Defence Research and Development Canada
Recherche et développement pour la défense Canada

DRDC | RDDC
technologysciencetechnologie

TA-35—Cyber Threat Data Model and Use Cases

Final Report

Dr. Antoine Lemay
International Safety Research (ISR)

Prepared by:
ISR
38 Colonnade Road North
Ottawa, Ontario
Canada K2E 7J6
Contractor's document number: ISR Report 6099-01-03 Version 2.0
PSPC Contract Number: W7714-156105-T35
Technical Authority: Melanie Bernier, Defence Scientist
Contractor's date of publication: September 2017

Defence Research and Development Canada

Contract Report
DRDC-RDDC-2017-C290
November 2017

Canada

IMPORTANT INFORMATIVE STATEMENTS

The information contained herein is proprietary to Her Majesty and is provided to the recipient on the understanding that it will be used for information and evaluation purposes only. Any commercial use including use for manufacture is prohibited.

Disclaimer: This document is not published by the Editorial Office of Defence Research and Development Canada, an agency of the Department of National Defence of Canada, but is to be catalogued in the Canadian Defence Information System (CANDIS), the national repository for Defence S&T documents. Her Majesty the Queen in Right of Canada (Department of National Defence) makes no representations or warranties, expressed or implied, of any kind whatsoever, and assumes no liability for the accuracy, reliability, completeness, currency or usefulness of any information, product, process or material included in this document. Nothing in this document should be interpreted as an endorsement for the specific use of any tool, technique or process examined in it. Any reliance on, or use of, any information, product, process or material included in this document is at the sole risk of the person so using it or relying on it. Canada does not assume any liability in respect of any damages or losses arising out of or in connection with the use of, or reliance on, any information, product, process or material included in this document.

This document was reviewed for Controlled Goods by Defence Research and Development Canada (DRDC) using the Schedule to the *Defence Production Act*.

TA-35 – Cyber Threat Data Model and Use Cases
Final Report

ISR Report 6099-01-03
Version 2.0
20 September 2017

Presented to:
Melanie Bernier
Defence Scientist, DGSTCO, Centre for Operational Research and Analysis
Defence Research and Development Canada / Government of Canada

Prepared by:

International Safety Research
38 Colonnade Road North
Ottawa, Ontario
Canada K2E 7J6

Abstract

This report documents the efforts related to the production of a data model based on the STIX 2.0 format to characterize cyber threats. The work produced four main outcomes:

- An analysis of the suitability of the STIX 2.0 standard to support the characterization of cyber threats;
- STIX 2.0 compliant data models to support automation or analysis;
- Profiles of Advanced Persistent Threat (APT) actors groups using the STIX 2.0 format; and
- Examples of exercise scenarios using the APT actor profiles to demonstrate use cases.

The main findings regarding the suitability of the STIX 2.0 standard are as follows:

- The standard is designed to represents threat information using graphs, with the various objects (threat actors, tools and malware, vulnerabilities, identities, etc.) modeled as nodes and the relationships between the objects represented as edges;
- The standard is designed around the concept of a minimum viable product, with a small number of rules and a large capacity for customization;
- The lack of enforced structures lends itself well to so-called “NOSQL” approaches, but makes automated processing more complex as the same information can be expressed in multiple forms; and
- The standard, at the time of writing, lacks in maturity with continually evolving documentation and only partial software support.

In terms of presenting a model, the report proposes to either embrace the unstructured nature of the standard in a NOSQL, or to enforce a certain structure to facilitate information retrieval. In the case of the NOSQL model, this would support the use of STIX 2.0 as a method to store indicators of compromise and provide some additional context in automated systems. In the case of the structured model, the use of predictable structures to store information would help analysts retrieve and cluster information, however at the cost of increased processing for storage.

Twelve different APT groups are profiled using a data model similar to the one proposed (due to lack of certain functionalities in the STIX 2.0 code base) covering a range of nation-state sponsors and a range of tools, techniques and procedures (TTPs). These profiles illustrate the breath of characteristics that can be modeled using STIX 2.0.

Finally, two exercise scenarios using the profiles demonstrate how cyber intelligence tasks would be performed. In the first scenario, participants perform a cyber-response, attributing the threat and extrapolating goals and potential impacts. In the second scenario, participants perform a cyber-intelligence planning process, generating tactical indicators and issuing warnings based on the threat level. The exercise description also provides indications on which particular object or object properties would be used in each step of the scenario.

Résumé

Ce rapport documente les efforts liés à la production d'un modèle de données basé sur le format STIX 2.0 afin de caractériser les cyber-menaces. Le projet a abouti à quatre éléments principaux:

- Une analyse de la pertinence de la norme STIX 2.0 pour soutenir la caractérisation des cyber-menaces;
- Des modèles de données conformes à STIX 2.0 pour prendre en charge l'automatisation ou l'analyse;
- Des profils des groupes d'acteurs APT (Advanced Persistent Threat) utilisant le format STIX 2.0; et
- Des exemples de scénarios utilisant les profils d'acteurs APT pour démontrer des cas d'application.

Les principales conclusions concernant la pertinence de la norme STIX 2.0 sont les suivantes:

- La norme est conçue pour représenter les informations sur les menaces à l'aide de graphiques, avec les différents objets (acteurs, outils et logiciels malveillants, vulnérabilités, identités, etc.) modélisés comme des nœuds et les relations entre les objets représentés comme des arêtes;
- La norme est conçue autour du concept de produit minimum viable, avec un petit nombre de règles et une grande capacité de personnalisation;
- Le manque de structures renforcées se prête bien aux approches dites «NOSQL», mais rend le traitement automatisé plus complexe car les mêmes informations peuvent être exprimées sous plusieurs formes; et
- La norme, au moment de la rédaction, manque de maturité avec une documentation en constante évolution et seulement un support logiciel partiel.

En termes de présentation d'un modèle, le rapport propose soit d'adopter la nature non structurée de la norme dans un NOSQL, soit d'imposer une certaine structure pour faciliter la recherche d'information. Dans le cas du modèle NOSQL, cela favoriserait l'utilisation de STIX 2.0 comme méthode de stockage des indicateurs de compromission et fournirait des informations supplémentaires dans des systèmes automatisés. Dans le cas du modèle structuré, l'utilisation de structures prévisibles pour stocker les informations aiderait les analystes à récupérer et à regrouper les informations, mais au prix d'un traitement accru pour le stockage.

Douze groupes APT différents sont profilés en utilisant un modèle de données similaire à celui proposé (en raison du manque de certaines fonctionnalités dans la base de code STIX 2.0) couvrant une gamme de parrains étatiques/nationaux et une gamme d'outils, techniques et procédures (TTPs). Ces profils illustrent la variété des caractéristiques qui peuvent être modélisées en utilisant STIX 2.0.

Enfin, deux scénarios d'exercices utilisant les profils de groupes APT démontrent comment les tâches de cyber-renseignement seraient exécutées. Dans le premier scénario, les participants effectuent une cyber-intervention, attribuant la menace et extrapolant les objectifs et les impacts potentiels. Dans le deuxième scénario, les participants exécutent un processus de planification de cyber-renseignement, générant des indicateurs tactiques et émettant des avertissements basés sur le niveau de menace. La description de l'exercice indique également quel objet particulier ou quelles propriétés d'objet serait à utiliser à chaque étape du scénario.

QUALITY ASSURANCE AND VERSION TRACKING

Authorization

Title	TA-35 - Cyber Threat Data Model and Use Cases Final Report	
Report number	6099-01-03	
Version	2.0	Signature
Prepared by	Dr. Antoine Lemay	OSB
Reviewed by	F. Kitching	
Approved by	I. Becking	
Approved for Corporate Release by	M. McCall	MMcCall

Version Tracking

Ver.	Action	By	Date
1.0	Release to Client	M. McCall	12 Sep 17
2.0			

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

This report was written as a reference document for the Defence Research and Development Canada (DRDC) Cyber Decision Making and Response (CDMR) project in support of the development of Canadian Armed Forces (CAF) computer network defence.

The report's primary author, Dr. Antoine Lemay, is a subject matter expert in the cyberspace domain with research experience in the study of cyber threats and attacks. Dr. Lemay identified and defined the data elements and structure required for a cyber-threat data model and developed use cases based on adversarial Tactics Techniques and Procedures (TTPs) that were used to populate the data model.

1.1 Objectives

The first objective of this report is to provide an evaluation of Structured Threat Information Expression (STIX) in the context of the high level model presented in the provided reference material [1]. Specifically it will:

1. Identify whether STIX can be used to implement the processes and characterization framework;
2. Identify and describe any gaps, and identify any alternative data structure as needed; and
3. Create the data model based on the results of 1 and 2.

The second objective of this report is to describe current Advanced Persistent Threat (APT) actors, their strategies and the tools they use in their normal operations. Finally, the third objective of the report is to present two scenarios for exercises that illustrate how the findings could be used in the context of cyber intelligence.

1.2 Scope

The scope of Section 2 of this document is limited to STIX 2.0, which is significantly different from its predecessor.

The scope of Section 3 is limited to current APT actors and does not speculate on potential future ATP actors within the cyber security field. It should be noted that, in the course of this work, the targets of the various APT groups was not addressed.

1.3 Contents

Section 1 contains a short introduction to the document.

Section 2 presents how STIX 2 could be used to model the threat component of *AD3 - Cyber Threat Data Model - High-level model and use cases* and how the implementation would tackle various use cases. This section starts by providing a brief overview of the STIX 2.0 standard. This is followed by the proposition for a STIX-inspired data model to represent cyber threats. Finally, a brief conclusion and suggestion for future work is provided.

Section 3 presents a summary of the Advanced Persistent Threat (APT) actors.

Section 4 presents two exercise scenarios that use the APT threat profiles to demonstrate the use cases described in section 2.

2. ANALYSIS OF THE STIX 2 FORMAT

As part of the Cyber Decision Making and Response (CDMR) project, work has been done to supplement automated cyber defence capabilities. One line of research is the investigation of how to leverage cyber intelligence for defence. This report covers the investigation of the STIX cyber threat intelligence sharing standard and the suitability of its data model to support the cyber threat model in development by DRDC.

2.1 Introduction to STIX 2.0

This subsection provides a brief overview of the STIX 2.0 standard. The contents of this subsection are based off the standard documentation available on the official STIX Github site [2]. This subsection starts by presenting an overview of what STIX is, then explains how STIX works. Finally, the subsection closes with a summary of the challenges and opportunities of using STIX in the context of intelligence analysis.

2.1.1 What is STIX?

STIX is a data format designed to standardize threat information in order to facilitate sharing and collaboration. As an example, let us consider an anti-virus company publishing research on a new virus from a known threat actor. If they publish the Indicators of Compromise (IoC) associated with this research in a PDF report, information consumers would have to retype the information in their information system or create a custom parser to automatically extract information from the report. Instead, if they publish the IoCs using the STIX format, the information could be imported via a generic STIX parser.

The ultimate goal of STIX is to enable the automation of tactical threat intelligence sharing, in particular IoC sharing. This goal is achieved through the standardization of data structures and the development of a limited common vocabulary. However, STIX is also developed to allow content providers to include contextual information by introducing a format to express relations between different objects. This makes the format more versatile than the use of a Comma Separated Values (CSV) file containing a list of indicators, the traditional method to share IoCs.

The STIX standard is accompanied by a sister protocol to exchange threat information stored in a STIX format. The Trusted Automated Exchange of Intelligence Information (TAXII) protocol enables the sharing of STIX information via HTTPS. While the use of TAXII is not required by STIX, STIX is optimized for use with the TAXII protocol.

While the standardization efforts have been started by the U.S. Department of Homeland Security (DHS), the STIX format is now under the guidance of the OASIS Cyber Threat Intelligence (CTI) Technical Committee. The standard is now in its second version.

2.1.1.1 *Design philosophy of STIX 2.0*

One of STIX's goals is standardizing message formats and developing a common vocabulary. However, this creates a dilemma. The more rigid the standard is, and the more keywords it uses, the less generally applicable it becomes. A particular required field might not be applicable for an indicator or the existing keywords might not be quite right to express a concept. This dilemma led to the design philosophy of STIX 2.0.

STIX 2.0 aims to be a standard where it is very hard to generate data that does not fit the standard, but very easy to import the data. As such, they have opted for a Minimum Viable Product (MVP) philosophy, meaning that the bare minimum is required by the standard. As such, a minimal adherence the STIX 2.0 standard represents the lowest common denominator in threat information sharing.

However, the standard is designed to be extensible, with a number of suggested best practices to include additional information and to structure the information according to a data model. Furthermore, the standard supports extensive customization of both the data model and the vocabulary. This extensibility ensures that threat intelligence providers are able to fully express the richness of their data, while keeping the requirements to produce STIX-compatible data minimal.

2.1.1.2 *Differences with STIX 1.0*

The design philosophy of STIX 2.0 is a departure from the design philosophy of the first version of STIX. As such, there are major differences between STIX 1.0 and STIX 2.0. These include:

- Streamlined model: STIX 1.0 had a wide range of features, most of which were never used or used improperly due to the complexity of the XML format and the presence of different standards for observable objects[1]. STIX 2.0 has streamlined the data model to standardize only the most commonly used features;
- Standard unification: In the first version of STIX, there was a different standard for threat information (STIX) and observations of threats (CybOX). STIX 2.0 directly includes the observations in the standard, reducing complexity;
- Use of JavaScript Object Notation (JSON): The first version of STIX used XML while STIX 2.0 uses JSON. This simplifies the use and declaration of the objects;
- Top level objects: Objects are no longer embedded within other objects, as was the case with the XML declaration of STIX 1.0, but are instead top level objects linked together with the new relationship objects;
- Relationship objects: Links between objects are now expressed via explicit relationship objects instead of being expressed through the XML structure; and
- Improved patterning: Data markings (e.g. marking data as confidential) no longer require a custom serialization specific language (the use of Xpath was required) and a specific language was developed to express indicator patterns (e.g. to express that observation of a specific hash is an indicator of compromise).

These changes are designed to increase ease of use and promote widespread adoption. This is done at the expense of a rigid data structure.

2.1.1.3 *Main use cases*

To determine the minimum viable functionality, a small number of use case examples have been developed, including:

[1] A more detailed justification for the move to a new standard can be found here: https://oasis-open.github.io/cti-documentation/stix/review

- Identifying a threat actor profile: Conveying threat intelligence about the identity behind a threat actor (e.g., the Mandiant exposé identifying APT1 as PLA unit 61398);
- Creating an indicator for malicious URL: Publishing a URL indicator of compromise for a particular malware (e.g. an anti-virus company publishing the URL of a malware command and control node);
- Creating an indicator for malware file hash: Publishing a hash indicator of compromise for a particular malware (e.g. an anti-virus company publishing the hash of the variant of a particular malware);
- Sighting of an indicator: A company disclosing that a particular indicator has been spotted on their network (e.g. a company uploading a suspicious file hash to check if it is malicious);
- Sighting of observed-data: A company reporting that is has successfully spotted an attack and the observation that made the spotting possible (e.g. a company disclosing that it has suffered a malware infection and adding the file hash of the particular malware that was used);
- Identifying threat actor leveraging attack pattern and malware: Conveying threat intelligence about the Tools Techniques and Procedures (TTPs) of a particular threat actor (e.g., the Mandiant exposé describing the toolset of APT and specific attack patterns used);
- Using marking definitions: Using custom labels to add labeling to information (e.g. adding custom labels to identify the threat level (green, orange, red) associated with a specific event); and
- Using granular marking: Using custom labels to add labeling to information contained within objects (e.g. labeling individual parts of an email indicator (e.g. from address, to address, presence of an attached file) with different threat levels).

STIX allows for a myriad of other use cases, but requires more customization than the main use cases that form the expected minimum functionality. As such, the main use cases can be used as a benchmark to estimate the amount of customization that is required.

2.1.2 How does STIX work

This subsection presents a summary of the inner workings of the STIX 2.0 format. It starts by describing its core object types, SDOs and SROs, then provides a more detailed description of the STIX data objects (SDOs) and STIX relationship objects (SROs). The subsection continues with a review of the STIX data model and of the tooling available to support STIX.

2.1.2.1 *SDOs and SROs*

The STIX data model can be interpreted as a directed graph. In that graph, the various pieces of information are the nodes of the graphs and the relationships between the objects are the edges of the graph.

The SDOs are the objects used to express the nodes of the graph. Each object represents a data point, be it an identity, a URL, an attack pattern or a campaign.

