Global Information Assurance Certification Paper

Automated Network Defense through Threat Intelligence and Knowledge Management

GIAC (GCIA) Gold Certification

Author: Christopher O'Brien, cobrien@cert.gov.uk
Advisor: Hamed Khiabani, Ph.D.

Accepted: 31 December 2015

Abstract

In order for threat intelligence to be leveraged effectively within organizations it must be well structured and enriched. These tasks are relatively trivial to accomplish but tend to be conducted ad-hoc or, worse still, manually. This paper sets out to demonstrate that automation and appropriate structuring of threat intelligence need not be prescriptive nor expensive. With current open source tools and languages an organization can dramatically increase the value of their threat intelligence through automated enrichment and fusion between the tactical and strategic sources, providing tangible network detection capability whilst answering broader contextual questions. This is achieved by using bulk data feed resource and open source reporting, processed using the open architecture STIX threat intelligence structures and deployed to test mechanisms using Bro and SNORT. The outcome is a network alerting capability that not only tells us when something that we already know has happened, but potentially who did it as well.

1. Introduction

Many organizations know that they should have cyber security threat intelligence, fewer know how to use it and fewer still are actually doing so. By far the fewest of these groups are those taking full advantage of the true nature of intelligence: "…acquisition and analysis of information to assess capabilities, intent and opportunities for exploitation by leaders at all levels" (UK Ministry of Defence, 2014). One of the most important aspects of intelligence is that it is actionable 'at all levels', so it is important to identify a technology and process for understanding and analyzing information to draw clarity on threats. The types of data analyzed to produce this intelligence in the field of cyber security has the advantage of being eminently structural and exploitable which presents great opportunity for automation.

Some data sets are more exploitable than others and there is a natural distinction between strategic and tactical intelligence (Poputa-Clean, 2015). Tactical intelligence is more quantifiable and, as such, is the first target of network defenders to fill their threat libraries. The assumption is that strategic intelligence – more qualitative data such as Threat Actors, and Tactics Techniques and Procedures (TTPs) – is less structured and, hence, less useful in tactical network defense. The separation of these data types can often lead to frustration for the network defender as the context provided by strategic data can greatly affect the deployment of tactical indicators – which must be the ultimate aim of this data. For example, observing a machine on your network beaconing with Dridex to the IP address '185.38.104.108' is a valuable piece of tactical intelligence that may trigger an incident response and deployment of network defenses. However, it gains even more value when coupled with the strategic intelligence that, as of writing, this IP address resolves to the domain 'sinkhole.cert.gov.uk'. Knowing that the infected machine appears to be connecting to a 'whitehat' sinkhole can allow for more accurate prioritization of incident response.

More recent developments in structured threat intelligence languages provide a means of rendering exploitable more qualitative data in order to maintain the additional context around the indicator. This logical ontology allows the indicator to hold an

inherent logic as part of the intelligence object which is better interpreted by automated systems allowing for faster, more efficient network defense deployments. Thus, this paper provides a hands-on method of automating network defenses through a structured threat intelligence language to deliver tactical intelligence enriched with strategic context.

1.1. Establishing a Common Language

The names we associate with specific cyber security threats vary greatly across the community. For example, the group commonly referred to as 'Havex' goes by many different aliases as addressed in Recorded Future's helpful summary (Recorded Future, 2014). In this article the author states that "As many as five different codewords have been given to…[these] cyber campaigns…" and they are often used inappropriately when attempting to classify observed cyber attacks. This poses difficulties at both the tactical and strategic levels; at the strategic level an incorrect reference to a TTP can lead to false assumptions on an incident's impact, thus skewing risk decisions. For example using the common name of 'Havex', used regularly to refer to a Threat Actor group, is more accurately a reference to the Havex Remote Access Trojan (RAT) (F-Secure, 2014). This tool has very specific functions and capabilities that are not necessarily reflective of other tools believed to be leveraged by this actor. These data conflicts made at the strategic level can reach to the tactical as a decision may be made to defend a network from 'Havex' (deploying signatures, blacklisting infrastructure, etc.) without taking in to account other vectors to which an organization may be vulnerable. In order to reconcile those conflicts it is advantageous to develop a common language for expressing threats which distinguishes between the TTPs being leveraged and the Threat Actors wielding them. To do so in a machine readable way is the ultimate goal.

However, this does not necessarily mean that everyone should use a single taxonomy. Organizations will naturally have different ways of interpreting threat intelligence to make it usable for their own 'levels' of response. Those interpretations will inherently hold additional value as they are bred from an organization's unique viewpoint of the threat with different analytical assertions. Provided we have a means of provenance and a method of recording the results of when those assertions are tested,

corroborated or refuted, the ability to render those findings to the wider community is advantageous.

Furthermore, malicious actors often deliberately attempt to conceal their identity and confuse the network defender to make the task of detecting them harder and the process of evasion easier. Hence, the notion of spending time and resource as a community to develop a single 'source of truth' taxonomy for all strategic intelligence structures is not only costly as we attempt to untangle a deliberately complex issue, but also a welcome result for the attacker as we spend more time worrying what we name something and less on actually defending against it.

In order to allow network defenders to manage their own strategic and tactical intelligence in accordance with their own internal tools and processes in a recognizably structured way, without the need for prescriptive 'single truth' definitions which are costly to develop and can limit the scope of community knowledge, suggests that the acceptance of an international 'standard' for the structuring of that threat intelligence be developed.

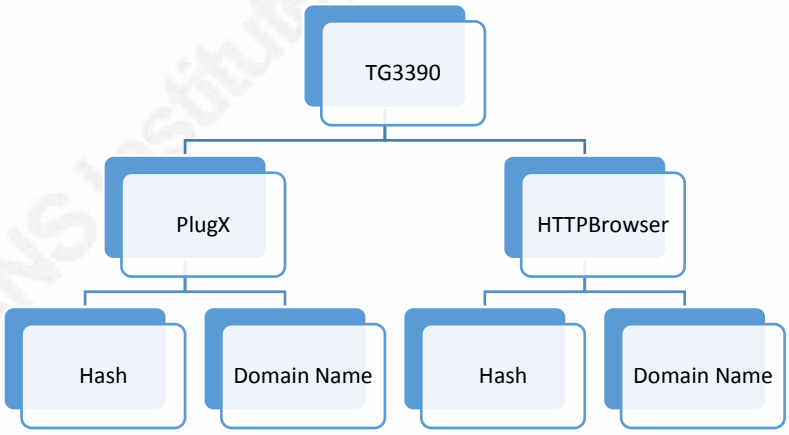


Figure 1. Diagram of a threat intelligence hierarchy. The structure allows for uniform data storage but lacks flexibility.

Hierarchies are naturally two-dimensional – as show in Figure 1 – and struggle with increased complexity. In the case of cyber security threat intelligence we will regularly see instances that attempt to break these models. As discussed previously, Havex is a RAT: a TTP leveraged by many Threat Actors. Whilst this specific TTP is commonly associated with a specific Threat Actor there is nothing stopping it being used

by another. As such, we are forced to investigate more multi-dimensional classification systems. Faceted Classification (Denton, 2009) introduces a means of sorting objects by facets, or features, of the object. For example, PlugX is a RAT and so is Havex. Instead of hierarchically linking PlugX as a subset of the TG3390 intrusion set (Dell SecureWorks, 2015) the facet of 'RAT' is applied to both. The data owner can then assert the Threat Actor groups of TG3390 and Energetic Bear as having a facet of 'Uses RATs in their attacks' which would lead to a common logic by which to assert a link. In this example we observe that a potentially incorrect assertion can be made that PlugX is used by Energetic Bear (and, indeed, that Havex is used by TG3390). This model recognizes that those assertions *could exist*, but it becomes clear that another object is required to link that relationship. In a cyber security context we would establish a link between an actor and their tools by observation of their usage – so we introduce the concept of an Incident or Indicator that demonstrates a sighting of the TTP being leveraged, linking it to a Threat Actor as an analytical assertion. This is illustrated in Figure 2. Facets can now become the primary means of querying the data set as they can be cut in multiple directions and in an automated way. This can be augmented with the idea of grouping facets into logical contexts (Messina, 2015) such as TTP, Threat Actor and Indicator. Enumerations are not necessarily fixed (to maintain flexibility) but provide a framework for a common language.

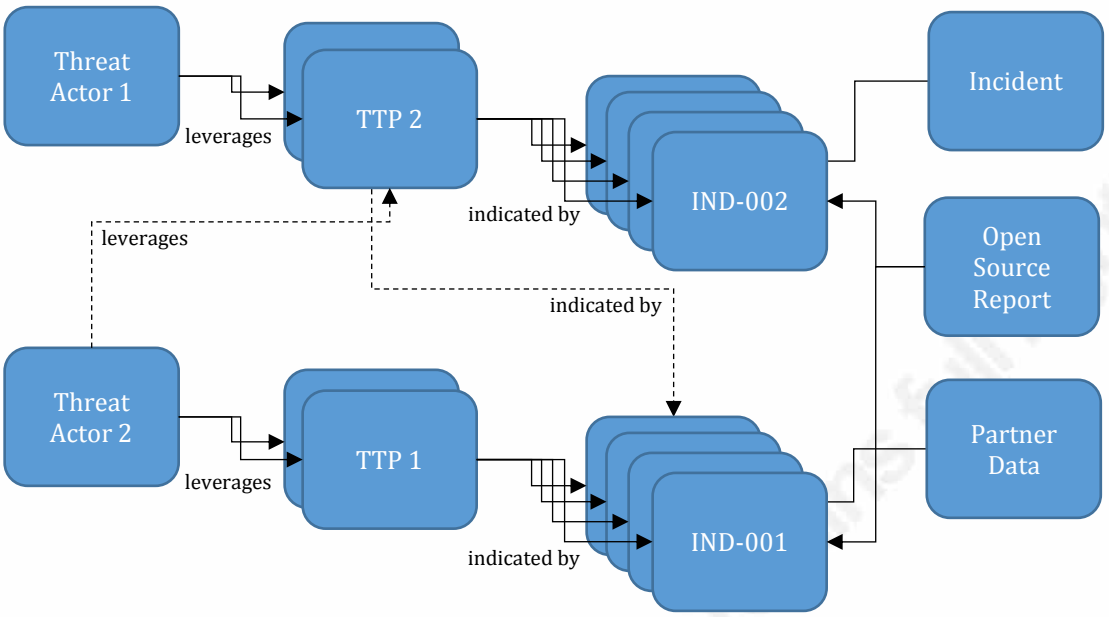


Figure 2. Diagram to show an object oriented framework using Faceted Classification. Dash-line arrows represent analytical assertions through cross-referencing.

Figure 2 illustrates what this logical, grouped-facet structure might look for the same data set. Several threat intelligence languages exist to share atomic data between organizations (Farnham, 2013) but few support the level of granularity required to develop such complex models. The OASIS Structured Threat Intelligence eXpression (STIX) language and associated supporting languages of Cyber Observable eXpression (CYBOX) and Trusted Automated eXchange of Indicator Information (TAXII) is one set of standards that allows for the depth of complexity to support such structures and also benefits from being an open standard, maintained and managed by an OASIS Technical Committee (OASIS, n.d.), to which other languages can refer. These higher-order languages not only make use of Faceted Classification but also group facets into logical cyber security incident object types to allow for more accurate classification of threat intelligence. Throughout this paper we will refer to the STIX family of languages as a means to demonstrate how complex strategic and tactical data sets can be rendered exploitable, shared and automatically deployed to increase network security at machine speeds.

2. Automated Deployment of Threat Intelligence

The following experiment demonstrates the value in automating the process of ingestion, exploitation, enrichment and deployment of threat intelligence. Through the use of open source data sets and tooling we aim to demonstrate the benefits of enrichment and the increased capability and the further enhancements offered by performing enrichment on structured data, allowing for the fusion of tactical and strategic intelligence. We focus on two common examples of input data: Open Source reporting, using 'Threat Group-3390 Targets Organizations for Cyberespionage' (Dell SecureWorks, 2015) as a use case; and bulk data feeds, in this case an aggregated data feed from AbuseSA (Codenomicon, n.d.). Whilst AbuseSA is a paid-for aggregation service, the process of data aggregation can be conducted with open source aggregation tools such as combine.py (MLSec Project, 2015).

Throughout the experiment the aim is to use low cost, open source and open architecture techniques to derive as much value from these feeds as possible for the purposes of automated network defense. This includes the deployment of indicators to a simulated network environment and sharing of the exploited and enriched data with other network defenders through automated threat intelligence sharing. The actual deployment of these assets to a network infrastructure greatly depends on the local requirements and the concept of network layout to support deployment of threat intelligence is addressed well in Poputa-Clean's paper (Poputa-Clean, 2015). In this experiment we focus on a generalized data flow process as illustrated in Figure 3 and implement the data sharing using the STIX/CYBOX/TAXII implementations of Soltra Edge (Soltra Edge, n.d.) which makes use of the OASIS open standards for Cyber Threat Intelligence (CTI) sharing.

