Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX™)

Sean Barnum
The MITRE Corporation

It is becoming increasingly necessary for organizations to have a cyber threat intelligence capability and a key component of success for any such capability is information sharing with partners, peers and others they select to trust. While cyber threat intelligence and information sharing can help focus and prioritize the use of the immense volumes of complex cyber security information organizations face today, they have a foundational need for standardized, structured representations of this information to make it tractable. The Structured Threat Information eXpression (STIX™) is a quickly evolving, collaborative community-driven effort to define and develop a language to represent structured threat information. The STIX language is meant to convey the full range of cyber threat information and strives to be fully expressive, flexible, extensible, automatable, and as human-readable as possible. Though relatively new and still evolving, it is actively being adopted or considered for adoption by a wide range of cyber threat-related organizations and communities around the world. All interested parties are welcome to participate in evolving STIX as part of its open, collaborative community via the STIX web site, email discussion lists and other collaborative forums.

Table of Contents

Introduction

This document reflects ongoing efforts to create, evolve, and refine the community-based development of sharing and structuring cyber threat information. STIX is built upon feedback and active participation from organizations and experts across a broad spectrum of industry, academia, and government. This includes consumers and producers of cyber threat information in security operations centers, CERTs, cyber threat intelligence cells, and security executives and decision makers, as well as numerous currently active information sharing groups, with a diverse set of sharing models. MITRE serves as the moderator of the STIX community on behalf of the Department of Homeland Security (DHS) and welcomes your participation.

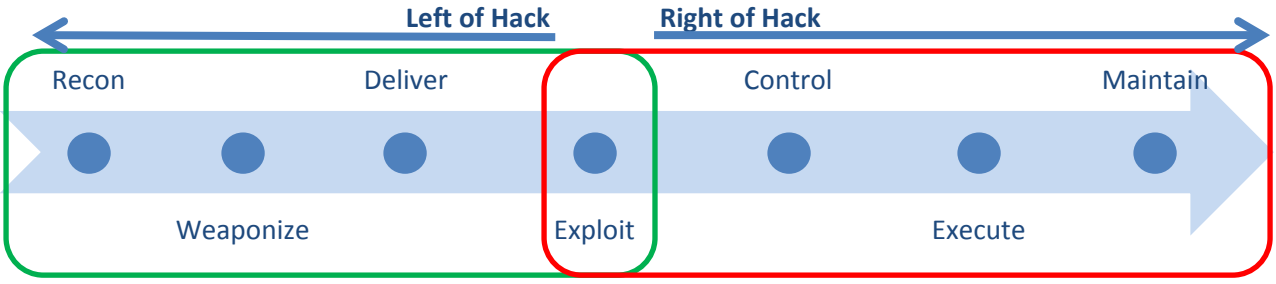
Background

Cyber security is a complex and multifaceted problem domain and continues to become more so. Our dependence on complex technology continues to grow and, at the same time, the threat environment continues to grow and evolve in dynamic and daunting ways. Traditional approaches for cyber security, that focus inward on understanding and addressing vulnerabilities, weaknesses and configurations are necessary but insufficient. Effective defense against current and future threats also requires the addition of a balancing, outward focus, on understanding the adversary's behavior, capability, and intent. Only through a balanced understanding of both the adversary and ourselves can we hope to understand enough about the true nature of the threats we face to make intelligent defensive decisions.

Today's evolving threat environment also brings with it far more complex attack scenarios. Alongside commoditized threats, more advanced capabilities that were rare in the past are now commonplace. Adversary behavior is not solely focused on widespread, disruptive activity, such as the Storm worm outbreak of years gone by, but rather it often involves more targeted, lower-profile multi-stage attacks that aim to achieve specific tactical objectives and establish a persistent foothold into our enterprises.

This newer attack scenario, and how to defend against it, can be effectively understood from the defensive perspective of a "kill chain"[1] showing the multiple steps in an attack. As shown below, the adversary's attack unfolds in a series of steps, ending with the attacker having an established foothold in the victim's network. This is the *modus operandi* of today's sophisticated advanced persistent threat, more commonly known as the APT. APT actors are typically assumed to be nation states but the same behaviors can also be exhibited by those engaged in conducting cyber crime, financial threats, industrial espionage, hacktivism, and terrorism.