Depending on the specific type of object, the SDO will contain different properties. For example, a campaign might have properties for its name, description, start time, end time and objectives, while a malware object would have a malware type, name and label properties. The properties are the main method to store information. Also, for each type of object, some properties are required and others are optional, ensuring minimal functionality while enabling the inclusion of additional information.

The SROs are the objects use to express the edges of the graph. Each relationship object represents a link between two SDOs. For example, a SRO between a threat actor SDO and a malware SDO might indicate that the threat actor uses that particular malware while a link between an observed object and an identity might signify that the object was observed by that particular entity.

2.1.2.2 *Main STIX objects*

There are 12 types of SDOs defined in STIX. These are illustrated in Table 1. In addition, STIX supports the creation of arbitrary custom objects and the creation of bundle objects. A bundle object is a collection of related or unrelated STIX objects in a single JSON file. A custom object is an object that supports the common STIX properties (see below) and contains other arbitrary information.

Table 1: SDO Types

SDO Type	Definition	Example
Attack Pattern	Method used to attack a target.	Spear phishing.
Campaign	A series of connected attacks spanning a finite amount of time and covering a specific target list.	The series of intrusions by APT28 to infiltrate the U.S. elections.
Course of Action	An action (or series of actions) taken to present or stop an attack, mainly used to automate defenses.	Patching a specific vulnerability on a target machine.
Identity	The identity of an individual, group or organization.	John Smith or the Finance sectors are two examples of identities.
Indicator	Pattern that can be observed and that is indicative of malicious activity.	The presence of a file hash corresponding to malware or a specific sequence of network packets associated with a vulnerability are two different types of indicator patterns.
Intrusion Set	A group of adversarial behaviours and/or resources with common properties that is believed to be attributable to a single actor.	A suspected link arising from common TTPs, but from an unknown threat actor.
Malware	A program that is designed and used for malicious activity.	The PlugX backdoor is an example of malware.
Observed Data	An assertion that specific data was observed.	A file with a given hash has been found on a machine.
Report	A document collecting threat	The APT1 exposé.

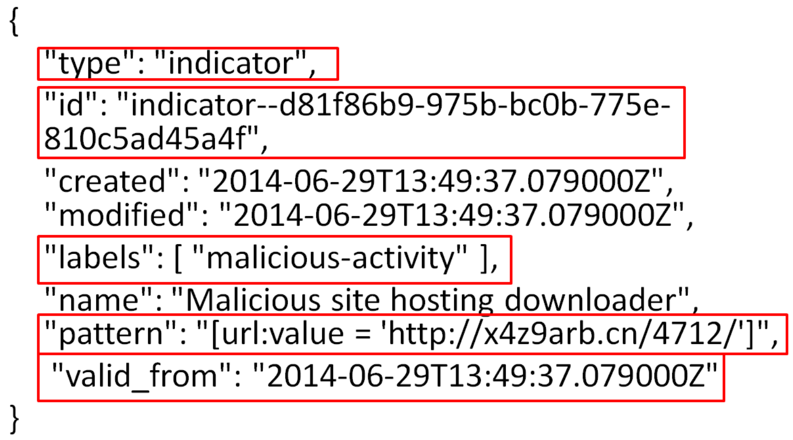
SDO Type	Definition	Example
	intelligence.	
Threat Actor	A threat actor is an entity engaging in malicious activity. This category is mainly a container for the characteristics of the entity, while the identity object is used to specify who they are.	The APT1 group is an example of a threat actor.
Tool	A legitimate program that is used for malicious activity.	The psexec tool, created by SysInternals for Windows administration, which is often abused by hackers, should be represented by a tool object.
Vulnerability	A software defect that can be used to gain access to a system or a network.	The CVE-2017-0143 Windows SMB code execution vulnerability.

In order to satisfy the requirements of the STIX standard, an object must support the following properties. Properties listed in red are required properties for all objects:

- **type**: Describes the object type;
- **id**: Provides a unique identifier for the object;
- created_by_ref: Is a reference to the identity of the object creator;
- **created**: Records the date and time of creation;
- **modified**: Records the date and time of modification;
- revoked: Indicates if the object has been revoked;
- labels: Is arbitrary labels associated with the object; each object type has a set of labels defined in the STIX specifications, but respect of this vocabulary is not mandatory;
- external_references: Links to external documentation describing the object;
- object_markings_ref: Is a reference to object markings objects; and
- granular_markings: Are markings related to other properties of the object.

It should be noted that the specifications of objects may contradict this general specification. As an example, the indicator object does not list the presence of the created and modified fields as required, but requires the presence of labels and adds new required fields for the pattern and date of validity. This illustrates a certain lack of maturity in the standard.

The SDO stores the information in a JSON structure containing a comma separated series of entries in the form of "field name":"field value". Figure 1 presents an example of JSON entry for an indicator SDO. The mandatory fields are identified by a red rectangle.



```
{
  "type": "indicator",
  "id": "indicator--d81f86b9-975b-bc0b-775e-
  810c5ad45a4f",
  "created": "2014-06-29T13:49:37.079000Z",
  "modified": "2014-06-29T13:49:37.079000Z",
  "labels": [ "malicious-activity" ],
  "name": "Malicious site hosting downloader",
  "pattern": "[url:value = 'http://x4z9arb.cn/4712/']",
   "valid_from": "2014-06-29T13:49:37.079000Z"
}
```

Figure 1: Example of an SDO

2.1.2.3 *Main STIX relationships*

STIX supports two main categories of SROs, relationships and sightings, which are defined as:

- Relationships: A link between two SDOs and the description of what the link is. As examples, a statement that APT28 (threat actor SDO) uses (link description) powershell (tool SDO) or that APT1 (threat actor SDO) targets (link description) the aerospace sector (identity SDO) are relationships; and
- Sightings: A sighting is a link between observed data as well as what the observed data represents and the source that observed it. For example, a statement that ACME bank (identity SDO) sighted (relation) the Sednit malware (malware SDO) in the form of a connection to a known command and control URL (URL indicator SDO) is a sighting.

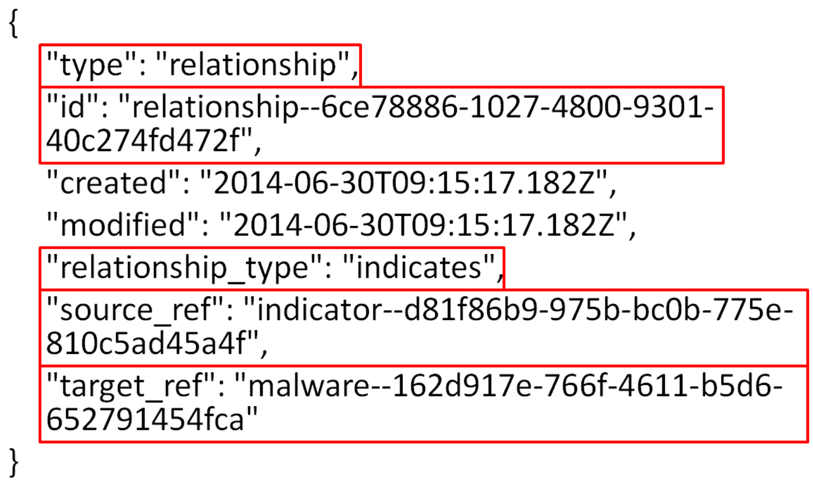
The links are established by referencing objects via their object ID. The links are directional, based on the descriptive vocabulary. In addition to sightings, the STIX standard defines a limited relationship vocabulary. The use of the vocabulary is not compulsory and arbitrary custom relationship keywords can also be used. Table 2, extracted from the standard, summarizes the types of relationships that are suggested by the STIX standard.

Table 2: Standardized relationships in STIX [2]

Source	Type	Target	Source	Type	Target
attack-pattern	targets	vulnerability	intrusion-set	attributed-to	threat-actor
attack-pattern	targets	identity	intrusion-set	targets	identity
attack-pattern	uses	malware	intrusion-set	targets	vulnerability
attack-pattern	uses	tool	intrusion-set	uses	attack-pattern
campaign	attributed-to	intrusion-set	intrusion-set	uses	malware
campaign	attributed-to	threat-actor	intrusion-set	uses	tool
campaign	targets	identity	malware	targets	identity
campaign	targets	vulnerability	malware	targets	vulnerability

Source	Type	Target	Source	Type	Target
campaign	uses	attack-pattern	malware	uses	tool
campaign	uses	malware	malware	variant-of	malware
campaign	uses	tool	threat-actor	attributed-to	identity
course-of-action	mitigates	attack-pattern	threat-actor	impersonates	identity
course-of-action	mitigates	malware	threat-actor	targets	identity
course-of-action	mitigates	tool	threat-actor	targets	vulnerability
course-of-action	mitigates	vulnerability	threat-actor	uses	attack-pattern
indicator	indicates	attack-pattern	threat-actor	uses	malware
indicator	indicates	campaign	threat-actor	uses	tool
indicator	indicates	intrusion-set	tool	targets	identity
indicator	indicates	malware	tool	targets	vulnerability
indicator	indicates	threat-actor			
indicator	indicates	tool			

The SRO stores the information in a JSON structure containing a comma separated series of entries in the form of "field name":"field value". Figure 2 presents an example of a JSON entry for a relationship SRO. The mandatory field as specified by the object description are identified by a red rectangle.



```
{
  "type": "relationship",
  "id": "relationship--6ce78886-1027-4800-9301-
  40c274fd472f",
  "created": "2014-06-30T09:15:17.182Z",
  "modified": "2014-06-30T09:15:17.182Z",
  "relationship_type": "indicates",
  "source_ref": "indicator--d81f86b9-975b-bc0b-775e-
  810c5ad45a4f",
  "target_ref": "malware--162d917e-766f-4611-b5d6-
  652791454fca"
}
```

Figure 2: Example of an SRO

2.1.2.4 *STIX data model*

The pre-defined relationships form an implicit data model that is used in STIX 2.0. The explicit data model from the documentation is presented in Figure 3.

In the figure, only the "variant-of" relationship is not represented. As can be seen from the figure, most of the effort deployed in the STIX data model is to describe the threat side (adversaries and TTPs sections) while limited effort is made to describe the victim side (identity, vulnerability, course of actions and sightings).

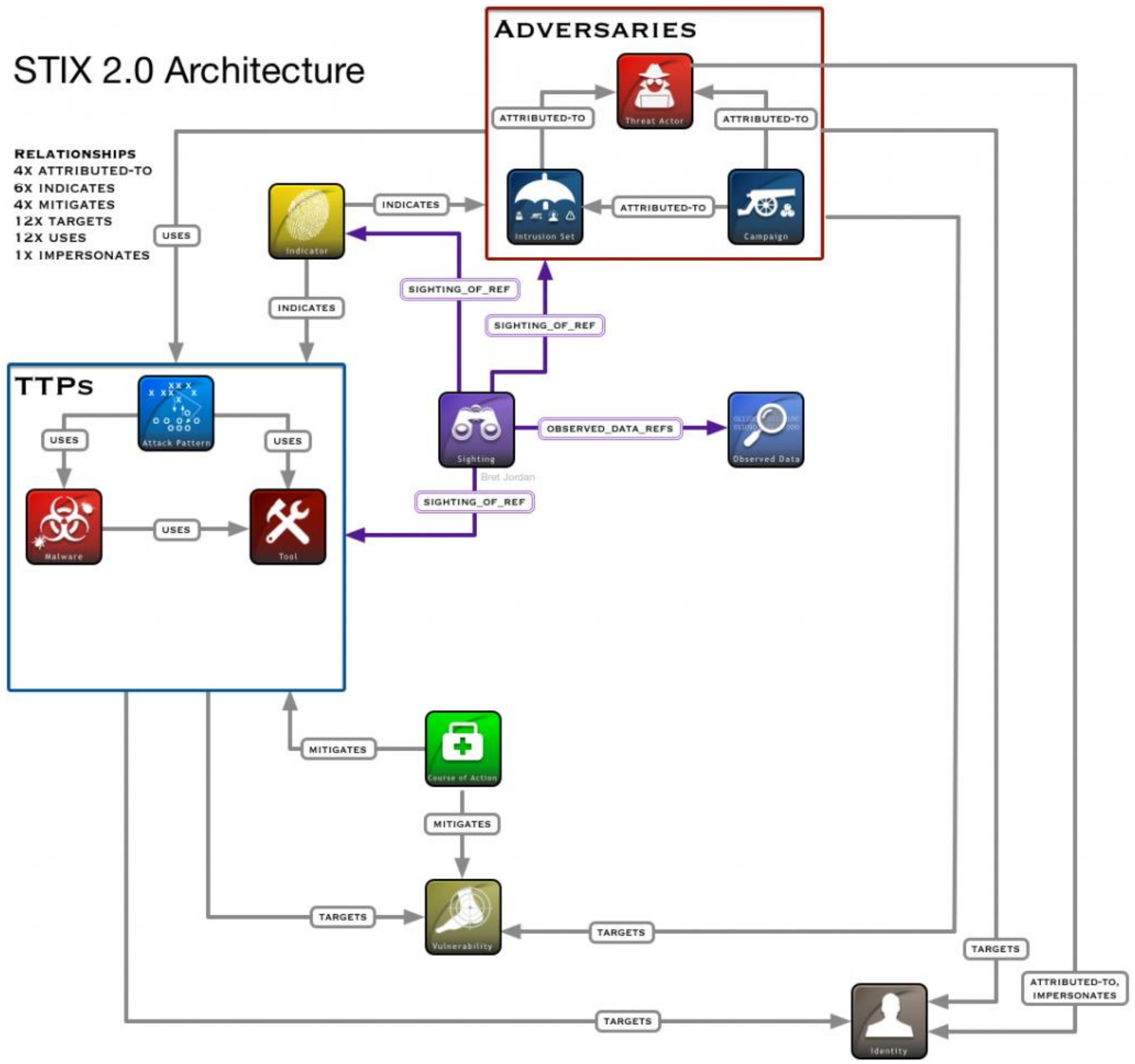


Figure 3: Explicit STIX 2.0 data model [2]

2.1.2.5 *STIX tooling*

The standardization team also provides a number of open source tools to support the adoption of STIX. However, it should be noted that, while these tools are offered in the official STIX Github, they are not authoritative and may not even be fully compliant with the standard. The following tasks are supported by official tools:

- Object generation: generate STIX compliant objects and JSON files;
- Visualisation: visualize STIX objects in a graphical environment. It should be noted that the tooling provided does not properly interpret arbitrary STIX and even some of the official examples are not supported;
- Pattern-matching: support for the indicator pattern language; and
- STIX elevator: convert files from the STIX 1 XML format to the STIX 2.0 JSON format.

2.1.3 Challenges and opportunities

There are many opportunities that arise from using STIX as an information model. There are, however, challenges as well. This subsection starts by presenting the opportunities, then present the challenges associated with working with STIX 2.0.

2.1.3.1 *Opportunities*

Opportunities that arise from using STIX include:

- Can benefit from threat information exchange: One of the more compelling reasons to align with the STIX format is the ability to draw information from a number of STIX providers. The use of these third party providers enhances the information gathering capabilities of the organization;
- Can incorporate more than just indicators: STIX offers the ability to be more expressive, by enabling the ability to express not only indicators, but the relations between them. This capability is critical for clustering and graph-based analysis; and
- Standard is actively supported: STIX is being actively supported by OASIS-CTI which provides guidance, updates the standards and promotes adoption of the standard. Furthermore, there is an active community that produces open source tools and libraries to help users integrate STIX in their environment.

2.1.3.2 *Challenges*

Challenges that result from using STIX include:

- Lack of maturity of the standard and tooling: The revision of the standard to the 2.0 mark is fairly recent and the standard is currently being revised. The official standards documents, at the time of this writing, have been versioned as working drafts. The last revision was June 19, 2017. As such, the standard may change in the future, which may negatively impact solutions developed with the current version of the standard.

  A related problem is the lack of maturity for the tooling. The current official version does not guarantee standard compliance and some may not even work. As an example, the STIX visualisation tool does not currently work with some of the examples from the constructor tool. Also, not all of the supporters of the STIX/TAXII standard have moved to the 2.0 standard. As an example, Alien Vault has made their OTX threat exchange service a STIX/TAXII server in April 2017 (after the release of the 2.0 standard), but used STIX version 1.

