious actor may regularly change their tools and infrastructure (and thus the indicator intrusion set) deployed across different campaigns, but their access vectors may remain consistent and well-known. In this case, this access TTP can be recognised and proactively defended against even while there is uncertainty of the intended subsequent activity. For example, if their access vector consistently exploits a vulnerability in software, regular and estate-wide patching can prevent the attack from taking place. Should these pre-emptive measures fail however, other IoCs observed across multiple campaigns may be able to prevent the attack at later stages in the kill chain.

4.1.2. IoCs can be used even with limited resources

IoCs are inexpensive, scalable, and easy to deploy, making their use particularly beneficial for smaller entities, especially where they are exposed to a significant threat. For example, a small manufacturing subcontractor in a supply chain producing a critical, highly specialised component may represent an attractive target because there would be disproportionate impact on both the supply chain and the prime contractor if it were compromised. It may be reasonable to assume that this small manufacturer will have only basic security (whether internal or outsourced) and while it is likely to have comparatively less resources to manage the risks it faces compared to larger partners, it can still leverage IoCs to great effect. Small entities like this can deploy IoCs to give a baseline protection against known threats without having access to a well-resourced, mature defensive team and the threat intelligence relationships necessary to perform resource-intensive investigations. One reason for this is that use of IoCs does not require the same intensive training as needed for more subjective controls, such as those based on manual analysis of tipped machine learning events. In this way, a major part of the appeal of IoCs is that they can afford some level of protection to organisations across spectrums of resource capability, maturity, and sophistication.

4.1.3. IoCs have a multiplier effect on attack defence effort

Individual IoCs can provide widespread protection that scales effectively for defenders. Within a single organisation, simply blocking one IoC may protect thousands of users and that blocking may be performed (depending on the IoC type) across multiple security controls monitoring numerous different types of activity within networks, endpoints, and applications. While discovering one IoC can be intensive, once shared via well-established routes (as discussed in Section 3.2.2) that individual IoC may, further, protect thousands of organisations and so all of their users. The prime contractor from our earlier example can supply IoCs to the small subcontractor and so further uplift that smaller entity’s defensive capability and at the same time protect itself and its interests.

Not only may multiple organisations benefit through directly receiving shared IoCs, but they may also benefit through the IoCs' application in services they utilise. In the case of an ongoing email phishing campaign, IoCs can be monitored, discovered, and deployed quickly and easily by individual organisations. However, if they are deployed quickly via a mechanism such as a protective DNS filtering service, they can be more effective still - an email campaign may be mitigated before some organisations' recipients ever click the link or before some malicious payloads can call out for instructions. Through such approaches other parties can be protected without additional effort.

4.1.4. IoCs are easily shared

There is significant benefit to be had from the sharing of IoCs and they can be easily shared for two main reasons: firstly, indicators are easy to distribute as they are textual and so in small numbers are frequently exchanged in emails, blog posts, or technical reports; secondly, standards such as MISP Core [MISPCORE], OpenIOC [OPENIOC], and STIX [STIX] provide well-defined formats for sharing large collections or regular sets of IoC along with all the associated context. Quick and easy sharing of IoCs gives blanket coverage for organisations and allows widespread mitigation in a timely fashion - they can be shared with systems administrators, from small to large organisations and from large teams to single individuals, allowing them all to implement defences on their networks.

4.1.5. IoCs can provide significant time savings

Not only are there time savings from sharing IoCs, saving duplication of investigation effort, but deploying them automatically at scale is seamless for many enterprises. Where automatic deployment of IoCs is working well, organisations and users get blanket protection with minimal human intervention and minimal effort, a key goal of attack defence. The ability to do this at scale and at pace is often vital when responding to agile threat actors that may change their intrusion set frequently and so the relevant IoCs also change. Conversely, protecting a complex network without automatic deployment of IoCs could mean manually updating every single endpoint or network device consistently and reliably to the same security state. The work this entails (including locating assets and devices, polling for logs and system information, and manually checking patch levels) introduces complexity and a need for skilled analysts and engineers. While it is still necessary to invest effort to eliminate false positives when widely deploying IoCs, the cost and effort involved can be far smaller than the work entailed in reliably manually updating all endpoint and network devices - for example, particularly on legacy systems that may be particularly complicated, or even impossible, to update.

