CCDCOE

NATO Cooperative Cyber Defence Centre of Excellence Tallinn, Estonia

Mauno Pihelgas (ed.)

Mitigating Risks arising from False-Flag and No-Flag Cyber Attacks

www.ccdcoe.org
publications@ccdcoe.org

Contents

Preface

Difficulty of attribution is one of the main challenges for nations in reducing the overall insecurity coming from cyberspace and addressing specific malicious actors. Lack of accurate attribution creates legal, doctrinal, operational and practical difficulties when responding to cyber attacks. Conversely, misattribution is a problem when an attack is made to appear to have attributed from another source.

NATO and other bodies have discussed various information-sharing approaches that might also contribute to reducing these risks. Additionally, international organisations such as the UN and OSCE have proposed confidence-building measures (CBMs) to also reduce the risk that misattribution could lead to conflict or escalation.

1 Introduction

This report on mitigating risks arising from false-flag and no-flag cyber attacks handles issues related to establishing proper attribution following cyber attacks, in which the entity responsible for launching the attack is unknown (*no-flag*) or considered falsified (*false-flag*) for any other reasons.

1.1 Terminology

As the reader has probably noticed, we have already used two somewhat ambiguous terms which might require some further explaining.

False-flag - A diversionary or propaganda tactic of deceiving an adversary into thinking that an operation was carried out by another party. [1]

No-flag - An operation conducted while being undeclared and when the operatives are either scantily marked or entirely unmarked. [2]

Historically, the term *false-flag* originated in the maritime domain, where a ship of one country would fraudulently sail under the flag of another country. [2]

1.2 Assumed background knowledge

When we consider attribution, and especially misattribution, from a technical perspective, it involves many technical details. Since explaining the technical details is out of the scope of this report, some previous understanding of computer networks, anonymisation techniques, and tracking back malicious users is expected from the reader. To help others, an article from Pihelgas [3] describes these topics in a simple manner.

1.3 Document outline

The document has been divided into several sections. The first chapter, about cyber information exchange, describes various programmes, collaboration initiatives and protocols for attribution of malicious cyber activity. It also discusses what the barriers to effective collaborative data exchange are, as well as what some of the common issues with attribution through data exchange can be. Then, a chapter describes the information-sharing between security practitioners that takes place in Operational Security Communities. Following that, a third section describes the process and the discussions that took place during a workshop that was specifically designed to test procedures in situations where attribution is lacking. The participants were asked to take part in the scenarios while discussing procedures that should be followed and offering their recommendations on potential courses of action. Finally, a summary of the project and some conclusions are given.

1.4 Project description

This project will propose technical means for reducing misattribution among parties with different levels of trust as well as proposals for procedures to employ during a crisis situation in order to establish attribution and reduce misattribution.

The aim of the project is to provide nations and international organisations with:

- An overview of challenges in countering false-flag and no-flag attacks;
- An analysis of current information-sharing initiatives, and a technical analysis of what type of information-sharing would have a demonstrably positive effect on the efficacy of attribution;

- recommendations for technical means for reducing misattribution among different levels of trust; and
- proposals for procedures to employ during a crisis situation.

The primary parts described in this report roughly follow the aims listed above in the Document outline.

1.5 Original request for support

This report is based on a Request for Support to the NATO Cooperative Cyber Defence Centre of Excellence (NATO CCD COE) dated 1 April 2013, initiated by a request from the Estonian Information System Authority's Cyber Security Branch (the Originating Organisation). This request was submitted through the NATO CCD COE Steering Committee and was approved for implementation.

Further communication with the Originating Organisation in order to clarify the scope revealed that their interest has grown from a primarily technical report to more of a multi-disciplinary report. During our work, an idea grew in the project team that we should hold a one-day table-top exercise with all relevant stakeholders to actually test some of the primary concepts that we have developed.

1.6 Acknowledgements

The author would like to thank everyone who provided their ideas and input to help develop this project: Anna-Maria Osula, Graeme Park, MAJ Harry Kantola, Henry Rõigas, Hillar Aarelaid, Johannes Tammekänd, Lauri Luht, Liina Areng, Liisa Past, and Teemu Uolevi Väisänen.

2 Cyber Information Exchange – Collaboration for Attribution of Malicious Cyber Activity

Graeme Park, Mauno Pihelgas

2.1 Introduction

As cyberspace grows in size and complexity, the cyber threat continues to increase apace. Once the domain of individuals with an interest in nothing more than notoriety, cyber threats have now become militarised, politicised and monetised. This shift in the paradigm has led to adversaries who are well resourced, more determined and more dangerous than the prolific hackers of yesteryear.

Whilst legislation and other legal tools play catch-up with an ever-changing digital landscape, one thing is certain: without sufficient attribution, it is impossible to enforce regulations, laws or treaties. Technical means exist, but many are easily duped, and competing priorities of non-repudiation versus privacy and freedom of speech create a division in the requirements of internet users. Additionally, as the internet is globally interconnected with traffic crossing multiple national boundaries, malicious actors are often well beyond the jurisdiction of the victim.

The issue is further complicated when attribution is defined. It is not enough to just locate a source IP address (unless looking solely at active defence): the identity of the attackers must be determined, as well as the parties they were acting on behalf of must also be unmasked.

In order to supplement the solely technical means of attribution, collaborative data exchange must ensure that when large amounts of data are brought together, data mining techniques and statistical analysis can afford us additional clues as to the author of such tools with a higher degree of certainty than technical means or independent data alone. By correlating the shared information, a more effective method for a community to detect potential risks and prevent cyber attacks at an early stage can be developed. [4]

This document will discuss the most prevalent protocols aimed at standardising information about malicious cyber activity, compare a number of programmes for cyber information-sharing partnerships, and finally discuss the issues surrounding collaborative data exchange.

2.2 Barriers to effective collaborative data exchange

There is an ever-growing consensus that Cyber Information Data Exchange should happen, yet whilst both governments and private organisation are sponsoring and generating the standards for cyber information data exchange, the rate of data exchange is not near the level it needs to be in order to effectively stem cyber incidents, let alone to assist with attribution. The ENISA report 'Detect, SHARE, Protect' warned around 200 major CERTs across Europe – including 21 in the UK – that 'the ever-increasing complexity of cyber-attacks requires more effective information-sharing' but that those involved were showing a 'lack of interest' in doing so. There are a number of reasons why organisations, partnerships or governments may be reluctant to share this kind of information; and understanding these reasons is the beginning of overcoming these obstacles and implementing an inclusive system.

2.2.1 Trust between partners

Trust between partners is essential. If information is not self-generated, then you must be able to implicitly trust the author, and a lack of trust in the integrity of the data received means that an information-sharing programme is destined to fail. This is one of the biggest issues facing data exchange; whilst there may be a level of trust within partnerships and multi-lateral franchises, pertinent information should often be shared sector-wide and not many organisations will implicitly trust their adversaries. Malicious intent aside, trust can be developed on sharing of useful information, although if information is considered to be less than useful, this can have a negative effect on any sharing relationship with other customers less likely to reciprocate high quality information. The crux of the issue revolves around the fact that the wider spread of the contributors to an information exchange, the better the overall fidelity of the information. Conversely, the more partners there are, the less likely this is to engender trust, leaving data exchanges in a paradoxical position.

In order to combat this, federated domains or trust partnerships can be established within a programme. This would allow for a number of communities with particular interests or relationships to share information whilst maintaining the Confidentiality and Integrity elements of the CIA triad.[1]

2.2.2 Reputational damage

Whilst there may be a long list of organisations and governments wishing to consume cyber threat information, sharing programmes in any form suffers from two kinds of parasitic users, and despite the content of these exchanges, both *lurkers* – those who observe but do not necessarily contribute – and *leechers* – those who maintain a negative ratio of downloaded or uploaded data – are common. It may be assumed that, due to the nature of this domain, this type of activity would be limited; however, as Johnson *et al*. state in their guide to cyber threat information-sharing: 'Knowledge of an adversary's TTPs[2] is advantageous … but sharing of this information may put the contributor at risk by exposing the protective or detective capabilities of the organization and result in threat shifting by the adversary.' No rational state is likely to dispose of its strategic advantages in these areas by making actionable information available to every other nation in the world through a multilateral organisation. [5]

In order to combat these issues and generate quantifiable data sets, there must be an incentive to data share. This would encourage those partaking in any network to contribute to the data sources and thus the continued success of any party. Yet with information-sharing comes the requirement for non-attribution. Effective trust partnerships limit the ability to contribute without naming the author but the only way this could be truly implemented is by way of an information broker who would act as a cut-out for any organisation wishing to contribute sensitive information without fear of attribution. The impact of this must be understood, as without an author, contextual understanding is more difficult and the data begins to lose an element of integrity and thus its credibility.

[1] Confidentiality, Integrity and Availability
[2] Tactics, Techniques and Procedures

2.2.3 Cost benefit

One of the final hurdles to overcome is the cost. Participation in any one of these programmes is likely to require a financial outlay. Programmes may cost money to join, and there are implementation fees associated with new hardware and software, as well as organisational change that may need to happen to implement these tools and which will in turn require training on new processes. Upfront costs may be self-evident, but additional costs are not always tangible or foreseen. Taking part in an information exchange programme such as the ones identified below requires cultural change and backing from the organisation's hierarchy. Therefore, the benefits of such a scheme must be demonstrated to the non-technical manager in order to secure financial backing.

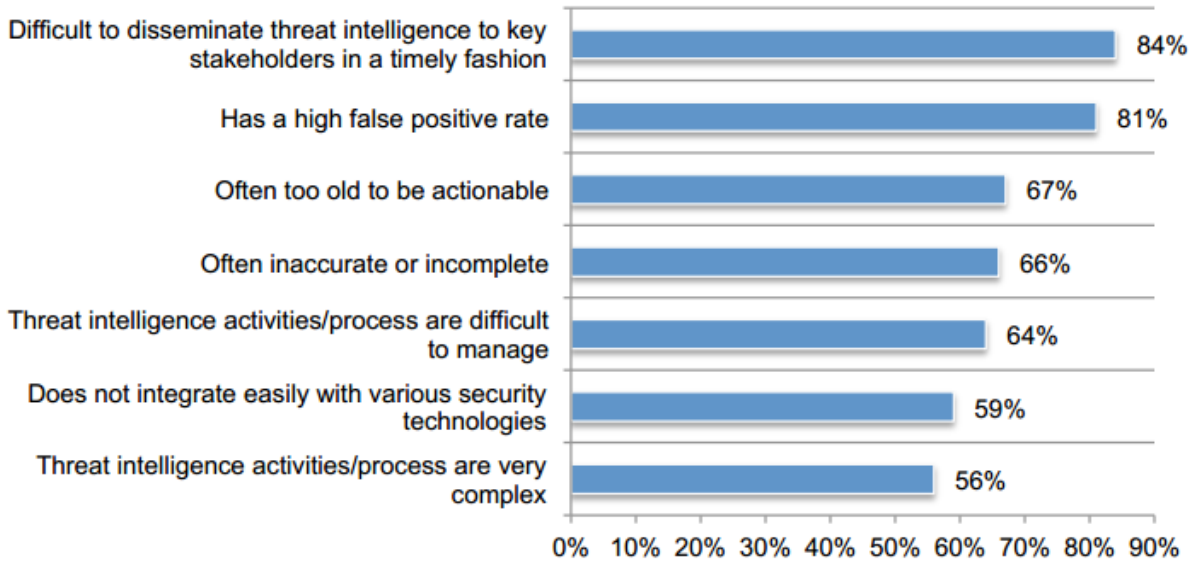

Figure 1 - Perceptions about problems with cyber threat intelligence [6]