- Lack of object unicity: One of the most important problems of the use of STIX as a data model for analysis rather than transport is that there is no guarantee of object unicity. The unique identity of the object is reflected in the object ID. The object ID is randomly generated at the time of the object creation by the threat intelligence provider. As an example, if Mandiant generates information on APT1 and Alien Vault generates information on APT1, the two APT1 threat actor objects will have different IDs. In fact, the same threat actor may even have two different IDs in two different reports by the same intelligence providers. This greatly increases the complexity of graph analysis, one of the main benefits of using STIX, as each bundle is essentially an unconnected sub-graph. Additional post-processing is required to reconnect the elements coming from disparate sources;

- Unreliability of patterns: Because the object ID cannot be relied upon to reference objects as the same object can have multiple IDs, it is necessary to perform matching on other object properties, such as name or tag. Similarly, when attempting to perform graph-based analysis, the edges of the graph only exist if the relevant relationship object is included. As the standard does not specify a mandatory formal syntax for declaring objects, when performing processing, we cannot reliably expect these objects to be present. For example, a provider may have left the name field empty or linked all indicators directly to a threat actor eschewing campaigns and attack patterns.

  Let us consider the following examples. Threat actor A employs the PlugX malware. One intelligence provider may report that fact as "PlugX indicates threat actor A" while another may report it as "actor A uses PlugX". Similarly, a provider may report on the group APT28 with the name object property set to "APT28", while another may use the name "Fancy Bear". This problem is particularly complex when dealing with public reporting of advanced persistent threat actors as there is no naming convention used by the industry. For example, the group APT28 may be referred to by one of the groups' numerous aliases (e.g. APT28, Fancy Bear), by some of the tools it uses (e.g. Sednit, Sofacy) or by the name of some of their campaigns (e.g. Pawn Storm). As these names are created for marketing purposes, different threat intelligence providers may not mention the other aliases of the group or may not even be aware of them. As such, post-processing is required to normalize patterns if any sort of pattern analysis or graph-based analysis is to be performed; and

- No database support: The JSON approach adopted by the STIX standard is the preferred model for the non-relational databases that are commonly used in web development today. These types of database eschew a rigid and predictable format in exchange for agility, scalability and speed. Considering the volume of indicators of compromise generated by threat intelligence providers, this appears to be a reasonable decision. However, there is currently no official support for any type of database format and no tooling to help support this. At the same time, the graph-based data structure makes it ill-suited to be used in a relational database. While these characteristics are in-line with the transport-oriented design, it makes more complex the use of STIX as a data model for storing data in a database.

2.2 Proposed Models

This subsection discusses how STIX could be used to implement a cyber-threat data model. It starts by presenting a model which best leverages the current STIX ecosystem. Then, a model that more closely follows the cyber threat model proposed by DRDC [1] will be described.

2.2.1 STIX-based model

In the default use of the STIX/TAXII stack, threat intelligence data from multiple providers is stored in a NoSQL database without any particular structure. When an incident occurs, the organization creates a "Sighting" event that contains the data that was observed. Observable data from the sighting event is then compared to indicators in the threat database using STIX pattern matching language. The machine then outputs contextual information related to threat actors based on relationships present in the database. Figure 4 illustrates the process.

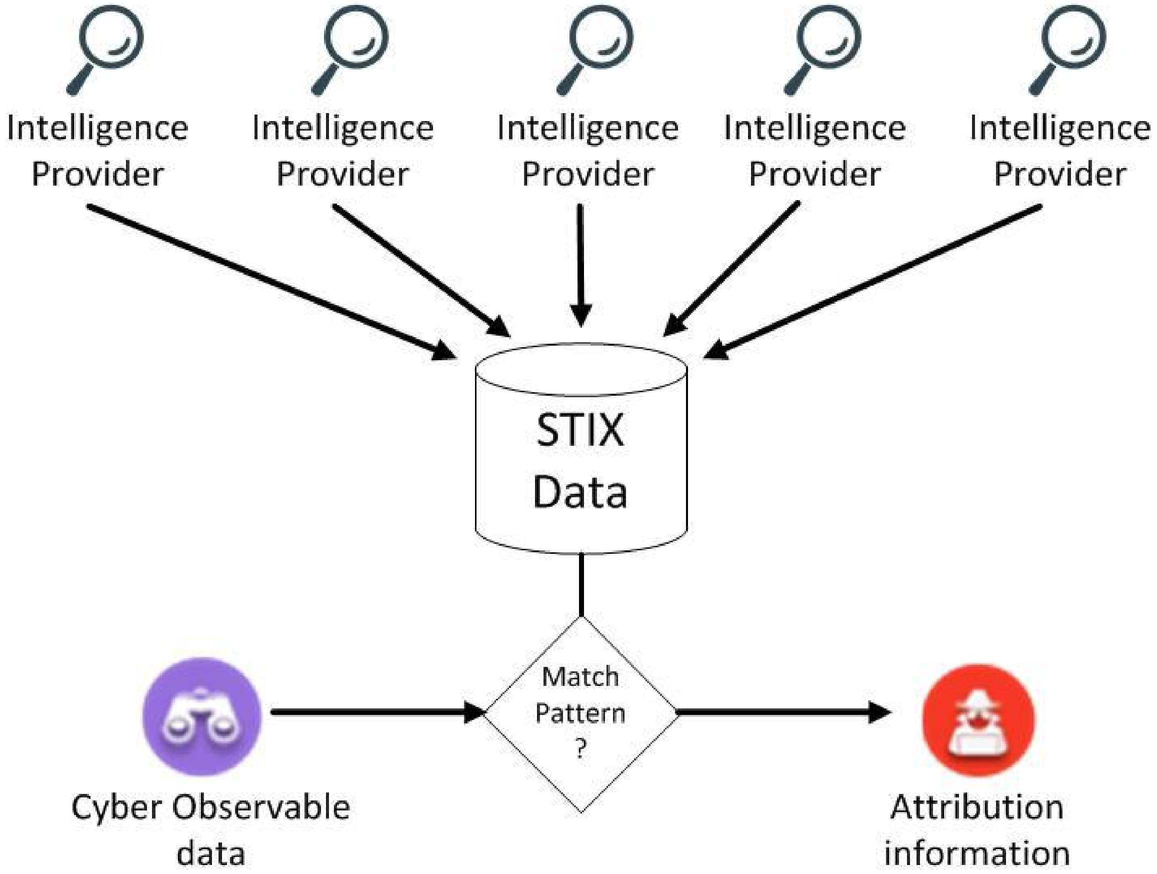


Figure 4: Unstructured model

The pseudo-code illustrating the process would be as follows:

1. Extract Campaign/Intrusion Set/Threat actor/Attack Pattern/Malware/Tool objects referenced by "sighting_of" references with the Sighting object ID as source;
2. Extract Observed_data objects referenced by "observed_data_ref" relationships with the Sighting object ID as source;
3. For each Observed_data object, find Indicator objects in the database that match the Observed_data;

4. For each Indicator object found, find Campaign/Intrusion Set/Threat actor/Attack Pattern/Malware/Tool objects referenced by "indicates" relationships with the Indicator Object ID as source; and
5. Output Campaign/Intrusion Set/Threat actor/Attack Pattern/Malware/Tool objects found.

This application is fairly straightforward, leveraging the more commonly used elements of the STIX data model. In that sense, it is easily automatable as it does not require any manual intervention on the part of analysts, works with the raw data in the form provided by intelligence providers and does not rely on any data structure that is not recommended for use in the standard. However, this data model supports an application that is mainly reactive in nature and which provides minimal context to tactical intelligence. While it could technically be used in a more proactive manner, the lack of a mandatory data format would not guarantee results, even if the data is present in the database. As an example, an analyst investigating a spear-phishing event might overlook email indicators that are not properly filed under the phishing attack pattern, but are instead directly linked to a threat actor. Also, it is likely that the various intelligence providers will generate the information independently from each other. This would mean that similar objects will have different objects ID, custom naming conventions and an inconsistent use of non-mandatory fields and vocabulary. This would make clustering and graph-based analysis very difficult.

2.2.2 Analysis-based model

In this data model, we use the basic STIX data model as an inspiration to create a more rigid data model that is aligned with intelligence tasks. It could be argued that any data model that does not fit STIX exactly requires data transformation to ensure that data matches the model. However, by closely aligning the data model with STIX, the differences between the normalized data and the raw data from the intelligence provider would be smaller. This may simplify the normalization process.

If we follow closely the STIX model, we will obtain a data model that is focused on the threat component. When looking at the cyber threat model proposed by DRDC, illustrated in Figure 5, it will provide a data model that is mostly concerned with the "Adversary" component.

However, the inclusion of additional data in the form of relationships included in STIX will enable a partial modelling of some of the relationships illustrated in Figure 5.

.

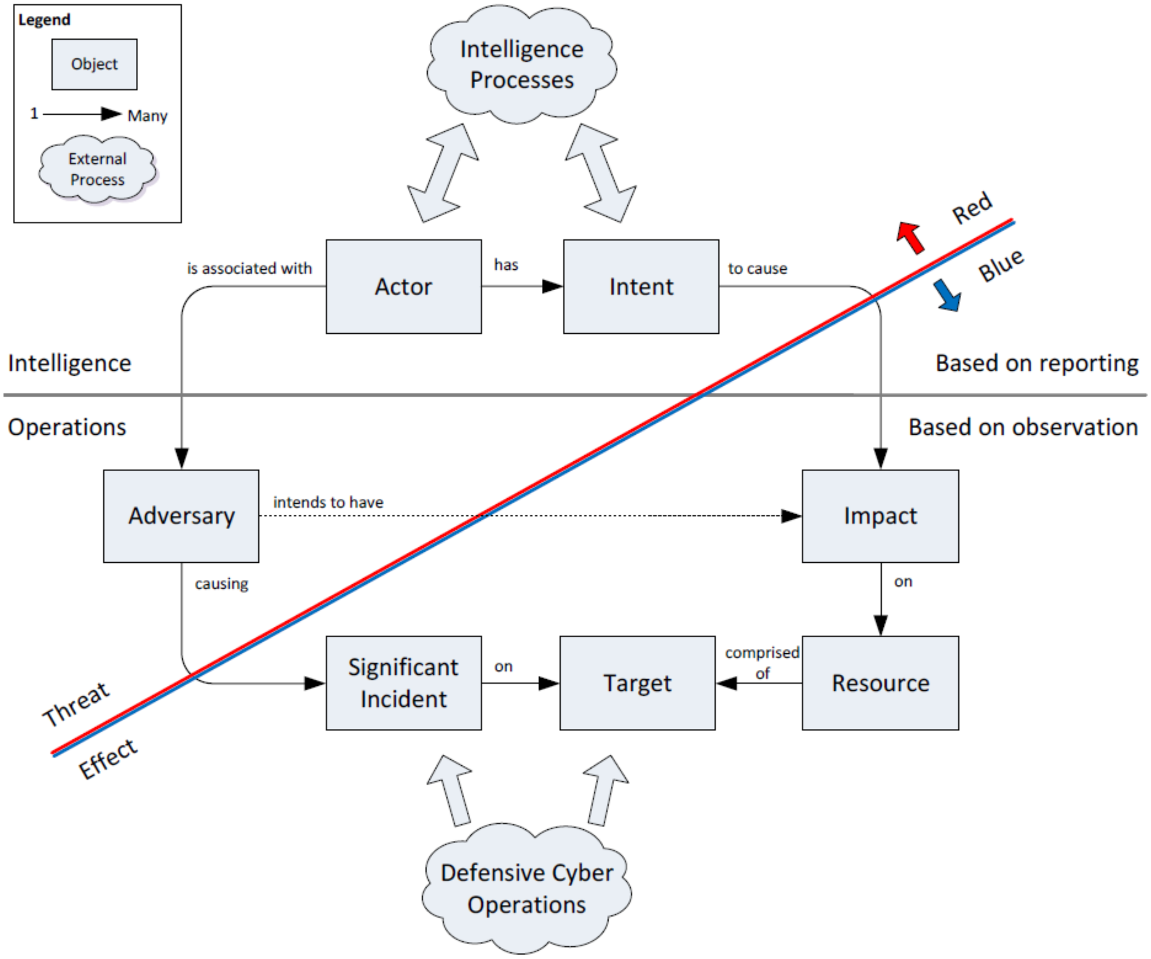


Figure 5: DRDC proposed cyber threat model [1][2]

2.2.2.1 *Modifications to the STIX default model*

In order to maximize the usefulness of the model for intelligence while preserving the interoperability of STIX, we want to remain standard compliant. In other words, we want to create a model that, if exported in a JSON format, could still be interpreted as STIX data. However, because we want more structured data to perform analysis, we do not expect arbitrary STIX data to conform to our specifications. To achieve this result, we will add a number of mandatory constraints to the existing STIX standard, as illustrated in Figure 6.

[2] The diagram is used with permission from the original authors

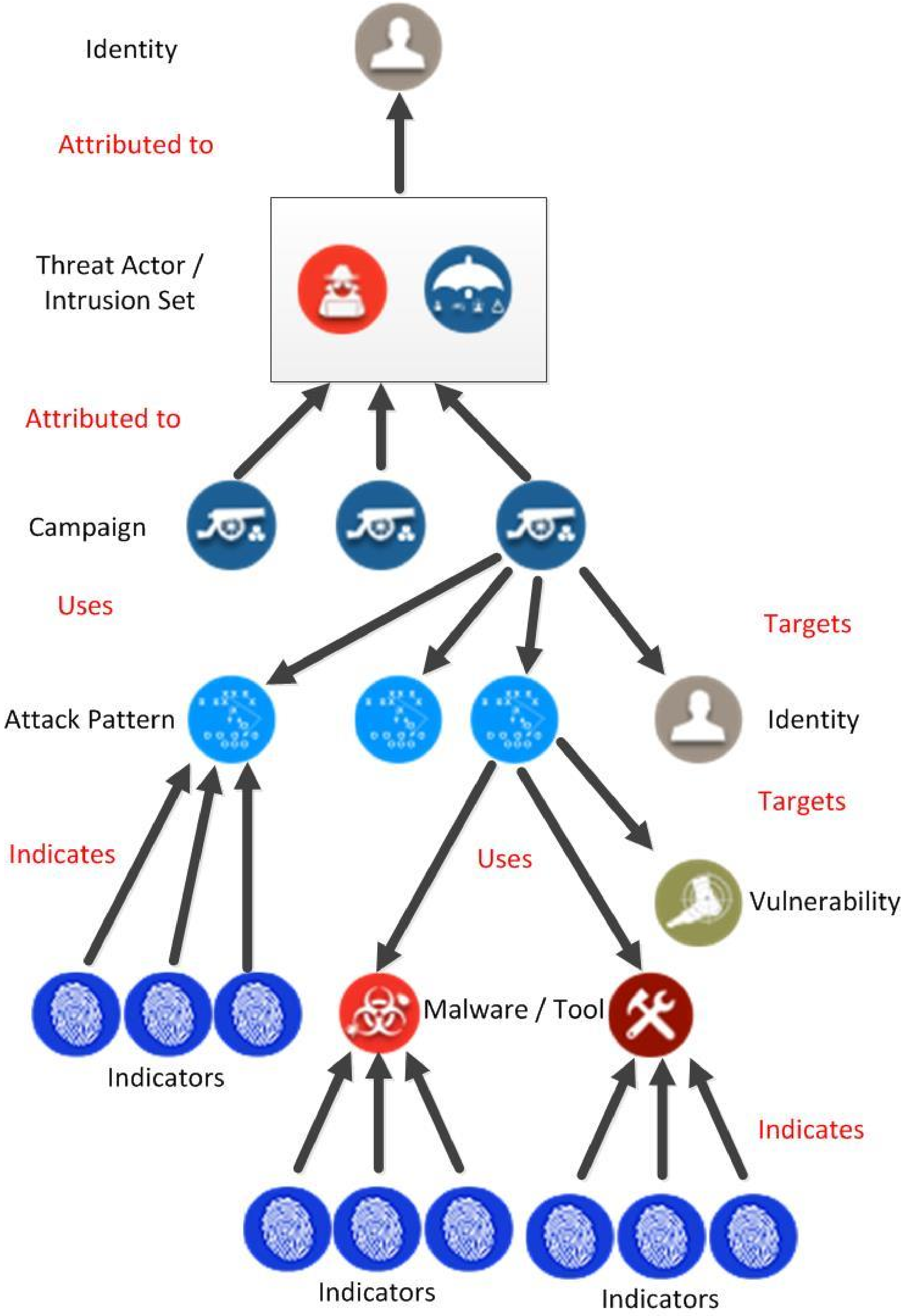


Figure 6: Proposed object structure

The following constraints would be added:

- Enforce unicity for the adversary SDOs. There should not be two objects referring to the same threat actors, intrusion set or campaign;
- Use intrusion set objects only for unknown threat actors. As threat actors and intrusion set objects are very similar, we propose the use of intrusion set objects to represent as of yet unattributed intrusions following a similar pattern. The intrusion set object should be deleted when the campaigns are attributed to an actor;
- Enforce unicity of attack patterns on a per campaign basis. While attack patterns can be replicated across multiple campaigns, there should not be two objects referring to the same attack pattern for a given campaign. If no attribution to a particular campaign exists, we recommend the creation of umbrella campaigns to act as a placeholder for unattributed activity;
- Enforce unicity of malware and tools on a per attack pattern basis. While the use of specific tools and malware can be replicated in multiple attack patterns, there should not be two objects referring to the same tool or malware for a given attack pattern;
- Enforce unicity of identity, vulnerability and indicators SDOs. As these represent specific instances of individuals, observable patterns or software defects, there should not be two objects referring to the same identity, vulnerability or indicator;
- Make the "aliases" property of threat actor objects mandatory (may be an empty list);
- Make use of the optional field "goals" of threat actor objects to record intent;
- Create custom properties for tool and malware objects to store information regarding the capabilities provided by the tool. This custom properties would follow the "impact" data model vocabulary described in the proposed DRDC cyber threat model [1]; and
- Create a mandatory hierarchy for relationships. To avoid dealing with relationship chains of arbitrary length, enforce a tree-like structure with threat actors at the top for objects. Figure 6 illustrates the proposed structure.