[1] http://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf



Example Cyber Kill Chain

The APT's desire and apparent capability to persist and cause ongoing damage is motivating the need to move beyond today's traditional reactive approaches to cyber security and become more proactive. Responding to incidents after the exploit has already occurred is very costly, both in the effective impact and in the level of effort necessary to root out the adversary's established foothold. To be proactive, cyber defenders need to fundamentally change the nature of the game by stopping the adversary's advance, preferably before the exploit stage of the attack illustrated in the kill chain (that is, moving left of the hack). Moving left of the hack requires defenders to evolve from a defensive strategy based primarily on after-the-fact incident investigation and response to one driven by cyber threat intelligence.

Just as traditional intelligence ascertains an understanding of adversaries' capabilities, actions, and intent, the same value carries over to the cyber domain. Cyber intelligence seeks to understand and characterize things like: what sort of attack actions have occurred and are likely to occur; how can these actions be detected and recognized; how can they be mitigated; who are the relevant threat actors; what are they trying to achieve; what are their capabilities, in the form of tactics, techniques, and procedures (TTP) they have leveraged over time and are likely to leverage in the future; what sort of vulnerabilities, misconfigurations, or weaknesses they are likely to target; what actions have they taken in the past; etc.

A holistic understanding of the threat posed by the adversary enables more effective decision support, prioritization of courses of action and a potential opportunity to fundamentally affect the balance of power between the defender and the adversary. According to Hutchins, Cloppert, and Amin [9]:

> "The effect of intelligence-driven CND [computer network defense] is a more resilient security posture. APT actors, by their nature, attempt intrusion after intrusion, adjusting their operations based on the success or failure of each attempt. In a kill chain model, just one mitigation breaks the chain and thwarts the adversary, therefore any repetition by the adversary is a liability that defenders must recognize and leverage." "Through this model, defenders can develop resilient mitigations against intruders and intelligently prioritize investments in new technology or processes." "If defenders implement countermeasures faster than adversaries evolve, it raises the costs an adversary must expend to achieve their objectives. This model shows, contrary to conventional wisdom, such aggressors have no inherent advantage over defenders."

Cyber threat intelligence itself poses a challenge in that no organization in and of itself has access to an adequate scope of relevant information for accurate situational awareness of the threat landscape. The

way to overcome this limitation is via sharing of relevant cyber threat information among trusted partners and communities. Through information sharing, each sharing partner can potentially achieve a more complete understanding of the threat landscape not only in the abstract but also at the instantial level of what specifics they can look for to find the attacker. The threat that organization A is facing today may very well be one that organization B will face tomorrow. This is especially true when A and B both fall within the targeting scope of a given adversary's campaign. If organization A can share their personal instantial knowledge of what they learned/saw about a given threat in the form of a cyber threat indicator[2], organization B may be able to take advantage of this knowledge to address the threat while it's attack on them is still left of the hack (that is, pre-exploit).

Given the evolving complexities of the threat landscape, the speed at which events occur, and the vast quantities of data involved in cyber threat intelligence and threat information sharing, establishing automation to aid human analysis or execute defensive actions at machine-speed is a prerequisite for any effective approach. Automation will require a feed of quality information and most defensive capabilities will typically be built not from homogenous architectures but rather from a diverse set of differing products and systems. The combination of all of these factors will require standardized, structured threat information representations so the widest possible community of information sources can contribute and leverage information without knowing ahead of time who will be providing what information.

One of the challenges threat-sharing organizations face is the ability to structure cyber threat information, yet not lose the human judgment and control involved in sharing. In many cases, organizations have a desire to exchange information in a way that is both human-readable as well as machine-parsable. This requirement is largely an artifact of many information sharing programs where organizations consume not just the data but also assess the data as part of an intelligence collection process. This intelligence process is largely driven by human intelligence analysts that are focused on types of analysis that are either inappropriate for automation or focused on making decisions that require a human-in-the-loop where the analyst directly benefits from reading threat information for situational awareness and context. In addition, because of the wide range in quality of the shared threat information, the intelligence analyst is often also assessing the fidelity based upon the sources and methods used to produce the threat information.

Given all of these factors, there exists a need for structured representations of threat information that are expressive, flexible, extensible, automatable and readable. This paper outlines a community-driven solution to this problem known as the Structured Threat Information eXpression, or STIX.

