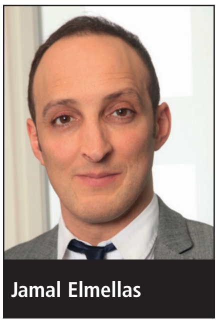
Knowledge is power: the evolution of threat intelligence

Jamal Elmellas, Auriga


Jamal Elmellas

Every day, organisations are inundated with data, whether generated internally or acquired from external feeds, on current and emerging security threats. Sifting through the data and transforming it into meaningful information for the organisation requires considerable time, effort and expertise. Enter threat intelligence (TI), a technical solution that has been sold as the ultimate form of proactive security for the past 15 years.

TI services are widely used by organisations to analyse and filter raw data about emerging threats from several sources to produce usable information in the form of management reports and data feeds for automated security control systems. The primary purpose of TI is to help organisations understand the risks of zero-day threats, advanced persistent threats (APTs) and exploits, especially those most likely to affect their particular environments. A leaner, meaner form of security, TI was supposed to win the arms war against the attacker by enabling the business to be responsive rather than defensive. And yet attacks have continued unabated.

> ***"Threat intelligence can and should be used to analyse and advise upon business strategy and not just as a means of collating and identifying potential attacks"***

The emphasis on TI as a technical solution has failed to capitalise on the full potential of TI as a more holistic strategic form of security, encompassing threat hunting and business intelligence. If we go back to the true definition of TI, as identified by analyst house Gartner, it describes it as, "evidence-based knowledge, including context, mechanisms, indicators, implications and actionable advice, about an existing or emerging menace or hazard to assets that can be used to inform decisions regarding the subject's response to that menace or hazard". That suggests TI can and should be used to analyse and advise upon business strategy and not just as a means of collating and identifying potential attacks.

TI is born

TI is still a nascent industry. De-perimeterisation, the obvious inadequacies of anti-virus solutions and other point solutions all conspired to force the security sector to seek out a new way of detecting and preventing threat realisation. It came to prominence in the mid-2000s as a solution to analysing and filtering data about emerging threats from several sources in real time to address the ever-increasing cyberthreat landscape. Open-source/open-pattern TI streams have since seen the concept evolve and the industry has now established a TI Cycle of clear processes. To understand the progression of TI, it's worth looking at some key milestones in its development.

> ***"The framework makes use of a script-based command line client that can be easily leveraged for automating the use of data, with data stored in the IODEF-defined standard format"***

The first TI solution was the Incident Object Description and Exchange Format (IODEF) which was released in December 2007.[1] The IODEF framework is an XML-based standard used to share incident information by computer security incident response teams, and the IODEF Data Model provides over 30 classes and sub-classes that are used to define incident data. These cover a wide range of information including contact, monetary impact, time, operating system and application. The framework also provides data-handling labels such as 'sensitive' and 'confidential'. It has been widely used in a number of projects and vendor products: for instance, a successful implementation of IODEF is used by the Anti-Phishing Working Group. There have been extensions of the IODEF standard to support the reporting of phishing, and IODEF is also used in products from DFLabs, Arcsight and Foundstone.[2]

The value of sharing

In 2009 the Research and Education Network Information Sharing and Analysis Centre developed the Collective Intelligence Framework (CIF).[3] CIF includes a server component that collects and stores TI data. That can be IP addresses, ASN (Autonomous System Numbers) numbers, email addresses, domain names and uniform resource locators (URLs) and other attributes. CIF data also includes information on types of threats, severity of an attack and the confidence of the data. CIF is novel in that it provides the ability to both control access through the use of an API key and to place restriction levels on the data. It has a robust set of features and supports all the required data types. The framework makes use of a script-based command line client that can be easily leveraged for automating the use of data, with data stored in the IODEF-defined standard format.