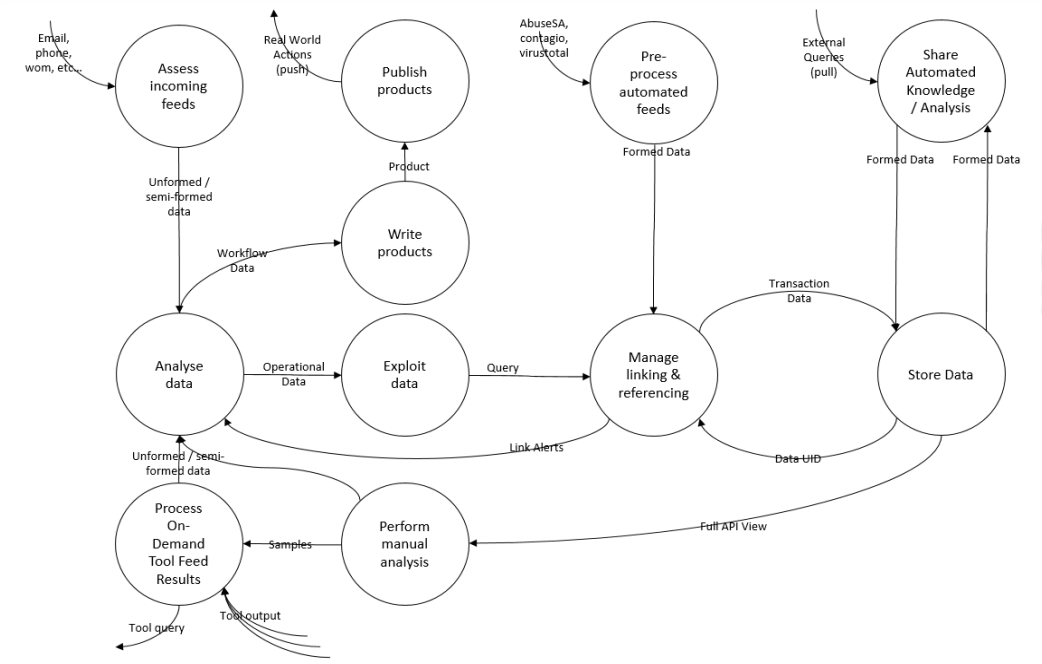


Figure 3. Non-tool specific data flow diagram for implementation of manual and automated data processing to a CTI knowledge base.

The experiment includes the development of Python scripting and the publically available 'python-stix' libraries (python-stix, n.d.) to ingest, exploit, enrich and deploy threat intelligence to common network defense tools. The custom scripts produced in the

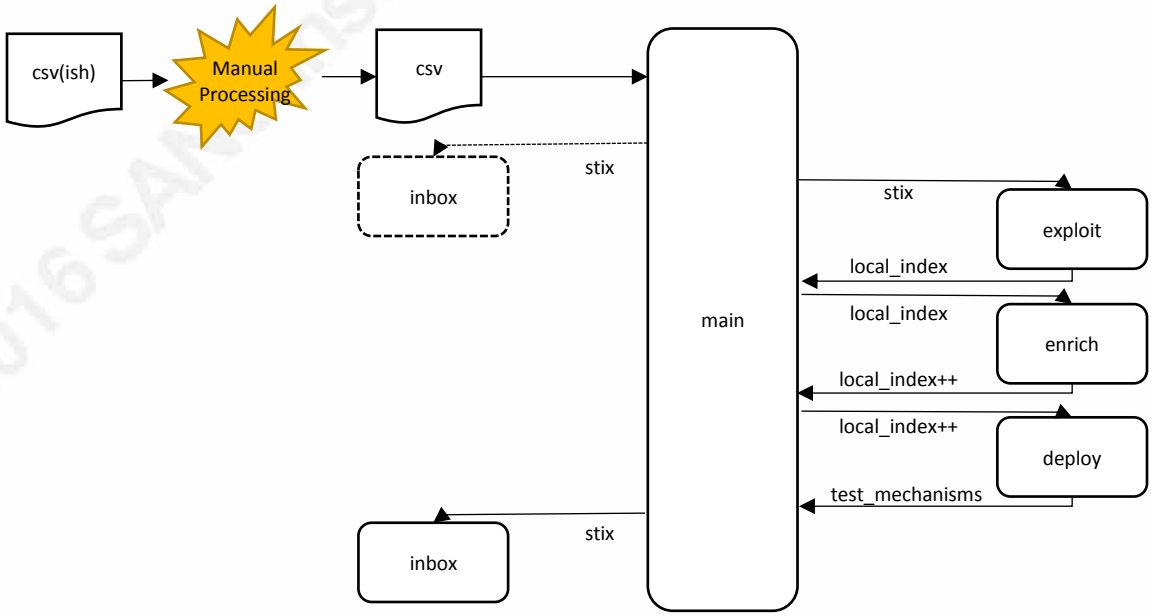


Figure 4. Data flow diagram for TG3390 ingest, exploit, enrich and deploy functions. This variant makes use of a non-STIX custom index for ease of enrichment.

Christopher O'Brien, cobrien@cert.gov.uk

experiment are now available at https://github.com/cobsec/pickup-stix.

Figure 4 illustrates the data flow for the custom script which was written for this experiment. The following custom functions were created:

inbox.py – using the original example produced by Soltra, the inbox script accepts a STIX object, expresses it in XML format and wraps it in XML formatted TAXII messaging protocols, allowing it to be sent or 'inboxed' to a specific TAXII server

exploit.py – this module contains data transforms into and out of STIX. The purpose of this module was to demonstrate the capability in python-stix to easily transform between data specifications. This is of particular importance when working with existing tools that use specific, sometimes proprietary, data protocols that are not automatically compatible with STIX. In this experiment we convert into / out of STIX to use a custom 'local index' using a Python dictionary as a means of testing cross-compatibility between STIX and other data structures. The exploitation of the data also establishes the implied links between tactical and strategic data by making use of the detailed STIX CTI structure.

enrich.py – once ingested, the structured tactical threat intelligence data is then enriched using two common enrichment sources: VirusTotal (VirusTotal, n.d.) for file hash and behavioral enrichment, and domaintools (domaintools, n.d.) for network detail pivoting. Both are accessed by means of API calls through paid-for services, though similar free services may also be used.

deploy.py – the enriched tactical data is then deployed to network defense mechanisms in order to assess the validity of the new data. In this experiment we use the SNORT Intrusion Detection System (SNORT, n.d.) and Bro Network Security Monitor (Bro, 2015). By combining the signature detection capabilities of SNORT with the more heuristic approach of Bro scripting we are able to demonstrate that tactical intelligence can quickly be related to strategic context for greater situational awareness.

config.py / config.ini – makes use of the Python library 'ConfigParser' to collate basic information required by the various modules in order to successfully run the experiment. Settings can be changed in config.ini in order to customize the process including the

specification of the inbox TAXII server, server login credentials and enrichment API keys.

2.1. Ingesting Feeds

The first step is to capture the data through a pre-processer. This often overlooked part of the process is arguably the most crucial as the tactical intelligence chosen to be retained will determine the granularity and accuracy of the strategic context. It is important to note that cross-referencing the incoming data is difficult at this stage as the ingest pre-processor will not have full visibility of the data set. High processor capability ingest methods may be able to actively query a database and remove incoming entries that may, for example, be duplicated in other feeds – however, this is more logically done after the data has been ingested in to the knowledge base so that Map-Reduce algorithms can be run on the back end. The process of enrichment is covered in greater detail in section 2.3.

2.1.1. Bulk Ingest

Bulk data feeds have the advantage of coming in some form of structured data format. Whilst they may not hold full granularity, there is often additional strategic intelligence to be gleaned from the context through which the data was acquired.

Figure 5 illustrates a flat JSON structure that can be readily manipulated through code or script. Part of the additional context is that this particular event was captured through a legitimate sinkhole activity. This gives us some key analytical assertions:

1. The destination IP of 204.95.99.204 is not necessarily ‘attacker infrastructure’
2. The source IP (as redacted), while likely infected by the malware, is not necessarily actively being ‘exploited’ in the Cyber Kill Chain sense (Hutchins, Cloppert, & Amin, 2010)
3. There is enough data to deduce that a complete Network Connection was established (source IP/port to destination IP/port)

4. As this event 'has occurred' sometime in the past, it can be considered an


```
{
        "feed":                 ["<redacted>"],
        "source time":          ["2015-09-30 23:59:54Z"],
        "cc":                   ["GB"],
        "ip":                   ["<redacted>"],
        "bgp prefix allocated":         ["2001-04-02"],
        "source asn":           ["<redacted>"],
        "http method":          ["post"],
        "source longitude":     ["<redacted>"],
        "source metro code":    ["0"],
        "http version":         ["1.1"],
        "destination ip":       ["204.95.99.204"],
        "malware":              ["caphaw"],
        "uuid":                 ["0f8e1af0-fd04-4a08-98c1-85493ee7d888"],
        "risk level estimate": ["high"],
        "source area code":     ["0"],
        "source city":          ["<redacted>"],
        "source cc":            ["gb"],
        "as allocated":         ["2013-08-08"],
        "latitude":             ["51.5"],
        "user agent":           ["mozilla/4.0 (compatible; msie 6.0; windows nt 5.1;
sv1; .net clr 1.0.4419)"],
        "information source":   ["sinkholemessage"],
        "source region":        ["<redacted>"],
        "source postal code":   ["<redacted>"],
        "description":          ["Indicates a Caphaw infection, where victim PC is
making an attempt to download a configuration file."],
        "geoip cc":             ["GB"],
        "http request":         ["/index.php"],
        "destination port":     ["443"],
        "source latitude":      ["<redacted>"],
        "http referrer":        ["0"],
        "registry":             ["ripencc"],
        "feeder":               ["<redacted>"],
        "asn":                  ["60339"],
        "type":                 ["botnet drone"],
        "observation time":     ["2015-10-01 00:07:04Z"],
        "longitude":            ["-0.13"],
        "source port":          ["54506"],
        "as name":              ["H3GUK Hutchison 3G UK Limited,GB"],
        "bgp prefix":           ["<redacted>"],
        "additional information": ["<redacted>"]
}
```


Figure 5. Redacted sample from the AbuseSA aggregator by Codenomicon. The flat structure does not lend itself to the structured language needed for granular threat intelligence, but knowledge of the context allows for enrichment.

Incident (Mitre Corporation, 2015)

The event by itself can be utilized straight away for network defense – for example, to search for the destination IP in enterprise traffic logs in an attempt to identify infected hosts. However, the event line by itself does not contain the above analytical assertions. If they are not captured somehow then the corporate knowledge of that event will be lost, along with its value.

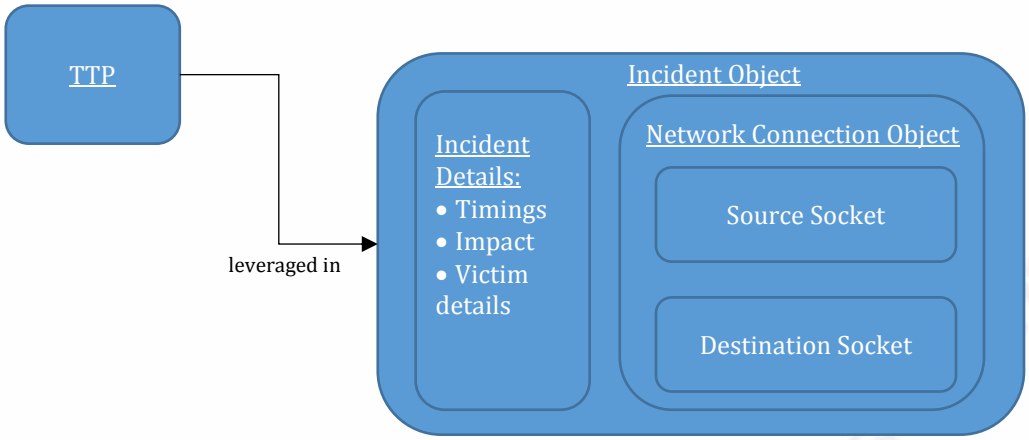


Figure 6. Single event from a bulk ingest feed. In this case there is enough data to form full 'socket' objects (IP/port pairs) which is a more accurate context than simply listing the IPs/ports. The Incident 'wrapper' allows for more detail including TTP.