Despite many international organisations talking about information-sharing, a survey from the Ponemon Institute reviewed perceptions (see Figure 1) about cyber threat intelligence and, if this is anything to go by, it is easy to see why organisation are not huge proponents of these programmes. The key issues identified by the Ponemon Institute are the requirement for better data sets, and an overall simpler mechanism for implementation and information-sharing. In order to address the data-sets issue, the correct level of granularity in data is required and this is often limited by the protocols involved. The next section will review some of the most prevalent protocols for cyber information data exchange. [6]

2.3 Data exchange protocols

Information exchange between interested parties is not uncommon; IT Security Officers and nowadays national and military CERTs have long exchanged information in order to assist with attribution, but this has taken place with methods as informal as email and attachments which, although direct, does not allow for effective information-sharing, management or exploitation. In 2008 it was stated that 'existing ontologies are not prepared for being reused and extended and the security community still needs a complete security ontology that solves these lacks and provides reusability, communication and knowledge sharing'. [7] In that context, collaborative programmes for information-sharing are on the rise and these are underpinned by protocols that are growing in effectiveness to allow for human-readable automation. While a single complete security ontology may be an unreachable goal, it is possible to make a set of ontologies interoperable, covering all aspects of security, and this would address the requirements. [8]

There are numerous data exchange protocols for sharing data within a collaborative environment, and all of them provide different levels of granularity. The three main protocols that will be reviewed are the Security Content Automation Protocol (SCAP), the Security Threat Information eXpression (STIX) and the Internet Object Description Exchange Format (IODEF). All proposed protocols have different levels of granularity, but have the common aim of seeking to codify atomic (strings, emails, FQDNs), behavioural (habits, profiles) and computed (hashes, IDS signatures) values with the aim of increasing information-sharing throughout relevant partnerships.

2.3.1 Security Content Automation Protocol (SCAP)

SCAP is a standard developed by the National Institute of Standards and Technology (NIST) which is aimed at automating vulnerability management, asset inventory and policy compliance. The protocol combines a number of existing standards, such as Common Vulnerabilities and Exposures (CVE), Common Configuration Enumeration (CCE), Common Platform Enumeration (CPE) and Open Vulnerability and Assessment Language (OVAL).

These existing standards are combined according to the SCAP control framework in order to perform initial assessments and continuous monitoring of software and systems. Whilst the concept is sound, the National Vulnerability Database (NVD), the US Government's repository for SCAP data, appears relatively immature and only a handful of tools have been validated by the NVD to conform to SCAP. Secpod.com and scaprepo.com host a large SCAP repository that is free to use and has a mature search tool able to process human-readable queries. Yet whilst SCAP has been used to good effect here, it is still not particularly useful for attribution. Nevertheless, these databases can assist with continual monitoring of best practice cyber defence.

The standard remains in constant development and the team specifically state 'we envision further expansion in compliance, remediation, and network monitoring, and encourage your contribution relative to these and additional disciplines'. [9]

SCAP has been criticised for its requirement of implicit rationale and definition relationship use, its unclear relationships and name ambiguity [10]. Whist SCAP brings together a repository of existing standards, its reliance on this information is also its weakness. In attempting to use many diverse yet pre-existing standards, there is duplication of effort and arbitrary information prevalent in all repositories. Whilst SCAP may have some uses, using it for threat identification and ultimately attribution would be futile. SCAP is ultimately a compliance system and although it incorporates a number of Independent Topic Ontologies (ITOs) such as CVE, CPE and CVSS, it does not address the key issues such as threats, actors and campaigns.

2.3.2 Structured Threat Information eXpression (STIX)

STIX is a collaborative programme developed by the MITRE corporation, aimed at designing and developing a standardised language to represent cyber threat information, it 'strives to be fully expressive, flexible, extensible, automatable, and as human-readable as possible'. [11] Relying on a number of existing standards, it

was initially designed as a threat-orientated data exchange protocol, but has since added numerous other fields in order to incorporate incidents to provide context to existing threats. The standard also adds specific information about threat actors, TTPs and other relevant information (Figure 1). Based on an XML framework, almost everything in this definitively-structured language is optional, such that any single use case could leverage only the portions of STIX that are relevant to it, from a single field to the entire language or anything in between without being overwhelmed by the rest. [12]

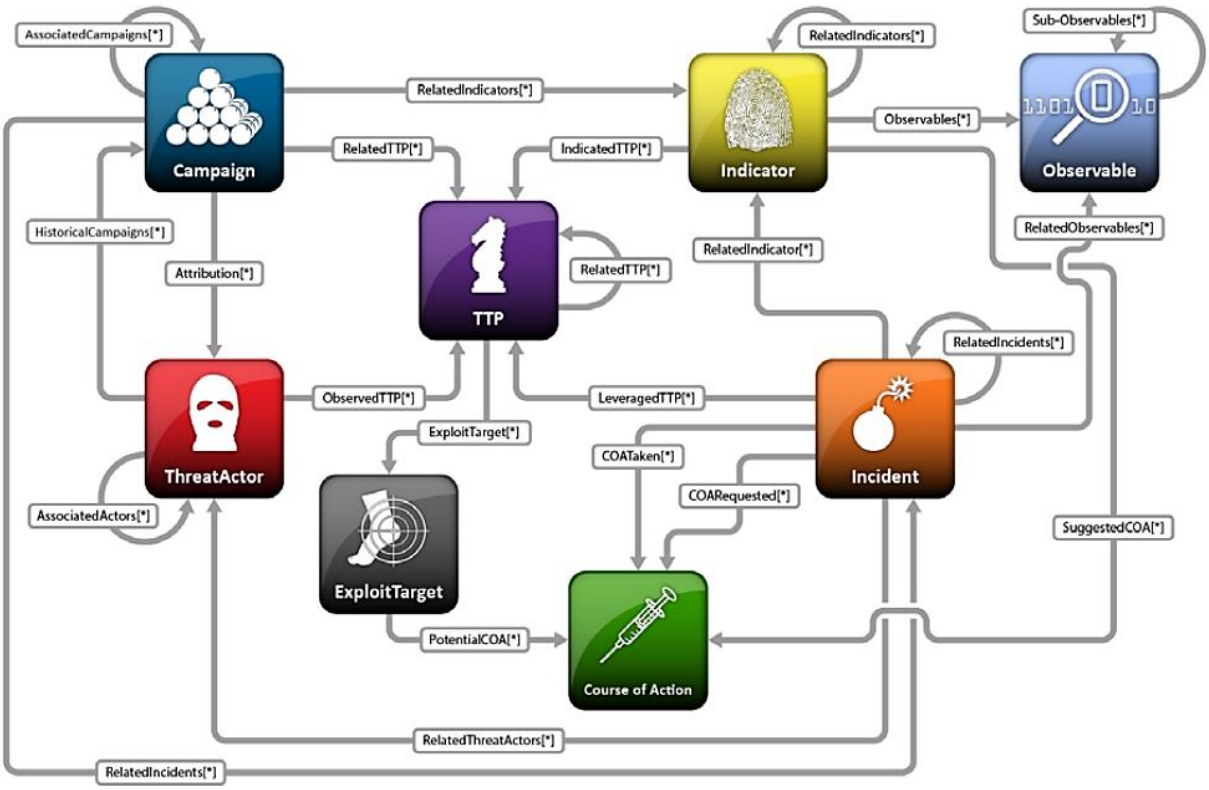

Figure 2 - STIX Architecture [12]

The STIX architecture seeks to go beyond any other expressive language in terms of its all-encompassing approach. This ambitious approach would allow for all of its core concepts to leverage existing standards (ITOs) where possible, but the architecture is certainly not a slave to these standards. The STIX whitepaper outlines 6 potential use cases which are represented in Figure 3 and range from assessment of cyber threats to response and mitigation, notably all defensive in nature.

The protocol itself provides a wide range of interoperability with ITOs, but also provides mechanisms for leveraging other critical schemas such as Common Attack Pattern Enumeration and Classification (CAPEC), Malware Attribute Enumeration and Characterisation (MAEC) and Common Vulnerability Reporting Framework (CVRF). STIX prides itself on its adaptability, and as a community effort it continues to gain real-world use and undergoes routine refinement based on user requirements.

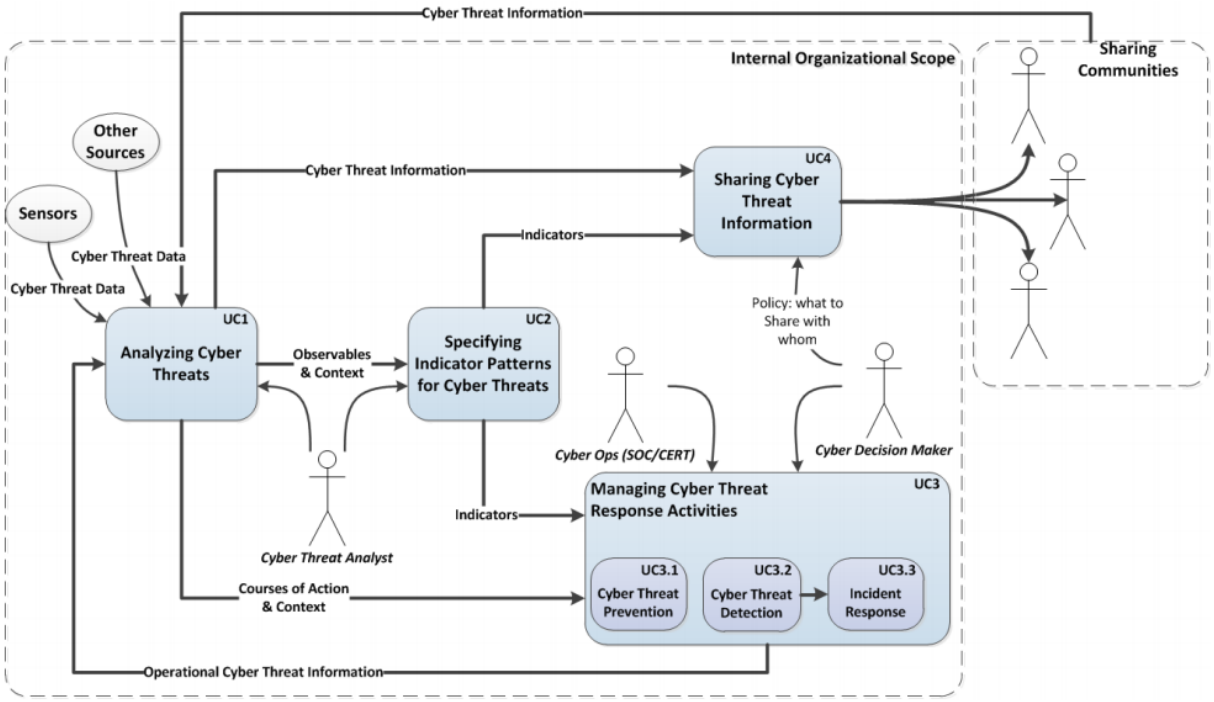

Figure 3 - STIX use cases [12]

The Trusted Automated Exchange of Indicator Information (TAXII) is the recommended format for exchanging STIX information. It is a project that was sponsored by the US Department of Homeland Security. It is not in itself a sharing initiative, but rather a means of sharing the information within a set programme or agreement. TAXII would provide three independent topologies (Figure 4) and it would be up to the particular sharing initiative to decide which ones to implement.