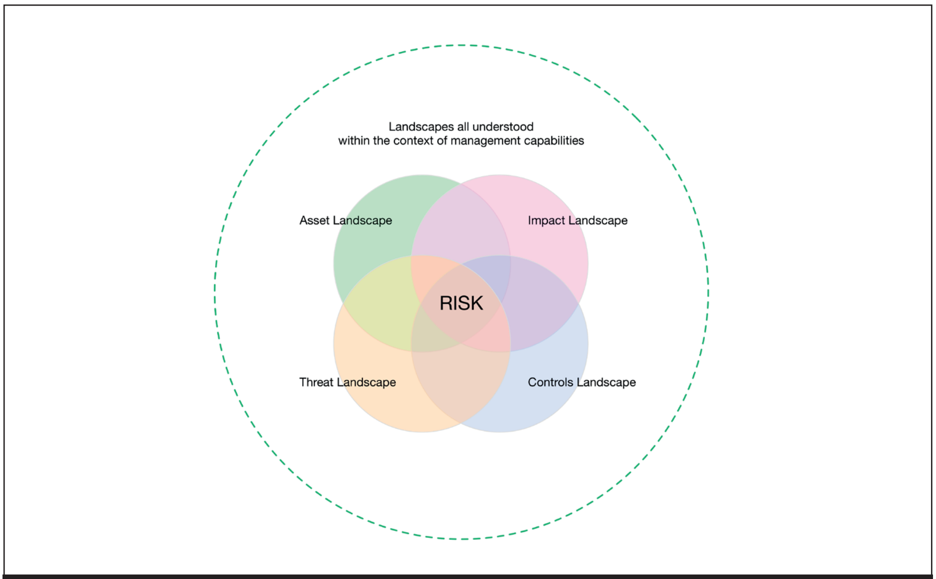
The Vocabulary for Event Recording and Incident Sharing (VERIS) framework was released by Verizon in March 2010 and provides a standard way for defining and sharing incident information.[4] Verizon releases the 'Data Breach Investigation Report' (DBIR) that leverages VERIS on an annual basis. For the 2013 report there was a total of 19 organisations supplying incident details. These organisations collected data using one of three methods: VERIS directly; re-entered in a VERIS application; or converted data from another schema.

The VERIS schema is divided into five sections: Incident Tracking, Victim Demographics, Incident Description, Discovery and Response and Impact Assessment. Each of the sections has multiple elements with specific data types and variables names. Some of the elements included are incident summary, confidence rating, primary industry and hacking variety. Some of these elements contain enumerated lists. For example, hacking variety is made up of an enumerated list of 46 hacking varieties. The varieties include things like brute force, buffer overflow, cache poisoning etc.

"VERIS marks a significant milestone as it saw the first step beyond merely tactical information towards strategic and risk-based information"

VERIS does have a limited ability to include Indicators of Compromise (IOC). This is done via a simple IOC element that stores an indicator and a comment about it. Interestingly, VERIS marks a significant milestone as it saw the first step beyond merely tactical information towards strategic and risk-based information.

With the VERIS framework, other organisations can contribute data in a standard format and vocabulary. This data can then be incorporated and used as a larger data set for analysis and reporting. A community database for VERIS data is available from Verizon, which contains over 1,200 incidents from the Department of Health and Human Services as well as other public incidents. The database is publicly available in JSON format. VERIS example files can be downloaded from a community site and there is also a Tableau-based interactive graph site available to view the data.


Landscapes all understood
within the context of management capabilities
Asset Landscape
Impact Landscape
RISK
Threat Landscape
Controls Landscape

According to VERIS: "The domain of information risk can be visually represented as four intersecting landscapes of Threat, Asset, Impact and Control. The organisation's capability to understand and manage risk requires information from each landscape."

The Open Indicators of Compromise (OpenIOC) framework, introduced by Mandiant in 2011, saw a swing back towards a tactical form of TI.[5] The framework is issued in Mandiant products, but has also been released as an open standard and provides definitions for specific technical details, including over 500 indicator terms. The framework allows new terms to be added easily because the terms are separate for the main schema. The majority of the terms are host-centric with titles beginning file, driver, disk, system, process or registry. This could include an indicative couple, like the notepad.exe file name and its MD5 hash.

Multiple IOCs can be combined using Boolean logic to define a specific malware sample or family. The combined logic can be used to look for items that should not be there as well as verifying expected items. For example, if a service runs a dynamic link library (DLL) file that is normally signed, finding a DLL file but not a valid signature could be an IOC. Examples are available for known malware. For instance, an example of the Nettraveler malware originally reported by Kaspersky is available on the Mandiant Blog as an IOC formatted XML file.[6] Examples of FileName, File Hash, IP Address and portable executable (PE) exports are included.

By now, the benefits of pooling TI were becoming apparent and this saw the creation of the Open Threat Exchange (OTX), developed by Alien Vault in February 2012. OTX cleanses aggregates, validates and publishes threat data streaming in from the broadest range of security devices across a community of more than 18,000 Open Source Security Information and Event Management (OSSIEM) and Alien Vault deployments.