The example in Figure 6 processes these types of events in to STIX objects which allows for that additional context to be captured. A threshold can also be applied at this stage for varying data qualities that will only create new objects if enough data exists to successfully make analytical assertions. This could easily be changed to make different assertions and produce different object models to represent them. For example if the data is lacking either end of the network connection (source or destination socket) the script can be adjusted to create a simple Indicator list. However, since the raw data has complete context it is important to capture as much of it as possible in the STIX object in order to increase data granularity. Such logic should be customized for the organizational intelligence requirements.

2.1.2. Open Source Reporting

Open source reports revealing indicators associated with so called Advanced Persistent Threats (APT) are somewhat of a trend of late. Unfortunately, most of these reports are either PDFs or online website-based reports which are not necessarily conveniently structured. In some cases even two reports from the same company may differ in how they structure their indicators. However, most reports will keep indicators in separate sections of a report at least making them available to be exploited through scripting as shown in Figure 7.

```
<TABLE class=tabularr cellSpacing=0 cellPadding=0 border=0><TBODY>
<TR>
<TH>Indicator</TH>
<TH>Type</TH>
<TH>Context</TH></TR>
<TR>
<TD>american.blackcmd.com</TD>
<TD>Domain name</TD>
<TD>TG-3390 infrastructure<BR>Confidence: High</TD></TR>
<TR>
<TD>api.apigmail.com</TD>
<TD>Domain name</TD>
<TD>TG-3390 infrastructure<BR>Confidence: High</TD></TR>
<TR>
<TD>apigmail.com</TD>
<TD>Domain name</TD>
<TD>TG-3390 infrastructure <BR>Confidence: High</TD></TR>
<TR>
<TD>backup.darkhero.org</TD>
<TD>Domain name</TD>
<TD>TG-3390 infrastructure <BR>Confidence: High</TD></TR>
<TR>
<TD>bel.updatawindows.com</TD>
<TD>Domain name</TD>
<TD>TG-3390 infrastructure <BR>Confidence: High</TD></TR>
<TR>
<TD>binary.update-onlines.org</TD>
<TD>Domain name</TD>
<TD>TG-3390 infrastructure <BR>Confidence: High</TD></TR>
<TR>
<TD>blackcmd.com</TD>
<TD>Domain name</TD>
<TD>TG-3390 infrastructure <BR>Confidence: High</TD></TR>
<TR>
<TD>castle.blackcmd.com</TD>
<TD>Domain name</TD>
<TD>TG-3390 infrastructure<BR>Confidence: High</TD></TR>
<TR>
<TD>ctcb.blackcmd.com</TD>
<TD>Domain name</TD>
<TD>TG-3390 infrastructure<BR>Confidence: High</TD></TR>
<TR>
<TD>darkhero.org</TD>
<TD>Domain name</TD>
```

Figure 7. Source code from the TG3390 online report. Whilst eminently scriptable due to the structured HTML table, a level of input is required from the analyst to ensure that the correct data is extracted, and in context.

This does, of course, require an analyst to read the report and identify the useful information, turning that logic in to a script which may not ever get used again. Whilst there are some impressive tools on the market to support with entity extraction and report markup (Bridges, Jones, Iannacone, Testa, & Goodall, 2014), they share a common requirement for training and/or human monitoring to achieve accurate results. The time invested in developing these sorts of capabilities may be better spent developing modular ingest scripts that can be applied to different report structures. The lack of repetition of structure across the industry is regrettable and it should be the aim of all publishers of cyber security threat intelligence to do so in the most predictable and repeatable structure possible whilst still retaining context. These data structures can always be rendered in to more aesthetically pleasing reports provided the data is structured well.

Even though a structured threat intelligence language adds value, it is sometimes preferably to use an alternate data structure for ease of compatibility with other tools. This concept is well explored in ‘Automated Defense - Using Threat Intelligence to

Augment Security' (Poputa-Clean, 2015) using the Nyx framework. For this experiment a simple Python dictionary was created from the more detailed threat intelligence data (STIX objects) to allow for rapid enrichment processing with other APIs. It is important to maintain the STIX generated UUID references with each Observable as it is enriched to ensure that future relational links can be established using the '`id_`' facet of the STIX objects. For example, the reference to observable IP Address in the TG3390 report of 72.11.141.133 was ingested through the Python script and assigned a randomly generated UUID of 719b088c-e2cc-43fa-9bd3-f34264f7e57f. This is then prepended with the object type name defined by the python-stix library (in this case an "Address" type) and a user-defined namespace (in this case, "certuk") resulting in the fully qualified UUID of certuk:Address-719b088c-e2cc-43fa-9bd3-f34264f7e57f. This object can be seen in Annex A on lines 91-99. During the experiment we create an 'enrichment' STIX file – a sample of which is at Annex B. This IP address was used during enrichment to identify a domain name, ECWARD.COM, which was added to this enrichment STIX file and then back-referenced to the original IP address at line 110 of Annex B.

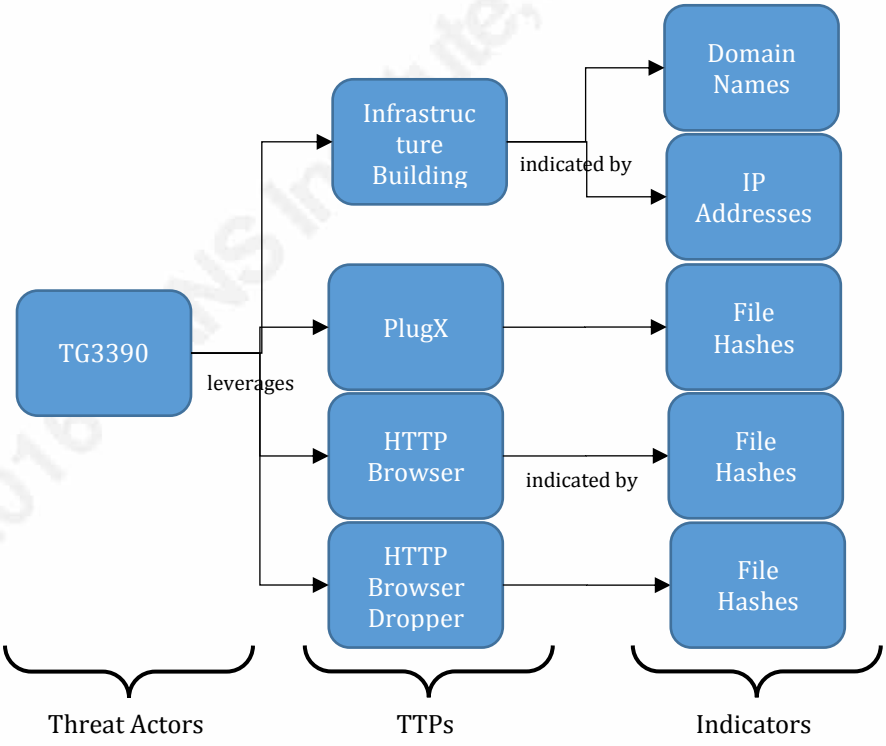


Figure 8. Diagram to show the resultant structure of a STIX implementation of the TG3390 report.

2.2. Data Structures

Once data is in a structure that accurately reflects both its tactical and strategic intelligence values, it then needs to be collated into a threat library or knowledge base for storage and future querying. Using structured languages such as STIX, where objects already exist with relationships between each other, means that more scalable non-relational databases including document stores such as mongoDB (mongoDB, n.d.), or graph stores such as Sqrrl (Sqrrl, n.d.). However, every deployment of a threat intelligence enabled network will be different and some network defenders may continue to find optimizations using other techniques such as traditional relational databases.

The point of using an open standard is that network owners should not be forced to be STIX compliant 'from the ground up'. Abstraction layers and data transforms are likely to be commonplace, particularly for well-established enterprises with pre-built data protocols. This may result in a loss of granularity as data is passed between languages, but the hope is that the open standard definition can be flexible enough that losses are minimalized. In this experiment we assume minimal dependence on existing languages and tools. Hence, we are able to take full advantage of the open architecture's full granularity.

Figure 9 shows an implementation of a STIX data model taking advantage of the inter-object relationships in order to minimize duplication of Observables. The aim is to hold all Observables in the knowledge base once and, when required as part of another STIX package, to reference it rather than create a duplicate copy. This provides more efficient data storage and a logical data relationship model which can be better exploited through automated analysis techniques. For example, we could consider the concept of 'quality' for a given Observable as being directly proportional to the number of relationships established between the Observable and other STIX objects. This makes the Observable more likely to provide pivoting capability and so, in analytical terms, may be considered a high quality Observable.

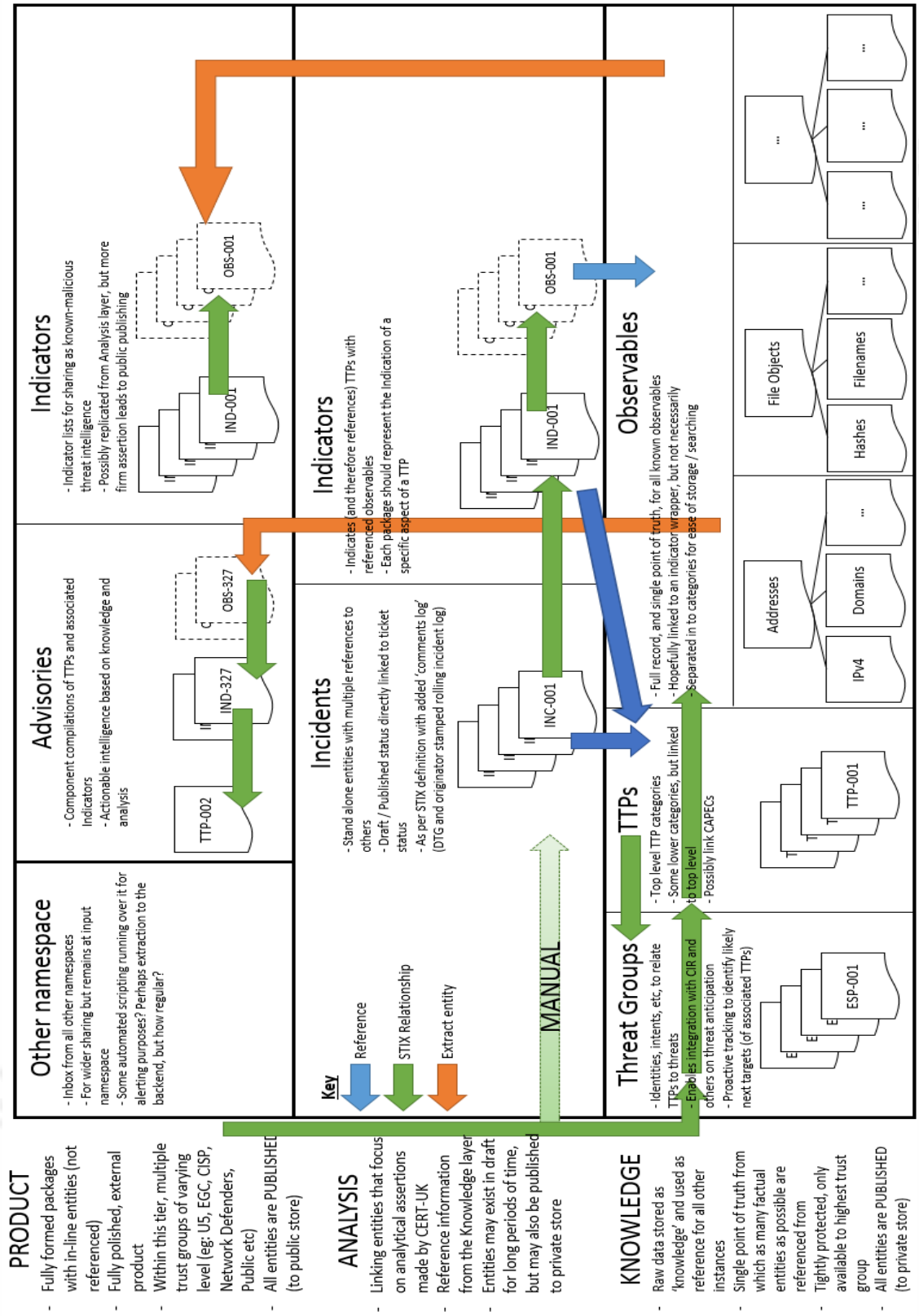


Figure 9. Data Flow model for CERT-UK showing the knowledge, analysis and product layers, along with the data processing function to support remote referencing wherever possible (to reduce duplication) and Observable inclusion where referencing is not sufficient.