4.1.6. IoCs allow for discovery of historic attacks

A network defender can use recently acquired IoCs in conjunction with historic data, such as logged DNS queries or email attachment hashes, to hunt for signs of past compromise. Not only can this technique help to build up a clear picture of past attacks, but it also allows for retrospective mitigation of the effects of any previous intrusion. This opportunity is reliant on historic data not having been compromised itself, by a technique such as Timestomp [Timestomp], and not being incomplete due to data retention policies, but is nonetheless valuable for detecting and remediating past attacks.

4.1.7. IoCs can be attributed to specific threats

Deployment of various modern security controls, such as firewall filtering or EDR, come with an inherent trade-off between breadth of protection and various costs, including the risk of false positives (see Section 5.2), staff time, and pure financial costs. Organisations can use threat modelling and information assurance to assess and prioritise risk from identified threats and to determine how they will mitigate or accept each of them. Contextual information tying IoCs to specific threats or actors and shared alongside the IoCs enables organisations to focus their defences against particular risks and so allows them the technical freedom and capability to choose their risk posture and defence methods. Producing this contextual information before sharing IoCs can take intensive analytical effort as well as specialist tools and training. At its simplest it can involve documenting sets of IoCs from multiple instances of the same attack campaign, say from multiple unique payloads (and therefore with distinct file hashes) from the same source and connecting to the same C2 server. A more complicated approach is to cluster similar combinations of TTPs seen across multiple campaigns over a period of time. This can be used alongside detailed malware reverse engineering and target profiling, overlaid on a geopolitical and criminal backdrop, to infer attribution to a single threat actor.

4.2. Case Studies

4.2.1. Introduction

The following two case studies illustrate how IoCs may be identified in relation to threat actor tooling (in the first) and a threat actor campaign (in the second). The case studies further highlight how these IoCs may be used by cyber defenders.

4.2.2. Cobalt Strike

Cobalt Strike [COBALT] is a commercial attack framework that consists of an implant framework (beacon), network protocol, and a C2 server. The beacon and network protocol are highly malleable, meaning the protocol representation 'on the wire' can be easily changed by an attacker to blend in with legitimate traffic. The proprietary beacon supports TLS encryption overlaid with a custom encryption scheme based on a public-private keypair. The product also supports other techniques, such as domain fronting [DFRONT], in attempt to avoid obvious passive detection by static network signatures.

4.2.2.1. Overall TTP

A beacon configuration describes how the implant should operate and communicate with its C2 server. This configuration also provides ancillary information such as the Cobalt Strike user's licence watermark.

4.2.2.2. IoCs

Tradecraft has been developed that allows the fingerprinting of C2 servers based on their responses to specific requests. This allows the servers to be identified and then their beacon configurations to be downloaded and the associated infrastructure addresses extracted as IoCs.

The resulting mass IoCs for Cobalt Strike are:

* IP addresses of the C2 servers

* domain names used

Whilst these IoCs need to be refreshed regularly (due to the ease of which they can be changed), the authors' experience of protecting public sector organisations show these IoCs are effective for disrupting threat actor operations that use Cobalt Strike.

These IoCs can be used to check historical data for evidence of past compromise, as well as deployed to detect or block future infection in a timely manner, thereby contributing to preventing the loss of user and system data.

4.2.3. APT33

In contrast to the first case study, this describes a current campaign by the threat actor APT33, also known as Elfin and Refined Kitten (see [Symantec]). APT33 has been assessed by industry to be a state-sponsored group [FireEye2], yet in this case study, IoCs still gave defenders an effective tool against such a powerful adversary. The group has been active since at least 2015 and is known to target a range of sectors including petrochemical, government, engineering, and manufacturing. Activity has been seen in countries across the globe, but predominantly in the USA and Saudi Arabia.

4.2.3.1. Overall TTP

The techniques employed by this actor exhibit a relatively low level of sophistication considering it is a state-sponsored group; typically, APT33 performs spear phishing (sending targeted malicious emails to a limited number of pre-selected recipients) with document lures that imitate legitimate publications. User interaction with these lures executes the initial payload and enables APT33 to gain initial access. Once inside a target network, APT33 attempts to pivot to other machines to gather documents and gain access to administrative credentials. In some cases, users are tricked into providing credentials that are then used with RULER, a freely available tool that allows exploitation of an email client. The attacker, in possession of a target's password, uses RULER to access the target's mail account and embeds a malicious script which will be triggered when the mail client is next opened, resulting in the execution of malicious code (often additional malware retrieved from the Internet) (see [FireEye]).