Current Approaches

STIX is relatively new, but the practice of cyber threat information sharing, particularly indicators, is not. The information being managed and exchanged today is typically very atomic, inconsistent, and very limited in sophistication and expressivity. Where standardized structures are used, they are typically

[2] Cyber Threat Indicator: A set of cyber observables combined with contextual information intended to represent artifacts and/or behaviors of interest within a cyber security context.

focused on only an individual portion of the overall problem, do not integrate well with each other, or lack coherent flexibility. Many existing indicator sharing activities are human-to-human exchanges of unstructured or semi-structured descriptions of potential indicators, conducted via web-based portals or encrypted email. A more recent trend is the machine-to-machine transfer of relatively simple sets of indicator data fitting already well-known attack models. Efforts fitting this description include the Research and Education Networking Information Sharing and Analysis Center's (REN-ISAC) Security Event System and its Collective Intelligence Framework component, the state of Washington's Public Regional Information Security Event Management (*PRISEM*), the Department of Energy's Cyber Federated Model (CFM), and CERT.FI's and CERT.EE's AbuseHelper.

STIX, however, aims to extend indicator sharing to enable the management and widespread exchange of significantly more expressive sets of indicators as well as other full-spectrum cyber threat information.

Currently, automated management and exchange of cyber threat information is typically tied to specific security product lines, service offerings, or community-specific solutions. STIX will enable the sharing of comprehensive, rich, "high-fidelity" cyber threat information across organizational, community, and product/service boundaries.

History

STIX originally evolved out of discussions among the security operations and cyber threat intelligence experts on the IDXWG email list (established by members of US-CERT and CERT.org in 2010 to discuss automated data exchange for cyber incidents) regarding the development of a standardized representation for cyber threat indicators. Out of these discussions a rough structured threat information architecture diagram was created. The original purpose of this architecture diagram was to clearly define the scope of what sorts of information should be included within a structured cyber threat indicator and what sorts of information should be defined in other related structures. This architecture diagram helped to clarify scope such that initial cuts at a structured representation for cyber threat indicators could be successfully drafted. As the concept and initial structure for cyber threat indicators matured, there was increasing interest from numerous parties in fleshing-out the rest of the structured threat information architecture. An XML Schema implementation of the full STIX architecture, incorporating the cyber threat indicator representation among others, is the result of those discussions and the vehicle for current ongoing development of the STIX language among a broad and dynamically growing community.

What is STIX?

STIX is a language, being developed in collaboration with any and all interested parties, for the specification, capture, characterization and communication of standardized cyber threat information. It does so in a structured fashion to support more effective cyber threat management processes and application of automation.

A variety of high-level cyber security use cases rely on such information including:

- Analyzing cyber threats
- Specifying indicator patterns for cyber threat

- Managing cyber threat response activities
- Sharing cyber threat information

STIX provides a common mechanism for addressing structured cyber threat information across and among this full range of use cases improving consistency, efficiency, interoperability, and overall situational awareness.

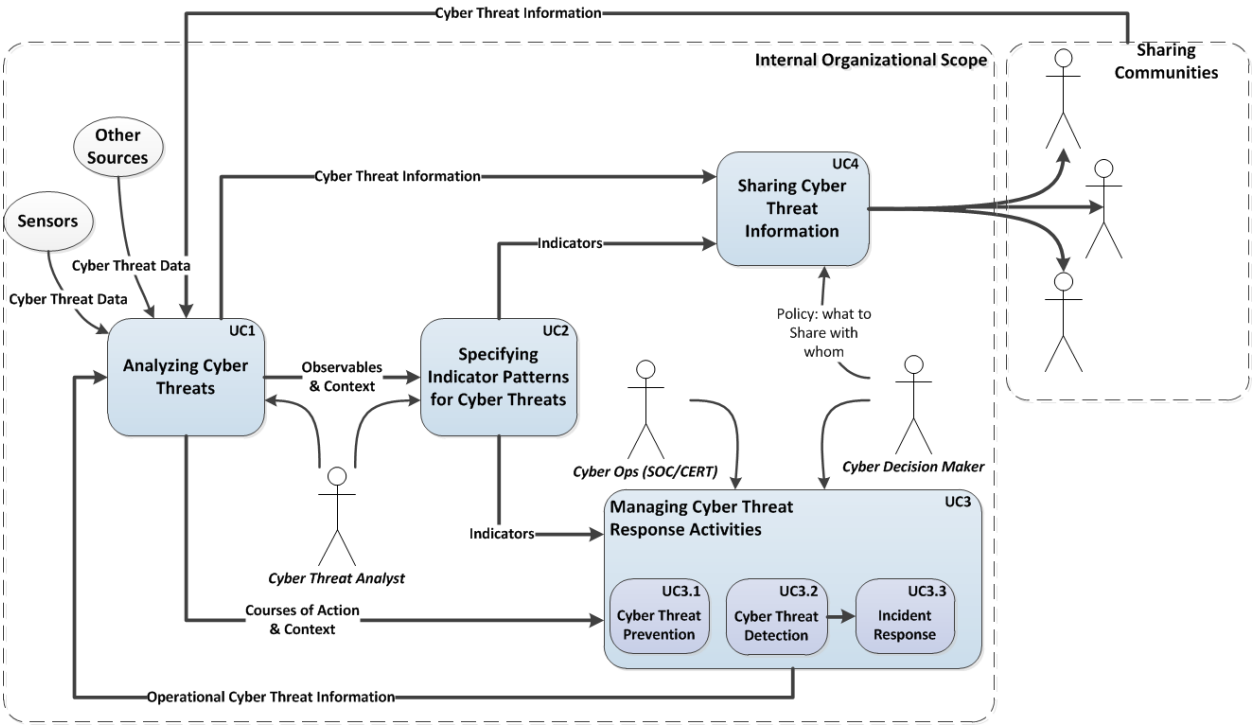
In addition, STIX provides a unifying architecture tying together a diverse set of cyber threat information including:

- Cyber Observables
- Indicators
- Incidents
- Adversary Tactics, Techniques, and Procedures (including attack patterns, malware, exploits, kill chains, tools, infrastructure, victim targeting, etc.)
- Exploit Targets (e.g., vulnerabilities, weaknesses or configurations)
- Courses of Action (e.g., incident response or vulnerability/weakness remedies or mitigations)
- Cyber Attack Campaigns
- Cyber Threat Actors

To enable such an aggregate solution to be practical for any single use case STIX is both flexible and extensible. Existing standardized languages may be leveraged as optional extensions where appropriate and numerous flexibility mechanisms are designed into the language. In particular, 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 for it (from a single field to the entire language or anything in between) without being overwhelmed by the rest. Specific subsets of STIX capabilities can be defined and agreed to beforehand in the form of profiles for use within sharing communities, by tools, etc.

Use Cases

STIX is targeted to support a range of core use cases involved in cyber threat management. Very simple overviews of these use cases are provided below.



Core Use Cases Targeted by STIX

(UC1) Analyzing Cyber Threats

A cyber threat analyst reviews structured and unstructured information regarding cyber threat activity from a variety of manual or automated input sources. The analyst seeks to understand the nature of relevant threats, identify them, and fully characterize them such that all of the relevant knowledge of the threat can be fully expressed and evolved over time. This relevant knowledge includes threat-related actions, behaviors, capabilities, intents, attributed actors, etc. From this understanding and characterization the analyst may then specify relevant threat indicator patterns, suggest courses of action for threat response activities, and/or share the information with other trusted parties. For example, in the case of a potential phishing attack, a cyber threat analyst may analyze and evaluate a suspected phishing email, analyze any email attachments and links to determine if they are malicious, determine if the email was sent to others, assess commonality of who/what is being targeted in the phishing attack, determine whether malicious attachments were opened or links followed, and keep a record of all analysis performed.

(UC2) Specifying Indicator Patterns for Cyber Threats

A cyber threat analyst specifies measurable patterns representing the observable characteristics of specific cyber threats along with their threat context and relevant metadata for interpreting, handling, and applying the pattern and its matching results. This may be done manually or with the assistance of

automated tooling and structured instantial threat information. For example, in the case of a confirmed phishing attack, a cyber threat analyst may harvest the relevant set of observables (e.g., to or from addresses, actual source, subject, embedded URLs, type of attachments, specific attachment, etc.) from the performed analysis of the phishing email, identify the relevant TTPs exhibited in the phishing attack, perform kill chain correlation of the attack, assign appropriate confidence for the indicator, determine appropriate handling guidance, generate any relevant automated rule patterns for the indicator (e.g. Snort, YARA, OVAL, etc.), assign any suggested courses of action, and package it all up as a coherent record for sharing (UC4, below) and future reference.

(UC3) Managing Cyber Threat Response Activities

Cyber decision makers and cyber operations personnel work together to prevent or detect cyber threat activity and to investigate and respond to any detected incidences of such activity. Preventative courses of action may be remedial in nature to mitigate vulnerabilities, weaknesses, or misconfigurations that may be targets of exploit. After detection and investigation of specific incidents, reactive courses of action may be pursued. For example, in the case of a confirmed phishing attack with defined indicators, cyber decision makers and cyber operations personnel work together to fully understand the effects of the phishing attack within the environment including malware installed or malware executed, to assess the cost and efficacy of potential courses of action, and to implement appropriate preventative or detective courses of action.