Christopher O'Brien, cobrien@cert.gov.uk

Figure 9 shows a suggested data model for representing cyber threat intelligence in a knowledge base. The key aspect is that knowledge – a fact of existence for a piece of information – is the core material on which we base analytical assertions. Each event, entry or Observable in a data feed can be considered 'knowledge' that we gain from a specific source. Equally, network defenders can create their own knowledge from Observables on a network either through manual investigation or as the byproduct of some other tool. These objects then form the atomic building blocks on which to construct our analytical assertions. Changes to objects in this layer are usually administrative only through processes such as de-duplication and reconciliation and tend not to hold context beyond what the object states.

The analysis layer allows the analyst to draw links between those building blocks which reside in the knowledge layer, develop analytical assertions and provide additional context to the knowledge. Data in this layer is more transient, often edited and added to by the analyst in order to improve the threat intelligence picture.

For example, an IP address may be recorded in the knowledge layer as having sent a DNS request for a given malicious domain. In an object oriented approach we would say that the IP address (address object) sent a DNS request packet (network connection object) containing reference to the malicious domain (domain name object). Each of those three objects can be considered knowledge and could be related to each other in the context of an Incident (which may also be considered knowledge depending on the source and local policies). An analyst may observe that there have been multiple other DNS requests made for the same domain name from other IP addresses and assess that there is a more wide-spread Incident. As such, an Indicator or Incident object can be created in the analysis layer which links the network connection objects of all DNS requests for that malicious domain as a potentially wide-spread Incident. It may also be prudent to add the domain name object to a 'malicious domain watchlist' in the analysis layer for others to refer to, perhaps for later publication to other network defenders.

Due to the potentially volatile nature of the data in the analysis layer, and the lack of context or value for the raw intelligence, the concept of the product layer is to allow confirmed analytical assertions to be shared with trust groups. The product layer needs to

accurately reflect current thinking for 'published' products but also ensure that useful context is provided wherever possible in order to increase the strategic intelligence value – the key advantage to a fully linked analytical layer. For example, having the same malicious domain watchlist from the analytical layer promoted to the product layer is useful, but including the references to the Incident objects associated with the network connection objects allows recipients to observe which domains have been seen most regularly in the knowledge base and, hence, provides a means of quantifying their risk.

2.3. Enrichment and Deployment

The main effort of this process is to enable automated deployment of threat intelligence for network defenders. As such, the objects that are now committed to the knowledge base should be enriched not just with other data sources, but also with test mechanisms to be leveraged by network defense tools. When selecting how and when to enrich data, it is important to consider both the volatility (the regularity with which the data is liable to change) and scope (the number of objects required to make valid assertions) of the enrichment. Figure 10 and Figure 11 demonstrate these two components in terms of their ease of implementation and accompanying examples.

As the size of our knowledge base increases the scope of querying does too. Hence, running queries against this large data set introduces performance concerns and the tendency with most big data approaches is to introduce local indexing to increase performance. Whilst a valid option, this can essentially result in duplication of data sets which, at high volumes, can be computationally and financially inefficient.

Increasing complexity of data models through unmanageable volatility is often tackled by holding replicated data sets for specific time periods. This can work well for 'investigations' which tend to focus on an occurrence at a specific point in time so that enrichment can be done in the context of the incident. However these time-bound models tend to work in discrete time and may lose the ability to conduct analysis through time as part of an automated analytical approach. Whilst these sorts of investigations tend to be done ad-hoc by analysts at the moment (for example, the historic registration details of infrastructure) it is important to avoid reducing our ability to do so by reverting to 'snapshots in time'.

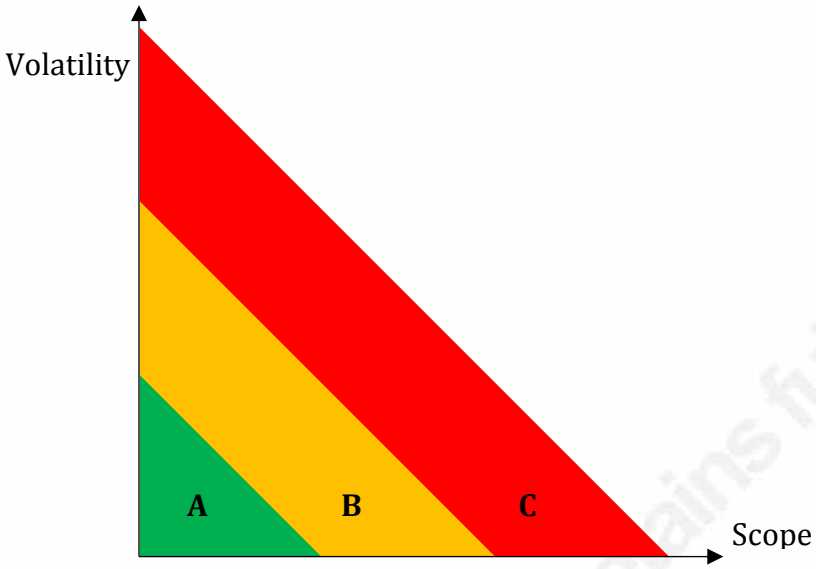


Figure 10. Diagram to show the increasing complexity of enrichment with increase in size of the data set (scope) and time-dependency of the enrichment source (volatility).

Area	Description	Example	Ease of Implementation
A	Object has little to no time dependency and can be enriched with other static data. Once enriched, these relationships are often static as well.	The hash of a malicious file being enriched with malicious domain names extracted from the binary through reverse engineering.	Simple
B	Enrichment sources may regularly change or the object may have several existing relationships that are subject to change. Results will be relatively stable but have a 'shelf life' when they may need to be reviewed.	Creating a SNORT rule to detect all network indicators associated with a TTP.	Moderate
C	Enrichment is dependent on querying a very large data set and comparing to dynamic, evolving enrichment sources. End results are often necessarily summarized and will need continual updates.	Showing the rate of detection of a specific Threat Actor on a network (identifying facets, live interpretation of source data and correlation to incident status)	Difficult

Figure 11. Table to show the types of enrichment using an object-oriented Facet Classification system as implemented during the experiment.

2.4. Automated Enrichment

Dell SecureWorks published a list of Observables with their report on a suspected APT which they refer to as TG3390. Included at Annex A is the trimmed contents of the report in a STIX structure plus some basic enrichment in the form of SNORT test mechanisms. Using the discussed principles and custom code as described earlier in Section 2, we processed the TG3390 report as a sample input and produced two STIX packages: The same data in a STIX format and an enrichment file containing additional Observables at Annexes A and B respectively. To demonstrate the value of the enrichment Figure 12 shows the output of the 'chain' function of exploit.py which shows the full logic chain that lead to the new Indicators.

```
SESEWUYU.COM ({{no_ref}}) --> 210.209.89.162 ({{no_ref}}) --> 728e5700a401498d91fb83159beec834 (d1e6920b-151d-40db-b83c-3b7635e5768e)
INFLATABLEKITE.NET ({{no_ref}}) --> 213.186.33.99 ({{no_ref}}) --> bbfd1e703f55ce779b536b5646a0cdc1 (43495fce-577b-4d59-8322-
3a0cf1c15e88)
LYONPLAGE.COM ({{no_ref}}) --> 213.186.33.99 ({{no_ref}}) --> bbfd1e703f55ce779b536b5646a0cdc1 (43495fce-577b-4d59-8322-3a0cf1c15e88)
BULKINGSTACKS.COM ({{no_ref}}) --> 190.123.36.113 ({{no_ref}}) --> 93e40da0bd78bebe5e1b98c6324e9b5b (35dd574a-9547-4039-9f41-
335a9242e37b)
CZFGV.EU ({{no_ref}}) --> 192.151.236.138 (6759e847-843c-44ee-9f4b-924661b7c97d) --> 86a05dcffe87caf7099dda44d9ec6b48 (befaf99c-cac5-
4caf-ba18-17231ad42f09)
JPCHS.COM ({{no_ref}}) --> 210.209.89.162 ({{no_ref}}) --> 728e5700a401498d91fb83159beec834 (d1e6920b-151d-40db-b83c-3b7635e5768e)
UCCHS.COM ({{no_ref}}) --> 210.209.89.162 ({{no_ref}}) --> 728e5700a401498d91fb83159beec834 (d1e6920b-151d-40db-b83c-3b7635e5768e)
VYBVG.EU ({{no_ref}}) --> 192.151.236.138 (6759e847-843c-44ee-9f4b-924661b7c97d) --> 86a05dcffe87caf7099dda44d9ec6b48 (befaf99c-cac5-
4caf-ba18-17231ad42f09)
www.download.windowsupdate.com ({{no_ref}}) --> b333b5d541a0488f4e710ae97c46d9c2 (2b474dd0-3f94-43d4-86d6-a7b6a95b25cf)
DROUVIN.COM ({{no_ref}}) --> 213.186.33.99 ({{no_ref}}) --> bbfd1e703f55ce779b536b5646a0cdc1 (43495fce-577b-4d59-8322-3a0cf1c15e88)
MICR0SOFT.ORG ({{no_ref}}) --> 49.143.205.30 (f5dc633e-1001-461b-8f38-834c99d32b33)
HOUSEOFCARS.NET.PL ({{no_ref}}) --> 213.186.33.99 ({{no_ref}}) --> bbfd1e703f55ce779b536b5646a0cdc1 (43495fce-577b-4d59-8322-
3a0cf1c15e88)
ECWARD.COM ({{no_ref}}) --> 72.11.141.133 (719b088c-e2cc-43fa-9bd3-f34264f7e57f)
APINK.LOVE ({{no_ref}}) --> 210.209.89.162 ({{no_ref}}) --> 728e5700a401498d91fb83159beec834 (d1e6920b-151d-40db-b83c-3b7635e5768e)
NHINF.EU ({{no_ref}}) --> 192.151.236.138 (6759e847-843c-44ee-9f4b-924661b7c97d) --> 86a05dcffe87caf7099dda44d9ec6b48 (befaf99c-cac5-
4caf-ba18-17231ad42f09)
ZLLQG.EU ({{no_ref}}) --> 192.151.236.138 (6759e847-843c-44ee-9f4b-924661b7c97d) --> 86a05dcffe87caf7099dda44d9ec6b48 (befaf99c-cac5-
4caf-ba18-17231ad42f09)
WEILINGYU.PW ({{no_ref}}) --> 210.209.89.162 ({{no_ref}}) --> 728e5700a401498d91fb83159beec834 (d1e6920b-151d-40db-b83c-3b7635e5768e)
LOPEZ-GARCIA.COM ({{no_ref}}) --> 213.186.33.99 ({{no_ref}}) --> bbfd1e703f55ce779b536b5646a0cdc1 (43495fce-577b-4d59-8322-
3a0cf1c15e88)
103.24.1.54 (30a10fe0-0cd0-4272-a54e-96f9def2e7d9) --> bbfd1e703f55ce779b536b5646a0cdc1 (43495fce-577b-4d59-8322-3a0cf1c15e88)
190.123.36.115 ({{no_ref}}) --> 2bec1860499aae1dbcc92f48b276f998 (5acaccc5-eeac-43a3-8d0c-e7ca2358274e)
190.123.36.113 ({{no_ref}}) --> 93e40da0bd78bebe5e1b98c6324e9b5b (35dd574a-9547-4039-9f41-335a9242e37b)
190.123.36.111 ({{no_ref}}) --> b333b5d541a0488f4e710ae97c46d9c2 (2b474dd0-3f94-43d4-86d6-a7b6a95b25cf)
2.18.213.208 ({{no_ref}}) --> b333b5d541a0488f4e710ae97c46d9c2 (2b474dd0-3f94-43d4-86d6-a7b6a95b25cf)
5.178.43.10 ({{no_ref}}) --> 4251aaf38a485b08d5562c6066370f09 (60dcaaac-dd21-4dae-b34e-a4ac7e77ef28)
64.4.10.33 ({{no_ref}}) --> 1cb4b74e9d030afbb18accf6ee2bfca1 (b6f11868-e483-46fb-9ab7-6642a4f3a701)
95.101.0.89 ({{no_ref}}) --> 86a05dcffe87caf7099dda44d9ec6b48 (befaf99c-cac5-4caf-ba18-17231ad42f09)
104.41.150.68 ({{no_ref}}) --> b333b5d541a0488f4e710ae97c46d9c2 (2b474dd0-3f94-43d4-86d6-a7b6a95b25cf)
88.221.14.115 ({{no_ref}}) --> 728e5700a401498d91fb83159beec834 (d1e6920b-151d-40db-b83c-3b7635e5768e)
65.55.56.206 ({{no_ref}}) --> 57e85fc30502a925ffed16082718ec6c (6aa043cf-7ff3-4d88-a4a0-8df604f64047)
23.99.222.162 ({{no_ref}}) --> f7a842eb1364d1269b40a344510068e8 (1b3fd223-5c45-4d8c-98b3-4ce6e309ecc6)
88.221.15.80 ({{no_ref}}) --> 12a522cb96700c82dc964197adb57ddf (da94cf08-2252-4238-af92-9012917a0813)
180.225.152.30 ({{no_ref}}) --> 1cb4b74e9d030afbb18accf6ee2bfca1 (b6f11868-e483-46fb-9ab7-6642a4f3a701)
104.209.134.106 ({{no_ref}}) --> 2bec1860499aae1dbcc92f48b276f998 (5acaccc5-eeac-43a3-8d0c-e7ca2358274e)
210.209.89.162 ({{no_ref}}) --> 728e5700a401498d91fb83159beec834 (d1e6920b-151d-40db-b83c-3b7635e5768e)
192.151.236.138 (6759e847-843c-44ee-9f4b-924661b7c97d) --> 86a05dcffe87caf7099dda44d9ec6b48 (befaf99c-cac5-4caf-ba18-17231ad42f09)
213.186.33.99 ({{no_ref}}) --> bbfd1e703f55ce779b536b5646a0cdc1 (43495fce-577b-4d59-8322-3a0cf1c15e88)
```

Figure 12. Output of the chain function showing newly discovered Indicators (marked with {{no_ref}}) and the links that lead to their discovery from the original input Indicators.