These modifications will facilitate clustering and graph analysis while remaining compliant with the STIX 2.0 standard. This would enable the export of this carefully structured data, used for intelligence analysis, to automated systems, including a system developed following the model described in section 2.2.1.

2.2.2.2 *Implementation of basic use cases*

Using the proposed modifications to the STIX data model, we can provide pseudo-code for some of the expected use cases.

The standard use case is the case of an analyst investigating an incident and wanting to provide attribution for the incident.

If the analyst is dealing with tactical information based on observed data, the analyst can follow the same process as the one described in section 3.1, with some slight modifications. The pseudo-code is as follows:

1. Extract Campaign/Intrusion Set/Threat actor/Attack Pattern/Malware/Tool objects

referenced by "sighting_of" references with the Sighting object ID as source;
2. Extract Observed_data objects referenced by "observed_data_ref" relationships with the Sighting object ID as source;
3. For each Observed_data object, find Indicator objects in the database that match the Observed_data;
4. For each Indicator object found:
    a. Find Attack Pattern/Malware/Tool objects referenced by "indicates" relationships with the Indicator Object ID as source;
    b. For each Malware/Tool object found, find Attack Pattern referenced as the source of "uses" relationship with the Malware/Tool object as destination and add to the attack patterns found;
    c. For each Attack Pattern found, find Campaign referenced as the source of "uses" relationship with the Attack Pattern object as a destination and add to campaign found;
5. For each Campaign found, find Intrusion Set or Threat actor referenced as the destination of "attributed to" relations with the Campaign object as Source; and
6. Output Intrusion Set/Threat actor objects found.

If the analyst is dealing with more abstract information, for example because the artefacts found do not correspond to any previously observed patterns, it is possible to use the data for clustering analysis.

As an example, let us consider the following incident. A piece of software was found on a machine, but the file hash does not match anything in the database. After some reverse engineering of the software, it is revealed that it is a form of rootkit. The following pseudo-code would represent the type of request performed by an analyst.

1. Extract all Attack pattern objects from the database where the name field includes "rootkit";
2. For each Attack Pattern object, find Campaign referenced as the source of "uses" relationship with the Attack Pattern object as a destination and add to campaign found;
3. For each Campaign found, find Intrusion Set or Threat actor referenced as the destination of "attributed to" relationships with the Campaign object as Source; and
4. Output Intrusion Set/Threat actor objects found. If the analyst then wants more information on the possible intent of the threat actors, the following additional steps could be added; and
5. For each Intrusion Set/Threat actor object found, output the content of the "goals" property.

Another possible use case for an intelligence database is the generation of indicators and warnings. Notably, as part of the Operation Planning Process (OPP), the generation of a series of indicators for enemy courses of actions (CoA) is required to build a response matrix [3].

Let us consider the following example. A force of Canadian troops are deployed as a peace keeping mission in Ukraine. The commander is worried about external interference from Russia, especially in his command and control systems. Intelligence is tasked with creating tactical indicators based off this mission statement. The following pseudo-code would enable the completion of that task:

1. Extract all Threat actors/Intrusion Sets that are referenced as the source of "attributed to" relationships with the "Russia" identity object as destination;
2. For each Threat Actor/Intrusion Set, find campaigns that are referenced as destination of "uses" relationship with the current actor object ID as the source;
3. For each Campaign, extract all Attack Patterns that are referenced as destination of "uses" relationship with the current campaign object ID as the source;
4. For each Attack Pattern, extract all Malware/Tools that are referenced as destination of "uses" relationship with the current Attack Pattern object ID as the source AND that contains "disruption" in its custom "capabilities" property, then extract indicators associated with the Malware/Tools found or their parent attack pattern; and
5. Output results.

Additional filtering could be made based on "targets" relationships to focus on campaigns that were previously used to target command and control networks and the TTPs associated with them.

In the case of the warnings use case, let us consider the following example. An incident occurred in a government agency, revealing a new espionage campaign from a group known as FUNKY BEAT. Based on the observation report, the analyst wants to send an alert to all potential victims. The following pseudo-code would enable the completion of the task:

1. Extract Threat actor object referenced by "sighting_of" references with the Sighting object ID as source;
2. Find campaigns that are referenced as destination of "uses" relationships with the current Threat Actor object ID as source;
3. For each campaign, find Identity objects that are referenced as destination of "targets" relationships with the current Campaign object ID as source; and
4. For each Identity found, send a message to the recipient identified in the "contact_information" (if present) property containing the warning.

Additional use cases could also be developed following a similar pattern.

2.3 Conclusion and future work

The STIX data model is a versatile model that can be adapted to a number of contexts. As such, it is possible to adopt a STIX compliant data model to support the cyber threat data model proposed by DRDC. Unfortunately, the various intelligence providers are unlikely to produce data that is fully compliant with the model, making it unsuitable for the type of unsupervised automation that was envisioned in the STIX standard. In that sense, we have proposed a model that supports unsupervised automation, but uses unstructured data, as well as a model that uses structured data, but does not support unsupervised automation. Data exported from the structured model could even be used as a data source to supplement the unstructured model.

Even though the STIX 2.0 standard contains a reasonable threat model, the standard is still evolving. As such, it is probably preferable to delay adoption until greater maturity is reached by the standard and its tooling support. However, additional research in related areas could improve the utility of the STIX standard in support of automated cyber intelligence analysis. In particular, the investigation of optimal use of NoSQL-type databases and best practices for developing this sort of data model would be interesting for the use of the vanilla version of STIX 2.0. Similarly, the investigation of semantic and graph database structures, which allow queries that return answers to semantic statements or graph-based queries, would allow for maximum exploitation of the relationship objects, one of the value-added element of the STIX format.

3. APT ACTORS

3.1 Chinese APT Actors

3.1.1 Comment Crew

The Comment Crew (a.k.a. APT1, Comment Panda, and The Shanghai Group) is a threat actor associated with China. It has been specifically attributed to the People's Liberation Army (PLA) unit 61398 by Mandiant. The threat actor seems primarily tasked with cyber espionage, but the attack of critical infrastructure targets for apparent battlefield shaping purpose alludes to a sabotage mission as well. One main campaign is documented: Operation ShadyRAT is an espionage campaign that is possibly related to the Comment Crew. However, a large volume of espionage activities in unnamed campaigns are documented in the Mandiant Report. Similarly, unnamed activities targeting critical infrastructure (CI) are documented by TrendMicro.

The main attack pattern used for internal compromise is spear phishing. This is often done in conjunction with email impersonation (impersonating a known sender) to entice victims to open attachments. The spear phishing will usually trigger the installation of the WEBC2 malware, a malware that hides its command and control communications in HTML comments (hence the namesake Comment Crew), but the HACKSFACE malware was also observed in the CI attack.

Once a foothold is established, the attackers will download additional tools on the machine to steal credentials or to load a more powerful Stage 2 implant that enables direct control of the machine or advanced espionage functionalities, such as capturing keystrokes and taking screenshots. Numerous tools that are either fully custom or that have been customized from publicly available hacker tools support this function.

With the help of stolen credentials, the group will continue access or expand access in the victim network. In addition to the access provided by the numerous backdoors, the Comment Crew will also abuse legitimate services to maintain access. In particular, web portals (including Outlook Web Access), VPN and remote desktop tools are used for continued access with legitimate credentials. In terms of lateral movement to expand access, the use of native windows task scheduling ("at" command) and SysInternal psexec are used with stolen credentials to install implants on other machines.

The use of native tools is also the primary vehicle for internal reconnaissance (i.e., for the fingerprinting of compromised hosts to determine their system specification and what level of network access they have). The group also uses native tools to hide their traces on Windows systems (e.g., through the deletion of Windows prefetch files in the CI attacks). The group also appears to want to avoid attribution by obfuscating their source IP. This is primarily done via the HTRAN tool that is installed on an intermediary machine to act as a proxy that obscures the attacker's source IP.

A mix of native tools and custom malware is used to exfiltrate data in the course of their espionage mission. The RAR file compressor is typically used to create password protected bundles of stolen information for exfiltration and the file transfer protocol (FTP) is used if available. Custom malware to export data to Google docs can be used if FTP is not available and most of the group's backdoors have the ability to exfiltrate files. A group of custom utilities to gather data from internal sources (e.g., PST files, mail servers or internal web servers) round out the group's arsenal.

3.1.2 Shell Crew

The Shell Crew (a.k.a. Deep Panda, WebMasters, KungFu Kittens, SportsFans, PinkPanther, and the Black Vine espionage group) is an APT group with ties to the Chinese Government, in particular through a Nanjing University nexus. The group is mainly involved in cyber espionage, participating in a number of well documented campaigns. In particular, a campaign targeting health insurance providers in the U.S. (e.g., Anthem, OPM) and a campaign targeting foreign policy think tanks have been detailed.

The primary attack pattern for gaining a foothold in the system used by this group is web compromise. Using a vulnerability present in the website, the group uploads one of their numerous webshells (hence their namesakes), which are scripts that enable remote access through a web interface. Using these stage 1 webshells, the group can then upload more traditional backdoors or additional tools required to perform lateral movement.

The group has also used other methods to infiltrate systems. For instance, their ability to hack web servers has enabled them to use watering hole attacks, compromising a public web server that they suspect their intended victim might visit. The compromised server is then used to host client-side exploits, including some 0-day exploits that are suspected to come from a developer supplying multiple Chinese APT groups (e.g., the Elderwood project). In the case of the health insurance provider campaign, they are suspected of using spear phishing instead of their traditional techniques which, presumably, had proven ineffective.

In contrast to groups relying exclusively on spear phishing, who must rotate their tools frequently to avoid detection, the Shell Crew relies on a relatively small number of attack tools, mainly webshells, backdoors and credential stealers. Crowdstrike has documented how they tend to rely on native Windows tools for lateral movements in their description of the campaign against think tanks. The groups appears to favor deploying powershell scripts through the Windows Management Instrumentation Console (WMIC), relying only on custom tools when these functionalities are not available. This is done using valid credentials stolen in the initial phases of the attack with credential stealing tools such as Mimikatz. Innocuous programs such as 7-zip and cftmon (compression utilities) are used to package data that will be exfiltrated.

To complement their dedication to increase the stealth of their operation from the use of hard to detect webshells and the reliance on native tools, the group also makes use of source obfuscation infrastructures. In particular, the group uses the HTRAN tool and the Terracotta VPN infrastructure to hide the source IP of their attacks.

3.1.3 Naikon

The Naikon group (a.k.a. APT30) is a group associated with China, in particular with the Kunming Technical Reconnaissance Bureau (TRB) designated as unit 78020. The group appears to be focused on South East Asia, with a particular focus on Chinese interests in the South China Sea. The group is less technically sophisticated than other groups, relying mostly on tools and exploits available on the Chinese underground, but is well organized.

The primary method of infection from the Naikon group is spear phishing. The group targets office vulnerabilities for which exploit code is widely available to trigger the download of their backdoor or simply uses social engineering to trick users into clicking on documents. They have notably used right-to-left-override (RTLO) tricks to hide file extensions or even just added a large number of blank spaces in the file name to camouflage the extension out of the visible area. The spear phishing technique has been used over the years to distribute versions of one of Naikon's backdoors that is customized for the victim.

The other method of propagation for the Naikon group is the distribution over removable media (e.g., USB). The group appears to have developed custom tools design to ferry malware updates and stolen documents from air gapped networks. However, only the SHIPSHAPE backdoor and the accompanying SPACESHIP and FLASHFLOOD tools appear to be used for that purpose.

Once the group has a toehold on the network, the group will perform internal reconnaissance of the victim network, steal passwords and perform lateral movement using a series of tools derived from the Chinese underground codebase. Naikon will also use a large number of freeware programs and default Windows utilities to perform the same tasks. Kaspersky hypothesize that the transition to legitimate tools is an evolution is due to the high level of detection of the Chinese underground malware previously used by the group. Instead of trying to develop anti-virus evasion techniques, the group evolved to software that does not trigger detection because it is recognized as legitimate software. Even then, the tools rely on relatively primitive command line utilities or SysInternal tools (e.g., netstat, psexec) instead of the more advanced WMI or Powershell tactics used by their peers.

Other evidence of the group's evolving tactics is that they are beginning to use the malware staging attack pattern. Instead of sending their backdoor directly as an email attachment, the group will send a dropper and a downloader first and then upgrade the system to a backdoor with more capabilities later.

3.1.4 Hurricane Panda

Hurricane Panda is a threat actor associated with China. Its main goal appears to be cyber espionage. The group is primarily affiliated with a single campaign called Operation Poison Hurricane.

To gain an initial foothold on the victim's network, the group favors web compromise attack patterns. In particular, targeting SQL injection and WebDAV vulnerabilities. The group then installs the China Chopper web shell, a web shell commonly available in the Chinese underground. The machine infected with the web shell can then be used as a pivot point to get further in the network.

Once inside a system, the group will use Mimikatz to steal credentials. They may also resort to exploiting privilege escalation vulnerabilities if the credentials stolen have insufficient privileges to pursue their mission. Using the stolen credentials, native tools such as WMIC and the "net use" command are used to move laterally. In particular, the group will use DLL side loading tricks to install Remote Access Tool (RAT) malware on remote systems. The specific malware used in Operation Poison Hurricane is PlugX/Kaba, a RAT commonly available in the Chinese underground.

The main particularity of this group is how they maintain access. To establish command and control (CnC), the group had performed DNS poisoning on Hurricane Electric DNS servers. This allowed the group to hijack the addresses of popular web sites (e.g., Github and Pinterest) making their CnC communications appear to go to legitimate web sites. The group also abused legitimate cloud services such as Google Code to store CnC information. This made their CnC communications appear to look like legitimate visits to the Google Code service.

3.2 Russian APT Actors

3.2.1 Fancy Bear

The Fancy Bear (a.k.a. Pawn Storm, Sednit, APT28, Stontium, Sofacy, or Tsar Team) actor is a group associated with Russia. It is mainly known for its role in the hacking of the Democratic National Committee (DNC) in the lead up to the 2016 U.S. elections. However, it took part in numerous other espionage campaigns against military and diplomatic targets, including TV5 Monde (A French news outlet) and the German Bundestag (government). In addition, dissidents and opposing government factions within Russia were also targeted. TrendMicro labels this espionage campaign the Pawn Storm campaign.