APT33 sometimes deploys a destructive tool which overwrites the master boot record (MBR) of the hard drives in as many PCs as possible. This type of tool, known as a wiper, results in data loss and renders devices unusable until the operating system is reinstalled. In some cases, the actor uses administrator credentials to invoke execution across a large swathe of a company's IT estate at once; where this isn't possible the actor may attempt to spread the wiper first manually or by using worm-like capabilities against unpatched vulnerabilities on the networked computers.

4.2.3.2. IoCs

As a result of investigations by a partnership of industry and the UK’s National Cyber Security Centre (NCSC), a set of IoCs were compiled and shared with both public and private sector organisations so network defenders could search for them in their networks. Detection of these IoCs is likely indicative of APT33 targeting and could indicate potential compromise and subsequent use of destructive malware. Network defenders could also initiate processes to block these IoCs to foil future attacks. This set of IoCs comprised:

* 9 hashes and email subject lines

* 5 IP addresses

* 7 domain names

5. Operational Limitations

The different IoC types inherently embody a set of trade-offs for defenders between the risk of false positives (misidentifying non-malicious traffic as malicious) and the risk of failing to identify attacks. The attacker’s relative pain of modifying attacks to subvert known IoCs, as discussed using the Pyramid of Pain (PoP) in Section 3.1, inversely correlates with the fragility of the IoC and with the precision with which the IoC identifies an attack. Research is needed to elucidate the exact nature of these trade-offs between pain, fragility, and precision.

5.1. Time and Effort

5.1.1. Fragility

As alluded to in Section 3.1, the Pyramid of Pain can be thought of in terms of fragility for the defender as well as pain for the attacker. The less painful it is for the attacker to change an IoC, the more fragile that IoC is as a defence tool. It is relatively simple to determine the hash value for various malicious file attachments observed as lures in a phishing campaign and to deploy these through AV or an email gateway security control. However, those hashes are fragile and can (and often will) be changed between campaigns. Malicious IP addresses and domain names can also be changed between campaigns, but this happens less frequently due to the greater pain of managing infrastructure compared to altering files, and so IP addresses and domain names provide a less fragile detection capability.

This does not mean the more fragile IoC types are worthless. Firstly, there is no guarantee a fragile IoC will change, and if a known IoC isn't changed by the attacker but wasn't blocked then the defender missed an opportunity to halt an attack in its tracks. Secondly, even within one IoC type, there is variation in the fragility depending on the context of the IoC. The file hash of a phishing lure document (with a particular theme and containing a specific staging server link) may be more fragile than the file hash of a remote access trojan payload the attacker uses after initial access. That in turn may be more fragile than the file hash of an attacker-controlled post-exploitation reconnaissance tool that doesn't connect directly to the attacker's infrastructure. Thirdly, some threats and actors are more capable or inclined to change than others, and so the fragility of an IoC for one may be very different to an IoC of the same type for another actor.

Ultimately, fragility is a defender's concern that impacts the ongoing efficacy of each IoC and will factor into decisions about end of life. However, it should not prevent adoption of individual IoCs unless there are significantly strict resource constraints that demand down-selection of IoCs for deployment. More usually, defenders researching threats will attempt to identify IoCs of varying fragilities for a particular kill chain to provide the greatest chances of ongoing detection given available investigative effort (see Section 5.1.2) and while still maintaining precision (see Section 5.2).

Finally, it is worth noting that fragility can apply to an entire class of IoCs for a range of reasons; for example, IPv4 addresses are becoming increasingly fragile due to addresses growing scarce, widespread use of cloud services, and the ease with which domain names can be moved from one hosting provider to another (thus changing IP range).

5.1.2. Discoverability

To be used in attack defence, IoCs must first be discovered through proactive hunting or reactive investigation. As noted in Section 3.1, IoCs in the tools and TTPs levels of the PoP require intensive effort and research to discover. However, it is not just an IoC's type that impacts its discoverability. The sophistication of the actor, their TTPs, and their tooling play a significant role, as does whether the IoC is retrieved from logs after the attack or extracted from samples or infected systems earlier.