In this experiment we only use two resources for enrichment. This can easily be extended to any other data resource which has an API. APIs are often rate-limited, even with paid-for services, so it is important to add appropriate controls to the use of the relevant keys during deployment. As discussed in Section 2.3, there is a temptation to

create local copies of useful resources in order to increase efficiency of querying and maximize the re-use of API response results. Whilst that may be considered efficient in terms of short term costs (over costly API query-limit increases), local indexes need additional resource to save to disk and are continuously under threat of becoming out of date due to a lack of regular updates. As such, organizations should consider the benefits of Query-in-Place (QiP) where remote resources are accessed dynamically on a just-in-time basis. This ensures that the data is as accurate as the primary source and reduces the burden of maintaining a local copy. The relevant advantages and disadvantages of each approach will be dependent on specific data feeds, though ensuring that both the QiP and local-copy solutions are accessible through an automated API enables flexibility in enrichment implementation.

2.5. Tactical-Strategic Fusion

Developing the test mechanisms is useful, but not novel. The value in containing them in a STIX object is in the ability to trace the tactical data (test mechanism associated with an Observable) through to its strategic context (association to a TTP or Threat Actor), pivot off that data and start to suggest other associated tactical data that may also be worth deploying. This allows for prioritization of other test mechanisms that may have not been considered previously (either for purposes of reducing load on the network sensors or because they were added in real time through dynamic enrichment) but also provides that strategic context to be able to understand the likely motivations of the attack due to the asserted link to a known Threat Actor.

During this experiment we obtain a copy of one of the samples referenced in the TG3390 report from VirusTotal. We then set up a victim on a WinXP machine and establish a shared folder from a networked Kali box. From the victim, we then browse to the network share and run the sample. Figure 13 shows a Wireshark representation of a packet capture from the experiment.

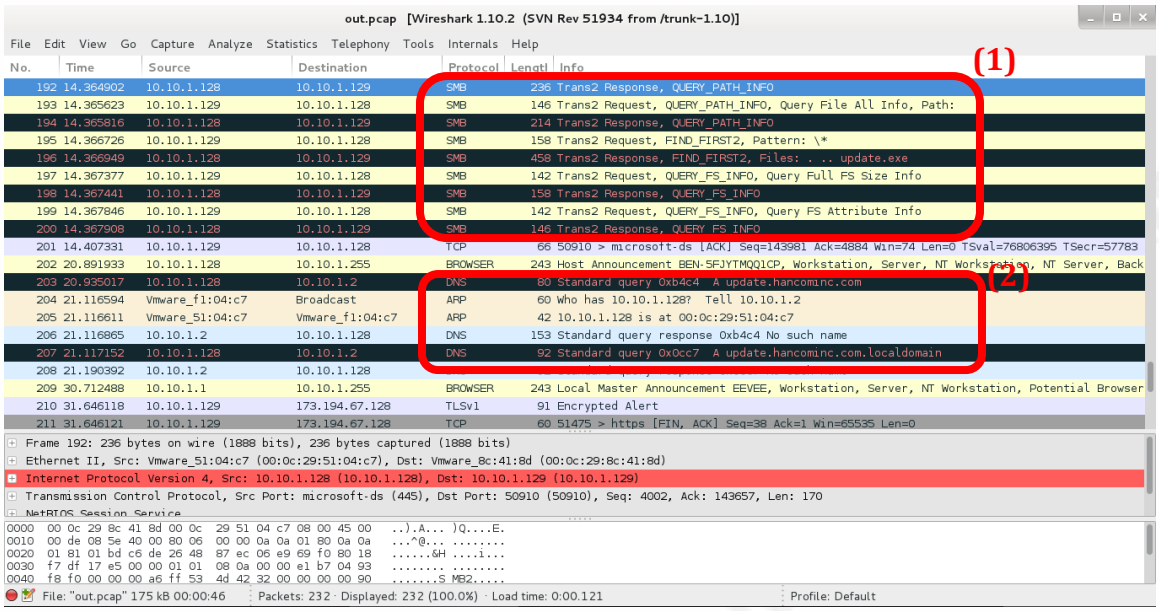


Figure 13. Screenshot showing packet capture during infection. Note highlighted sections showing SMB session (1) to transfer 'update.exe' which can later be carved out using Wireshark, and ensuing DNS request for update.hancominc.com (2).

Importantly, the domain name that is queried through DNS activity following the running of the sample is the same as one from the TG3390 report. By writing SNORT test mechanisms such as that at line 69 of Annex A we can ensure that this activity would be detected. We can also confirm that the file hash is preserved through transfer as illustrated in Figure 14.

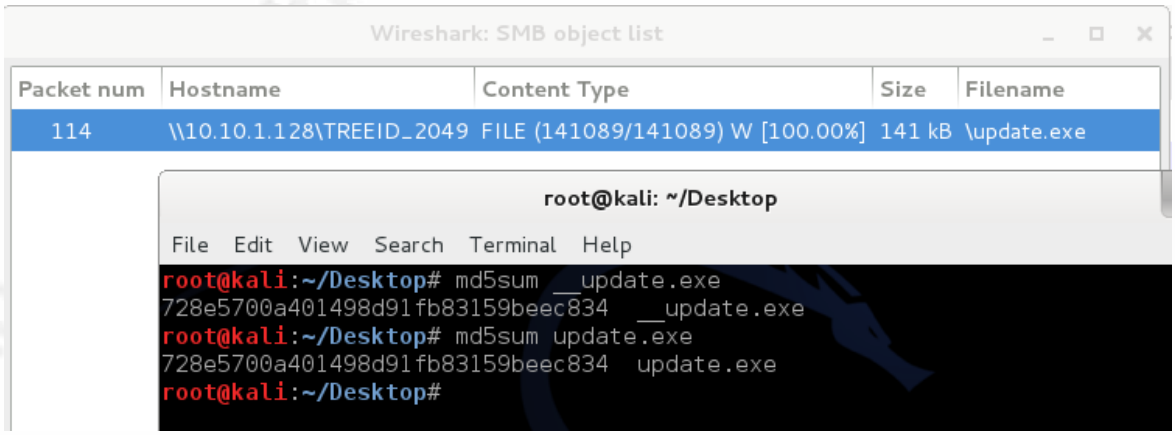


Figure 14. Screenshot to show the malware pre-infection (update.exe) and the file carved out of Wireshark pcap from the SMB transfer (__update.exe). The hashes are the same as shown by the MD5 hash comparison.

Both the file transfer and DNS request can be detected using a combination of SNORT and Bro.

Christopher O'Brien, cobrien@cert.gov.uk

2.5.1. File Hash Detection

Bro is a Network Security Monitor tool can implement heuristic detection and use of a native scripting language to enhance network detection. As of version 2.2, Bro comes equipped with the File Analysis Framework (FAF) (Bro, 2015). This allows for files of multiple different protocols to be detected in pcap and have analysis run against them without the need for custom protocol parsing. At time of writing the core FAF framework does not support SMB, thought the below is an output from the same

```
#path   files
#open   2015-12-13-13-56-22
#fields ts              tx_hosts        rx_hosts        analyzers       md5
#types  time            set[addr]       set[addr]       set[string]     string
1449069585.211766       166.78.45.16    10.10.1.129     MD5     eccb787a93193fbb4238bcd24520d75e
1449069585.211766       166.78.45.16    10.10.1.129     MD5     23bb075facfbfb026935b31fb6ee063b
1449069585.211766       166.78.45.16    10.10.1.129     MD5     2e7db2a31d0e3da4b25f49b9542a2e1a
#close  2015-12-13-13-56-22
```

Figure 15. Snipped extract from files.log as an output of Bro on the same executable (highlighted) rendered to the victim as a drive-by download. Note that the hash of the same file is actually different.

experiment using a 'drive-by download' setup (HTTP) which Bro does currently support.

Note that the same executable is used in Figure 15 as in Figure 14, however the file hash is different. On investigation it appeared that the HTTP transfer caused slight change to the executable, suspected to be due to URL Encoding of ASCII-characters however this was not confirmed. It is important to note that many of these file-carving techniques are still emerging and may not be consistent across implementations (Bro and Wireshark, in this case) and a potential direction for future work may be to investigate whether file hashes are consistently reproducible in network detection systems.

2.5.2. Network Connection Detection

A simple SNORT rule, as created in the TG3390 enrichment code, can then be deployed to the appropriate rule set to run over the same pcap in order to detect any follow on network connections. The SNORT rules are created by the custom deploy.py module and clearly marked as 'Automated deployment'. This is a critical part of the data marking process to ensure that any network defenders responding to a SNORT alert generated through this method are aware that it has not necessarily been validated by anything other than an automated process. This is highlighted through the experiment as one of the domains returned through VirusTotal enrichment as being beaconed to by a

known malicious file hash was 'www.download.windowsupdate.com'. Instinctively this is likely to be a false positive – however, the enrichment has successfully identified that malicious code of this type does attempt to communicate with this domain. Hence, while the Indicator itself is legitimate, the context within which it is detected is important.

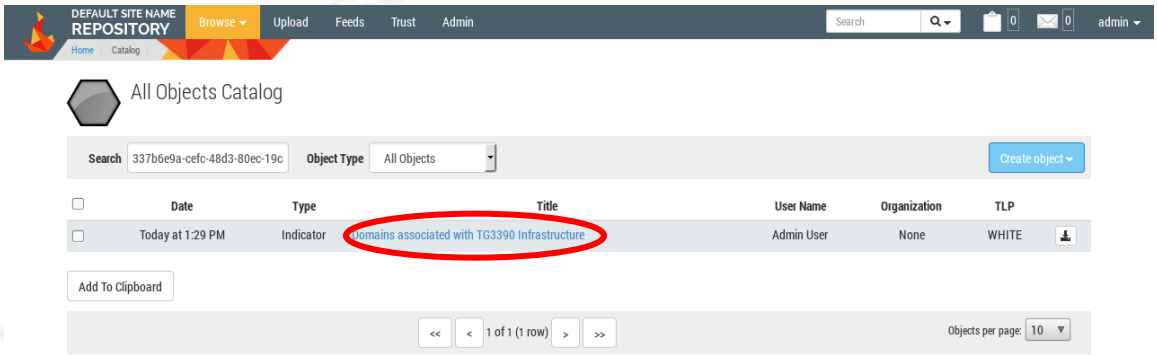
In the TG3390 example we deploy SNORT rules for all observables to a given rule file as determined by the config.ini file. Provided this rule file is incorporated as part of a SNORT deployment, a script can then be run to monitor for new alerts. The added benefit of this technique being that the SNORT rule is saved to a STIX file in the context of a Threat Actor and a TTP, so additional context can automatically be used to enrich

```
sonion@sonion1:/etc/nsm/templates/snort$ sudo snort --daq pcap -A console -q -c snort.conf
 -r /home/sonion/Desktop/out.pcap
12/02-15:19:19.835492  [**] [1:20000417:0] Automated STIX deployment - 337b6e9a-cefc-48d3-
80ec-19c8dbf3c7c8 [**] [Priority: 0] {UDP} 10.10.1.2:53 -> 10.10.1.128:1027
sonion@sonion1:/etc/nsm/templates/snort$
```

the alert and assist in incident triage.

Figure 16. Screenshot of SNORT running automatically generated rules from STIX over a pcap sample of the malicious executable SMB transfer reproduced in Figure 13.

The successful detection is merely an indication, and clear labelling is put in the SNORT rule's message to indicate that this is an automated deployment – warranting



further investigation, but not necessarily to be trusted as a true positive.