To gain an initial foothold in a victim's network, the group leverages its considerable resources. The group appears to have access to a large number of exploits, including a significant amount of 0-day exploits. These exploits are leveraged through either spear phishing containing malicious documents or spear phishing redirecting users to a watering hole web site hosting an exploit kit (SEDKit). A JavaScript trick is also leveraged for phishing of web service credentials (e.g., Outlook Web Access (OWA) credentials). The group displays a high level of sophistication, fingerprinting the victim through either the SEDkit web exploit kit or through a staged malware infection. The fingerprinting assures the use of the correct exploit, but also validates that the victim is a suitable victim for the group. In addition to these methods, the group also appears to bundle Sednit with legitimate applications to create Trojan horse versions of legitimate applications.

Once a victim has been infected by the Sednit stager and validated, the group will download additional tools on the machine to perform their mission. Notably, the group will typically install a custom version of the X-Agent malware that will provide the information stealing capabilities and handle the communications back to the attackers. The group can also, depending on their needs, upload keyloggers, credential stealers, tools to take screenshots, tools to enable a host to act as a proxy for other Sednit infections, rootkits or even a tool to infect USB drives in order to jump air gaps.

Additional attack patterns are likely, but not well documented in open source literature. For example, the group is known to perform internal reconnaissance using server-hosted nmap scripts to find open ports. However, all of the documented attack patterns rely on client-side exploits which do not require Internet facing open ports. This should be taken as another indicator of the group's high level of secrecy.

3.2.2 Dukes

The Dukes (a.k.a. APT29, Cozy Bear, CozyDukes, and CozyCzar) is an APT attack group that is affiliated with Russia. Its main goal appears to be espionage, especially against diplomatic targets. It has been implicated, alongside the Fancy Bear actor, in the hack of the DNC in the lead up to the 2016 U.S. presidential elections.

To gain an initial foothold in the network, the Dukes almost exclusively use spear-phishing campaigns. These campaigns use exploits (including 0-day exploits) when available, but often rely on users clicking a malicious self-extracting archive or screen saver program attached to the email. This has led some to believe that the group does not have access to a steady supply of vulnerabilities for their operations. The only case where the group did not rely on spear phishing to infect computers is for the distribution of the OnionDuke variant of the “Duke” series of malware. That variant was installed by being added to files transiting through a specific TOR network node to create a Trojan horse version of the file.

Once a victim is compromised, the group will load one of their backdoors, such as MiniDuke, PinchDuke, CosmicDuke, GeminiDuke, CozyDuke or CloudDuke in the case of spear phishing or OnionDuke through Trojan files. These backdoors have similar capabilities for enabling remote access and stealing confidential information. These backdoors also enable the attackers to steal credentials, escalate local privileges or download additional implants either via built-in functionality, or through downloading additional tools such as Mimikatz. In recent years, the Dukes have started to use this ability to download additional files to develop a series of stage 2 backdoors to provide better information stealing capabilities or more covert persistence. The SeaDuke and HammerDuke backdoors have been created for that purpose and are only deployed through implant staging.

A particularity of the group is that it abuses legitimate services to exfiltrate data and implement CnC. In particular, they make use of legitimate cloud storage services, which are unlikely to be blocked at the perimeter, to exfiltrate confidential data. For CnC, they often make use of Twitter and Github to communicate with their backdoors. In the case of HammerDuke, this is accompanied with the use of steganography to further disguise their communications as legitimate. This feature is fairly unique to this particular APT group.

3.2.3 Snake

The Snake group (a.k.a. Turla, Venomous Bear, Waterbug, and Agent.BTZ) is the group associated with Russia with the most technically advanced implant arsenal. This arsenal is based around the Snake rootkit. The group has a high degree of operational hygiene, leaving few clues as to the extent of their campaigns. However, they have been associated with a few well known victims. In particular, the group is deemed responsible for the infection of the classified network of the U.S. military with the Agent.BTZ malware via an infected thumb drive.

The group uses multiple attack patterns to gain entry into a network. Spear phishing, watering hole, transforming installers into Trojan horses, and breaking through third party vendors have all been documented. This is in addition to the infected removable media pattern observed in the case of the U.S. military infection.

When the victim's network is initially compromised, the group will install a first stage implant (Epic Turla/Wipbot/Tavdig) which will validate the compromised machine to ensure it is not a machine used for malware analysis. The first stage implant will also enable the attackers to fingerprint the system and the documents available on the network to determine its intelligence value. This internal reconnaissance is done through batch scripts running native Windows commands. Further lateral movement inside compromised networks also make use of Windows tools such as WMI and psexec, but also makes use of custom SMB scanners and Mimikatz.

High value targets will then be upgraded to the more powerful second stage implant (e.g., Snake, Uroburos, Carbon, Pfinet, Snark). The second stage backdoor makes use of rootkit technology and stores all additional modules, configuration, tasks, and task results in a virtual encrypted file system to prevent precise analysis of the actions it has taken by defenders. The advanced backdoor also support a form of peer-to-peer communication via Windows named pipes to allow data exfiltration from machines that are not directly connected to the CnC server.

This operational care can also be observed in other areas. A first example is the extensive tooling to fingerprint hosts before watering hole attacks. The group makes use of an infrastructure labelled WITCHCOVEN to track visitors with supercookies to identify targets. For example, an English speaking person living in the Netherlands that often visits diplomatic sites and U.K. news has a higher likelihood of being a U.K. diplomat to the E.U. The system then redirects the users to numerous fingerprinting scripts to identify precisely the software used by the target for exploitation. A second example is their use of satellite hijacking in addition to the more traditional use of proxy and VPN technologies to hide their source addresses.

3.2.4 Sandworm

The Sandworm group (a.k.a. Quedagh) is a group affiliated with Russia. The group is mostly known for its campaign using the Black Energy malware (also known as lancafdo or the Sandworm malware). In the first campaign, the goal was mainly cyber espionage, especially against European targets and energy infrastructure, but a file wiper module enabling sabotage was also found. The other campaign resulted in the black out in Ukraine in December 2015 due to a cyber-attack on a number of energy distribution companies.

The group uses a number of attack patterns to gain an initial foothold in their victim's network. The most common is speculated to be spear phishing as they have targeted a number of different client vulnerabilities over the years, including at least one 0-day. These vulnerabilities are more easily exploited via spear phishing, hence the speculation that this is their preferred method of entry. However, they have also packaged their malware as other executables (Trojan applications), disguised themselves as installers for legitimate software (fake installers) and installed them via other malware that was already present. All of these techniques install one of two versions of the Black Energy malware. The "Big" version of the malware is a customized version of a crimeware kit offered on the Russian underground. This version is installed as a rootkit on the machine. The "Lite" version of the malware does away with the rootkit component and is generally more lightweight, but is hypothesized to be more stable on a larger variety of systems. Both versions support the addition of several modules to supplement the basic backdoor functionalities.

The group relies mostly on Black Energy modules to perform their other tasks. The "si" module allows the attacker to steal information and fingerprint systems. The "vs" modules allows attackers to perform port scans for internal reconnaissance and run an embedded psexec to spread to other machines on the network. The "tv" modules enables the attack to infect existing version of team viewers in order to transform the team viewer in a Trojan horse malware and the "jn" executable enables the attacker to infect executables and installers to create fake installers. The "DSTR" plugin allows the destruction of system files and documents. Another file wiper, the "KillDisk" module was used in the Ukraine campaign to add the capability to affect files related to industrial control systems. A backdoor SSH server was also dropped in that campaign to allow remote access to attackers.

In addition to their standard Black Energy toolkit, the group also appears to use distributed denial of service (DDoS) attacks in specific cases. In the Ukraine campaign, a DDoS attack was launched to disrupt the phone lines of the victims to prevent them from discovering the attack. The tool used to perform this task is however unknown.

3.3 Middle Eastern APT Actors

3.3.1 AjaxTM and Rocket Kittens

Two series of intrusions, attributed to groups with different names, have been attributed to Iran. As the intrusions are similar in terms of techniques used and do not overlap, some have speculated that they are the product of the same group. As such, they are grouped under the label Iran groups.

The AjaxTM, or Ajax security team (a.k.a. Clever Kittens or the Operation Cleaver Team), is a group of patriotic hackers that have evolved from performing web defacement to state espionage. They are known for a series of intrusions labelled Operation Saffron Rose, or #OpCleaver, around 2012-2013.

One of the group’s main attack patterns to gain entry in a network is through web compromise. The group will hack a web server, install one of their web shells and perform lateral movement from there. The group is also known to use spear phishing, redirecting users to web sites that look legitimate and enticing them to install a Trojan application. The group has developed a tool named binder_1 to create these booby trapped executables.

The backdoor installed by the group are relatively primitive and are closer to crimeware than to espionage platforms. Only the PVZ tool has the capability to load modules post-infection. Even then, at least one module, the SYN flood module, is expected to be used for a criminal botnet but not for espionage. This may explain the need for the group to rely on numerous other hacking tools, such as Cain&Abel (man-in-the-middle tool), NetCrawler (remote execution in Windows network tool) and Mimikatz to perform post-exploitation tasks including credential stealing and lateral movement. However, their tooling does support data exfiltration by using FTP or SMTP communication. This communication is often independent of the main CnC channel for the backdoor, requiring an alternate infrastructure. However, the group may also use netcat or tunnel Remote Desktop protocol (RDP) sessions over SSH via a PuTTY utility to exfiltrate data as well. Overall, the tendency of the group is to repurpose commonly available tools, attempting to tweak them to evade anti-virus detection. The group creates custom Mimikatz wrappers, custom netcat implementation, custom exploits and so on. This indicates a limited proficiency on the part of the group.

The Rocket Kittens group (a.k.a. Flying Kittens) are known for a series of intrusions labelled Operation Woolen-Goldfish, which includes Operation Thamar Reservoir and Gholee. This series of intrusions is linked together from the use of a toolset developed by the same developer working under the pseudonym Wool3n.h4t and occurred around 2014-2015.

This series of intrusions rely mainly on spear phishing. Potential victims are sent a macro-enabled office document containing the implant. The implant is typically a repurposed agent from Core Impact, a commercial hacking platform, or Metasploit, an open source hacking platform. The group also uses a custom backdoor named MPK. The group also used web hacking, relying mostly on commercial grade vulnerability scanners such as Acunetix, SQLMap, NetSparker and Havij and web shells available on the underground.

The group also resorted to social engineering, using traditional phishing and pretexting. The phishing has been confirmed by the capture of their custom interface (Oyum management system) to automate the phishing process. The phone pretexting has been confirmed by some of the victims reporting heavily accented phone calls trying to get them to visit the phishing pages.

Once a machine has been compromised, the group will install a keylogger or a credential stealer to steal additional information. The CWoolger/.NETWoolger keylogger and the FireMalv credential stealer are the custom tools used by the team to perform this task.

The tendency of this group appears to be the repurposing of commercial grade tools, with limited development of some specialized functions. This observation and the apparent lack of operational security (i.e., unsecured CnC, data from internal testing, including the password of the main developer, on production CnC, etc.) speak to a limited level of proficiency from the group. This lends credence to the speculation that this group is the evolution of the Saffron Rose group.

3.3.2 MoleRats

The MoleRats group (a.k.a. Gaza Cyber Gang, Gaza Hacker Team, or DustySky) is a group of patriotic hackers from the Gaza strip that have evolved to cyber espionage. They appear to be mainly focused on diplomatic and military interests of the immediate region, targeting Israeli and other Palestinian victims, but victims in Europe, Asia and the United States have also been found. The group is unsophisticated and mostly relies on commonly available hacking tools that are easily detected by security software.

The attack pattern for initial compromise by the MoleRats group is relatively standard. Their usual method is spear phishing using topical stories as lure. However, they have also used less targeted "spray and pray" approaches. As they do not have access to vulnerabilities, they typically send the malware in an attached file or send a link to download the file containing the malware from the Internet. The group also makes use of traditional phishing to entice users to enter credentials on fake web pages. This activity seems to have limited tooling support as evidenced by the reuse of phishing infrastructure containing months-old news stories.

The malware used by the group appears to be off-the-shelf RATs malware providing basic backdoor and espionage capabilities such as njRAT, ExtremeRAT and Poison Ivy. However, in their latest campaign, they appear to have developed a custom RAT for their own use. As these malicious programs are well known by the anti-virus community, they have to rely on a multitude of tricks to load them in memory, such as using H-Worm, a Visual Basic Script worm, as a stager for their malware. The group also relies on other malware to perform post-exploitation tasks, for example using BrowserPasswordDump to steal cached browser credentials. This, again, facilitates the detection in comparison to the use of native tools for post exploitation tasks from more advanced groups.

Overall, the group currently has very limited capabilities and is not documented to have lateral movement capabilities. However, they appear to have started to acquire some maturity as demonstrated by beginning the development of their own tool chain.

3.4 Other ATP Actors

3.4.1 The Lazarus group

The Lazarus group (a.k.a. Silent Chollima) is a group affiliated with North Korea. The main tasking of the group appears to be cyber espionage, but the group has also leveraged its position within opposing networks for direct monetary gain (i.e., theft) and for disruption operations. Examples of Lazarus campaigns include Operation Troy, an espionage campaign which was transitioned in a massive disruption attack called Dark Seoul, the hack on Sony Pictures, an extortion campaign that turned into a disruption campaign, the WannaCry worm, a ransomware attack, and the cyber heist at the Bangladesh Central Bank. The group also has a lot of other documented activity that is not directly affiliated to a specific campaign. This is due in part to the fact that the group is unusually prolific for a targeted attack group, but also to its propensity to erase traces of its presence using data destruction software. This creates a large number of fragmentary reporting.

The group appears to be using an “Anything goes approach” to initial intrusion, varying its attack patterns based on available resources. The group has used spear phishing when client vulnerabilities were available, water hole techniques when browser exploits were available, or even stage payloads in machines compromised in prior operations. For example, a machine whose intelligence value has been depleted might use its espionage backdoor to load a data wiper for a disruption operation. The traces of compromise for espionage purposes would be conveniently deleted as the disruption operation proceeds.

The group appears to be mostly reliant on its multiple backdoors for all other tasks on a victim's machine from data exfiltration to credential stealing and secure deletion. Of the use of native tools that can be observed by more advanced groups, only the WMIC is used for lateral movement, and only via malware integration. This speaks to the group’s overall low level of technical expertise, although they appear to have a high level of operational expertise as attested by the multi-stage operations and a clear aversion to detection.

3.4.2 Careto

The Careto group (a.k.a. The Mask) is a group of unknown origin but which appears to have Spanish as their native language. Very little is known about them, but the high quality of their toolset and their focus on espionage appears to suggest nation state backing.

The attack pattern that has been used for known instances of Careto attacks is spear phishing. The victims is lured to a server which analyzes the victim to make sure it is an intended target and to serve one of the multiple exploits available to the group. The victim is then infected with one of the group's backdoors. The backdoors consist of the namesake Careto backdoor, which is a general purpose backdoor, the SGH backdoor, which is an advanced espionage platform with rootkit functionality and numerous versions of the SBD backdoor for more exotic systems.

Little is known of the methods used by the group in the post-exploitation phase. However, numerous plugin modules for the SGH backdoor have been found. A number of them, such as the module to intercept and record Skype conversations, the module to intercept file access and the module to intercept network traffic, appears to be geared toward data exfiltration. As these modules are loaded to expand the functionality as needed, they also attest to the use of implant staging by the attackers.

4. EXERCISE SCENARIOS

This section covers two exercise scenarios that illustrate the use cases presented in section 2 via the threat profiles presented in section 3. The first example illustrates use cases based on attribution and the second example illustrates use cases based on indicators and warnings.

4.1 Exercise 1 - Attribution

This exercise is designed to practice intelligence driven incident response.

4.1.1 Task 1: Intent and methods

In this task the exercise participants are expected to use the threat profiles to extrapolate intent from an initial incident.

4.1.1.1 ***Context***

In the context of the Ukraine crisis, a small force was deployed in Eastern Europe to help deter further aggression. To support field operations, a small network was deployed with the help of local forces. The network consists of an unclassified network, where most of the day-to-day operations occur, a classified network and a command and control network. The unclassified network can access the Internet, but a firewall prevents outside attackers from initiating connections. The unclassified network is also connected to an allied network to facilitate coordination. The command and control network is connected to the unclassified network, but protected by a very restrictive firewall. Only a small number of personnel can access machines in that network to perform remote maintenance. All of the personnel are located in the administrative subnet and are required to use a Virtual Private Network (VPN) to connect. The classified network is separated from all the other networks by an air gap. The following Figure illustrates the network.