For example, on an infected endpoint it may be possible to identify a malicious payload and then extract relevant IoCs, such as the file hash and its C2 server address. If the attacker used the same static payload throughout the attack this single file hash value will cover all instances. If, however, the attacker diversified their payloads, that hash can be more fragile and other hashes may need to be discovered from other samples used on other infected endpoints. Concurrently, the attacker may have simply hard-coded configuration data into the payload, in which case the C2 server address can be easy to recover. Alternatively, the address can be stored in an obfuscated persistent configuration either within the payload (e.g., within its source code or associated resource) or the infected endpoint's filesystem (e.g., using alternative data streams [ADS]) and thus requiring more effort to discover. Further, the attacker may be storing the configuration in memory only or relying on a domain generation algorithm (DGA) to generate C2 server addresses on demand. In this case, extracting the C2 server address can require a memory dump or the execution or reverse engineering of the DGA, all of which increase the effort still further.

If the malicious payload has already communicated with its C2 server, then it may be possible to discover that C2 server address IoC from network traffic logs more easily. However, once again multiple factors can make discoverability more challenging, such as the increasing adoption of HTTPS for malicious traffic - meaning C2 communications blend in with legitimate traffic, and can be complicated to identify. Further, some malwares obfuscate their intended destinations by using alternative DNS resolution services (e.g., OpenNIC [OPENNIC]) or by performing transformation operations on resolved IP addresses to determine the real C2 server address encoded in the DNS response [LAZARUS].

5.2. Precision

5.2.1. Specificity

Alongside pain and fragility, the PoP's levels can also be considered in terms of how precise the defence can be, with the false positive rate usually increasing as we move up the pyramid to less specific IoCs. A hash value identifies a particular file, such as an executable binary, and given a suitable cryptographic hash function the false positives are effectively nil; by suitable we mean one with preimage resistance and strong collision resistance. In comparison, IoCs in the upper levels (such as some network artefacts or tool fingerprints) may apply to various malicious binaries, and even benign software may share the same identifying characteristics. For example, threat actor tools making web requests may be identified by the user-agent string specified in the request header. However, this value may be the same as used by legitimate software, either by the attacker's choice or through use of a common library.

It should come as no surprise that the more specific an IoC the more fragile it is - as things change, they move outside of that specific focus. While less fragile IoCs may be desirable for their robustness and longevity, this must be balanced with the increased chance of false positives from their broadness. One way in which this balance is achieved is by grouping indicators and using them in combination. While two low-specificity IoCs for a particular attack may each have chances of false positives, when observed together they may provide greater confidence of an accurate detection of the relevant kill chain.

5.2.2. Dual and Compromised Use

As noted in Section 3.2.2, the context of an IoC, such as the way in which the attacker uses it, may equally impact the precision with which that IoC detects an attack. An IP address representing an attacker's staging server, from which their attack chain downloads subsequent payloads, offers a precise IP address for attacker-owned infrastructure. However, it will be less precise if that IP address is associated with a cloud hosting provider and it is regularly reassigned from one user to another; and it will be less precise still if the attacker compromised a legitimate web server and is abusing the IP address alongside the ongoing legitimate use.

In a similar manner, a file hash representing an attacker's custom remote access trojan will be very precise; however, a file hash representing a common enterprise remote administration tool will be less precise depending on whether the defender organisation usually uses that tool for legitimate systems administration or not. Notably, such dual use indicators are context specific both in whether they are usually used legitimately and in the way they are used in a particular circumstance. Use of the remote administration tool may be legitimate for support staff during working hours, but not generally by non-support staff, particularly if observed outside of that employee's usual working hours.

It is reasons such as these that context is so important when sharing and using IoCs.

5.3. Privacy

As noted in Section 3.2.2, context is critical to effective detection using IoCs. However, at times, defenders may feel there are privacy concerns with how much to share about a cyber intrusion, and with whom. For example, defenders may generalise the IoCs' description of the attack, by removing context to facilitate sharing. This generalisation can result in an incomplete set of IoCs being shared or IoCs being shared without clear indication of what they represent and how they are involved in an attack. The sharer will consider the privacy trade-off when generalising the IoC, and should bear in mind that the loss of context can greatly reduce the utility of the IoC for those they share with.