OTX adopts a centralised system for collecting TI that interoperates with an Open Source SIEM system, where OSSIM users can configure their system to upload their threat data to OTX. Collected data is validated by AlientVault and TI is then delivered to all OSSIM users that subscribe to OTX and is also available to those using the Collective Intelligence Framework (CIF) system. (OTX is used by any OSSIM users that have enabled it as well as any CIF users accessing the system.)

> ***"The STIX framework was designed to support analysing cyberthreats, specifying indicator patterns, managing response activities and sharing threat information"***

The focus of OTX is to provide data to the public but it lacks the ability to restrict access for community use. OTX does provide an automated mechanism for sharing CTI data although there does not appear to be any way to control who can access submitted data. Therefore, while OTX does provide a valuable service, its functionality is limited to publicly sharing data.

Coming of age

Mitre developed Structured Threat Information Expression (STIX) in April 2013.[7] STIX defines threat information by including not only threat detail but also the context of the threat. The STIX framework was designed to support analysing cyberthreats, specifying indicator patterns, managing response activities and sharing threat information. It makes use of XML to define threat-related constructs such as campaign exploit target, incident, indicator, threat actor and TTP. The standard includes a number of other standards, giving it a richness and completeness not found with other solutions, and such is its appeal that the STIX solution has quickly gained worldwide support from financial services, CERTS, vendors, governments, and industrial control systems, as well as enterprise users.

The success of STIX saw Mitre develop a complimentary standard in 2013 called the Trusted Automated Exchange of Indicator Information (TAXII) which defines a set of services and message exchanges for exchanging cyberthreat information.[8] TAXII was designed to be flexible enough to support multiple sharing models, including variations of 'hub and spoke' as well as 'peer to peer'. These models allow for push or pull transfer of TI data. The models are supported by four core services: Discovery, Feed Management, Inbox and Poll.

TAXII makes use of XML and HTTP for message content and transport. It also allows for custom formats and protocols and includes standard mechanisms for confidentiality, integrity and attribution.

There is a number of high profile groups that are using TAXII and it has been adopted as a planned standard by Microsoft as part of its Microsoft Active Protections Program. The standard will be used to share TI data with MAPP members. And it is also used by the Financial Services Information Sharing Analysis Centre (FS-ISAC) where members can leverage STIX and TAXII to access TI. APIs are available for discover and pull for the current FS-ISAC deployment.

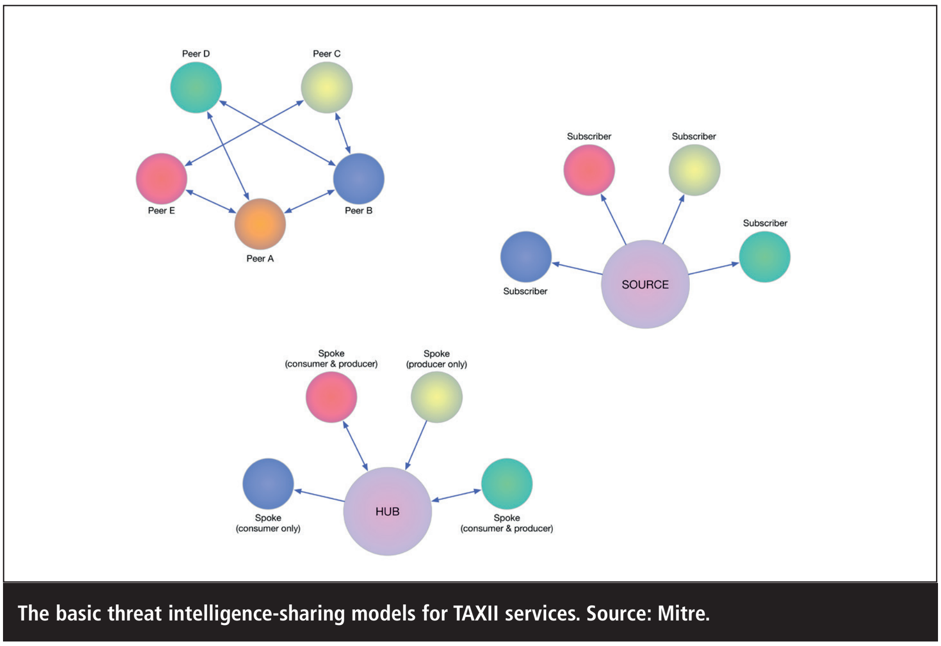
The TI cycle

An effective TI process will have a number of areas of focus. Breaking it down into specific functions allows it to be more scalable, as personnel are likely to be better skilled at specific aspects of intelligence: individual parts of the cycle can be focused on and developed, while it also makes it easier to track where insufficient results are generated in the process, or specific weaknesses.