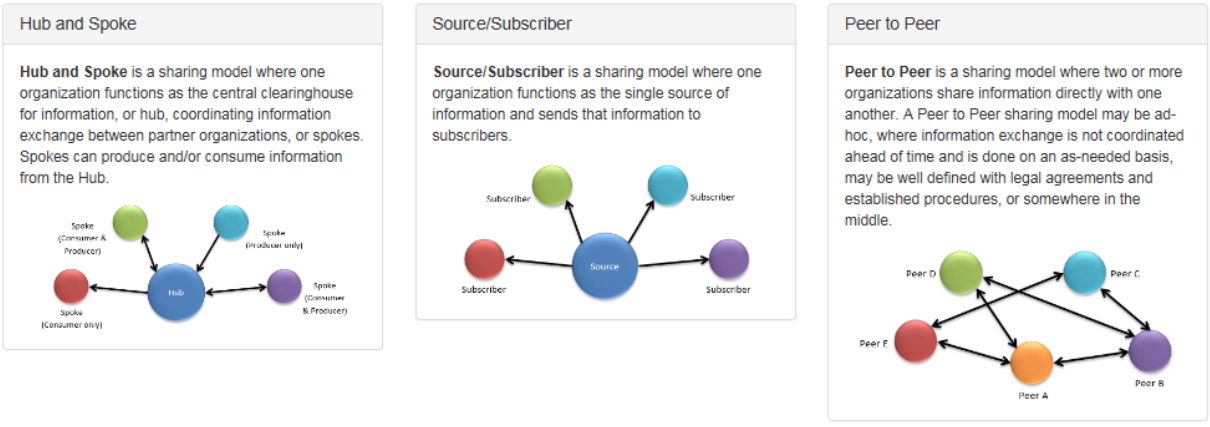

Figure 4 - TAXII sharing topologies [13]

The STIX standard certainly has potential and has already been adopted by a number of influential organisations, especially in the United States by organisations such as the Department of Homeland Security and Hewlett-Packard and it will be implemented into the NATO Cyber Data Information Exchange (CDXI).

2.3.3 The Incident Object Description Exchange Format (IODEF)

The Incident Object Description Exchange Format (IODEF) is an Internet Engineering Task Force (IETF) standard created as a data format for CERTs to exchange cyber incident information. Based on XML, IODEF provides the recommended format for exchanging alarms, alerts, incidents and other relevant security information between individual devices and local monitoring points to centralised analysis centres. [14] Initially designed as a reporting system for ongoing incidents, it has recently been expanded and amended.

The IODEF data model includes over 30 classes and sub-classes used to define incident data. The classes cover a wide range of information, including contacts, monetary impact, time, operating systems and applications, and also includes labels to comment upon such things as confidence and sensitivity. The language again allows for the leveraging of existing ITOs such as CAPEX, CPE, CVE and OVAL. [15]

The format is currently being used in a number of places such as the Anti-Phishing Working Group (APWG), and the Collective Intelligence Framework. The APWG has extended the IODEF standard to support additional elements regarding email incidents and phishing, which demonstrates its flexibility.

In 2014 the Managed Incident Lightweight Exchange (MILE) working group has also proposed an extension to the IODEF standard (RFC 2703) which would support additional data such as attack patterns, countermeasure instructions, event logs and severity. [16]. The IODEF-SCI extensions can be seen in Figure 5 below.

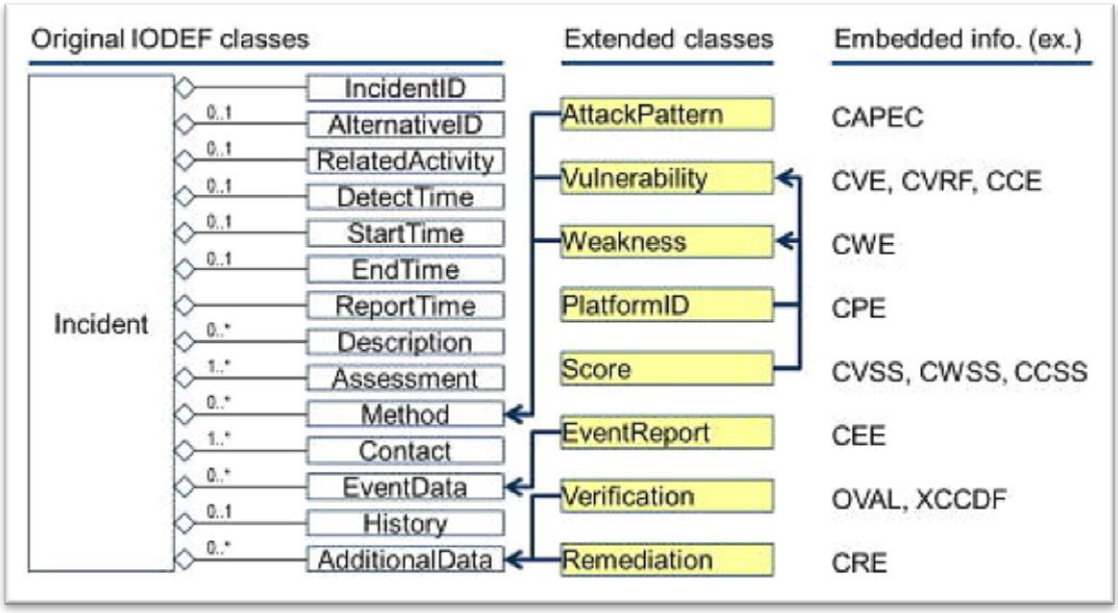

Figure 5 - IODEF-SCI – Extensions [16]

The biggest flaw with IODEF is that it was built initially to share incident data, not indicators of compromise (IoC) and was therefore reliant on other formats in order to describe TTPs or campaigns. IODEF does a good job at what it was initially designed for, but expanding something to fit a different requirement is never as good as defining a particular requirement from the outset. With attribution in mind, STIX is still a much more practical protocol (although not specifically designed for exchanging attribution information).

Real-time Inter-network Defence (RID) is a standard for communicating cyber threat information that could be defined in a number of different schemas. Used as the key transportation protocol for IODEF information, the goal of RID is to add elements of security (HTTPS/TLS), non-repudiation (digital signatures) and elements of

encryption and anonymity to the data exchange. It consists of a discrete number of message types to allow pulling and pushing of message data, and has a policy class to allow for the realisation of federated user groups. The proposed relationships that could be considered are client-to-vendor, inter/intra community and peer-to-peer.

All of the schemas and protocols above have a number of similar touch points, but were each enacted to serve slightly different purposes. Whilst they have a number of commonalities and rely upon a number of the same ITOs, the most fitting protocol is determined by the customer and their needs.

In terms of addressing attribution, the most suitable would be data exchanged within the STIX format, due to its high fidelity and information that would best assist intelligence assets at confirming the source of any attack or breach; yet a combination of IODEF/STIX and SCAP may represent the best format for increasing threat management, compliance and assurance.

The schemas and transport protocols are only a small element of the equation. The systems that are defined based upon these protocols are what make the difference, as it is not necessarily a matter of the format that data is exchanged in, but who is an active consumer and producer of the data. The next section will review a number of prominent information exchange programmes.

2.4 Information exchange programmes

There are numerous information-sharing initiatives currently live or in development. Each of them offers something slightly different, be it automation or lack thereof, context applied to data, forums for specific industries, or different modelling systems. Many initiatives go no further than providing a forum for discussion, with data exchanges offering nothing more than a web portal with access allowed to interested parties, but tools are beginning to emerge that focus more on information exchange rather than community development. More recently, subject-matter experts from the RSA organisation have stated that data standards for describing and transmitting threat information have advanced significantly, but much progress is needed to extend existing standards and drive wider adoption in vendor solutions. Threat information sharing and collaboration programs help organisations to augment their expertise and capabilities in detecting and remediating advanced threats, but most sharing programs are hindered by a heavy reliance on manually intensive, non-scalable processes and workflows. [8]

A number of different information exchanges are in development because one tool does not fit all situations. The tools do all have some common features, such as the dependency upon existing data ITOs and requirements for structured exchange, but the way to implement these requirements differs. There are also subtle differences in their purposes and the requirements of their customer bases. A number of programmes that seek to address the barriers to information exchange identified in the first chapter are reviewed below.

2.4.1 Cyber Defence Data Exchange and Collaboration Infrastructure (CDXI)

NATO's CDXI is a data exchange programme that has undergone significant development over the last few years. Still in the conceptual stages, it is hoped that it will offer a complete solution featuring an Agile Data Model (ADM). As no one standard for defining data suits all organisations, there is disagreement even within similar sectors about what standard to use; the ADM allows for the CDXI program to remain somewhat agnostic.

The proposed system would implement a range of schemas and have a flexible approach regarding cost models. The main selling points of the CDXI are:

1) The incorporation of an Agile Data Model (ADM) which would allow data providers to modify their data models and data consumers to adjust their automated applications independently and at their own pace.
2) The encouragement of dissention by allowing multiple possible values within one field in order to expose disagreement and encourage consensus to be sought. [8]

Currently in a very nascent form, it is undergoing concept validation with key stakeholders and an early prototype could be developed by George Washington University in 2015, with full project delivery over the next five years, subject to financial agreements.

At present the closest among NATO's offerings is the Multinational Cyber Defence Capability Development (MNCD2), currently consisting of three distinct work packages (WP1-3) spanning Technical Information-sharing, Cyber Defence Situational Awareness and Distributed Multi-sensor Collection and Correlation Infrastructure. It is a collaborative effort by NATO nations Canada, Denmark, Norway, The Netherlands and Romania. [17]

WP1 specifically looks at formalising Technical Information-sharing and has trialled the Cyber Information and Incident Coordination System (CIICS) that has been developed under this package. The system is an automated information exchange that is hosted in a web-based GUI, which leverages some of the classes defined in STIX. At present the information classes are not broad enough to store Indicators of Compromise and TTPs, but this will be added as the project matures. The project is a precursor to CDXI which will incorporate a broader range of data, add the human input element, and most importantly add the agile data model. This system is currently in use only by the contributing nations, but is expected to be available for free for all NATO nations.

NATO has also already developed the Malware Information-sharing Programme (MISP) which is a semi-automated searchable repository that is available to NATO member nations. Whilst they benefit from information-sharing, the context of the attack is removed. It has been well received by NATO members and already has over 500 entries after just one year of operation, signalling nations' willingness to contribute information if the environment is deemed conducive. [18]

As CIICS gains momentum within the NATO community and they begin to transition to an all-encompassing data model which allows industry to contribute to data sets, CDXI will provide a very useful tool for assisting with wholesale attribution alongside national technical means.

2.4.2 AbuseSA

AbuseSA is an automated data aggregation tool. It receives feeds from multiple different sources and 'harmonises' the data in order to represent it in different human-readable formats. Feeds can be aggregated from security companies, partners and other organisations to allow for a holistic and representative overview of vast amounts of data.

The tool is based on a distributed architecture that receives its feeds from different sources before sorting them into independent 'chat rooms' based on type of 'abuse' represented such as malware or spam. Information is exchanged within the program by using XMPP with TLS, and chat-room information can be represented in several visual formats including wikis, SQL reports or mail digests.

Figure 6 shows an example of the AbuseSA GUI representing a number of different rooms and incidents on a global scale, and Figure 7 shows some of the more detailed information that can be sought regarding a particular incident or event.