Self-censoring by sharers appears more prevalent and more extensive when sharing IoCs into groups with more members, into groups with a broader range of perceived member expertise (particularly the further the lower bound extends below the sharer's perceived own expertise), and into groups that do not maintain strong intermember trust. Trust within such groups appears often strongest where members: interact regularly; have common backgrounds, expertise, or challenges; conform to behavioural expectations (such as by following defined handling requirements and not misrepresenting material they share); and reciprocate the sharing and support they receive. Research opportunities exist to determine how IoC sharing groups' requirements for trust and members' interaction strategies vary and whether sharing can be optimised or incentivised, such as by using game theoretic approaches.

5.4. Automation

While IoCs can be effectively utilised by organisations of various sizes and resource constraints, as discussed in Section 4.1.2, automation of IoC ingestion, processing, assessment, and deployment is critical for managing them at scale. Manual oversight and investigation may be necessary intermittently, but a reliance on manual processing and searching only works at small scale or for occasional cases.

The adoption of automation can also enable faster and easier correlation of IoC detections across log sources, time, and space. Thereby, the response can be tailored to reflect the number and overlap of detections from a particular intrusion set, and the necessary context can be presented alongside the detection when generating any alerts for defender review. While manual processing and searching may be no less accurate (although IoC transcription errors are a common problem during busy incidents), the correlation and cross-referencing necessary to provide the same degree of situational awareness is much more time consuming.

A third important consideration when performing manual processing is the longer phase monitoring and adjustment necessary to effectively age out IoCs as they become irrelevant or, more crucially, inaccurate. Manual implementations must often simply include or exclude an IoC, as anything more granular is time consuming and complicated to manage. In contrast, automations can support a gradual reduction in confidence scoring enabling IoCs to contribute but not individually disrupt a detection as their specificity reduces.

6. Best Practice

6.1. Comprehensive Coverage and Defence-in-Depth

IoCs provide the defender with a range of options across the Pyramid of Pain's (PoP) layers, enabling them to balance precision and fragility to give high confidence detections that are practical and useful. Broad coverage of the PoP is important as it allows the defender to cycle between high precision but high fragility options and more robust but less precise indicators. As fragile indicators are changed, the more robust IoCs allow for continued detection and faster rediscovery. For this reason, it's important to collect as many IoCs as possible across the whole PoP.

At the top of the PoP, TTPs identified through anomaly detection and machine learning are more likely to have false positives, which gives lower confidence and, vitally, requires better trained analysts to understand and implement the defences. However, these are very painful for attackers to change and so when tuned appropriately provide a robust detection. Hashes, at the bottom, are precise and easy to deploy but are fragile and easily changed within and across campaigns by malicious actors.

Endpoint Detection and Response (EDR) or Anti-Virus (AV) are often the first port of call for protection from intrusion but endpoint solutions aren't a panacea. One issue is that there are many environments where it is not possible to keep them updated, or in some cases, deploy them at all. For example, the Owari botnet, a Mirai variant [Owari], exploited Internet of Things (IoT) devices where such solutions could not be deployed. It is because of such gaps, where endpoint solutions can't be relied on (see [EVOLVE]), that a defence-in-depth approach is commonly advocated, using a blended approach that includes both network and endpoint defences.

If an attack happens, then you hope an endpoint solution will pick it up. If it doesn't, it could be for many good reasons: the endpoint solution could be quite conservative and aim for a low false-positive rate; it might not have ubiquitous coverage; or it might only be able to defend the initial step of the kill chain [KillChain]. In the worst cases, the attack specifically disables the endpoint solution or the malware is brand new and so won't be recognised.

In the middle of the pyramid, IoCs related to network information (such as domains and IP addresses) can be particularly useful. They allow for broad coverage, without requiring each and every endpoint security solution to be updated, as they may be detected and enforced in a more centralised manner at network choke points (such as proxies and gateways). This makes them particular useful in contexts where ensuring endpoint security isn't possible such as "Bring Your Own Device" (BYOD), Internet of Things (IoT) and legacy environments. It's important to note that these network-level IoCs can also protect against compromised endpoints when these IoCs used to detect the attack in network traffic, even if the compromise passes unnoticed. For example, in a BYOD environment, enforcing security policies on the device can be difficult, so non-endpoint IoCs and solutions are needed to allow detection of compromise even with no endpoint coverage.