> ***"The board or senior decision-makers need to identify what they specifically want to know and what the TI programme should be telling them"***

The model used most predominantly is the Intelligence Cycle, which features the following functions: requirements, collection, analysis, production and evaluation.[9]

The requirements function, which can also be referred to as 'tasking', is the key to a successful TI programme. The board or senior decision-makers need to identify what they specifically want to know and what the TI programme should be telling them. A typical example of a requirement might be: 'to be informed of all publicly known, widely exploited vulnerabilities within one day



The basic threat intelligence-sharing models for TAXII services. Source: Mitre.

of them becoming known'. Moreover, TI teams need to work with the board or key decision-makers to agree on requirements that are not only feasible but, crucially, will supply products on which the organisation will be able to act.

Collection comprises a large variety of sources. These include news feeds, paid-for services or feeds, forums, white papers, or even human sources. Therefore the collection step can dominate much of a TI budget. Almost all paid-for TI from vendors comes under this category and will require some form of analysis. This will provide a greater understanding of which sources are most likely to produce the desired information, be reliable and provide information that can be consumed in a timely manner.

There are cases when analysis will be relatively simple. A typical example of a straightforward analysis is parsing a feed into a firewall deny-and-alert ruleset. However, in other cases it will require extracting the relevant information from a larger source such as a report and understanding which elements apply to the organisation's assets.

An interplay between collection and analysis often occurs. This can be due to collection not producing the required raw materials, or perhaps different information needs to be collected for appropriate analysis.

The outcome is an intelligence 'product' that is created and disseminated to the client. The product will depend upon the category of the TI and the client. For example, it might see the generation of a report to the board, a white paper to defenders, or simply an approved rule added to the defence hardware.

The evaluation of TI product to ensure it meets the original requirements is frequently neglected in the cycle. In the case that the requirements have been met, the product can further feed the requirements to help develop new, deeper requirements that build upon the intelligence product and the TI cycle can repeat. On the other hand, if the produced TI does not meet requirements, then it suggests a failure at some point and the cycle model can be used to establish where the failure occurred.

The future of TI

TI is still regarded as something of a passive medium and for this reason there's an increasing emphasis on threat hunting. This describes a more aggressive focused tactic that sees data sourced from dynamic sources, such as social media and the dark web, before being analysed by various algorithms and filtered by sector, geography, regulations etc which make the data pertinent to the business.

This type of next-generation TI harnesses all of the data at its disposal – both inside and outside the organisation – to create a map in real time of potential attacks, helping to steer resource allocation and inform future security spend. The downside is it is very specific to the organisation and needs to be tailored to the business. But once in place, it will enable the company to monitor and even possibly predict sector-specific attacks in real time.

> ***"The first step in predictive TI is acquiring sources of good TI and adopting a productive methodology of analysing them"***

In the ever-evolving landscape of cyberthreats, for many organisations, simple detection and remediation is no longer enough. Some cyber-security companies are now going one step further in providing predictive TI. Predictive TI, as identified by Prakash, Chen and Truve, is largely about determining the probability of an attack happening.11-13 It is part of an intelligence-led security revolution that has accompanied the emergence of big data and behavioural analytics. The same basic principles used in these data management disciplines are now being ported over to the business world to help organisations better correlate likely attack patterns and entry points in their computer networks.

The first step in predictive TI is acquiring sources of good TI and adopting a productive methodology of analysing them. This can be achieved by building a dedicated team of researchers and analysts in-house to identify and monitor these sources, or it can see the organisation outsource by subscribing to a TI service provided by a third party. Or the organisation could test the waters by joining a vertical-specific, non-profit information-sharing and analysis centre that serves as a central resource for gathering and sharing information about cyberthreats.

Shift in thinking

There has been a shift in the way senior security personnel are thinking, away from a post-incident to pre-incident TI. The future of TI is predictive: however, the trend is in its infancy and has a long way to go before it is fully baked. But it's generating substantial interest because it could help close the gap between malicious actors and those who seek to defend against attacks. By focusing on the motivation, the target and the intent, we will have a better chance of mitigating the impact of the attack.

The next step for cyber-security companies will be to not only anticipate what will happen and when, but also what should we do. This will be achieved by incorporating prescriptive business intelligence into TI. Prescriptive business intelligence is largely about recommending the course of action, so the business decision-makers can take this information and act upon it.