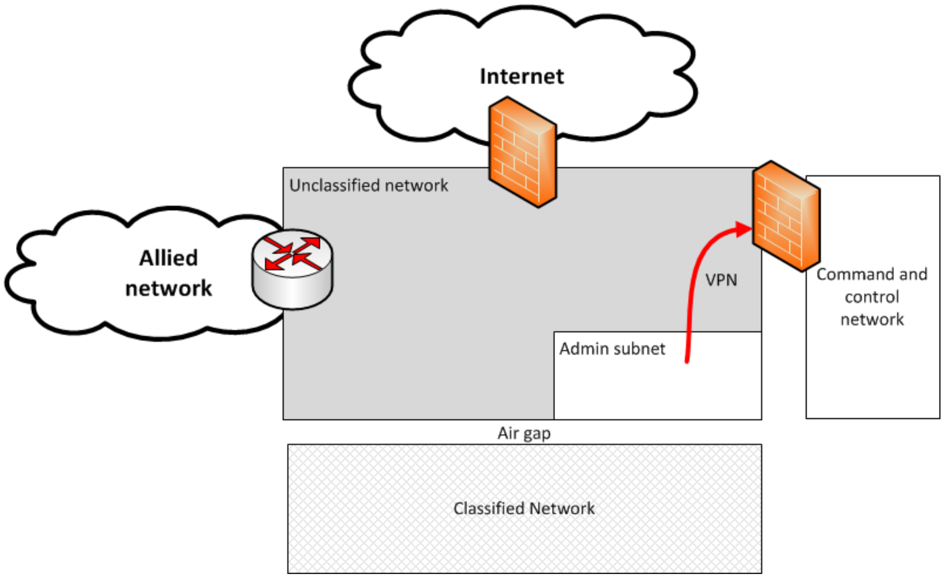


Figure 7: Network

An anti-virus alarm was issued last night after a signature update. It appears the malware found comes from a new version of the Backdoor.W32.Tavdig signature. This alarm was generated for a host in the unclassified network, but not in the admin subnet. Help desk support has already cleared the infection by re-imaging the machine as no other computer was available to perform the tasks performed on the compromised machine. This most likely erased any artefacts that could be gleaned from host forensics. However, the incident was ultimately reported to you, mostly to keep you informed.

4.1.1.2 *Actions*

The first step would be to identify the threat actor. The Tavdig backdoor is an alternate name for the Epic Turla backdoor, the first stage tool used by the Snake actor, presented in section 3.2.3. The Snake actor is affiliated with Russia, which should be considered a serious threat given the current context of deployment. Furthermore, the Snake actor has been associated with high profile compromises in military systems.

To obtain that information, the participants would be expected to use the attribution process presented as pseudo-code in section 2.2.2.2. To summarize, he would extract the record for the Tavdig malware object from the Stix database and follow the links to the threat actor associated (for example Tavdig < (uses) Spear Phishing <(uses) Unnamed Cyber Espionage (is attributed to)> Snake).

In terms of intent, their primary goal is espionage, as revealed by the goal attribute of the Snake threat actor object. This means that it is unlikely that the adversary intends to create disruption in the command and control network. Using the effect-driven reasoning presented in the cyber threat model proposed in [1], three reasonable hypothesises could be put forward:

- The compromise is used as an entry point in the allied network;
- The goal is to exfiltrate classified information; and
- The goal is to exfiltrate keying material (e.g., cryptographic keys or data used to generate keys or credentials such as VPN password and seeds for two factor authentication) to enable further operations.

Regarding the hypothesis of using the compromise as an entry point in the allied network, this is in line with the third party compromise attack pattern displayed by the group. However, it is at least equally likely that the allied network was used to compromise us via this attack pattern as it is yet unclear what their target is. Unless the intelligence value of the allied network is higher than our own, this appears to be an unlikely hypothesis.

In terms of exfiltrating classified information, this is consistent with previous activities of the group, as evidence by the Unnamed Cyber Espionage campaign object. Furthermore, the group possesses capabilities specifically designed to jump air gaps (infect removable media attack pattern), which they have used in the past, as evidenced by the infect removable media attack pattern object. So, at first glance, this is a likely hypothesis.

As for exfiltrating keying material, this is also consistent with the typical goals of the group (goal property of the Snake threat actor object) and the VPN credentials have inherent value for future cyber operations aimed at disrupting command and control, which would be a likely goal of the adversary if we use the threat-driven reasoning presented in the statement of work (SOW) [1]. Given the deployment context, this also appears to be a likely hypothesis.

4.1.2 Inject 1: Responding to the incident

In this task, the participants are expected to use the likely intent to extrapolate which resources would be impacted in order to direct collection of further evidence.

4.1.2.1 *Context*

Using your initial report, it was possible to convince the commander that this incident should be further looked into. However, as the machine was already re-imaged, all the intelligence from that machine was lost. At first glance, the only thing that can be done is to be on watch for any new infection and to advise first level responders to preserve forensics evidence. Unless a clear collection plan for further evidence of a compromise is put forward, that is the course of action that will be taken by the commander.

4.1.3 Actions

In order to determine what was the ultimate goal, we need to extrapolate which resources are likely to be impacted if either (or both) of the likely hypotheses is true. This would be done using a method similar to the pseudo-code to create tactical indicators presented in Section 2.2.2.2, but creating unique identifiers for the likely hypotheses rather than for actors.

If the Snake group intends to exfiltrate classified information, they will be required to jump the air gap. To do so, they are likely to use the ‘infect removable media’ attack pattern. It would be possible to inspect removable media for signs of the Agent.BTZ malware or other similar tools.

If the Snake group intends to gather keying material, they are more likely to use the lateral movement attack pattern to reach the admin subnet, where the VPN credentials can be stolen. Looking for signs of lateral movement could be an avenue for collection. This would be achieved by selecting all objects that are the target of a “uses” relationship object with the Lateral Movement attack pattern object from the Snake group. This would create a list of malware and tools that would be indicators of this attack pattern. Unfortunately, the specific tools used by the Snake group for lateral movement are mostly valid Windows tools (e.g. psexec and WMIC) used with stolen credentials. This leaves no traces on local machines and is very difficult to distinguish from regular activity in network logs. The tool used to steal credentials (Mimikatz) could be used as an artefact that would support this theory, but the traces have likely been erased when the machine was reimaged.

The best evidence of the goal to steal keying material would be to find stage 2 implants (e.g. Turla or Carbon) on the admin subnet. As these represent high value targets, the attackers are likely to "upgrade" the machines to a high value backdoor. As such, using the observational reasoning presented in the SOW references [1], we can infer from the presence of the backdoors that the machines where these would be found are the true target. As these backdoors use rootkit technology, it may be very difficult to find them. Furthermore, the use of encrypted file systems techniques do not allow defenders to know for certain what has been executed on the machine or what has been stolen. However, the presence of these high value backdoors on the admin machines indicate that VPN keying material was likely the target.

Collection efforts should be directed to the evidence that strongly confirm the hypotheses put forward.

4.1.4 Inject 2: Teamwork

In this task, the participants are presented with evidence of cooperation between two attack groups. Furthermore, since the new group involved (Sandworm), also includes sabotage as its goals, the participants are expected to notice a change in the overall goals of the operation from a pure espionage campaign to a campaign that might involved sabotage.

4.1.4.1 *Context*

After further collection was performed, an instance of the Carbon backdoor was found on one of the workstation in the admin subnet. It was not possible to crack the encrypted file system to determine what the actions of the attackers were. However, a thorough review of the VPN logs reveal an unauthorized access using stolen credentials. The attackers appear to have used the credentials to install a version of the Dropbear SSH server for an unknown purpose.

4.1.4.2 *Actions*

The Dropbear SSH server is not a tool usually found in the Snake's group arsenal. This is attested by the lack of a Dropbear SSH malware object in their profile. However, the Sandworm group deploys versions of the Dropbear SSH server that contains a hard-coded password to give administrative access. Looking at the goals property of the Sandworm threat actor object, the group appears to have a tasking to perform sabotage as well as espionage. Using the effect-driven reasoning proposed in the SOW references [1], their presence in the command and control network is indicative of the intent to create an availability impact on the command and control network. As the Sandworm group, presented in Section 3.2.4, is also affiliated with Russia, it seems reasonable that they could have profited from an earlier intelligence gain from the Snake group.

4.2 Exercise 2 - Indicators and Warnings

This exercise is designed to practice intelligence tasks related to cyber defense planning and incident response.

4.2.1 Task 1: Planning process

In this task, the exercise participants are expected to perform the planning process to identify indicators and warnings.

4.2.1.1 *Context*

A series of diplomatic incidents have inflamed the relations between the U.S. and Iran. The Supreme Leader of Iran has called upon patriotic Iranians and any friend of the Republic to cause harm in any way possible to the U.S. and their allies. As the flare up in tension was sudden, cyber was deemed a likely avenue of attack as the Internet provides instant access to foreign individuals.

The commander has conveyed his intent to be particularly vigilant for cyber-attacks related to this threat. He expects a series of indicators and warnings that would enable defenders to:

a) Identify likely avenues of attacks to prepare the defence; and
b) Prioritize cyber incidents related to this threat

4.2.1.2 *Actions*

One of the first steps is to identify attack groups linked to nation states that are susceptible to be called to arms by the Supreme Leader. To do so, the participants would select all the sources of attributed-to relationship objects where the target is Iran, or Iranian allies. The most likely candidates are the Iran Attack group(s) (Clever Kitten and Rocket Kitten), presented in Section 3.3.1. However, the MoleRats group, presented in Section 3.3 is affiliated with Hamas, which receives significant funding from Iran. They could also take part in the offensive.

Based on that information, we can look at the attack patterns used for initial compromise as likely avenues of attack by selecting attack patterns that are the target of uses relationship objects that has the relevant threat actors are sources. For the MoleRats, we can be confident that the initial avenue of attack would be through spear phishing or untargeted phishing as they have limited capabilities to use other attack patterns. The Iran groups on the other hand have a propensity for web compromise and social engineering (e.g. phone pretexting and credential harvesting through phishing) in addition to spear phishing. This represents the likely avenues of attacks.

Based on these likely avenues of attack, using the effect-driven reasoning presented in the SOW references [1], we infer the likely targets based on the expected impacts on resources of these attack patterns. So, more resources should be allocated to secure web servers, watch incoming mail communication, and employee awareness programs for phishing and social engineering. On the other hand, less resources should be allocated to patching (non-web) vulnerabilities as none of the groups appear to target these vulnerabilities for their initial infection or lateral movement.

In terms of indicators to identify cyber incidents related to this threat, we can follow the pseudo-code for the production of tactical indicators presented in Section 2.2.2.2. However, both groups appear to favour tools that are widely available, which would yield few unique identifiers, although Iran attack groups typically also add some customized tools. In both cases, the techniques and procedures used are too common to create additional signatures on attack patterns or particular sequences of events.

4.2.2 Inject 1: A web compromise

In this task, the exercise participants are expected to use the threat profiles at their disposition to identify that a web compromise does not come from a group close to Iran, but rather from a Chinese group. This would lead the participants to label that incident as low priority.

4.2.2.1 *Context*

As the defenders strengthen the web servers, they notice the presence of an unexpected script on one of the servers. After a quick technical investigation, it is revealed to be the CFM backdoor by UFO web shell, indicating a web compromise. As the machine appears to have been compromised for some time, the attackers could have severely expanded their access. As such, investigating this compromise would take significant resources. The commander requires an assessment to help him decide if he should prioritize the investigation or the additional shoring up of defenses.

4.2.2.2 *Actions*

The first step is to evaluate if the attack pattern is consistent with the current adversary by looking at attack pattern objects. As the Iran groups have a predisposition for web compromise, it seems consistent with their TTP. Furthermore, the tool used is a hacker tool rather than a custom implant. This is also consistent with previous behaviour. However, and does not figure as the target of a uses relationship object starting from the web compromise attack pattern, but comparing it with other groups known to use web compromise attack patterns, we see that the Shell Crew attack group has used that specific tool in the past. This should urge participants to conclude that the incident is of Chinese origins, albeit with low confidence.

To increase confidence in the estimate, the participants should be encouraged to suggest a limited investigation of the machine to dig for more artefacts. A quick integrity scan of the machine compared to a clean image indicates that the sticky key executable has been modified, corresponding to another malware object found in the Shell Crew profile. This would increase the confidence of the Shell Crew actor.

4.2.3 Inject 2a: A malicious email

In this task, the participants are expected to identify a spear phishing email containing malware attributed to the MoleRats group. This event should be prioritized as it relates to the current threat. However, the MoleRats profile should reveal their limited proficiency and a concurrent incident (inject 2b) should be considered a higher threat.

4.2.3.1 *Context*

Now that the defensive preparations have been completed, you are just waiting for new incidents. As you come in to work in the morning, it appears that numerous incidents have occurred during the previous watch. For these incidents, a new set of signatures pushed the previous day has identified an email that contained a Poison Ivy variant. This email was sent to multiple recipients. Initial investigations by the previous watch have revealed that at least one individual has opened the file attachment. The commander requires an assessment to help him decide how he should prioritize investigating this incident.

4.2.3.2 *Actions*

In a manner similar to inject 1, the participants are expected to compare the attack pattern to determine consistency with the current adversaries. The spear phishing attack pattern appears similar to the one used by the MoleRats, as the MoleRats profile includes a Poison Ivy malware object that is linked to the Spear Phishing attack pattern object. However, the attack pattern could be replicated by multiple other groups. This should lead the participants to conclude with low confidence that this is the MoleRats group. Asking for additional investigation, based on the technique used to generate tactical indicators previously described, to increase confidence in the assessment would likely involve investigating whether password dumping tools have been uploaded to the system as this would be a required step to expand access.

As this group possesses limited capabilities, the incident should be prioritized as an incident related to the Iranian situation to the commander, but with the caveat that this is the lesser threat within the adversaries.

4.2.4 Inject 2b: Connection error?

In this task, the participants are expected to investigate the consequences of a phishing attack to identify an attack by an Iranian group. This event should be given the highest priority as it relates to the current threat and is associated with the groups with the greatest capabilities.

4.2.4.1 *Context*

Now that the defensive preparations have been completed, you are just waiting for new incidents. As you come in to work in the morning, it appears that numerous incidents have occurred during the previous watch. For these incidents, a remote employee has called in to the help desk in a previous shift to complain about having issues connecting to the network. The troubleshooting revealed that this was because someone else was already connected with his user name. The help desk has terminated the connection and reset the password.

4.2.4.2 *Actions*

The first step is to determine if the attack pattern is consistent with known groups. Here, a legitimate connection is used. At first glance, this does not appear to be linked to known attack patterns of groups of interest. However, this is actually consistent with the consequence of someone falling having their credentials stolen, for example through phishing. Once the credentials are stolen, the attacker can now log in without any tools. The phishing attack pattern is present in the profiles of both actors of interest and many more. Further cursory investigation, based on the generation of tactical indicators, could look for phishing emails or further investigate the connection end point inside our network to see what additional tools were dropped. A custom version of Mimikatz (e.g. zhMimikatz), which is a malware object only present in the Iran groups profile, dropped on the machine to start expanding access would be a telltale sign of the involvement of Iran groups.

As the Iranian groups have greater capabilities than MoleRats groups, this incident should get the highest priority.

5. CONCLUSION

This report, while accurate and pertinant to cyber defence in the short term, should be understood in the context of rapidly evolving technology, the ever-changing tools available to APT actors, and the security measures that will continue to be developed to address these threats. Cyber defence must be an ongoing investment, with strategies and tactics being updated continuously to meet the newly developed capabilities of the APT actors. This include the continual modification and improvement of tools like STIX 2.0.

REFERENCES

[1] DRDC Ottawa. *TA-35 – CYBER THREAT DATA MODEL AND USE CASES.* Ottawa: DRDC, 2017

[2] OASIS Cyber Threat Intelligence Committee, "Cyber Threat Intelligence Committee Github page," [Online]. Available: https://oasis-open.github.io/cti-documentation/. [Accessed July 2017]

[3] A. Lemay, S. Knight and J. M. Fernandez, "Preparation of the Cyber Environment (IPCE): Finding the High Ground in Cyberspace," Journal of Information Warfare, vol. 13, no. 3, pp. 47-56, 2014

The following documents were used as source material for their associated sections.

Section 3.1.1 - Common Crew

[1] Mandiant, "APT1 - Exposing One of China's Cyber Espionage Units," February 2013. [Online]. Available: http://intelreport.mandiant.com/ Mandiant_APT1_Report.pdf. [Accessed 8 August 2013].