Figure 17. Screenshot of the uploaded STIX file as viewed in Soltra Edge, showing the cross-referenced Observable ID as detected in the SNORT alert.

This script 'inboxes' both the original TG3390 (STIX format) and the enrichment data in a Soltra Edge server which can then be queried in the ensuing incident response. As shown in Figure 17, the recorded contextual data can then be rendered to other users of (in this case) Soltra Edge so that they can quickly benefit from the corporate

knowledge of this structured data. Whilst this example shows a manual usage of a graphical user interface, the back end query can be further enhanced to make use of a remote access API for the TAXII server. This is another potential route for future work.

3. Conclusion

Threat Intelligence is complex and we need models to be able to conceptualize the problem space and be able to assimilate bulk data in an automated way. The use of structured threat intelligence languages, such as STIX, allow network defenders to structure data in such a way as to allow automated defense deployments without compromising on strategic context.

There are limitations with these languages as high flexibility often leads to infinite coding possibilities. With so many varied interpretations of 'standard' CTI structures it is difficult to agree on a standardized structure. However, previous languages have made the mistake of being too hierarchical and not allowing for faceted object structures. A compromise is required to retain repeatability for automated systems while allowing analytical flexibility. This may be in the form of 'templates' for standard threat intelligence reporting structures such as the ones developed through the course of this experiment.

To make better use of the intelligence there is a need for further development of interoperable analytical tools. Whilst some conversion languages do exist for commonly used tools, few take full advantage of the conceptual object model in languages such as STIX. This interoperability would benefit from an improved querying functionality for STIX. TAXII currently does not have the level of querying capability to act as a true API to STIX-structured data making QIP nearly impossible without replication of a RESTful (or similar) overlay.

All of the techniques demonstrated come from open source and open architecture techniques. Even those network defenders on limited budgets can use techniques in this paper to dramatically increase their usage of threat intelligence data to protect their networks. Capability can be increased even further if those publishing reports and creating intelligence feeds could leverage the same techniques in publishing as they do

for their own internal knowledge management. I look forward to the first vendor to deliver an 'APT outing' report in full STIX format of their own volition.