One example of how IoCs provide a layer of a defence-in-depth solution is Protective DNS (PDNS), a free and voluntary DNS filtering service provided by the UK NCSC for UK public sector organisations [PDNS]. In 2018, this service blocked access to 57.4 million DNS queries for 118,527 unique reasons (out of 68.7 billion total queries) for the organisations signed up to the service [ACD2019]. 28 million of them were for domain generation algorithms (DGAs) [DGAs], including 15 known DGAs which are a type of TTP.

IoCs such as malicious domains can be put on PDNS straight away and can then be used to prevent access to those known malicious domains across the entire estate of over 460 separate public sector entities that use NCSC's PDNS [Annual2019]. Coverage can be patchy with endpoints, as the roll-out of protections isn't uniform or necessarily fast - but if the IoC is on PDNS, a consistent defence is maintained. This offers protection, regardless of whether the context is a BYOD environment or a managed enterprise system. Other IoCs, like Server Name Indicator values in TLS or the server certificate information, also provide IoC protections.

Similar to the AV scenario, large scale services face risk decisions around balancing threat against business impact from false positives. Organisations need to be able to retain the ability to be more conservative with their own defences, while still benefiting from them. For instance, a commercial DNS filtering service is intended for broad deployment, so will have a risk tolerance similar to AV products; whereas DNS filtering intended for government users (e.g. PDNS) can be more conservative, but will still have a relatively broad deployment if intended for the whole of government. A government department or specific company, on the other hand, might accept the risk of disruption and arrange firewalls or other network protection devices to completely block anything related to particular threats, regardless of the confidence, but rely on a DNS filtering service for everything else.

Other network defences can make use of this blanket coverage from IoCs, like middlebox mitigation, proxy defences, and application layer firewalls, but are out of scope for this draft. Note too that DNS goes through firewalls, proxies and possibly to a DNS filtering service; it doesn't have to be unencrypted, but these appliances must be able to decrypt it to do anything useful with it, like blocking queries for known bad URIs.

Covering a broad range of IoCs gives defenders a wide range of benefits: they are easy to deploy; they provide a high enough confidence to be effective; at least some will be painful for attackers to change; their distribution around the infrastructure allows for different points of failure, and so overall they enable the defenders to disrupt bad actors. The combination of these factors cements IoCs as a particularly valuable tool for defenders with limited resources.

6.2. Security Considerations

This draft is all about system security. However, when poorly deployed, IoCs can lead to over-blocking which may present an availability concern for some systems. While IoCs preserve privacy on a macro scale (by preventing data breaches), research could be done to investigate the impact on privacy from sharing IoCs, and improvements could be made to minimise any impact found. The creation of a privacy-preserving IoC sharing method, that still allows both network and endpoint defences to provide security and layered defences, would be an interesting proposal.

7. Conclusions

IoCs are versatile and powerful. IoCs underpin and enable multiple layers of the modern defence-in-depth strategy. IoCs are easy to share, providing a multiplier effect on attack defence effort and they save vital time. Network-level IoCs offer protection, especially valuable when an endpoint-only solution isn't sufficient. These properties, along with their ease of use, make IoCs a key component of any attack defence strategy and particularly valuable for defenders with limited resources.

For IoCs to be useful, they don’t have to be unencrypted or visible in networks - but crucially they do need to be made available, along with their context, to entities that need them. It is also important that this availability and eventual usage copes with multiple points of failure, as per the defence-in-depth strategy, of which IoCs are a key part.

8. IANA Considerations

This draft does not require any IANA action.

9. Acknowledgements

Thanks to all those who have been involved with improving cyber defence in the IETF and IRTF communities.

10. Informative References

[ACD2019] Levy, I. and M. S, "Active Cyber Defence - The Second Year", 2019, <https://www.ncsc.gov.uk/report/active-cyber-defence-report-2019>.

[ADS] Microsoft, "File Streams (Local File Systems)", 2018, <https://docs.microsoft.com/en-us/windows/win32/fileio/file-streams>.

[ALIENVAULT] AlienVault, "AlienVault", 2021, <https://otx.alienvault.com/>.

[Annual2019] NCSC, "Annual Review 2019", 2019, <https://www.ncsc.gov.uk/annual-review/2019/ncsc/docs/ncsc_2019-annual-review.pdf>.

[COBALT] Cobalt Strike, "OVERRULED: Containing a Potentially Destructive Adversary", 2021, <https://www.cobaltstrike.com/>.