[2] D. Alperovitch, "Revealed: Operation Shady RAT," 2011. [Online]. Available: http://www.mcafee.com/us/resources/white-papers/wp-operation-shady-rat.pdf. [Accessed 11 January 2017].

[3] P. Coogan, "Targeted Attacks Make WinHelp Files Not So Helpful," Symantec, 15 October 2012. [Online]. Available: https://www.symantec.com/connect/blogs/targeted-attacks-make-winhelp-files-not-so-helpful. [Accessed 11 January 2017].

[4] G. Hoglund, "Inside an APT Covert Communications Channel," Fast Horizon, 16 August 2011. [Online]. Available: http://fasthorizon.blogspot.ca/2011/08/inside-apt-comment-crew-covert.html. [Accessed 11 January 2017].

[5] K. Wilhoit, "The SCADA That Didn't Cry Wolf - Who's Really Attacking Your ICS Equipment? (Part 2)," 2013. [Online]. Available: http://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-the-scada-that-didnt-cry-wolf.pdf. [Accessed 11 January 2017].

Section 3.1.2 - Shell Crew

[1] Federal Bureau of Investigations, "FBI Liaison Alert System #A-000049-MW," February 2015. [Online]. Available: http://krebsonsecurity.com/wp-content/uploads/2015/02/FBI-Flash-Warning-Deep-Panda.pdf. [Accessed 11 January 2017].

[2] "RSA Incident Response: Emerging Threat Profile Shell_Crew," January 2014. [Online]. Available: https://www.emc.com/collateral/white-papers/h12756-wp-shell-crew.pdf. [Accessed 11 January 2017].

[3] RSA Research, "TERRACOTTA VPN - Enabler of Advanced Threat Anonymity," 4 August 2015. [Online]. Available: https://blogs.rsa.com/wp-content/uploads/2015/08/Terracotta-VPN-Report-Final-8-3.pdf. [Accessed 11 January 2017].

[4] ThreatConnect Research Team, "The Anthem Hack: All Roads Lead to China," ThreatConnect, 27 February 2015. [Online]. Available: https://www.threatconnect.com/the-anthem-hack-all-roads-lead-to-china/. [Accessed 11 January 2017].

[5] D. Alperovitch, "Deep in Thought: Chinese Targeting of National Security Think Tanks," Crowdstrike, 7 July 2014. [Online]. Available: https://www.crowdstrike.com/blog/deep-thought-chinese-targeting-national-security-think-tanks/. [Accessed 11 January 2017].

[6] J. DiMaggio, "The Black Vine cyberespionage group," 6 August 2016. [Online]. Available: http://www.symantec.com/content/en/us/enterprise/ media/security_response/ whitepapers/the-black-vine-cyberespionage-group.pdf. [Accessed 11 January 2017].

[7] RyanJ, "Mo' Shells Mo' Problems – Deep Panda Web Shells," Crowdstrike, 20 February 2014. [Online]. Available: https://www.crowdstrike.com/blog/mo-shells-mo-problems-deep-panda-web-shells/. [Accessed 3 July 2017].

Section 3.1.3 - Naikon

[1] FireEye Labs / FireEye Threat Intelligence, "APT30 AND THE MECHANICS OF A LONG-RUNNING CYBER ESPIONAGE OPERATION - How a Cyber Threat Group Exploited Governments and Commercial Entities across Southeast Asia and India for over a Decade," FireEye, 2015.

[2] ThreatConnect inc. and Defense Group Inc., "Project CAMERASHY Closing the aperture on China's unit 78020," 2015. [Online]. Available: https://www.threatconnect.com/camerashy-resources/. [Accessed July 2017].

[3] K. Baumgartner and M. Golovkin, "The MsnMM Campaigns - The Earliest Naikon APT Campaigns," May 2015. [Online]. Available: https://securelist.com/files/2015/05/TheNaikonAPT-MsnMM1.pdf. [Accessed 11 January 2017].

[4] K. Baumgartner and M. Golovkin, "The Naikon APT," Kaspersky, 14 May 2015. [Online]. Available: https://securelist.com/analysis/publications/69953/the-naikon-apt/. [Accessed 4 February 2017].

[5] C. Raiu and M. Golovkin, "The Chronicles of the Hellsing APT: the Empire Strikes Back," Kaspersky, 15 April 2015. [Online]. Available: https://securelist.com/analysis/publications/69567/the-chronicles-of-the-hellsing-apt-the-empire-strikes-back/. [Accessed 11 January 2017].

Section 3.1.4 - Hurricane Panda

[1] D. Alperovitch, "CrowdStrike Discovers Use of 64-bit Zero-Day Privilege Escalation Exploit (CVE-2014-4113) by Hurricane Panda," Crowdstrike, 14 October 2014. [Online]. Available: https://www.crowdstrike.com/blog/crowdstrike-discovers-use-64-bit-zero-day-privilege-escalation-exploit-cve-2014-4113-hurricane-panda/. [Accessed 11 January 2017].

[2] D. Alperovitch, "Cyber Deterrence in Action? A story of one long HURRICANE PANDA campaign," Crowdstrike, 13 April 2015. [Online]. Available: https://www.crowdstrike.com/blog/cyber-deterrence-in-action-a-story-of-one-long-hurricane-panda-campaign/. [Accessed 11 January 2017].

[3] N. Moran, J. Homan and M. Scott, "Operation Poisoned Hurricane," FireEye, 6 August 2014. [Online]. Available: https://www.fireeye.com/blog/threat-research/2014/08/operation-poisoned-hurricane.html. [Accessed 11 January 2017].

[4] A. Schworer and J. Liburdi, "Storm Chasing: Hunting Hurricane Panda," Crowdstrike, 26 January 2015. [Online]. Available: https://www.crowdstrike.com/blog/storm-chasing/. [Accessed 11 January 2017].

Section 3.2.1 - Fancy Bear

[1] Bitdefender, "APT28 Under the Scope A Journey into Exfiltrating Intelligence and Government Information," December 2015. [Online]. Available: https://download.bitdefender.com/resources/media/materials/white-papers/en/Bitdefender_In-depth_analysis_of_APT28%E2%80%93The_Political_Cyber-Espionage.pdf. [Accessed 12 January 2017].

[2] FireEye, "APT28: a Window Inot Russia’s Cyber Espionage Operations?," FireEye, October 2014. [Online]. Available: http://www2.fireeye.com/rs/fireye/images/rpt-apt28.pdf. [Accessed 12 January 2017].

[3] Microsoft, "Microsoft Security Intelligence Report Volume 19 | January through June, 2015," 2015. [Online]. Available: http://download.microsoft.com/download/4/4/C/44CDEF0E-7924-4787-A56A-16261691ACE3/Microsoft_Security_Intelligence_Report_Volume_19_A_Profile_Of_A_Persistent_Adversary_English.pdf. [Accessed 12 January 2017].

[4] ESET Research, "Sednit espionage group now using custom exploit kit," ESET, 8 October 2014. [Online]. Available: http://www.welivesecurity.com/2014/10/08/sednit-espionage-group-now-using-custom-exploit-kit/. [Accessed 27 January 2017].

[5] Kaspersky Lab's Global Research & Analysis Team, "Sofacy APT hits high profile targets with updated toolset," Kaspersky, 4 December 2015. [Online]. Available: https://securelist.com/blog/research/72924/sofacy-apt-hits-high-profile-targets-with-updated-toolset/. [Accessed 12 January 2017].

[6] D. Alperovitch, "Bears in the Midst: Intrusion into the Democratic National

Committee," Crowdstrike, 15 June 2016. [Online]. Available: https://www.crowdstrike.com/blog/bears-midst-intrusion-democratic-national-committee/. [Accessed 12 January 2017].

[7] M. Bailey, "MATRYOSHKA MINING Lessons from Operation RussianDoll, January 2016," January 2016. [Online]. Available: http://paper.seebug.org/papers/APT/APT_CyberCriminal_Campagin/2016/2016.03.09.Operation_RussianDoll/wp-mandiant-matryoshka-mining.pdf. [Accessed 12 January 2017].

[8] J. Calvet, "Sednit Espionage Group Attacking Air-Gapped Networks," ESET, 11 November 2014. [Online]. Available: http://www.welivesecurity.com/2014/11/11/sednit-espionage-group-attacking-air-gapped-networks/. [Accessed 12 January 2017].

[9] J. Calvet, "The Sednit Group: "Cyber" Espionage in Eastern Europe," in NorthSec, Montreal, 2015.

[10] J. Calvet, J. Campos and T. Dupuy, "Visiting The Bear Den A Journey in the Land of (Cyber-)Espionage," in RECon, Montreal, 2016.

[11] G. Cluley, "New ESET research paper puts Sednit under the microscope," ESET, October 2016. [Online]. Available: http://www.welivesecurity.com/2016/10/20/new-eset-research-paper-puts-sednit-under-the-microscope/. [Accessed 12 January 2017].

[12] D. Creus, T. Halfpop and R. Falcone, "Sofacy's 'Komplex' OS X Trojan," PaloAlto Networks, 26 September 2016. [Online]. Available: http://researchcenter.paloaltonetworks.com/2016/09/unit42-sofacys-komplex-os-x-trojan/. [Accessed 12 January 2017].

[13] R. Falcone and B. Lee, "New Sofacy Attacks Against US Government Agency," PaloAlto Networks, 14 June 2016. [Online]. Available: http://researchcenter.paloaltonetworks.com/2016/06/unit42-new-sofacy-attacks-against-us-government-agency/. [Accessed 12 January 2017].

[14] L. Kharouni, F. Hacquebord, N. Huq, J. Gogolinski, F. Mercês, A. Remorin and D. Otis, "Operation Pawn Storm Using Decoys to Evade Detection," October 2014. [Online]. Available: http://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-operation-pawn-storm.pdf. [Accessed 12 January 2017].

Section 3.2.2 - Dukes

[1] Symantec Security Response, ""Forkmeiamfamous": Seaduke, latest weapon in the Duke armory," Symantec, 13 July 2015. [Online]. Available: https://www.symantec.com/connect/blogs/forkmeiamfamous-seaduke-latest-weapon-duke-armory. [Accessed 12 January 2017].

[2] F-Secure Labs Security Response, "COSMICDUKE - Cosmu with a twist of MiniDuke," 2014. [Online]. Available: https://www.f-

secure.com/documents/996508/1030745/cosmicduke_whitepaper.pdf. [Accessed 11 January 2017].

[3] F-Secure labs Security Response, "COZYDUKE," 2015. [Online]. Available: https://www.f-secure.com/documents/996508/1030745/CozyDuke. [Accessed 11 January 2017].

[4] Fire Eye Threat Intelligence, "HAMMERTOSS: Stealthy Tactics Define a Russian Cyber Threat Group," July 2015. [Online]. Available: https://www2.fireeye.com/rs/848-DID-242/images/rpt-apt29-hammertoss.pdf. [Accessed 12 January 2017].

[5] ESET Research, "Miniduke still duking it out," ESET, 20 May 2014. [Online]. Available: http://www.welivesecurity.com/2014/05/20/miniduke-still-duking/. [Accessed 30 January 2017].

[6] F-Secure Labs Threat Intelligence, "THE DUKES - 7 years of Russian cyberespionage," September 2015. [Online]. Available: https://www.f-secure.com/documents/996508/1030745/dukes_whitepaper.pdf. [Accessed 12 January 2017].

[7] D. Alperovitch, "Bears in the Midst: Intrusion into the Democratic National Committee," Crowdstrike, 15 June 2016. [Online]. Available: https://www.crowdstrike.com/blog/bears-midst-intrusion-democratic-national-committee/. [Accessed 12 January 2017].

[8] J. Grunzweig, "Unit 42 Technical Analysis: Seaduke," PaloAlto Networks, 14 July 2015. [Online]. Available: http://researchcenter.paloaltonetworks.com/2015/07/unit-42-technical-analysis-seaduke/. [Accessed 12 January 2017].

[9] B. Levene, R. Falcone and R. Wartell, "Tracking MiniDionis: CozyCar's New Ride Is Related to Seaduke," PaloAlto Networks, 14 July 2015. [Online]. Available: researchcenter.paloaltonetworks.com/2015/07/tracking-minidionis-cozycars-new-ride-is-related-to-seaduke/. [Accessed 12 January 2017].

[10] C. Raiu, I. Soumenkov, K. Baumgartner, V. Kamluk and G. R. a. A. Team, "The MiniDuke Mystery: PDF 0-day Government Spy Assembler 0x29A Micro Backdoor," February 2013. [Online]. Available: https://kasperskycontenthub.com/ wp-content/uploads/sites/43/vlpdfs/ themysteryofthepdf0-dayassemblermicrobackdoor.pdf. [Accessed 12 January 2017].

Section 3.2.3 - Snake

[1] GovCERT.ch, "APT Case RUAG Technical Report," 23 May 2016. [Online]. Available: https://www.melani.admin.ch/dam/melani/de/dokumente/2016/ technical%20report%20ruag.pdf.download.pdf/Report_Ruag-Espionage-Case.pdf. [Accessed 12 January 2017].

[2] FireEye Threat Intelligence, "PINPOINTING TARGETS: Exploiting Web Analytics to Ensnare Victims," November 2015. [Online]. Available: https://www2.fireeye.com/rs/848-DID-242/images/rpt-witchcoven.pdf. [Accessed

12 January 2017].

[3] BAE Systems Apllied Intelligence, "SNAKE CAMPAIGN. CYBER ESPIONAGE TOOLKIT," 2014. [Online]. Available: http://paper.seebug.org/papers/APT/APT_CyberCriminal_Campagin/2014/snake_whitepaper.pdf. [Accessed 12 January 2017].

[4] Kaspersky Lab's Global Research & Analysis Team, "The Epic Turla Operation," Kaspersky, 7 August 2014. [Online]. Available: https://securelist.com/analysis/publications/ 65545/the-epic-turla-operation/. [Accessed 12 January 2017].

[5] Security Response, "The Waterbug attack group," 14 January 2016. [Online]. Available: https://www.symantec.com/content/en/us/enterprise/media/security_response/whitepapers/waterbug-attack-group.pdf. [Accessed 12 January 2017].

[6] B. Bartholomew, "KopiLuwak: A New JavaScript Payload from Turla," Kaspersky, 2 February 2017. [Online]. Available: https://securelist.com/blog/research/77429/kopiluwak-a-new-javascript-payload-from-turla/. [Accessed 2 February 2017].

[7] A. Dereszowski, "Andrzej Dereszowski – Turla:Development & Operations [Rooted CON 2015 - ENG]," Spain, 2015.

[8] C. Raiu and K. Baumgartner, "The ‘Penquin’ Turla A Turla/Snake/Uroburos Malware for Linux," Kaspersky, 8 December 2014. [Online]. Available: https://securelist.com/blog/research/67962/the-penquin-turla-2/. [Accessed 12 January 2017].

[9] S. Tanase, "Satellite Turla: APT Command and Control in the Sky," Kaspersky, 9 September 2015. [Online]. Available: https://securelist.com/blog/research/72081/satellite-turla-apt-command-and-control-in-the-sky/. [Accessed 12 January 2017].

Section 3.2.4 - Sandworm

[1] F-Secure Labs Security Response, "BLACKENERGY & QUEDAGH," 2014. [Online]. Available: https://www.f-secure.com/documents/996508/1030745/blackenergy_whitepaper.pdf. [Accessed 27 January 2017].

[2] Symantec Security Response, "Sandworm Windows zero-day vulnerability being actively exploited in targeted attacks," Symantec, 14 October 2014. [Online]. Available: https://www.symantec.com/connect/blogs/sandworm-windows-zero-day-vulnerability-being-actively-exploited-targeted-attacks. [Accessed 12 January 2017].

[3] Brod, "Beware BlackEnergy If Involved In Europe/Ukraine Diplomacy," F-Secure, 30 June 2014. [Online]. Available: https://www.f-secure.com/weblog/archives/00002721.html. [Accessed 27 January 2017].

[4] A. Cherepanov, "BlackEnergy by the SSHBearDoor: attacks against Ukrainian news media and electric industry," ESET, 3 January 2016. [Online]. Available: http://www.welivesecurity.com/2016/01/03/blackenergy-sshbeardoor-details-2015-

attacks-ukrainian-news-media-electric-industry/. [Accessed 12 January 2017].