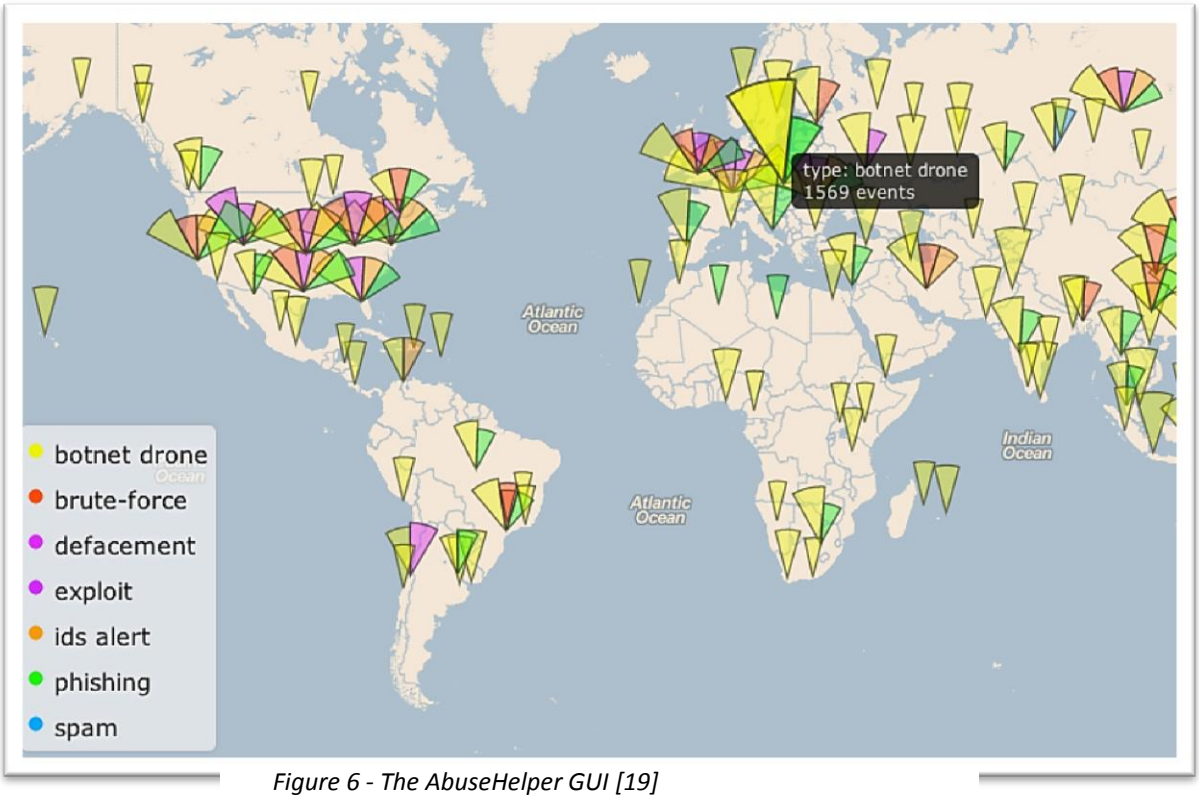

Figure 6 - The AbuseHelper GUI [19]

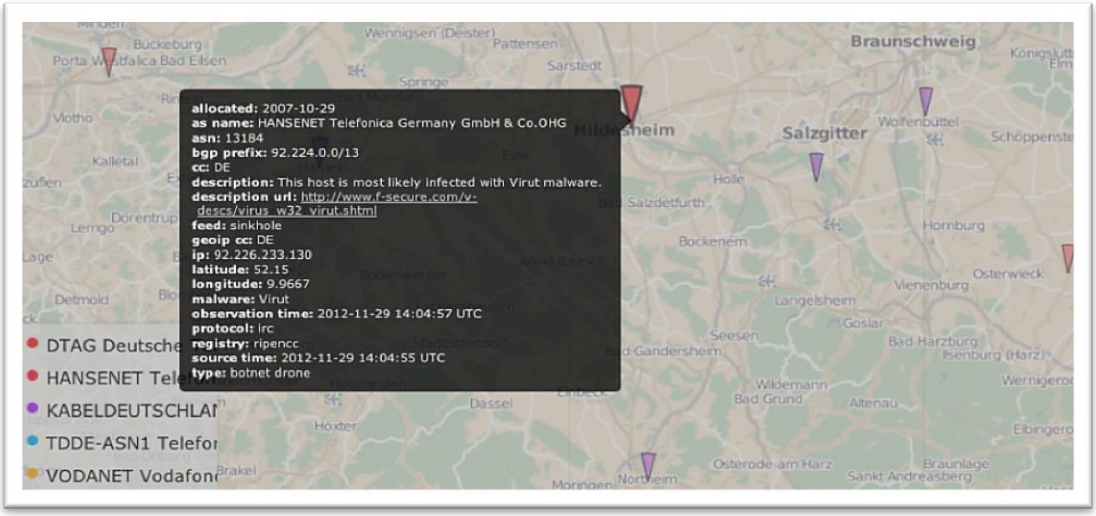

Figure 7 - AbuseHelper Extended Information Report [19]

What is quickly apparent is that the tool does a good job of saving human resources by filtering, harmonising and normalising vast amounts of data, and the graphical representations allow lower lever of technically qualified people to interpret the sources.

Data is aggregated from both external and internal sources in a multitude of formats, making AbuseSA completely protocol and format agnostic. This aids with rapid integration into already existing systems and, as Codenomicon state, 'the solution for automation, reporting, and visualisation of related information must not be hardcoded to the current problem set. The ability to deal with new data has been built in to the system and its design. As a result, AbuseSA is able to adapt to the constant change which is characteristic to the domain.' [19]

Whilst the system may be able to represent any XML data such as IODEF or STIX, it would require an additional level of configuration in order to understand some of the less common data expressed in these protocols. Without this configuration it is likely that when the harmonisation takes place this data will be truncated. In addition to the threat representation and automation, AbuseSA offers a series of internal network sensors which take the AbuseSA feeds and apply them against network traffic to highlight known malicious activity and reduce false positives.

The system allows for collaboration in smaller communities as it is up to the end user to define the information feeds that they will sign up to. For example CERT-UK has recently begun consuming NCSC-Fi sources and will reciprocate in the near future. [20] It remains the prerogative of the information source to share as little or as much information as they see fit as long as they meet the minimum criteria.[3] That said, whilst it is possible to remove some elements of the information when organisations are concerned about reputation or secretive information (the feed field can be obscured), it still remains apparent which source has provided the information as AbuseSA works directly to their server. The lack of anonymity afforded to its users may create a significant barrier to sharing certain types of information outside of trusted communities, and may make some organisations parsimonious with the granularity of data.

[3] The minimum requirements for an abuse report to be actionable include: Identity Key (IP, Domain, URL, and/or email address), Time Stamp (Observation and source), Classification (Type and Taxonomy), and Feed. The Feed can be obscured for sources to be anonymized.

2.4.3 The Cybersecurity Information Exchange Framework (X.1500)

The Cybersecurity Information Exchange Framework known as CYBEX is an attempt by the Technology Arm of the International Telecommunication Union (ITU), a specialised UN agency, to document a structured language for cyber security information exchange. A proponent of the SCAP protocol, the system aims to incorporate existing ITOs in a coherent manner via security automation tools in order to assist with baseline configuration, and leverages IODEF (and custom IODEF extensions) for information capture and a number of industry standards for evidence collection. The framework focuses on information exchange rather than information acquisition or use.

Figure 8 demonstrates the scope of CYBEX, including protection, detection and response, whilst Figure 9 demonstrates how existing ITOs could be leveraged within each domain.

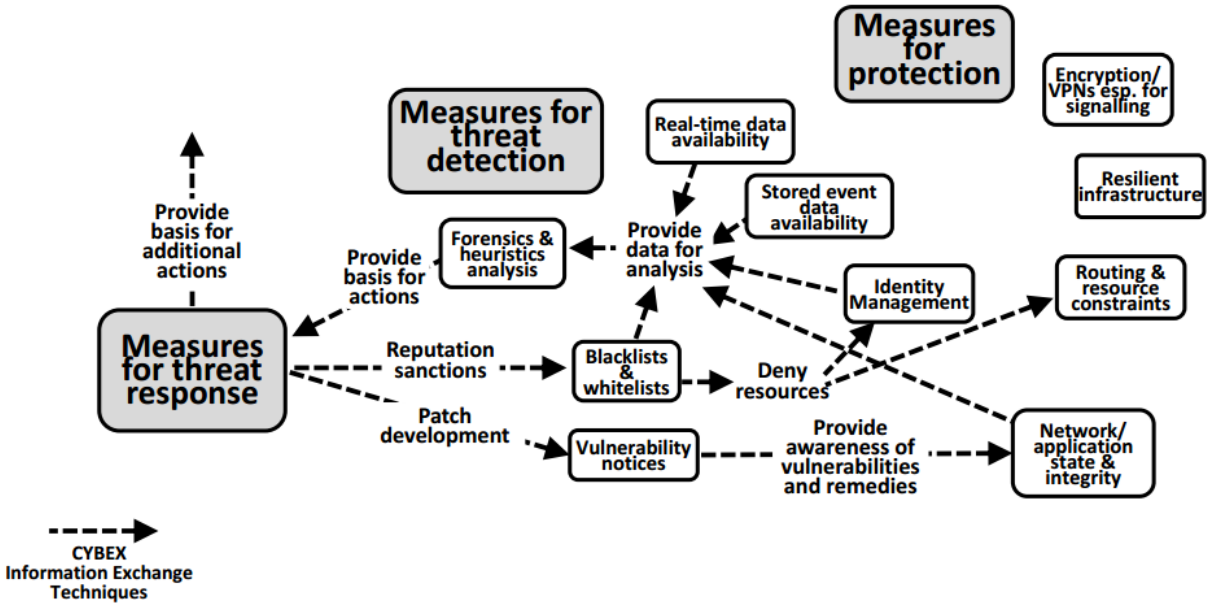

Figure 8 - CYBEX Global Security Model [21]

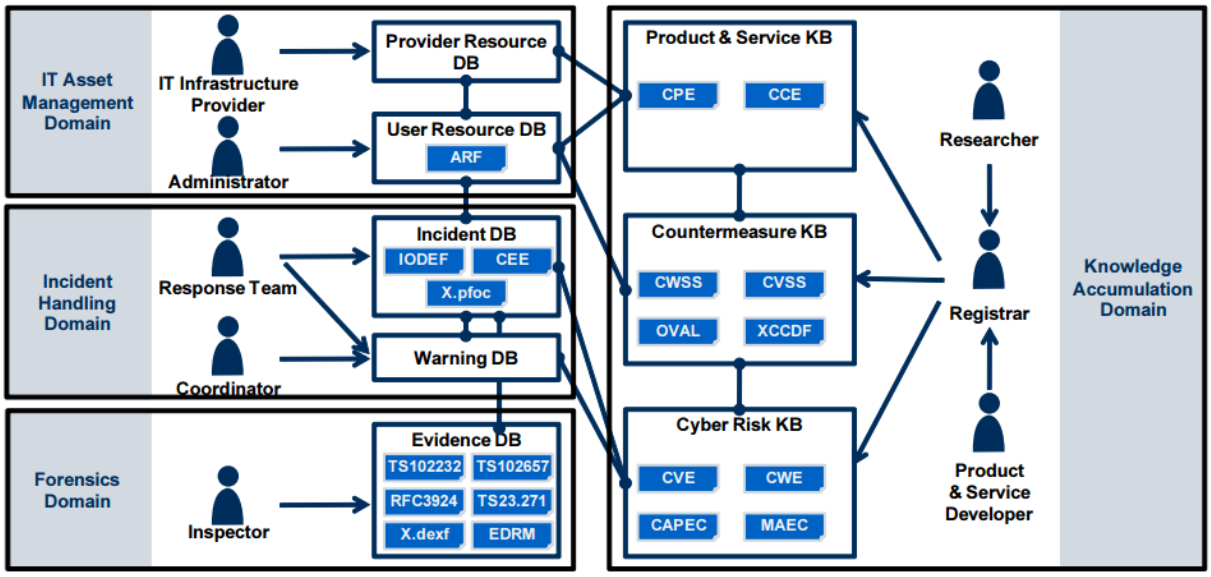

Figure 9 - Cyber Security Information Specifications in CYBEX [22]

The data query model allows for centralised hierarchical sharing, where an entity can select a specific registry to search, or decentralised sharing using search engines. It seems that there is no attempt within this standard to facilitate specific communities of trust or obscure data sources, but there was only limited example data available during this research. Apparently nascent in nature, all the literature reviewed suggests that CYBEX is still evolving. The standard seeks to ensure singularity of information on a given cybersecurity matter, but unless globally recognised it will fail to realise its full potential.