Annex A

Output from TG3390 exploitation process generated a full STIX interpretation including enrichment through SNORT rules. Note: Domains are added as distinct test mechanisms as part of a list whilst groups of IP addresses are assumed to be ‘or’ indicators of activity associated with TG3390.

```
<stix:STIX_Package
        xmlns:AddressObj="http://cybox.mitre.org/objects#AddressObject-2"
        xmlns:DomainNameObj="http://cybox.mitre.org/objects#DomainNameObject-1"
        xmlns:FileObj="http://cybox.mitre.org/objects#FileObject-2"
        xmlns:certuk="https://cert.gov.uk"
        xmlns:cybox="http://cybox.mitre.org/cybox-2"
        xmlns:cyboxCommon="http://cybox.mitre.org/common-2"
        xmlns:cyboxVocabs="http://cybox.mitre.org/default_vocabularies-2"
        xmlns:indicator="http://stix.mitre.org/Indicator-2"
        xmlns:marking="http://data-marking.mitre.org/Marking-1"
        xmlns:snortTM="http://stix.mitre.org/extensions/TestMechanism#Snort-1"
        xmlns:stix="http://stix.mitre.org/stix-1"
        xmlns:stixCommon="http://stix.mitre.org/common-1"
        xmlns:stixVocabs="http://stix.mitre.org/default_vocabularies-1"
        xmlns:tlpMarking="http://data-marking.mitre.org/extensions/MarkingStructure#TLP-1"
        xmlns:ttp="http://stix.mitre.org/TTP-1"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" id="certuk:Package-e07d604c-0b6a-4b28-8ddc-e22561a6ca18" version="1.2">
    <stix:STIX_Header>
        <stix:Title>TG3390</stix:Title>
        <stix:Description>Dell SecureWorks Counter Threat Unit(TM) (CTU) researchers investigated activities associated with Threat
Group-3390[1] (TG-3390) - http://www.secureworks.com/cyber-threat-intelligence/threats/threat-group-3390-targets-organizations-for-
cyberespionage/</stix:Description>
        <stix:Handling>
            <marking:Marking>
                <marking:Controlled_Structure>../../../../descendant-or-self::node()</marking:Controlled_Structure>
                <marking:Marking_Structure xsi:type='tlpMarking:TLPMarkingStructureType' color="WHITE"/>
            </marking:Marking>
        </stix:Handling>
    </stix:STIX_Header>
    <stix:Indicators>
        <stix:Indicator id="certuk:indicator-337b6e9a-cefc-48d3-80ec-19c8dbf3c7c8" timestamp="2015-12-13T23:25:35.094930+00:00"
xsi:type='indicator:IndicatorType'>
            <indicator:Title>Domains associated with TG3390 Infrastructure</indicator:Title>
            <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">Domain Watchlist</indicator:Type>
            <indicator:Observable id="certuk:Observable-5af24b0c-4f62-4bc1-bb7b-5a47ff6a8e4d">
                <cybox:Observable_Composition operator="OR">
                    <cybox:Observable id="certuk:Observable-1be5d43e-8a40-4169-af95-8b758387de67">
                        <cybox:Object id="certuk:DomainName-6e7c9b86-c559-4921-b162-1572c3465c95">
                            <cybox:Properties xsi:type="DomainNameObj:DomainNameObjectType">
                                <DomainNameObj:Value>american.blackcmd.com</DomainNameObj:Value>
                            </cybox:Properties>
                        </cybox:Object>
                    </cybox:Observable>
                    <snip></snip>
                    <cybox:Observable id="certuk:Observable-2c4d555f-05bf-43b5-ad00-f1f18e6097e4">
                        <cybox:Object id="certuk:DomainName-9bb9dac0-c155-4b06-b65b-dbd834e18aef">
                            <cybox:Properties xsi:type="DomainNameObj:DomainNameObjectType">
```

```
                            <DomainNameObj:Value>update.hancominc.com</DomainNameObj:Value>
                        </cybox:Properties>
                    </cybox:Object>
                </cybox:Observable>
            </cybox:Observable_Composition>
        </indicator:Observable>
        <indicator:Indicated_TTP>
            <stixCommon:TTP idref="certuk:ttp-05bcd4b6-392d-42d1-80b0-ce9340b5bd3c" xsi:type='ttp:TTPType'/>
        </indicator:Indicated_TTP>
        <indicator:Test_Mechanisms>
            <indicator:Test_Mechanism xsi:type='snortTM:SnortTestMechanismType'>
                <indicator:Efficacy timestamp="2015-12-13T23:25:36.555749+00:00">
                    <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">High</stixCommon:Value>
                </indicator:Efficacy>
                <indicator:Producer>
                    <stixCommon:Identity>
                        <stixCommon:Name>Auto</stixCommon:Name>
                    </stixCommon:Identity>
                </indicator:Producer>
                <snortTM:Rule><![CDATA[alert udp $HOME_NET 53 -> any any (msg:"Automated STIX deployment - 337b6e9a-cefc-48d3-
80ec-19c8dbf3c7c8";content:"|08|american|08|blackcmd|03|com|00|"; sid: 20059624;)]]></snortTM:Rule>
                <snortTM:Rule><![CDATA[alert udp $HOME_NET 53 -> any any (msg:"Automated STIX deployment - 337b6e9a-cefc-48d3-
80ec-19c8dbf3c7c8";content:"|06|update|09|hancominc|03|com|00|"; sid: 20000417;)]]></snortTM:Rule>
            </indicator:Test_Mechanism>
        </indicator:Test_Mechanisms>
        <indicator:Confidence timestamp="2015-12-13T23:25:35.095606+00:00">
            <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">High</stixCommon:Value>
        </indicator:Confidence>
    </stix:Indicator>
    <stix:Indicator id="certuk:indicator-79ae6553-9bec-4d2d-a93c-318950151a18" timestamp="2015-12-13T23:25:35.109011+00:00"
xsi:type='indicator:IndicatorType'>
        <indicator:Title>[H] IP Addresses associated with TG3390 Infrastructure</indicator:Title>
        <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">IP Watchlist</indicator:Type>
        <indicator:Observable id="certuk:Observable-0f261ec6-b84a-4402-a932-0edafb93331d">
            <cybox:Observable_Composition operator="OR">
                <cybox:Observable id="certuk:Observable-79ed3f57-43f6-4578-910c-df128d3afb8a">
                    <cybox:Object id="certuk:Address-a534bd67-d4ef-413e-bc03-595d0f187867">
                        <cybox:Properties xsi:type="AddressObj:AddressObjectType" category="ipv4-addr">
                            <AddressObj:Address_Value>208.115.242.36</AddressObj:Address_Value>
                        </cybox:Properties>
                    </cybox:Object>
                </cybox:Observable>
                <snip></snip>
                <cybox:Observable id="certuk:Observable-719b088c-e2cc-43fa-9bd3-f34264f7e57f">
                    <cybox:Object id="certuk:Address-d005526f-15a5-4d6f-9fb7-62fa681469f2">
                        <cybox:Properties xsi:type="AddressObj:AddressObjectType" category="ipv4-addr">
                            <AddressObj:Address_Value>72.11.141.133</AddressObj:Address_Value>
```

```
                        </cybox:Properties>
                    </cybox:Object>
                </cybox:Observable>
            </cybox:Observable_Composition>
        </indicator:Observable>
        <indicator:Indicated_TTP>
            <stixCommon:TTP idref="certuk:ttp-05bcd4b6-392d-42d1-80b0-ce9340b5bd3c" xsi:type='ttp:TTPType'/>
        </indicator:Indicated_TTP>
        <indicator:Test_Mechanisms>
            <indicator:Test_Mechanism xsi:type='snortTM:SnortTestMechanismType'>
                <indicator:Efficacy timestamp="2015-12-13T23:25:36.546388+00:00">
                    <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">High</stixCommon:Value>
                </indicator:Efficacy>
                <indicator:Producer>
                    <stixCommon:Identity>
                        <stixCommon:Name>Auto</stixCommon:Name>
                    </stixCommon:Identity>
                </indicator:Producer>
                <snortTM:Rule><![CDATA[alert tcp $HOME_NET any ->
[208.115.242.36,208.115.242.37,208.115.242.38,66.63.178.142,72.11.148.220,72.11.141.133,74.63.195.236,74.63.195.236,74.63.195.237,74.6
3.195.238,103.24.0.142,103.24.1.54,106.187.45.162,192.151.236.138,192.161.61.19,192.161.61.20,192.161.61.22,103.24.1.54,67.215.232.179
,96.44.177.195] any (msg:"Automated STIX deployment - 79ae6553-9bec-4d2d-a93c-318950151a18"; sid: 20080767;)]]></snortTM:Rule>
            </indicator:Test_Mechanism>
        </indicator:Test_Mechanisms>
        <indicator:Confidence timestamp="2015-12-13T23:25:35.121027+00:00">
            <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">High</stixCommon:Value>
        </indicator:Confidence>
    </stix:Indicator>
    <stix:Indicator id="certuk:indicator-26cb2ff5-e474-4987-9aed-621232c786d6" timestamp="2015-12-13T23:25:35.121260+00:00"
xsi:type='indicator:IndicatorType'>
        <indicator:Title>[M] IP Addresses associated with TG3390 Infrastructure</indicator:Title>
        <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">IP Watchlist</indicator:Type>
        <indicator:Observable id="certuk:Observable-e9cc38e5-f89e-48b5-9cab-7e00e4f32bb0">
            <cybox:Observable_Composition operator="OR">
                <cybox:Observable id="certuk:Observable-35769669-ac3e-4ad8-ad60-2b52e963ad41">
                    <cybox:Object id="certuk:Address-abea616e-370b-4d47-a0a0-14ece1151b77">
                        <cybox:Properties xsi:type="AddressObj:AddressObjectType" category="ipv4-addr">
                            <AddressObj:Address_Value>49.143.192.221</AddressObj:Address_Value>
                        </cybox:Properties>
                    </cybox:Object>
                </cybox:Observable>
                <snip></snip>
                <cybox:Observable id="certuk:Observable-7d30568f-d07c-4c0b-96c3-ba6912d40895">
                    <cybox:Object id="certuk:Address-cd77b023-198f-4d11-b024-ac692af7861c">
                        <cybox:Properties xsi:type="AddressObj:AddressObjectType" category="ipv4-addr">
                            <AddressObj:Address_Value>49.143.205.30</AddressObj:Address_Value>
                        </cybox:Properties>
```

```
                        </cybox:Object>
                    </cybox:Observable>
                </cybox:Observable_Composition>
            </indicator:Observable>
            <indicator:Indicated_TTP>
                <stixCommon:TTP idref="certuk:ttp-05bcd4b6-392d-42d1-80b0-ce9340b5bd3c" xsi:type='ttp:TTPType'/>
            </indicator:Indicated_TTP>
            <indicator:Test_Mechanisms>
                <indicator:Test_Mechanism xsi:type='snortTM:SnortTestMechanismType'>
                    <indicator:Efficacy timestamp="2015-12-13T23:25:36.546563+00:00">
                        <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">Medium</stixCommon:Value>
                    </indicator:Efficacy>
                    <indicator:Producer>
                        <stixCommon:Identity>
                            <stixCommon:Name>Auto</stixCommon:Name>
                        </stixCommon:Identity>
                    </indicator:Producer>
                    <snortTM:Rule><![CDATA[alert tcp $HOME_NET any ->
[49.143.192.221,67.215.232.181,67.215.232.182,96.44.182.243,96.44.182.245,96.44.182.246,49.143.205.30] any (msg:"Automated STIX
deployment - 26cb2ff5-e474-4987-9aed-621232c786d6"; sid: 20065749;)]]></snortTM:Rule>
                </indicator:Test_Mechanism>
            </indicator:Test_Mechanisms>
            <indicator:Confidence timestamp="2015-12-13T23:25:35.121359+00:00">
                <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">Medium</stixCommon:Value>
            </indicator:Confidence>
        </stix:Indicator>
        <stix:Indicator id="certuk:indicator-cb377a6e-70ab-470b-85ba-9fa499a05258" timestamp="2015-12-13T23:25:35.124570+00:00"
xsi:type='indicator:IndicatorType'>
            <indicator:Title>File hashes for HTTP Browser Dropper</indicator:Title>
            <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">File Hash Watchlist</indicator:Type>
            <indicator:Observable id="certuk:Observable-4f5af697-c110-4fd4-a243-cd2fa59d0998">
                <cybox:Observable_Composition operator="OR">
                    <cybox:Observable id="certuk:Observable-44ea94c0-6e91-4748-9148-3c1822912be4">
                        <cybox:Object id="certuk:File-34ec4dab-d8c4-4161-8826-f2c788c03327">
                            <cybox:Properties xsi:type="FileObj:FileObjectType">
                                <FileObj:Hashes>
                                    <cyboxCommon:Hash>
                                        <cyboxCommon:Type xsi:type="cyboxVocabs:HashNameVocab-1.0">MD5</cyboxCommon:Type>

<cyboxCommon:Simple_Hash_Value>1cb4b74e9d030afbb18accf6ee2bfca1</cyboxCommon:Simple_Hash_Value>
                                    </cyboxCommon:Hash>
                                </FileObj:Hashes>
                            </cybox:Properties>
                        </cybox:Object>
                    </cybox:Observable>
                    <snip></snip>
                    <cybox:Observable id="certuk:Observable-a75b7d05-ab53-4c33-82de-9f67f93911c1">
```

```
                        <cybox:Object id="certuk:File-162dde0b-fef5-46b1-9227-2bb06d6aebc9">
                            <cybox:Properties xsi:type="FileObj:FileObjectType">
                                <FileObj:Hashes>
                                    <cyboxCommon:Hash>
                                        <cyboxCommon:Type xsi:type="cyboxVocabs:HashNameVocab-1.0">MD5</cyboxCommon:Type>
<cyboxCommon:Simple_Hash_Value>2bec1860499aae1dbcc92f48b276f998</cyboxCommon:Simple_Hash_Value>
                                    </cyboxCommon:Hash>
                                </FileObj:Hashes>
                            </cybox:Properties>
                        </cybox:Object>
                    </cybox:Observable>
                </cybox:Observable_Composition>
            </indicator:Observable>
            <indicator:Indicated_TTP>
                <stixCommon:TTP idref="certuk:ttp-a7a8a618-f372-4520-871c-cf8ccf4939a8" xsi:type='ttp:TTPType'/>
            </indicator:Indicated_TTP>
            <indicator:Confidence timestamp="2015-12-13T23:25:35.124644+00:00">
                <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">High</stixCommon:Value>
            </indicator:Confidence>
        </stix:Indicator>
        <stix:Indicator id="certuk:indicator-2c150e19-9435-4ebf-9097-f3d0431e1541" timestamp="2015-12-13T23:25:35.124085+00:00"
xsi:type='indicator:IndicatorType'>
            <indicator:Title>File hashes for HTTP Browser</indicator:Title>
            <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">File Hash Watchlist</indicator:Type>
            <indicator:Observable id="certuk:Observable-44a589b5-9990-4436-8c20-856f6ddb375d">
                <cybox:Observable_Composition operator="OR">
                    <cybox:Observable id="certuk:Observable-d1e700da-26df-4bee-b0b2-90178cbcc6e6">
                        <cybox:Object id="certuk:File-80f3d451-2625-4eef-97cd-60036fbfac11">
                            <cybox:Properties xsi:type="FileObj:FileObjectType">
                                <FileObj:Hashes>
                                    <cyboxCommon:Hash>
                                        <cyboxCommon:Type xsi:type="cyboxVocabs:HashNameVocab-1.0">MD5</cyboxCommon:Type>
<cyboxCommon:Simple_Hash_Value>014122d7851fa8bf4070a8fc2acd5dc5</cyboxCommon:Simple_Hash_Value>
                                    </cyboxCommon:Hash>
                                </FileObj:Hashes>
                            </cybox:Properties>
                        </cybox:Object>
                    </cybox:Observable>
                    <snip></snip>
                    <cybox:Observable id="certuk:Observable-0913c9ff-7dc8-4484-9f06-46e4392215e6">
                        <cybox:Object id="certuk:File-47e3aee3-2bfa-49a6-9c4c-f23a5c36dbe6">
                            <cybox:Properties xsi:type="FileObj:FileObjectType">
                                <FileObj:Hashes>
                                    <cyboxCommon:Hash>
                                        <cyboxCommon:Type xsi:type="cyboxVocabs:HashNameVocab-1.0">MD5</cyboxCommon:Type>
```

```
                                    <cyboxCommon:Simple_Hash_Value>f627bc2db3cab34d97c8949931cb432d</cyboxCommon:Simple_Hash_Value>
                                        </cyboxCommon:Hash>
                                    </FileObj:Hashes>
                                </cybox:Properties>
                            </cybox:Object>
                        </cybox:Observable>
                    </cybox:Observable_Composition>
                </indicator:Observable>
                <indicator:Indicated_TTP>
                    <stixCommon:TTP idref="certuk:ttp-02bd5584-4b78-4006-8276-9c9c17eeade5" xsi:type='ttp:TTPType'/>
                </indicator:Indicated_TTP>
                <indicator:Confidence timestamp="2015-12-13T23:25:35.124191+00:00">
                    <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">High</stixCommon:Value>
                </indicator:Confidence>
            </stix:Indicator>
            <stix:Indicator id="certuk:indicator-139edb52-2086-42d5-9dcb-3f4916b77f3b" timestamp="2015-12-13T23:25:35.150474+00:00"
xsi:type='indicator:IndicatorType'>
                <indicator:Title>File hashes for PlugX Dropper</indicator:Title>
                <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">File Hash Watchlist</indicator:Type>
                <indicator:Observable id="certuk:Observable-46f13f0f-a8c7-44aa-9132-dbe3c06c52b4">
                    <cybox:Observable_Composition operator="OR">
                        <cybox:Observable id="certuk:Observable-a8d65cc1-718b-4bd6-88be-0bf4d75ef96a">
                            <cybox:Object id="certuk:File-223764fe-a61a-4387-a340-9ebd96071d8e">
                                <cybox:Properties xsi:type="FileObj:FileObjectType">
                                    <FileObj:Hashes>
                                        <cyboxCommon:Hash>
                                            <cyboxCommon:Type xsi:type="cyboxVocabs:HashNameVocab-1.0">MD5</cyboxCommon:Type>

<cyboxCommon:Simple_Hash_Value>b313bbe17bd5ee9c00acff3bfccdb48a</cyboxCommon:Simple_Hash_Value>
                                        </cyboxCommon:Hash>
                                    </FileObj:Hashes>
                                </cybox:Properties>
                            </cybox:Object>
                        </cybox:Observable>
                        <snip></snip>
                        <cybox:Observable id="certuk:Observable-b309b661-c592-429f-a7fd-5dbdf175caa4">
                            <cybox:Object id="certuk:File-635c8764-6ca9-40d7-a0da-8813c041db78">
                                <cybox:Properties xsi:type="FileObj:FileObjectType">
                                    <FileObj:Hashes>
                                        <cyboxCommon:Hash>
                                            <cyboxCommon:Type xsi:type="cyboxVocabs:HashNameVocab-1.0">MD5</cyboxCommon:Type>

<cyboxCommon:Simple_Hash_Value>462fd01302bc40624a44b7960d2894cd</cyboxCommon:Simple_Hash_Value>
                                        </cyboxCommon:Hash>
                                    </FileObj:Hashes>
                                </cybox:Properties>
```

```
                        </cybox:Object>
                    </cybox:Observable>
                </cybox:Observable_Composition>
            </indicator:Observable>
            <indicator:Indicated_TTP>
                <stixCommon:TTP idref="certuk:ttp-e241bd75-eab8-491b-8e7c-74780ab40414" xsi:type='ttp:TTPType'/>
            </indicator:Indicated_TTP>
            <indicator:Confidence timestamp="2015-12-13T23:25:35.150556+00:00">
                <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">High</stixCommon:Value>
            </indicator:Confidence>
        </stix:Indicator>
    </stix:Indicators>
    <stix:TTPs>
        <stix:TTP id="certuk:ttp-05bcd4b6-392d-42d1-80b0-ce9340b5bd3c" timestamp="2015-12-13T23:25:35.085516+00:00"
xsi:type='ttp:TTPType'>
            <ttp:Title>Infrastructure Building</ttp:Title>
            <ttp:Intended_Effect timestamp="2015-12-13T23:25:35.085779+00:00">
                <stixCommon:Value xsi:type="stixVocabs:IntendedEffectVocab-1.0">Unauthorized Access</stixCommon:Value>
            </ttp:Intended_Effect>
            <ttp:Behavior>
                <ttp:Attack_Patterns>
                    <ttp:Attack_Pattern>
                        <ttp:Description>Infrastructure Building</ttp:Description>
                    </ttp:Attack_Pattern>
                </ttp:Attack_Patterns>
            </ttp:Behavior>
        </stix:TTP>
        <stix:TTP id="certuk:ttp-02bd5584-4b78-4006-8276-9c9c17eeade5" timestamp="2015-12-13T23:25:35.121598+00:00"
xsi:type='ttp:TTPType'>
            <ttp:Title>HTTP Browser</ttp:Title>
            <ttp:Intended_Effect timestamp="2015-12-13T23:25:35.123986+00:00">
                <stixCommon:Value xsi:type="stixVocabs:IntendedEffectVocab-1.0">Theft - Intellectual Property</stixCommon:Value>
            </ttp:Intended_Effect>
            <ttp:Behavior>
                <ttp:Malware>
                    <ttp:Malware_Instance>
                        <ttp:Name>HTTP Browser</ttp:Name>
                    </ttp:Malware_Instance>
                </ttp:Malware>
            </ttp:Behavior>
        </stix:TTP>
        <stix:TTP id="certuk:ttp-a7a8a618-f372-4520-871c-cf8ccf4939a8" timestamp="2015-12-13T23:25:35.124416+00:00"
xsi:type='ttp:TTPType'>
            <ttp:Title>HTTP Browser Dropper</ttp:Title>
            <ttp:Intended_Effect timestamp="2015-12-13T23:25:35.124468+00:00">
                <stixCommon:Value xsi:type="stixVocabs:IntendedEffectVocab-1.0">Theft - Intellectual Property</stixCommon:Value>
            </ttp:Intended_Effect>
```

```
            <ttp:Behavior>
                <ttp:Malware>
                    <ttp:Malware_Instance>
                        <ttp:Name>HTTP Browser Dropper</ttp:Name>
                    </ttp:Malware_Instance>
                </ttp:Malware>
            </ttp:Behavior>
        </stix:TTP>
        <stix:TTP id="certuk:ttp-e241bd75-eab8-491b-8e7c-74780ab40414" timestamp="2015-12-13T23:25:35.150306+00:00"
xsi:type='ttp:TTPType'>
            <ttp:Title>PlugX Dropper</ttp:Title>
            <ttp:Intended_Effect timestamp="2015-12-13T23:25:35.150373+00:00">
                <stixCommon:Value xsi:type="stixVocabs:IntendedEffectVocab-1.0">Theft - Intellectual Property</stixCommon:Value>
            </ttp:Intended_Effect>
            <ttp:Behavior>
                <ttp:Malware>
                    <ttp:Malware_Instance>
                        <ttp:Name>PlugX Dropper</ttp:Name>
                    </ttp:Malware_Instance>
                </ttp:Malware>
            </ttp:Behavior>
        </stix:TTP>
    </stix:TTPs>
</stix:STIX_Package>
```