There are already vendors attempting to materialise on the aforementioned and encapsulating cutting-edge research results in order to maximise the efficiency of their detection and prediction capabilities. For example, scalable outsourced next-generation Security Operations Centre (SOC) services are emerging that are capable of providing alerts and warnings of specific pending attacks that threaten each organisation they serve. In order to achieve that, the SOC has to quickly tailor and deliver proactive TI by collecting and analysing the motivations, intentions, objectives and capabilities of the specific threat actors most likely to launch an attack against the client. Furthermore, these results must be shaped adequately in order to facilitate a broad spectrum of actors and decision-makers, from first responders seeking appropriate responses to ongoing threats, to C-level executives researching where to invest their cyber-security budget.

Thus, TI predictive analytics will be further optimised by business intelligence analytics as identified by Glass and the use of various social media feeds and real-time feeds across blogs, micro-blogs and news sources both from the observable and the deep web.[13] These emerging methodologies highlight that TI is a far from prescribed technical solution but is rather a constantly evolving, analytical methodology that promises to forewarn and forearm the organisation.

About the author

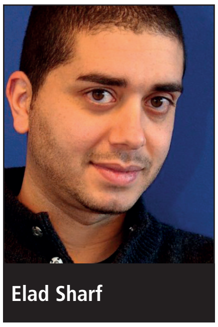
Jamal Elmellas is technical director at Auriga (www.aurigaconsulting.com), the data security consultancy, and is a qualified CLAS security consultant. He has advised on a number of high-profile public and private sector security projects. In addition to his familiarity with compliance standards and security architecture, Elmellas also specialises in the transition of traditional based infrastructures into virtualised environments.

References

1. Danyliw, R. 'The incident object description exchange format'. Network Working Group, IETF, Dec 2007. Accessed Jun 2016. www.ietf.org/rfc/rfc5070.txt.
2. Cain, P; Jevans, D. 'Extensions to the IODEF-Document Class for Reporting Phishing'. Internet Engineering Task Force, Jul 2010. Accessed Jun 2016. https://tools.ietf.org/html/rfc5901.
3. 'Collective Intelligence Framework'. The CSIRT Gadgets Foundation. Accessed Jun 2016. http://csirtgadgets.org/collective-intelligence-framework/.
4. VERIS home page. Accessed Jun 2016. http://veriscommunity.net/.
5. 'Sophisticated Indicators for the Modern Threat Landscape: An Introduction to OpenIOC'. OpenIOC, 2011. Accessed Jun 2016. http://openioc.org/resources/An_Introduction_to_OpenIOC.pdf.
6. Gibb, W. 'NetTraveler in OpenIOC format'. FireEye, 18 Jun 2013. Accessed Jun 2016. www.fireeye.com/blog/threat-research/2013/06/nettraveler-openioc.html.
7. 'TAXII: An Overview'. Mitre, 2013. Accessed Jun 2016. http://taxii.mitre.org/about/documents/TAXII_Overview_briefing_July_2013.pdf.
8. 'Use cases'. Stix project, Mitre. Accessed June 2016. http://stixproject.github.io/usecases/.
9. 'Integrating Threat Intelligence: Defining an intelligence driven cyber-security strategy'. CPNI, 11 Jun 2015. Accessed Jun 2016. www.cpni.gov.uk/Documents/Publications/2015/11-jUNE-2015-CONTEXT_CPNI_Threat_Intelligence_FINAL.pdf.

Making security awareness training work

Elad Sharf, Performanta



Elad Sharf

The incoming EU General Data Protection Regulation (GDPR) requires companies to notify the EU authorities of data breaches within 72 hours. Currently, just 28% of cyber-attacks are reported to the police, cites the Institute of Directors, despite half of these attacks causing disruption to business operations.[1,2] This demonstrates a clear need for increased transparency across the EU, and is heralded as an excellent move by the European Commission. By not reporting these crimes, businesses are depriving both law enforcement agencies and other companies of insights that could be gained from each incident. This is one of many pan-European initiatives, designed to increase data sharing, co-operation and regulation that benefit the UK.

Within the UK we have seen large recruitment drives to try to increase the available cyber-security talent, growing the UK's cyber-abilities and therefore the level of protection available. Within the last year GCHQ has hired 800 people and BT has hired 900 for entry-level cyber-security jobs, with the goal of building a strong cyber-security skillset. Controversially, in the same timeframe the UK Government has taken steps to weaken some forms of encryption, compel Internet Service Providers (ISPs) to store sensitive user browsing history and has even stated that where possible it intends to be 'flexible' with the implementation of the GDPR in favour of the