The most prevalent implementation of Cybex to date comes in the form of the Japan vulnerability information portal site (JVN). It consists of three main entities: MyJVN which implements SCAP in order to allow customers to check their software version compliance levels; JVN, which provides vulnerability countermeasure information through an information security early warning partnership (a public and private enterprise) and allows for collaboration with software vendors; and JVNiPedia (see Figure 10), a 'dirty' database of daily updated information that has not yet been added to the JVN approved database.

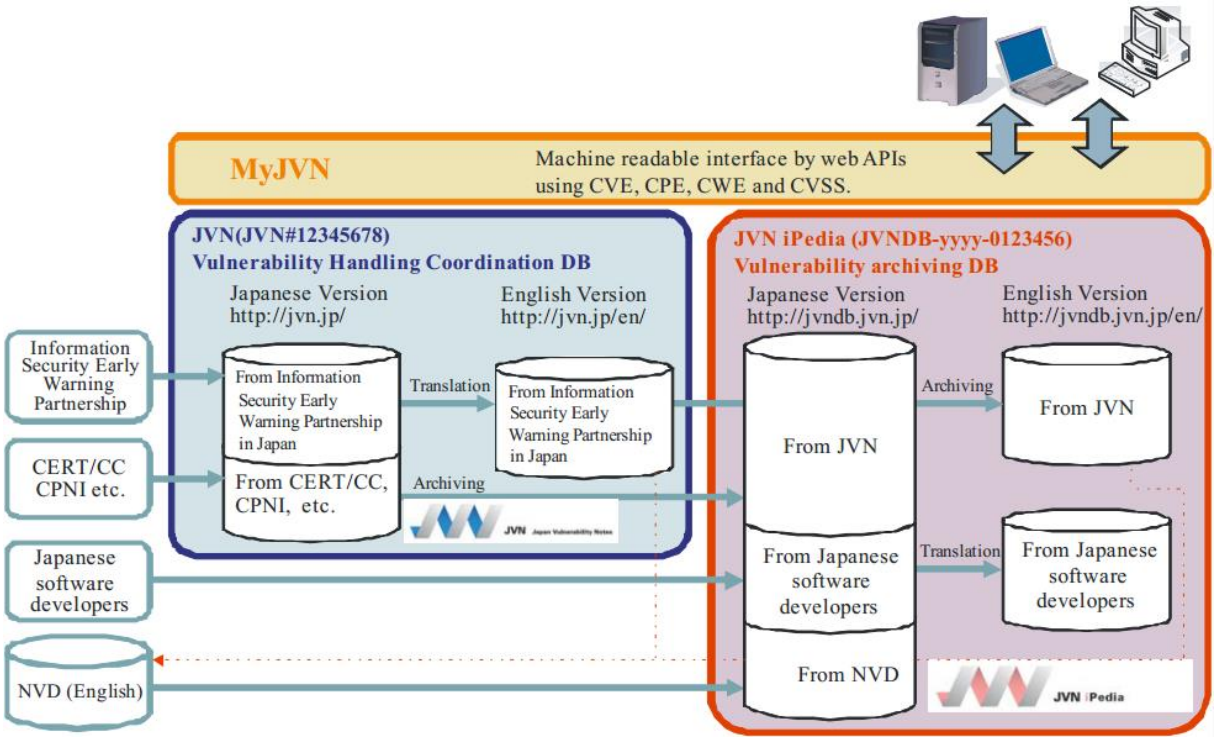

Figure 10 - My JVN Database relationships [23]

From the information reviewed it was difficult to ascertain the usefulness of Cybex for information exchange in attribution. The JVN framework allows for correct software configuration and allows for exchange of vulnerabilities with specific versions or configurations, but does not seem to exchange significant IoCs or allow for specific threats or campaign information to be represented. Cybex does leverage IODEF to pass incident information, and specific forensic information is recorded regarding incidents, but the standard has not seen widespread uptake as of yet. Despite this, with a significant backer such as the ITU, this standard may well develop at a later date into something more suited to the function of attribution.

2.5 Issues with attribution through data exchange

Data exchange may increase an organisation's ability to defend itself against known attacks and exploits, but even when armed with a fully functional and fit-for-purpose data exchange that is populated with high fidelity information, attribution of cyber attacks may still be difficult.

Data exchange is only one element of successful attribution. Whilst it is important to exchange quality information in a timely manner, there are a number of other considerations of how data exchange can assist with attribution. Data can only be populated into an exchange programme if attacks or vulnerabilities are detected. Each of the participants in an exchange must have mature intrusion detection systems and highly trained security professionals to ensure that information given to exchanges is as detailed and accurate as possible.

Although governments and the military were once at the forefront of technological development, this stage has long passed. Industry now leads in technology and as such is a key component of any cyber information exchange. One of the key issues is that, whilst industry may have a well-developed system, their requirements for a data exchange differ vastly to those of a nation state. In all of the example programmes reviewed their primary concern was to increase defence of partners' networks, not to attribute attacks.

When an organisation (or even a government department to some extent) suffers a cyber-attack they would first want to mitigate the attack, then mitigate any after effects such as negative press, falling share prices, or loss of stakeholder confidence. By seeking attribution they prolong the negative effects of an attack, which in turn perpetuates all of these other issues, and for little gain because those attacking are often outside of the host country's jurisdiction and an arrest provides little in the way of compensation to the company. Put simply, attribution is not wholly in their interest.

By contrast, a nation state must be able to attribute, because without attribution retaliation is improbable and a state or cooperative security organisation is rendered ineffective. Collaborative data exchanges may assist with attribution, but unless a source is owned by a nation state, it is difficult to have complete confidence in its data. The only way to overcome this is to know exactly how data was gathered and who it was shared by, and to have the complete evidential chain. This is where data collaboration has some issues to overcome. The secretive nature of national technological agencies makes it is unlikely that the level of detail required to effectively attribute cyber activity will ever be shared or fully opened up to scrutiny by national security elements. The other factor that needs to be considered is that all this information, even when represented in human-readable format, needs to be correctly changed from information to intelligence. This requires significant expertise and cannot be automated. Organisations and governments need to ensure that they have sufficient trained cyber intelligence experts who can interpret not only the technical information represented in data exchange reports, but also the digital forensic traits and geo-political factors that would assist with attribution.

That is not to say that information exchange is futile. As with any form of intelligence gathering, additional information sources are always of benefit to an intelligence analyst. Whilst a nation state or organisation may not be able to definitively point the finger by way of collaborative cyber information, the process of effective data exchange will certainly assist and, figuratively speaking, create an extra string to the intelligence analyst's bow. It is also very useful for nation states and smaller businesses that do not yet have a mature technological authority capable of determining attribution on their own.

2.6 Different dimension of information exchange – Operational Security Communities

In addition to the formal information exchange protocols and programmes described in the previous chapter, there are also effective and more private *Operational Security Communities*. This compendious chapter describing these communities is based on the materials of Kaeo [24], Greene [25], and Aarelaid [26].

2.6.1 Motivation

In order for a cyber security specialist to carry out effective incident response, cyber-risk management, and investigations, it is necessary to actively participate and collaborate with these operational security communities. These communities have rules, expectations, trust networks, and paranoia that makes them difficult to find and difficult to gain access to. Various groups operate differently and have different policies. For example, some are open to all, some are highly peer vetted, and some are personality driven, whereas others are interest driven. [24]

The motivation for these groups stems from the fact that security engineers need to find their colleagues in other service provider networks. For example, previously if a large attack happened, there was no way for the people who needed to work with each other to find each other, let alone work collectively to mitigate the attack. The solution to this was *Aggressive Collaboration.* [24]

2.6.2 Aggressive Collaboration

Key principles of Aggressive Collaboration are the following: [24]

- Chain of Trust – 'If I trust you and you trust him, then I can also trust him.'
- Sphere of Trust – 'The group together can be seen as a sphere, realm, zone, of trust.'
- Need to Know – 'I trust you. You are someone I can depend on, but you don't really need to know about the details of this incident. Not being in a Need to Know Sphere does not mean you are not trusted.'
- Chain of Action – 'You trust someone, but will they be able to do something, be responsive, and/or make something happen?'

From the principles above, one can derive a few common sense expectations that have to be followed by all members. One has to bear in mind that *lurking* is considered a bad behaviour in such communities, as is not taking action when trusted to do so. Each individual is also responsible the information posted and discussed within the community, and one should never forward information from an operational security group without the explicit permission of the person who posted the information. Inability to meet these expectations erodes trust and also your reputation. However, being able to meet those expectations does pay off. [24] [26]

2.6.3 Gaining access

In addition to having different expectations towards their members, many of the groups are also difficult to find and gain access. The communities usually accept practitioners who have the ability to act and influence something within their span of control. In order to gain access, the interested person either has to know someone already in the community, or *when* the individual has met all the prerequisites he or she will be invited to participate in the group. Being from the government or CERT does not qualify for or override any of the requirements. [26]

For some named examples and more detailed descriptions of Operational Security Communities, see the referenced sources by Kaeo [24] and Greene [25].

2.7 Summary

Collaborative data exchanges of cyber event information and indicators of compromise have been conducted for many years, but often in an *ad hoc* way between security professionals. Standards have been developed to report very specific pieces of information and these have been widely adopted by the cyber security community, but what has remained lacking is a holistic data exchange drawing on all of these elements of information.

Some major bodies and organisations have developed (with government backing in many cases) data exchange types and protocols. These have been adopted into a number of programmes which are beginning (or promise) to deliver a large degree of uptake and success. These programmes have increased the ability of subscribers to defend their networks and become more aware of security events and IoCs, but they have been developed with security in mind, not attribution.

As a result, whilst collaborative information exchange may be able to assist in attribution, it is unlikely that it will ever become a reliable and standalone means of attribution, but more an additional factor to assist with technical and intelligence methods when defining the provenance of malicious activity.

Active private industry to private industry cooperation is critical before any successful public sector to private industry partnership. Achieving effective processes in your work area requires active participation and collaboration in the respective Operational Security Communities. These communities have rules and expectations that all the members have to follow in order to get in and also to maintain their membership. However, in the end, the investment does turn into actual results.

3 The workshop

Mauno Pihelgas, Johannes Tammekänd

3.1 Introduction

The workshop was held on 21st of January 2015 at the NATO CCD COE premises. The aim of the event was to promote efficient communication between people who are representing various organisations in Estonia. Although different in nature, the organisations are still facing similar threats emanating from cyber space. Fighting these threats in an efficient manner requires reliable communications channels to be established before such attacks happen.