Annex B

STIX package showing the structure of the newly identified data through enrichment. This package is a standalone package but contains references to the original in order to maintain the provenance chain. This is shown in the sections labelled "Source of enrichment for…" such as the one at line 103 for domain enrichment.

```
<stix:STIX_Package
        xmlns:AddressObj="http://cybox.mitre.org/objects#AddressObject-2"
        xmlns:DomainNameObj="http://cybox.mitre.org/objects#DomainNameObject-1"
        xmlns:certuk="https://cert.gov.uk"
        xmlns:cybox="http://cybox.mitre.org/cybox-2"
        xmlns:cyboxCommon="http://cybox.mitre.org/common-2"
        xmlns:cyboxVocabs="http://cybox.mitre.org/default_vocabularies-2"
        xmlns:indicator="http://stix.mitre.org/Indicator-2"
        xmlns:marking="http://data-marking.mitre.org/Marking-1"
        xmlns:stix="http://stix.mitre.org/stix-1"
        xmlns:stixCommon="http://stix.mitre.org/common-1"
        xmlns:stixVocabs="http://stix.mitre.org/default_vocabularies-1"
        xmlns:tlpMarking="http://data-marking.mitre.org/extensions/MarkingStructure#TLP-1"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" id="certuk:Package-2d6668c4-1f39-46ba-9839-8b1009bf7256" version="1.2">
    <stix:STIX_Header>
        <stix:Title>TG3390 - Enrichment</stix:Title>
        <stix:Description>Enrichment stix file to the Dell SecureWorks Counter Threat Unit(TM) (CTU) researchers investigated
activities associated with Threat Group-3390[1] (TG-3390) - http://www.secureworks.com/cyber-threat-intelligence/threats/threat-group-
3390-targets-organizations-for-cyberespionage/</stix:Description>
        <stix:Handling>
            <marking:Marking>
                <marking:Controlled_Structure>../../../../descendant-or-self::node()</marking:Controlled_Structure>
                <marking:Marking_Structure xsi:type='tlpMarking:TLPMarkingStructureType' color="WHITE"/>
            </marking:Marking>
        </stix:Handling>
    </stix:STIX_Header>
    <stix:Indicators>
        <stix:Indicator id="certuk:indicator-29883c0b-8f0f-47a0-867e-0648e982a816" timestamp="2015-12-13T23:28:12.449687+00:00"
xsi:type='indicator:IndicatorType'>
            <indicator:Title>Suspected TG3390 IP Addresses obtained through automated enrichment</indicator:Title>
            <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">IP Watchlist</indicator:Type>
            <indicator:Observable id="certuk:Observable-23ec76af-e827-4eac-a4a4-f8401679bb11">
                <cybox:Observable_Composition operator="OR">
                    <cybox:Observable id="certuk:Observable-6a52a217-65d6-4a0c-b417-b3a92f22fd5c">
                        <cybox:Object id="certuk:Address-fb76e105-4ab4-400f-b216-f58e3ebf82af">
                            <cybox:Properties xsi:type="AddressObj:AddressObjectType" category="ipv4-addr">
                                <AddressObj:Address_Value>64.4.10.33</AddressObj:Address_Value>
                            </cybox:Properties>
                        </cybox:Object>
                    </cybox:Observable>
                    <snip></snip>
                    <cybox:Observable id="certuk:Observable-4487850d-e534-4dc2-8f9b-4b461e32247d">
                        <cybox:Object id="certuk:Address-81554538-ccc8-439e-bb0a-7bba350c7bb1">
                            <cybox:Properties xsi:type="AddressObj:AddressObjectType" category="ipv4-addr">
                                <AddressObj:Address_Value>2.18.213.208</AddressObj:Address_Value>
                            </cybox:Properties>
                        </cybox:Object>
```

```
                    </cybox:Observable>
                </cybox:Observable_Composition>
            </indicator:Observable>
            <indicator:Confidence timestamp="2015-12-13T23:28:12.449866+00:00">
                <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">Low</stixCommon:Value>
            </indicator:Confidence>
            <indicator:Related_Indicators>
                <indicator:Related_Indicator>
                    <stixCommon:Relationship>Source of enrichment for IPs</stixCommon:Relationship>
                    <stixCommon:Indicator id="certuk:indicator-9bc483a0-4650-4f3f-a851-a80c686a80ef" timestamp="2015-12-
13T23:28:12.449974+00:00" xsi:type='indicator:IndicatorType'>
                        <indicator:Title>Related indicator wrapper for source of enrichment</indicator:Title>
                        <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">IP Watchlist</indicator:Type>
                        <indicator:Observable id="certuk:Observable-b4a4465b-5df9-4639-960b-a1bfd560c969">
                            <cybox:Observable_Composition operator="OR">
                                <cybox:Observable id="certuk:File-635c8764-6ca9-40d7-a0da-8813c041db78">
                                    <cybox:Description>Source of enrichment for: certuk:Observable-6a52a217-65d6-4a0c-b417-
b3a92f22fd5c, </cybox:Description>
                                </cybox:Observable>
                                <snip></snip>
                                <cybox:Observable id="certuk:File-b846e425-150b-4d4c-9a3e-9c91e0912f02">
                                    <cybox:Description>Source of enrichment for: certuk:Observable-9e0b3d47-9d12-4deb-8b3b-
f3c5737095d1, certuk:Observable-ea811097-c6cb-458c-9ece-fe0a371b4da9, certuk:Observable-4487850d-e534-4dc2-8f9b-4b461e32247d,
certuk:Observable-7165151f-e631-4ae0-a958-d1a8347ca278, certuk:Observable-7165151f-e631-4ae0-a958-d1a8347ca278, </cybox:Description>
                                </cybox:Observable>
                            </cybox:Observable_Composition>
                        </indicator:Observable>
                        <indicator:Confidence timestamp="2015-12-13T23:28:12.450050+00:00">
                            <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">Medium</stixCommon:Value>
                        </indicator:Confidence>
                    </stixCommon:Indicator>
                </indicator:Related_Indicator>
            </indicator:Related_Indicators>
        </stix:Indicator>
        <stix:Indicator id="certuk:indicator-7d8bd59b-f9fa-4807-9972-4ed24fc7cd46" timestamp="2015-12-13T23:28:12.450145+00:00"
xsi:type='indicator:IndicatorType'>
            <indicator:Title>Suspected TG3390 Domains obtained through automated enrichment</indicator:Title>
            <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">Domain Watchlist</indicator:Type>
            <indicator:Observable id="certuk:Observable-2b394a29-5533-4a10-b58b-9a480cec9a2e">
                <cybox:Observable_Composition operator="OR">
                    <cybox:Observable id="certuk:Observable-5063b55a-601b-4361-a12d-94f5daaccc28">
                        <cybox:Object id="certuk:DomainName-75a10cf0-01b5-4ac0-b452-25e6f379312f">
                            <cybox:Properties xsi:type="DomainNameObj:DomainNameObjectType">
                                <DomainNameObj:Value>ECWARD.COM</DomainNameObj:Value>
                            </cybox:Properties>
                        </cybox:Object>
                    </cybox:Observable>
```

```
                    <snip></snip>
                </cybox:Observable_Composition>
            </indicator:Observable>
            <indicator:Confidence timestamp="2015-12-13T23:28:12.450209+00:00">
                <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">Low</stixCommon:Value>
            </indicator:Confidence>
            <indicator:Related_Indicators>
                <indicator:Related_Indicator>
                    <stixCommon:Relationship>Source of enrichment for Domains</stixCommon:Relationship>
                    <stixCommon:Indicator id="certuk:indicator-5f77d0b0-c032-48d7-907c-1633acaaadf4" timestamp="2015-12-
13T23:28:12.450296+00:00" xsi:type='indicator:IndicatorType'>
                        <indicator:Title>Related indicator wrapper for source of enrichment</indicator:Title>
                        <indicator:Type xsi:type="stixVocabs:IndicatorTypeVocab-1.1">Domain Watchlist</indicator:Type>
                        <indicator:Observable id="certuk:Observable-d6f201cd-e326-4797-b2be-1a4f5c93f509">
                            <cybox:Observable_Composition operator="OR">
                                <cybox:Observable id="certuk:Address-719b088c-e2cc-43fa-9bd3-f34264f7e57f">
                                    <cybox:Description>Source of enrichment for: certuk:Observable-5063b55a-601b-4361-a12d-
94f5daaccc28, </cybox:Description>
                                </cybox:Observable>
                                <snip></snip>
                            </cybox:Observable_Composition>
                        </indicator:Observable>
                        <indicator:Confidence timestamp="2015-12-13T23:28:12.450358+00:00">
                            <stixCommon:Value xsi:type="stixVocabs:HighMediumLowVocab-1.0">Medium</stixCommon:Value>
                        </indicator:Confidence>
                    </stixCommon:Indicator>
                </indicator:Related_Indicator>
            </indicator:Related_Indicators>
        </stix:Indicator>
    </stix:Indicators>
</stix:STIX_Package>
```

References

Bridges, R. A., Jones, C. L., Iannacone, M. D., Testa, K. M., & Goodall, J. R. (2014, June 9). *Automatic Labeling for Entity Extraction in Cyber Security*. Retrieved from Cornell University Library: http://arxiv.org/abs/1308.4941

Bro. (2015, December 1). *File Analysis Framework*. Retrieved from Bro Network Security Monitor: https://www.bro.org/sphinx-git/frameworks/file-analysis.html

Codenomicon. (n.d.). *AbuseSA Product*. Retrieved from Codenomicon: http://www.codenomicon.com/products/abusesa/

Dell SecureWorks. (2015, August 5). *Threat Group-3390 Targets Organizations for Cyberespionage*. Retrieved from Dell SecureWorks Cyber Threat Intelligence: http://www.secureworks.com/cyber-threat-intelligence/threats/threat-group-3390-targets-organizations-for-cyberespionage/

Denton, W. (2009, March 28). *How to Make a Faceted Classification and Put It On the Web*. Retrieved from Miskatonic University Press: https://www.miskatonic.org/library/facet-web-howto.html

domaintools. (n.d.). Retrieved from https://www.domaintools.com/

Farnham, G. (2013, October 14). *Tools and Standards for Cyber Threat Intelligence Projects.* Retrieved from SANS Reading Room: https://www.sans.org/reading-room/whitepapers/warfare/tools-standards-cyber-threat-intelligence-projects-34375

F-Secure. (2014, June 23). *Havex Hunts For ICS/SCADA Systems*. Retrieved from F-Secure Labs: https://www.f-secure.com/weblog/archives/00002718.html

Hutchins, E. M., Cloppert, M. J., & Amin, R. M. (2010). *Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains*. Retrieved from Lockheed Martin Reports.

Messina, C. (2015, August 25). *Groups for Twitter; or A Proposal for Twitter Tag Channels*. Retrieved from FactoryCity: http://factoryjoe.com/blog/2007/08/25/groups-for-twitter-or-a-proposal-for-twitter-tag-channels/

Mitre Corporation. (2015). *Incident vs Indicator*. Retrieved from STIX Project Idioms: http://stixproject.github.io/documentation/idioms/incident-vs-indicator/

MLSec Project. (2015, April 25). *MLSec Project*. Retrieved from Github Repository: https://github.com/mlsecproject/combine

mongoDB. (n.d.). *mongoDB Documentation*. Retrieved from https://docs.mongodb.org/manual/?_ga=1.104701184.1278358839.1405025465

OASIS. (n.d.). *OASIS Cyber Threat Intelligence Technical Committee*. Retrieved from OASIS Open Standards: https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=cti

Poputa-Clean, P. (2015). *Automated Defense - Using Threat Intelligence to Augment Security*, 2-24. Retrieved from SANS Reading Room.

python-stix. (n.d.). *python-stix*. Retrieved from GitHub: https://github.com/STIXProject/python-stix

Recorded Future. (2014, November 20). *Breaking the code on Russian Malware*. Retrieved from https://www.recordedfuture.com/russian-malware-analysis/

SNORT. (n.d.). Retrieved from https://www.snort.org/

Soltra Edge. (n.d.). *Soltra Edge*. Retrieved from https://soltra.com/

Sqrrl. (n.d.). *Sqrrl Enterprise*. Retrieved from http://sqrrl.com/product/sqrrl-enterprise/

UK Ministry of Defence. (2014). *Understanding and Intelligence Support to Joint Operations* (3rd ed.). Shrivenham: Development, Concepts and Doctrine Centre.

VirusTotal. (n.d.). Retrieved from https://www.virustotal.com/