[5] P. Ducklin, "The “Sandworm” malware – what you need to know," Sophos, 15 October 2014. [Online]. Available: https://nakedsecurity.sophos.com/2014/10/15/the-sandworm-malware-what-you-need-to-know/. [Accessed 12 January 2017].

[6] R. Lipovsky, "Back in BlackEnergy *: 2014 Targeted Attacks in Ukraine and Poland," ESET, 22 September 2014. [Online]. Available: http://www.welivesecurity.com/2014/09/22/back-in-blackenergy-2014/. [Accessed 27 January 2017].

[7] R. Lipovsky, "CVE-2014-4114: Details on August BlackEnergy PowerPoint Campaigns," ESET, 14 October 2014. [Online]. Available: http://www.welivesecurity.com/2014/10/14/cve-2014-4114-details-august-blackenergy-powerpoint-campaigns/. [Accessed 27 january 2017].

[8] R. Lipovsky and A. Cherepanov, "Last-minute paper: Back in BlackEnergy: 2014 targeted attacks in the Ukraine and Poland," in Virus Bulletin, Seattle, 2014.

Section 3.3.1 - AjaxTM and Rocket Kittens

[1] Cylance, "#OPCLEAVER," 2014. [Online]. Available: https://cdn2.hubspot.net/hubfs/270968/assets/Cleaver/Cylance_Operation_Cleaver_Report.pdf. [Accessed 11 January 2017].

[2] Clearsky, "Gholee – a “protective edge” themed spear phishing campaign," Clearsky, 4 September 2014. [Online]. Available: http://www.clearskysec.com/gholee-a-protective-edge-themed-spear-phishing-campaign/. [Accessed 1 February 2017].

[3] Checkpoint Software Technologies, "Rocket Kitten: A Campaign with 9 Lives," 2015. [Online]. Available: http://blog.checkpoint.com/wp-content/uploads/2015/11/rocket-kitten-report.pdf. [Accessed 11 January 2017].

[4] Clearsky, "Thamar Reservoir An Iranian cyber-attack campaign against targets in the Middle east," June 2015. [Online]. Available: http://www.clearskysec.com/wp-content/uploads/2015/06/Thamar-Reservoir-public1.pdf. [Accessed 2 February 2017].

[5] M. Dahl, "Cat Scratch Fever: CrowdStrike Tracks Newly Reported Iranian Actor as FLYING KITTEN," Crowdstrike, 13 May 2014. [Online]. Available: https://www.crowdstrike.com/blog/cat-scratch-fever-crowdstrike-tracks-newly-reported-iranian-actor-flying-kitten/. [Accessed 11 January 2017].

[6] G. Evron and T. Werner, "Rocket Kitten: Advanced Off-the-Shelf Targeted Attacks Against Nation States," in 31c3 Chaos Communication Congress, Hamburg, 2014.

[7] A. Meyers, "Whois Clever Kitten," Crowdstrike, 4 April 2013. [Online]. Available: https://www.crowdstrike.com/blog/whois-clever-kitten/. [Accessed 11 January 2017].

[8] C. Pernet and K. Lu, "Operation WOOLEN-GOLDFISH," 18 March 2015. [Online]. Available: https://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-operation-woolen-goldfish.pdf. [Accessed 11 January 2017].

[9] C. Pernet and E. Sela, "The Spy Kittens Are Back: Rocket Kitten 2," September 2015. [Online]. Available: https://www.trendmicro.com/cloud-content/us/pdfs/security-intelligence/white-papers/wp-the-spy-kittens-are-back.pdf. [Accessed 11 January 2017].

[10] J. Scott-Railton and K. Kleemola, "London Calling: Two-Factor Authentication Phishing From Iran," Citizenlab, 27 August 2015. [Online]. Available: https://citizenlab.org/2015/08/iran_two_factor_phishing/. [Accessed 11 January 2017].

Section 3.3.2 - MoleRats

[1] Clearsky - Cyber security, "Operation DustySky," January 2016. [Online]. Available: http://www.clearskysec.com/wp-content/uploads/2016/01/Operation%20DustySky_TLP_WHITE.pdf. [Accessed 11 January 2017].

[2] T. Dahms, "Molerats, Here for Spring!," FireEye, 2 June 2014. [Online]. Available: https://www.fireeye.com/blog/threat-research/2014/06/molerats-here-for-spring.html. [Accessed 11 January 2017].

[3] C. Doman, "Moonlight – Targeted attacks in the Middle East," Vectra, 26 October 2016. [Online]. Available: https://blog.vectranetworks.com/blog/moonlight-middle-east-targeted-attacks. [Accessed 3 February 2017].

[4] S. Fagerland, "Systematic cyber attacks against Israeli and Palestinian targets going on for a year," November 2012. [Online]. Available: http://cyber-peace.org/wp-content/uploads/2014/01/ Cyberattack_against_Israeli_and_Palestinian_targets.pdf. [Accessed 11 January 2017].

[5] A. Kasza and E. Idrizovic, "Houdini's Magic Reappearance," Palo Alto Networks, 25 October 2016. [Online]. Available: http://researchcenter.paloaltonetworks.com/2016/10/unit42-houdinis-magic-reappearance/. [Accessed 3 February 2017].

[6] B. Parys, "MoleRats: there's more to the naked eye," PwC, 21 November 2016. [Online]. Available: http://pwc.blogs.com/cyber_security_updates/2016/11/molerats-theres-more-to-the-naked-eye.html. [Accessed 11 January 2017].

[7] N. Villeneuve, T. Haq and N. Moran, "Operation Molerats: Middle East Cyber Attacks Using Poison Ivy," FireEye, 23 August 2013. [Online]. Available: https://www.fireeye.com/blog/threat-research/2013/08/operation-molerats-middle-east-cyber-attacks-using-poison-ivy.html. [Accessed 11 January 2017].

Section 3.4.1 - The Lazarus group

[1] @zerosum0x0, "DoublePulsar Initial SMB Backdoor Ring 0 Shellcode Analysis," Blogspot, 21 April 2017. [Online]. Available: https://zerosum0x0.blogspot.ca/2017/04/doublepulsar-initial-smb-backdoor-ring.html. [Accessed 25 August 2017].

[2] Kaspersky Lab Global Research and Analysis Team, "Lazarus Under The Hood," 2016. [Online]. Available: https://securelist.com/files/2017/04/Lazarus_Under_The_Hood_PDF_final.pdf. [Accessed 25 August 2017].

[3] Novetta, "Operation Blockbuster," 2015. [Online]. Available: https://www.operationblockbuster.com/wp-content/uploads/2016/02/Operation-Blockbuster-Report.pdf. [Accessed 25 August 2017].

[4] D. Alperovitch, "Unprecedented Announcement by FBI Implicates North Korea in Destructive Attacks," Crowdstrike, 19 December 2014. [Online]. Available: https://www.crowdstrike.com/blog/unprecedented-announcement-fbi-implicates-north-korea-destructive-attacks/. [Accessed 11 January 2017].

[5] J. Blasco, "Operation BlockBuster unveils the actors behind the Sony attacks," Alien Vault, 24 February 2016. [Online]. Available: https://www.alienvault.com/blogs/labs-research/operation-blockbuster-unveils-the-actors-behind-the-sony-attacks. [Accessed 11 January 2017].

[6] J. Genwei and J. Kimble, "Hangul Word Processor (HWP) Zero-Day," 2015. [Online]. Available: https://www.fireeye.com/content/dam/fireeye-www/global/en/blog/threat-research/FireEye_HWP_ZeroDay.pdf. [Accessed 25 August 2017].

[7] K. Kochetkova, "What is known about the Lazarus Group: Sony hack, military espionage, attacks on Korean banks and other crimes," Kaspersky Lab, 24 February 2016. [Online]. Available: https://blog.kaspersky.com/operation-blockbuster/11407/. [Accessed 11 January 2017].

[8] M. Lee, W. Mercer, P. Rascagneres and C. Williams, "Player 3 Has Entered the Game: Say Hello to 'WannaCry'," TALOS, 12 May 2017. [Online]. Available: http://blog.talosintelligence.com/2017/05/wannacry.html. [Accessed 25 August 2017].

[9] C. Raiu, K. L. G. R. &. A. Team and J. A. Guerrero-Saade, "Operation Blockbuster revealed," Kaspersky, 24 February 2016. [Online]. Available: https://securelist.com/blog/incidents/73914/operation-blockbuster-revealed/. [Accessed 11 January 2017].

[10] R. Sherstobitoff, I. Liba and J. Walter, "Dissecting Operation Troy: Cyberespionage in South Korea," July 2013. [Online]. Available: http://www.mcafee.com/ca/resources/white-papers/wp-dissecting-operation-troy.pdf. [Accessed 11 January 2017].

[11] D. Tarakanov, "The “Kimsuky” Operation: A North Korean APT?," 11 September 2013. [Online]. Available: https://securelist.com/analysis/publications/57915/the-kimsuky-operation-a-north-korean-apt/. [Accessed 11 January 2017].

Section 3.4.2 - Careto

[1] McAfee Labs, "Careto Attack – The Mask," 12 February 2014. [Online]. Available: https://kc.mcafee.com/resources/sites/MCAFEE/content/live/PRODUCT_DOCUMENTATION/25000/PD25037/en_US/McAfee_Labs_Threat_Advisory_Careto_Attack_The%20Mask_3.pdf. [Accessed 11 January 2017].

[2] Kaspersky labs, "Unveiling "Careto" - The Masked APT," February 2014. [Online]. Available: https://kasperskycontenthub.com/wp-content/uploads/sites/43/vlpdfs/unveilingthemask_v1.0.pdf. [Accessed 11 January 2017].

Section 4 - EXERCISE SCENARIOS

[1] M. Kellet and M. Bernier, "Cyber threat data model - High-level model and use cases (DRDC-RDDC-2016-D080)," DRDC, 2016

DOCUMENT CONTROL DATA

(Security markings for the title, abstract and indexing annotation must be entered when the document is Classified or Designated)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g., Centre sponsoring a contractor's report, or tasking agency, are entered in Section 8.)

International Safety Research (ISR)
38 Colonnade Road North
Ottawa, Ontario
Canada K2E 7J6

2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.)

CAN UNCLASSIFIED

2b. CONTROLLED GOODS

NON-CONTROLLED GOODS
DMC A

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

TA-35—Cyber Threat Data Model and Use Cases: Final Report

4. AUTHORS (last name, followed by initials – ranks, titles, etc., not to be used)

Lemay, A.

5. DATE OF PUBLICATION (Month and year of publication of document.)

September 2017

6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)

56

6b. NO. OF REFS (Total cited in document.)

96

7. DESCRIPTIVE NOTES (The category of the document, e.g., technical report, technical note or memorandum. If appropriate, enter the type of report, e.g., interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Contract Report

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

DRDC – Centre for Operational Research and Analysis
Defence Research and Development Canada
101 Colonel By Drive
Ottawa, Ontario K1A 0K2
Canada

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)

Project: 05ac - Cyber Decision Making and Response (CDMR)

9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)

W7714-156105-T35

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)

DRDC-RDDC-2017-C290

10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

ISR Report 6099-01-03, Version 2.0

11a. FUTURE DISTRIBUTION (Any limitations on further dissemination of the document, other than those imposed by security classification.)

Public release

11b. FUTURE DISTRIBUTION OUTSIDE CANADA (Any limitations on further dissemination of the document, other than those imposed by security classification.)

12. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

This report documents the efforts related to the production of a data model based on the STIX 2.0 format to characterize cyber threats. The work produced four main outcomes:

- An analysis of the suitability of the STIX 2.0 standard to support the characterization of cyber threats;
- STIX 2.0 compliant data models to support automation or analysis;
- Profiles of Advanced Persistent Threat (APT) actors groups using the STIX 2.0 format; and
- Examples of exercise scenarios using the APT actor profiles to demonstrate use cases.

The main findings regarding the suitability of the STIX 2.0 standard are as follows:

- The standard is designed to represents threat information using graphs, with the various objects (threat actors, tools and malware, vulnerabilities, identities, etc.) modeled as nodes and the relationships between the objects represented as edges;
- The standard is designed around the concept of a minimum viable product, with a small number of rules and a large capacity for customization;
- The lack of enforced structures lends itself well to so-called “NOSQL” approaches, but makes automated processing more complex as the same information can be expressed in multiple forms; and
- The standard, at the time of writing, lacks in maturity with continually evolving documentation and only partial software support.

In terms of presenting a model, the report proposes to either embrace the unstructured nature of the standard in a NOSQL, or to enforce a certain structure to facilitate information retrieval. In the case of the NOSQL model, this would support the use of STIX 2.0 as a method to store indicators of compromise and provide some additional context in automated systems. In the case of the structured model, the use of predictable structures to store information would help analysts retrieve and cluster information, however at the cost of increased processing for storage.

Twelve different APT groups are profiled using a data model similar to the one proposed (due to lack of certain functionalities in the STIX 2.0 code base) covering a range of nation-state sponsors and a range of tools, techniques and procedures (TTPs). These profiles illustrate the breath of characteristics that can be modeled using STIX 2.0.

Finally, two exercise scenarios using the profiles demonstrate how cyber intelligence tasks would be performed. In the first scenario, participants perform a cyber-response, attributing the threat and extrapolating goals and potential impacts. In the second scenario, participants perform a cyber-intelligence planning process, generating tactical indicators and issuing warnings based on the threat level. The exercise description also provides indications on which particular object or object properties would be used in each step of the scenario.

Ce rapport documente les efforts liés à la production d'un modèle de données basé sur le format STIX 2.0 afin de caractériser les cyber-menaces. Le projet a abouti à quatre éléments principaux:

- Une analyse de la pertinence de la norme STIX 2.0 pour soutenir la caractérisation des cyber-menaces;
- Des modèles de données conformes à STIX 2.0 pour prendre en charge l'automatisation ou l'analyse;
- Des profils des groupes d'acteurs APT (Advanced Persistent Threat) utilisant le format STIX 2.0; et
- Des exemples de scénarios utilisant les profils d'acteurs APT pour démontrer des cas d’application.

Les principales conclusions concernant la pertinence de la norme STIX 2.0 sont les suivantes:

- La norme est conçue pour représenter les informations sur les menaces à l'aide de graphiques, avec les différents objets (acteurs, outils et logiciels malveillants, vulnérabilités, identités, etc.) modélisés comme des noeuds et les relations entre les objets représentés comme des arêtes;
- La norme est conçue autour du concept de produit minimum viable, avec un petit nombre de règles et une grande capacité de personnalisation;
- Le manque de structures renforcées se prête bien aux approches dites «NOSQL», mais rend le traitement automatisé plus complexe car les mêmes informations peuvent être exprimées sous plusieurs formes; et
- La norme, au moment de la rédaction, manque de maturité avec une documentation en constante évolution et seulement un support logiciel partiel.

En termes de présentation d'un modèle, le rapport propose soit d’adopter la nature non structurée de la norme dans un NOSQL, soit d'imposer une certaine structure pour faciliter la recherche d'information. Dans le cas du modèle NOSQL, cela favoriserait l'utilisation de STIX 2.0 comme méthode de stockage des indicateurs de compromission et fournirait des informations supplémentaires dans des systèmes automatisés. Dans le cas du modèle structuré, l'utilisation de structures prévisibles pour stocker les informations aiderait les analystes à récupérer et à regrouper les informations, mais au prix d'un traitement accru pour le stockage.

Douze groupes APT différents sont profilés en utilisant un modèle de données similaire à celui proposé (en raison du manque de certaines fonctionnalités dans la base de code STIX 2.0) couvrant une gamme de parrains étatiques/nationaux et une gamme d'outils, techniques et procédures (TTPs). Ces profils illustrent la variété des caractéristiques qui peuvent être modélisées en utilisant STIX 2.0.

Enfin, deux scénarios d'exercices utilisant les profils de groupes APT démontrent comment les tâches de cyber-renseignement seraient exécutées. Dans le premier scénario, les participants effectuent une cyber-intervention, attribuant la menace et extrapolant les objectifs et les impacts potentiels. Dans le deuxième scénario, les participants exécutent un processus de planification de cyber-renseignement, générant des indicateurs tactiques et émettant des avertissements basés sur le niveau de menace. La description de l'exercice indique également quel objet particulier ou quelles propriétés d'objet serait à utiliser à chaque étape du scénario.

13. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g., Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

threat, data model, threat model, cyber intelligence, Cyber Threat, Advanced Persistent Threat, Use Cases