3.2 Objectives

The general objective was to heighten awareness by identifying and discussing proper response activities, plans and procedures, available resources, and communication strategies. The emphasis was on efficient communication between different parties (IT specialists, managers, media, etc.) at all levels of command.

As an added bonus, we were also planning to take advantage of the situation and collect real-world knowledge and gather feedback for our project from experts (i.e., the participants) who are already working with similar issues on a daily basis.

3.3 Participants

The participants were representing following Estonian organisations:

- Defence Forces;
- Defence League;
- Information Board;
- Information System Authority;
- Internal Security Service;
- Ministry of Economic Affairs and Communications;
- Ministry of Foreign Affairs; and
- Ministry of the Interior.

3.4 Execution

In each of the discussion sessions we presented a different case scenario and we expected the participants to tackle the problems and questions that were encountered within the scenarios. As it turned out, the first of the two proposed scenarios seemed to capture the essence of attribution-related issues more precisely and also offered a controversial twist in the scenario that fuelled discussion for quite some time. Overall, the workshop was a success in terms of bringing together people who are already working on similar topics, but might not always have the most relevant contacts for efficient information-sharing that would aid proper and timely attribution when a cyber attack occurs.

Outside of the workshop the project development group was mainly working in the scope of international law, but since the group of workshop participants consisted only of Estonians, we took Estonia and Estonian law as an example. This was done purely for the sake of clarity and efficiency of the group discussions. During the

development, we used references like State A, State B, State X, etc. We are hoping that this would allow every reader to make their own associations and not constrain their imagination. The only assumption we are making is that State A is a NATO nation, while State B and State X are non-NATO nations.

The responses described below are likely to have been dependent on the audience that we had for the workshop, and would differ in other circles. Unfortunately, this time we did not have a national CERT representative nor the representative of the Government Office Communication Unit present within the group. The responses below are not the official statements of the participants nor participating organisations. They are rather meant to describe the general discussion and ideas regarding the topics handled in the scenarios.

To explain the scenario representation in the following chapters, the scenario storyline is aligned to the left and the responses of the participants are aligned to the right, presented similarly to a conversation between two parties.

3.4.1 Scenario 1 - NATO member massively attacks cyber infrastructure in State B

Introduction
State A and State B have had tension in their relations for quite a long time.

In State A's public sector IT infrastructure there are always at least a few computers that are infected with malware. Most of the time it is just an incidental infection that will be cleaned up within a reasonable time and without any consequences. In a way this is considered a normal everyday situation.

Saturday - Day 1 - Malicious cyber activities
On one unfortunate day those infected computers in State A started a cyber attack and cause denial-of-service on several high-profile systems in State B. The web site of a leading political party in State B was rendered unavailable for several hours, because the server crashed as a result of the attack. Other governmental websites were also attacked, but they were able to recover after about an hour.

Later that day...
CERT of State B sends an e-mail to CERT in State A, describing the attack and asking State A to put a stop to it. The strange network traffic is also confirmed by the Duty Officer at the CERT of State A. He notified network administrators who were able to isolate most of the offending workstations. State A's CERT responds to State B's CERT that ISP was notified and the attack should now be stopped.

Sunday - Day 2
Next day a similar incident takes place. Again a list of offending IP addresses originating from State A is provided. CERT in State A handles the case and the attacks stop.

Monday - Day 3
First thing in the morning State B sends a verbal note[4] about the cyber attacks to State A.

Roughly around the same time an article in an internationally read newspaper under the influence of State B publishes Monday morning news stating 'NATO member massively attacks cyber infrastructure in State B'. The article blames State A for launching the attacks. The media is 'on fire' and leaders in State A are put in a difficult situation with frequent calls from reporters asking for comments on those claims.

[4] Verbal note (also: third-person note) - An unsigned diplomatic note written in the third person, of the nature of a memorandum but sometimes considered to be more formal. [30]

First round of discussion and reactions

Initial actions

If such events were to happen, the initial actions of the participants would not differ much from everyday work situations involving communication and giving advice at the political level.

The first requests for more information would be towards the CERT. They should start actively blocking and examining the infected hosts and find out whether there is any initial knowledge, intelligence, or known motivation regarding the attackers. Is the state really able to mitigate the attacks? The CERT should act as a universal adviser when it comes to problems regarding governmental computer systems.

Internal communication and coordinated information-sharing are extremely important in these situations. There were some opinions that information-sharing inside the state should be organised in a more efficient way. For example, there should be a central information manager who is responsible for gathering and disseminating this information from and to relevant parties. Anyone who is authorised can go and ask for relevant information from this central source.

Externally, let State B know that the situation is being dealt with and that the hosts are blocked. Also stress the fact that State A was not behind the actual attack. It would be reasonable to make a press release to counter the false accusations, however, care should be taken not to start an altercation that could have a snowball effect in the media. We do not have all the details about the incident yet, so being concise in the press release would probably be a wise choice.

Since the story has already made it to international news, our NATO allies should be reassured that State A has not sanctioned this activity. Ministry of Foreign Affairs should be briefed about the situation and made aware of the official posture and line to take of State A towards these events (e.g., what to tell when someone requests information?).

Later, evening of day 3

Techies in State A take the infected machines offline and reveal that they were infected via a link in an e-mail to a malicious website. Furthermore, network traffic analysis reveals that the attack command originated from an IP address somewhere in a distant State X. The address is most likely some exotic VPS[5] service running a C&C proxy. State A has requested information about the proxy server, but all the communication is hindered by the fact that State A has not established a strong relationship with State X. After taking the identified infected computers offline no further cyber attacks follow, at least for the time being.

Second round of discussion and reactions

Depending on the diplomatic relationship between State A and State X, and the urgency of the situation, State A has three apparent courses of action:

1. Slow – Send a verbal note concisely describing the situation and requesting more information about the IP addresses.

[5] VPS (Virtual Private Server) – A private virtual machine that is often sold as a service by an internet hosting provider. The owner has administrative access to the server's operating system and thus can install and run software of their own choosing.

2. Fast – Try to make contact by calling the embassy or other institutions directly.
3. Going the extra mile – Asking to send a representative/expert to State X to assist and oversee the situation. This could, but rarely does, expedite getting the requested help from State X. At least the general public would see that the situation is being dealt with.

Another option would be to contact the service provider of the VPS service, but due to the fact that the service provider resides in another country there might be legal constraints. However, this is still a viable option and it does not take much effort to explore this opportunity.

As usual in such attacks, the network traffic analysis did not reveal much besides the host where the command was distributed to the infected machines. From the network perspective, there is not much more information to gather at this stage. If possible, at this point it would make sense to find malware experts who would be able to analyse and reverse engineer the malware that infected the computers involved in the attack.

Tuesday - Day 4

Digital forensics experts have analysed the attack information and the malware that was involved. While trying to take apart and reverse engineer the malware, the forensic investigators revealed some suspicious references to a group called H4ck3Z. Most likely a VPS service in State X was used to obfuscate the track back to the actual source of the attack.

Intelligence suggests that H4ck3Z is a hacktivist group running its operations mostly out of State B and is rumoured to be partially funded by State B. Open-Source Intelligence and social networks reveal that the group has become increasingly active of late, and is making strong claims against the political decisions of State A.

Third round of discussion and reactions

Next steps and decisions should not be rushed, because the attacks have stopped. There is a high possibility that more information will become available over time and as the investigation proceeds. That is why it is not always the best idea to judge any evidence too quickly. For example, blaming state X on day 3, when the information was revealed that the attack coordination came from a server on their network. In case of a press release one might make a statement that the attacks are over, describe them technically and reassure that everything is under control. After such incidents media monitoring is essential so that reactions can be made to any new claims, thus preventing any news spiralling out of control.

Since this information is not verified and only known to a small group of people, at this point it would not be wise to publicly assign any blame on anyone. Even mentioning that the potential attackers are known or there are suspects might fuel further speculation and pressure from the media.

If the technical analysis has completed then a small group of selected people should gather and decide what data to release. Also taking into account what is already known within the governmental structures and to the public (e.g., whether this information is still under their control).

Depending on how much information we are willing to share, there should potentially be an evaluation of the information from an impartial 3rd party, although in this case trust (or the lack of it) is the primary inhibitor of information-sharing. State A could request verification of the evidence and assistance in bringing the hacktivist group to justice, but this is most likely out of the influence of State A.

In the end, it will most likely be a political decision whether to announce that the hackers operating out of State B were responsible for the malware and potentially also for the attacks.

End of scenario 1

3.4.2 Scenario 2 - Defaced websites and hacked cyber identities of key military personnel

Introduction

The defence forces of State A operate a number of public websites, to inform the citizens of State A about its organisation and activities, issue press releases, inform their own personnel, and for recruitment purposes. The Minister of Defence, the Commander of the Defence Forces and other high ranking personnel use social media to communicate with the public.

Day 1 - Malicious cyber activities

A group of hacktivists in State B, that does not approve State A's foreign and defence policy, conducts cyber attacks against State A's defence forces' websites. Some websites have been defaced several times during the past weeks. Currently the defence forces are facing a series of DDoS attacks that has been causing intermittent downtime of their websites for the past 24 hours.

A popular international news portal publishes an article stating 'Hackers take down State A's defence forces websites'. The hacktivist group claims that they have additional attack vectors already in sight.

News portals in State A are also publishing their own articles while referring to the claims in the original publication. The news portals are contacting the defence forces' PR representative to ask for explanations about the incident.

First round of discussion and reactions

Defending public-facing websites is of course important, but rather than having websites defaced and unavailable, sites could be made easier to protect by replacing them temporarily with static content (no forms, search functionality, database queries, etc.). This holds true especially for sites that host critical data. Although various protection methods are available, in case of a DDoS attack cooperation with the ISP is essential.

Communication personnel should give a press statement that such actions will not be tolerated, but nothing too critical. The statement should be short and concrete – workflow is slightly disturbed and experts are dealing with it. Most importantly calmness should be maintained and reassured. No hasty accusations should be made towards State B.

Since the hackers warned about acquiring new targets, other governmental organisations should be alerted and should take precautionary measures to protect their computer systems. This information should be proactively distributed among relevant organisations.

Day 2

The DDoS attacks continued overnight for another 12 hours. In the morning it is revealed that the social media accounts of the Minister of Defence and the Commander of the Defence Forces have been hacked. Attackers have posted wrong and publicly shaming information using the hacked profiles.

Meanwhile, State B has not responded to any requests made by State A. Neither have they confirmed or denied the accusations.

Second round of discussion and reactions

State A should make another request to State B requiring them to fulfil their duty of due diligence stemming from international law. State B should find the systems causing the attacks and make them stop the attacks.

Social media has received quite a lot of attention during recent years. This holds true for the government institutions as well. Attacks against government institutions' social media pages around the world are not uncommon. Steps should be taken to protect social media accounts and raise the general awareness of all employees who have or plan to open a social media account.

If the social media accounts have been compromised, first priority should be restoring effective control over the accounts. Next would be to figure out how the hackers gained control over the accounts and make sure it does not happen again. For example, on one hand, if the password was too simple, then it is the matter of changing the password and making sure that the attacker has not created any additional access vectors (e.g., by changing the account's e-mail address). If the access to the account was breached by gaining access to the e-mail account connected to the social media account in the first place, then the issue boils down to securing the other accounts and making sure any devices used to connect to these accounts are not infected with malware.