[DFRONT] InfoSec Resources, "Domain Fronting", 2017, <https://resources.infosecinstitute.com/topic/domain-fronting/>.

[DGAs] MITRE, "Dynamic Resolution: Domain Generation Algorithms", 2020, <https://attack.mitre.org/techniques/T1483/>.

[EVOLVE] McFadden, M., "Evolution of Endpoint Security - An Operational Perspective", 2021, <https://datatracker.ietf.org/doc/draft-mcfadden-opsec-endp-evolve/>.

[FireEye] O'Leary, J., Kimble, J., Vanderlee, K., and N. Fraser, "Insights into Iranian Cyber Espionage: APT33 Targets Aerospace and Energy Sectors and has Ties to Destructive Malware", 2017, <https://www.fireeye.com/blog/threat-research/2017/09/apt33-insights-into-iranian-cyber-espionage.html>.

[FireEye2] FireEye, "OVERRULED: Containing a Potentially Destructive Adversary", 2018, <https://www.fireeye.com/blog/threat-research/2018/12/overruled-containing-a-potentially-destructive-adversary.html>.

[GoldenTicket] Soria-Machado, M., Abolins, D., Boldea, C., and K. Socha, "Kerberos Golden Ticket Protection", 2014, <https://cert.europa.eu/static/WhitePapers/UPDATED - CERT-EU_Security_Whitepaper_2014-007_Kerberos_Golden_Ticket_Protection_v1_4.pdf>.

[KillChain] Lockheed Martin, "The Cyber Kill Chain", 2020, <https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html>.

[LAZARUS] Kaspersky Lab, "Lazarus Under The Hood", 2018, <https://media.kasperskycontenthub.com/wp-content/uploads/sites/43/2018/03/07180244/Lazarus_Under_The_Hood_PDF_final.pdf>.

[Mimikatz] Mulder, J., "Mimikatz Overview, Defenses and Detection", 2016, <https://www.sans.org/reading-room/whitepapers/detection/mimikatz-overview-defenses-detection-36780>.

[MISP] MISP, "MISP", 2019, <https://www.misp-project.org/>.

[MISPCORE] MISP, "MISP Core", 2020, <https://github.com/MISP/misp-rfc/blob/master/misp-core-format/raw.md.txt>.

[NCCGroup] Jansen, W., "Abusing cloud services to fly under the radar", 2021, <https://research.nccgroup.com/2021/01/12/abusing-cloud-services-to-fly-under-the-radar/>.

[OPENIOC] Gibb, W., "OpenIOC: Back to the Basics", 2013, <https://www.fireeye.com/blog/threat-research/2013/10/openioc-basics.html>.

[OPENNIC] OpenNIC Project, "OpenNIC Project", 2021, <https://www.opennic.org/>.

[Owari] NCSC, "Owari botnet own-goal takeover", 2018, <https://www.ncsc.gov.uk/report/weekly-threat-report-8th-june-2018>.

[PDNS] NCSC, "Protective DNS", 2019, <https://www.ncsc.gov.uk/information/pdns>.

[PoP] Bianco, D.J., "The Pyramid of Pain", 2014, <https://detect-respond.blogspot.com/2013/03/the-pyramid-of-pain.html>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[STIX] OASIS Cyber Threat Intelligence, "STIX", 2019, <https://oasis-open.github.io/cti-documentation/stix/intro>.

[Symantec] Symantec, "Elfin: Relentless", 2019, <https://www.symantec.com/blogs/threat-intelligence/elfin-apt33-espionage>.

[TAXII] OASIS Cyber Threat Intelligence, "TAXII", 2021, <https://oasis-open.github.io/cti-documentation/taxii/intro.html>.

[Timestomp] OASIS Cyber Threat Intelligence, "Timestomp", 2019, <https://attack.mitre.org/techniques/T1099/>.

[TLP] FIRST, "Traffic Light Protocol", 2021, <https://www.first.org/tlp/>.

Authors' Addresses

Kirsty Paine
Splunk Inc.

Email: kirsty.ietf@gmail.com

Ollie Whitehouse
NCC Group

Email: ollie.whitehouse@nccgroup.com

James Sellwood
Twilio

Email: jsellwood@twilio.com

Andrew Shaw
UK National Cyber Security Centre

Email: andrew.s2@ncsc.gov.uk