During the discussion, a question was raised about how long it takes to deactivate and regain control of various social media accounts when they have been hacked. There is no clear answer, because it depends on many aspects; for example, whether the user still has control over the e-mail or phone number that was associated with the account, and the attacker has not changed them. In such cases Twitter allows the hacked user to enter any of the previously associated phone numbers, so even if the hacker replaces the phone number, this does not necessarily mean that the user is locked out. [27]

Facebook also has a process for handling cases where the account has been compromised and the user no longer has access to the associated e-mail account. For safety and misuse reasons, in such cases there is a mandatory 24-hour waiting period when access is given back to the user. This would allow legitimate users to be notified if someone else has initiated the process. [28]

Following such an incident, a public press statement should be released stating that the account has been compromised and detailing that the recent messages have been fraudulently posted. If possible, the account should already have been disabled or the posts in question deleted by the time this statement is released to the press, because such a statement would definitely lead more visitors to the site. Extreme care should be taken if the accounts were actually used to disseminate links to malicious websites or malware.

Users should also pay attention to fake accounts that deliberately mimic or impersonate their legitimate accounts. This tactic could be used by hackers to trick people into thinking that they are communicating with the actual person. Such accounts should be proactively monitored and, in case of violations, reported to the maintainer of the web site.

Day 3

Intelligence suggests that the hacktivist group is actively supported by State B authorities, but this is not enough to prove that the group is acting under the instructions of or under the effective control of State B.

Third round of discussion and reactions

Defacement of government websites by a hacktivist group is not a breach of sovereignty, since no substantial physical damage was done. Furthermore, proving any kind of relationships between individual groups' and state actions is difficult, since someone can always claim that the evidence has been falsified or wrongly interpreted.

On the question of what could be considered substantial evidence for proving attribution, there is no general agreement and it has to be established on a case-by-case basis.

Nevertheless, some think that State A should definitely pressure State B to start an investigation regarding the suspicion towards the hacktivist group. However, State A does not have direct influence over this process, especially if the two countries do not have good diplomatic relations.

End of Scenario 2

3.5 International law and strategic communication perspective

When developing the scenarios, the project team was working with a broader scope and not particularly with the dynamics and requirements of the workshop in mind. In addition to not focusing on specific countries, we were working within the scope of international law. Note that the legal analysis relies strongly on the Tallinn Manual Process publications. [29]

3.5.1 Scenario 1 - NATO member massively attacks cyber infrastructure in State B

Introduction

State A and B have had tension in their relations for quite a long time. In IT infrastructure of State A's public sector there are always at least a few computers that are infected with malware. Most of the time it is just an incidental infection that will be cleaned up within a reasonable time without much consequence. In a way this is considered a normal everyday situation.

Malicious cyber activities

On one unfortunate day those infected computers in State A start a cyber attack and cause DoS on several high-profile systems in State B. The web site of a leading political party in State B was rendered unavailable for several hours, because the server crashed as a result of the attack. Other governmental websites were also attacked, but they were able to recover in about an hour.

Aftermath

Now, B publicly blames A for launching the attack. Coincidentally, all of the IPs involved are from State A's public organisations and nowhere else. Techies in State A take the infected machines offline and analysis reveals that the attack command originated from a C&C proxy somewhere in a distant state X. State A has requested information about the proxy server, but all the communication is hindered by the fact that A has not established a strong relationship with State X. No further cyber attacks follow after taking the identified infected computers offline.

Responses

Technical analysis

The infected computers are collected and analysed. Forensic experts take a look at the malware that was involved. This needs to be followed up, and a report received from the analysts. Meanwhile, State A's CERT can confirm that the attack was not initiated from State A – the computers involved were infected and acted on the instructions from the C&C situated in State X.

The information about the C&C server and the affected targets is shared among other organisations in State A, so they will be on the lookout for any related activity. Limited information can be shared with trusted parties in relevant Operational Security Communities (see chapter about Different dimension of information exchange – Operational Security Communities).

A plan is needed to prevent similar incidents in the future by isolating any other reported infections in a more timely manner.

Legal analysis

The legal analysis begins with asking whether the attacks against State B can be attributable to State A. The mere fact that the attacks are originating from a state is not sufficient evidence for attributing the attacks to that state, but at the same time it is an indication that this state may be indirectly associated with the attacks (TM rule 6 [29]). Based on the scenario, State B does not at the moment have enough evidence to (legally) claim that the attacks are attributable to State A. By repeating the same analyses for State X we see that also there we are lacking evidence to confirm State X's involvement.

However, according to international law, a state should not knowingly allow the cyber infrastructure located in its territory or under its exclusive governmental control to be used for acts that adversely and unlawfully affect other States (TM rule 5 [29]), also called the due diligence obligation. Based on this scenario, and lacking further evidence, we assume that neither the State X nor State A were aware of the attacks before they were informed by the victim state. Also, the attacks from both states ceased after informing them. Accordingly, both states have fulfilled their obligation in putting an end to the attacks deriving from their infrastructure.

Should State B seek compensation for the damage caused (i.e. reparation), it would first need to determine attribution, e.g. who was launching the attacks, as well as determine the breach of an international obligation.

Legality of technical responses

There are no immediate legal issues with the proposed responses.

Strategic communication analysis

Political actors do need information about the incident and what has been done to mitigate the activity. This information should not include technical information, but should include the type of incident and method of mitigation of the incident. No further information should be distributed unless specifically asked. Too much technical information will only disturb the strategic level of analysis and might be forwarded in an inappropriate way or misunderstood. Leaders also need an estimate of the impact of the attack in the attacked environment; how serious the incident was or could have become, whether it impacted on military, economic, social or SCADA systems. This will allow them to use the use correct rhetoric while discussing it with their counterparts. Verified information about the affected systems and how false flag or no flag affects the economic system are of vital importance.

It is important that the technical and legal issues surrounding attribution be communicated to the high-level political decision-makers in State A and State B as in the case of disruptive attacks that influence the public, there will be public pressure to find the guilty party. In such a situation political players will want to be able to point fingers. However, State B pointing fingers to State A with no or weak attribution or proven malicious intent can lead to undesirable political and diplomatic consequences and should be avoided.

Information-sharing by leaders or commanders should, towards the attacked party, consist of information on the mitigating actions taken; that the infected computer has been taken off the net and investigation is going on and that there has been contact initiated with State X in regard of investigating the C&C issues. Information of the non-compliance from State X should also be brought forward as in the legal analysis. In particular, care should be taken that State A sets up clear, open and regular communication, with messages emphasising their lack of fault or malign intent. This public communication (including through the media) has to include the steps taken to improve the situation and possible timeline. State X should only be included in if clear attribution can be made.

Towards allies and bi-/multilateral companions information-sharing regarding the type of infection and actions taken should be shared. This means that the leaders/commanders should ensure that the technical people collaborate and have the possibilities to share needed information. This should be the case in normal situations and therefore supervision and support for the information exchange is the key issue here. Especially if this is a reoccurring issue, it is particularly important that direct and lasting specialist-to-specialist relationships be established, almost on a personal level. This allows them to respond and cooperate in the most effective way without information and time loss. These relationships are one of the critical elements in effective international response to cyber incidents.

Towards State A's own organisations there should be revision (check-up, not necessary rewrite) of the supervision processes and tools on how to mitigate infected systems. If the processes and tools are still valid, it should be also informed. State A and the appropriate

bodies within State A should refrain from speculating on the future effects of the incidents while clearly communicating the known scope.

Toward the public audience there should be coherent messaging accordingly to what are told to the attacked party and what are shared on general level towards allies and other companions. The information should be limited to that there have been an infected system and it has been excluded from the network. The public audience should also be informed of possible implications to which they might be subject, i.e. loss of a specific service or information no longer available. There is no need to elaborate on C&C systems or the non-compliance of State X; information that the infected system might have been infected from outside is enough. However, at no point should any party rule out the possibility of further incidents.

Later in the investigation new information is revealed...

Digital forensics experts have analysed the attack information and the malware that was involved. While trying to take apart and reverse engineer the malware, the forensic investigators revealed some suspicious references to a group called H4ck3Z. Most likely a VPS service in State X was used to obfuscate the track back to the actual source of the attack.

Intelligence suggests that H4ck3Z is a hacktivist group running its operations mostly out of State B. It is rumoured to be partially funded by State B. Open-Source Intelligence and social networks reveal that the group has become increasingly active lately and making strong claims against the political decisions of State A.

Responses

Technical analysis

Digital forensics and cyber reconnaissance offered some new information to this case. However, technical attribution, as well as human attribution, always has a degree of uncertainty associated with it. We can never be 100% sure whether the information giving credit to H4ck3Z was genuine or simply planted for the investigation to find.

Legal analysis

Based on the limited information we have, we cannot attribute the activities of H4ck3Z to State B, since ‘partial funding’ is not enough to reach the threshold of ‘effective control’, neither do we know of specific instructions from State B. Hence, we cannot determine state responsibility.

A State should not knowingly allow the cyber infrastructure located in its territory (or under its exclusive governmental control) to be used for acts that adversely and unlawfully affect other states (TM rule 5 [29]). Based on the scenario, and due to lack of further evidence, we assume that State B is not aware of the attacks, and should therefore be informed. Since the scenario does not mention any continuation of the attacks we assume that State B is not violating this obligation.

Any legal reaction to H4ck3Z's activities should be taken based on national criminal law.

Strategic communication analysis

Towards the attacking state, information concerning the new facts should be presented at a working group level and it should be pointed out that, according to international law, they should be active in investigating the incident. It could also be beneficial to let the attacked party know that this information will be also shared with others.

Towards allies and bi- or multilateral companions, newly found information should be shared at all levels, political, operational and technical. This information is vital for understanding and finding similar attacks or approaches within their systems. This information could also bring new insight to current assessments on operational and strategic level and thus come up with improved or new situational awareness issues.

Towards the state's own organisations, it should be assured that information found is shared among critical information and infrastructure organisations so as to eliminate possible similar threats. The main burden is on the technical staff, but others should support the spread of information and knowledge.

If the attribution is clear, and particularly if some of the evidence can be released to public, public steps outlined in the previous public response outlined above should be undertaken for clear attribution. As there is likely to be significant public interest at this point, evidence pointing to a hacker group in State A should be released.

Additionally, public communication implicating State B in the international arena might influence State B to use their resources to stop the attacks. This, however, has to be balanced with the likelihood of escalating tensions over connecting hacker group activity to a state.

End of Scenario 1

3.5.2 Scenario 2 - Defaced websites and hacked cyber identities of key military personnel

Introduction

The defence forces of State A operate a number of public websites, to inform the citizens of State A about its organisations and activities, issue press releases, inform their own personnel and for recruitment purposes. The Minister of Defence, the Commander of the Defence Forces and other high ranking personnel are using social media to communicate with the public.

Malicious cyber activities

A group of state-funded hacktivists in State B, which does not approve State A's foreign and defence policy, conducted cyber attacks against State A's defence forces' websites and hacked the social media accounts of the Minister of Defence and the Commander of the Defence Forces. The websites were defaced several times. During the last weeks two defence forces' websites experienced downtime by DDoS attacks for 24 hours altogether. Based on multiple sources of intelligence, State A has concluded that the hacktivists are actively supported, organisationally and financially, by State B authorities. State B has neither confirmed nor denied the accusations.

Responses

Technical analysis

The defence forces' cyber unit is considering the following measures:

Gather evidence by monitoring the hacktivists group's activities and mapping their cyber infrastructure and modes of operation; collecting information about the group's key personnel and their 'cyber identities' (IP addresses, e-mail accounts, use of social media etc.).

The first step would be to deactivate the social media accounts and regain control over them. Although it is not yet certain how access to the accounts was gained, social network accounts should be better protected – e.g., using strong passwords together with two factor authentication. Awareness in this area should be promoted and encouraged. Cooperation with the social network site administrators is required to find out where the attack originated from, and how the attackers gained access.

Legal analysis

State B vis-a-vis State A

In order to discuss possible international law measures, state attribution must be confirmed. In this case, there is evidence of the hacktivists being supported financially and organisationally, but this is not enough to prove that the group is acting under the instructions of or under the effective control of State B. However, should there be more evidence to confirm state attribution, we could discuss the possible breach of sovereignty of State A. The threshold of breach of sovereignty in the case of defacing government websites of State A is not clear. Some authors claim that any 'virtual presence' in other state's networks is a breach of sovereignty. Also, the prohibition of intervention should be assessed, but given the lack of a coercive nature of these activities, or 'intervening directly or indirectly in the internal or external affairs of other States', it would likely not apply (TM rule 10 [29]). The legal interpretation of these principles should be carefully done by State A before taking any further action. The determination of an internationally wrongful act is a prerequisite for countermeasures under the state responsibility law.

If we cannot attribute the operations to State B, we should check the applicability of the due diligence operation. However, before being able to apply the obligation, we would need to determine that State B was not aware of the attacks. If it was not, State B would have the due diligence obligation to stop the operations coming from its territory.

Legality of technical responses

In responding to the operations without state attribution, national legal measures and international cooperation apply. The breach of national law must be taken into account both in State A and State B frameworks.

Strategic communication analysis

Political and higher military commanders should address the actions but are limited by legality.

Towards the states from where the attacks origin from there should be clear messaging on evidence gathering of the hacking group in State B. Since there seems to be evidence that State B supported the activities, not much cooperation is to be expected. Nevertheless, the lines of communication should be kept open and facts on illegalities stressed. No technical information should be conversed.

Towards allies and bi- and multilateral companion, technical evidence and facts should be shared within chosen and appropriate forums. TTPs and other information of documented attack vectors should be shared and solutions on how to mitigate the attacks and defacements should be discussed. These actions are mostly for technical personnel, but higher level authorities should initiate the contacts and choose with whom to cooperate and to share the information with.

Towards their own organisations, higher level authorities could search, suggest and validate possible cooperation parties and thus coordinate the fight against the attacker. One option could be that, together with allies, places could be found to temporarily establish a mirroring site for the attacked sides. This would probably need some agreements between states or other special arrangements. Third parties such as ISPs should be helped in patching up their system to prevent further exploitation.

Toward the public audience the messaging should include facts, but no technical details, about the attack and deliver accurate information. This information has to be backed up with reliable evidence or references in order to mitigate the messaging produced by the defacing group. Information that secure systems are still intact could be presented, if it does not reveal sensitive information about the system or its content.

End of Scenario 2

3.6 Summary

Overall, the workshop could be considered successful and we achieved the aims we had set for this event. Although, there were a few rough edges in the details of the scenarios, we hope that in general the participants found the event useful. Based on the feedback, it seems that having the event was a good idea and it brought together people who should be actively collaborating and sharing information to protect the nation from cyber attacks that lack attribution. Having such events would bring people together in a way similar to Operational Security Communities.

From the perspective of the organisers, we were able to get a lot of valuable information and insights from the participants. Whilst we presented the participants with our work that was the basis of the previous two chapters in this report, they were able to provide real-world knowledge that nicely complemented our theoretical work.

4 Project summary and conclusions

The project was dealing with the issues of attribution. The primary challenge for nations is that often, especially in the case of cyber attacks, there is not enough information available to achieve proper attribution (no-flag attacks). The situation is even worse when the attacker tries to encumber attribution by impersonating someone else (false flag attacks). This project looked at some of the ways to help reduce the insecurity about misattribution, one of the key methods being effective information exchange.

The project plan was designed to consist of a theoretical part and practical part to test some of the ideas that were developed during the project lifecycle. The theoretical part was handled in the first two chapters in the current report, whilst the more practical workshop is described in the third and final chapter.

In the chapter Cyber Information Exchange – Collaboration for Attribution of Malicious Cyber Activity, we looked at various information exchange protocols and official programmes that are being used, developed and funded by various organisations. It was revealed that whilst collaborative information exchange may be able to assist in attribution, it is unlikely that it will ever become a reliable and standalone means of attribution, but more an additional factor to assist with technical and intelligence methods when defining the provenance of malicious activity.

The following chapter focused on a Different dimension of information exchange – Operational Security Communities. These groups are often closed to the public with access given only to a selected few meeting all the prerequisites. However, sources claimed that these groups are necessary for security specialist to do effective incident response, cyber-risk management, and investigations. It is essential for them to actively participate in these communities to find their colleagues in other service provider networks.

The final chapter described the scenarios and discussions during The workshop that was held in the NATO CCD COE on the 21st of January 2015. The workshop welcomed specialists from various government institutions who deal with such events and information exchange on a daily basis. The workshop revealed that some of our proposed ideas conform to the opinions from the participants. Furthermore, one of the key observations is that the feedback from participants revealed that such an event is very useful in forming closer personal relationships among people dealing with similar issues – somewhat similar to the Operational Security Communities that were described in the second chapter.

It is obvious that attribution can be and usually is difficult. Whilst the request for support noted primarily technical methods, after discussing this with the Originating Organisation, we broadened the scope of the project to also include legal, strategic and communication aspects of attribution. Note that relying solely on one aspect (e.g., technical) does not provide a viable solution and is most likely unable to provide confidence required for achieving attribution. Combining multiple aspects would definitely be a better solution.

5 References

[1] G. Hughes, The Military's Role in Counterterrorism: Examples and Implications for Liberal Democracies, Carlisle: Strategic Studies Institute, 2011.

[2] Essential Intelligence, "No-Flag Operations," 28 January 2012. [Online]. Available: http://essential-intelligence-network.blogspot.de/2012/01/no-flag-operations.html. [Accessed 20 January 2015].

[3] M. Pihelgas, "Back-Tracing and Anonymity in Cyberspace," in *Peacetime Regime for State Activities in Cyberspace*, Tallinn, NATO CCD COE Publications, 2013, pp. 31-60.

[4] W. Zhao and G. White, "Designing a Formal Model Facilitating Collaborative Information Sharing for Community Cyber Security," The University of Texas at San Antonio, 2014.

[5] A. M. Rutowski, W. A. Foster and S. E. Goodman, "Multilateral Cyber Security Solutions: Contemporary Realities," *Federation of American Scientist,* 2012.

[6] Ponemon Institute LLC, "Intelligence Driven Cyber Defense," February 2015. [Online]. Available: http://cyber.lockheedmartin.com/hs-fs/hub/444799/file-2508671409-pdf/Documents/Reports/ponemon-intelligence-driven-cyber-defense-survey-report-2015.pdf. [Accessed 28 February 2015].

[7] C. Blanco, J. Lasheras, R. Valencia-Garcia, E. Fernandez-Medina, A. Toval and M. Piattini, "A Systematic Review and Comparison of Security Ontologies," *The third Interntional Conference on Availability, Reliability and Security,* pp. 813-820.

[8] L. Dandurand and O. Serrano Serrano, "Towards Improved Cyber Security Information Sharing," in *5th International Conference on Cyber Conflict*, Tallinn, 2013.

[9] NIST, "The Security Content Automation Protocol (SCAP)," 2014. [Online]. Available: http://scap.nist.gov/. [Accessed 8 December 2014].

[10] W. M. Fitzgerald and S. M. Foley, "Avoiding Inconsistencies in the Security Content Automation Protocol," Ireland, 2013.

[11] MITRE, "Structured Threat Information eXpression - A Structured Language for Cyber Threat Intelligence Information," [Online]. Available: stix.mitre.org. [Accessed 8 December 2014].

[12] S. Barnum, "Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX™)," *The MITRE Corporation,* 2013.

[13] The MITRE Corporation, "TAXII Introduction," [Online]. Available: http://taxiiproject.github.io/getting-started/intro/. [Accessed 15 October 2014].

[14] ArcSight, "ArcSight Incorporates XML Messaging Standards in Its Distributed Security Architecture," 2014. [Online]. Available: http://www.prnewswire.com/news-releases/arcsight-incorporates-xml-messaging-standards-in-its-distributed-security-architecture-70837872.html. [Accessed 8 December 2014].

[15] G. Farnham, “Tools and Standards for Cyber Threat Intelligence Projects,” *The SANS Institute,* 2012.

[16] K. M. Moriarty, “IETF MILE: Improving Incident Information Sharing Standards,” 2011.

[17] MNCD2, “MultiNational MNCD2 - Cyber Defence Capability Development,” 2014. [Online]. Available: https://mncd2.ncia.nato.int/. [Accessed 09 December 2014].

[18] NCIA, “Malware Information Sharing Platform,” NCI Agency, Mons, 2014.

[19] Codenomicon, “AbuseSA,” 2014. [Online]. Available: https://www.clarifiednetworks.com/AbuseSA. [Accessed 5 December 2014].

[20] CERT-UK, “CERT-UK adds National Cyber-Security Centre Finland IoC feed to its list of AbuseSA feeds,” 2014. [Online]. Available: https://www.cert.gov.uk/resources/news/2014/09/cert-uk-adds-national-cyber-security-centre-finland-ioc-feed-to-its-list-of-abusesa-feeds/. [Accessed 08 December 2014].

[21] M. Euchner, “ICT Security – Cybersecurity – CYBEX,” *TSB Briefing to the Regional Offices,* 2011.

[22] A. Rutowski, Y. Kadobayashi, I. Furey, D. Rajnovic, R. Martin and T. Takahashi, “CYBEX – The Cybersecurity Information Exchange Framework (X.1500),” *ACM SIGCOMM Computer Communication Review,* vol. 40, no. 5, 2010.

[23] ITU-T, *SERIES X: DATA NETWORKS, OPEN SYSTEM COMMUNICATIONS AND SECURITY - Cybersecurity information exchange – Overview of cybersecurity,* International Telecommunication Union, 2011.

[24] M. Kaeo, “Inter Network Cooperation,” 2014. [Online]. Available: https://nsrc.org/workshops/2014/apricot14-security/raw-attachment/wiki/Agenda/4-2-2.inter-network-cooperation.pdf. [Accessed 26 February 2015].

[25] B. R. Greene, “Techniques, Tools and Processes to Help Service Providers Clean Malware from Subscriber Systems,” M3AAWG, 22 October 2012. [Online]. Available: https://www.m3aawg.org/system/files/M3AAWG-Malware-Greene-Seg4-Turning-Point.pdf. [Accessed 26 February 2015].

[26] H. Aarelaid, “The Big World - What are Operational Security Communities?,” 2014. [Online]. Available: http://slid.es/hillar/the-big-world. [Accessed 26 February 2015].

[27] Twitter, “My account has been hacked,” [Online]. Available: https://support.twitter.com/articles/185703-my-account-is-compromised-hacked-and-i-can-t-log-in. [Accessed 25 February 2015].

[28] Avinash S, “How to Recover a Hacked Facebook Account,” [Online]. Available: http://www.tricks99.com/2014/12/how-to-recover-my-hacked-facebook-account.html. [Accessed 15 February 2015].

[29] M. N. Schmitt, Tallinn Manual on the International Law Applicable to Cyber Warfare, Cambridge: Cambridge University Press, 2013.

[30] Wiktionary, “note verbale,” [Online]. Available: http://en.wiktionary.org/wiki/note_verbale. [Accessed 19 January 2015].