(UC3.1) Cyber Threat Prevention

Cyber decision makers evaluate potential preventative courses of action for identified relevant threats and select appropriate actions for implementation. Cyber operations personnel implement selected courses of action in order to prevent the occurrence of specific cyber threats whether through general prophylactic application of mitigations or through specific targeted mitigations initiated by predictive interpretation of leading indicators. For example, in the case of a confirmed phishing attack with defined indicators, a cyber decision maker may evaluate a suggested preventative course of action (e.g., implementing a blocking rule at the email gateway) defined within an indicator for the phishing attack, determine its relevant cost and risk, and decide whether or not to implement it. If it is decided to implement the suggested course of action, cyber operations personnel carry out the implementation.

(UC3.2) Cyber Threat Detection

Cyber operations personnel apply mechanisms (both automated and manual) to monitor and assess cyber operations in order to detect the occurrence of specific cyber threats whether in the past through historical evidence, currently ongoing through dynamic situational awareness, or apriori through predictive interpretation of leading indicators. This detection is typically via cyber threat indicator patterns. For example, in the case of a confirmed phishing attack with defined indicators, cyber operations personnel may harvest any specified observable patterns from defined indicators of the attack and apply them appropriately within the operational environment to detect any evidence of the phishing attack occurring.

(UC 3.3) Incident Response

Cyber operations personnel respond to detections of potential cyber threats, investigate what has occurred or is occurring, attempt to identify and characterize the nature of the actual threat, and potentially carry out specific mitigating or corrective courses of action. For example, in the case of a confirmed phishing attack, cyber operations personnel may conduct investigative activities to determine whether the phishing attack was successful in carrying out negative effects within the target environment (e.g., was malware installed or run) and if so, attempt to characterize in detail those effects (e.g., which systems were affected by malware, what data was exfiltrated, etc.). Once the effects are understood, cyber operations personnel would implement appropriate mitigating or corrective courses of action (e.g. wipe and restore systems, block exfiltration channels, etc.).

(UC4) Sharing Cyber Threat Information

Cyber decision makers establish policy for what sorts of cyber threat information will be shared with which other parties and how it should be handled based on agreed to frameworks of trust in such a way as to maintain appropriate levels of consistency, context and control. This policy is then implemented to share the appropriate cyber threat indicators and other cyber threat information. For example, in the case of a confirmed phishing attack with defined indicators, the policies predefined by cyber decision makers could enable the relevant indicators to be automatically or manually shared with trusted partners or communities such that they could take advantage of the knowledge gained by the sharing organization.

Guiding Principles

In its approach to defining a structured representation for cyber threat information the STIX effort strives to adhere to and implement a core set of guiding principles that community consensus has deemed necessary. These principles are as follows:

Expressivity

In order to fully support the diversity of threat-relevant use cases within the cyber security domain STIX is targeted to provide aggregated expressive coverage across all of its targeted use cases rather than specifically targeting one or two. STIX is intended to provide full expressivity for all relevant information within the cyber threat domain.

Integrate rather than Duplicate

Wherever the scope of STIX encompasses structured information concepts for which adequate and available consensus standardized representations already exist, the default approach is to provide appropriate mechanisms to integrate these representations into the overall STIX architecture rather than attempt to unnecessarily duplicate them.

STIX directly leverages the following constituent schema:

- Cyber Observable eXpression (CybOX™)
  http://cybox.mitre.org/

Version 1.0 of STIX provides loose-coupling mechanisms and default implementations for leveraging the following constituent schemas as appropriate:

- Common Attack Pattern Enumeration and Classification (CAPEC™)
  http://capec.mitre.org/
- Malware Attribute Enumeration and Characterization (MAEC™)
  https://maec.mitre.org/
- Common Vulnerability Reporting Framework (CVRF)
  http://www.icasi.org/cvrf
- OASIS Customer Information Quality (CIQ) xPRL
  https://www.oasis-open.org/committees/ciq/

Flexibility

In order to support a wide range of use cases and information of varying levels of fidelity, STIX is intentionally designed to offer as much flexibility as possible. STIX adheres to a policy of allowing users to employ any portions of the standardized representation that are relevant for a given context and avoids mandatory features wherever possible.

Extensibility

In order to support a range of use cases with potentially differing representation details and to ensure ease of community-driven refinement and evolution of the language, the STIX design intentionally builds in extension mechanisms for domain specific use, for localized use, for user-driven refinement and evolution, and for ease of centralized refinement and evolution.

Automatability

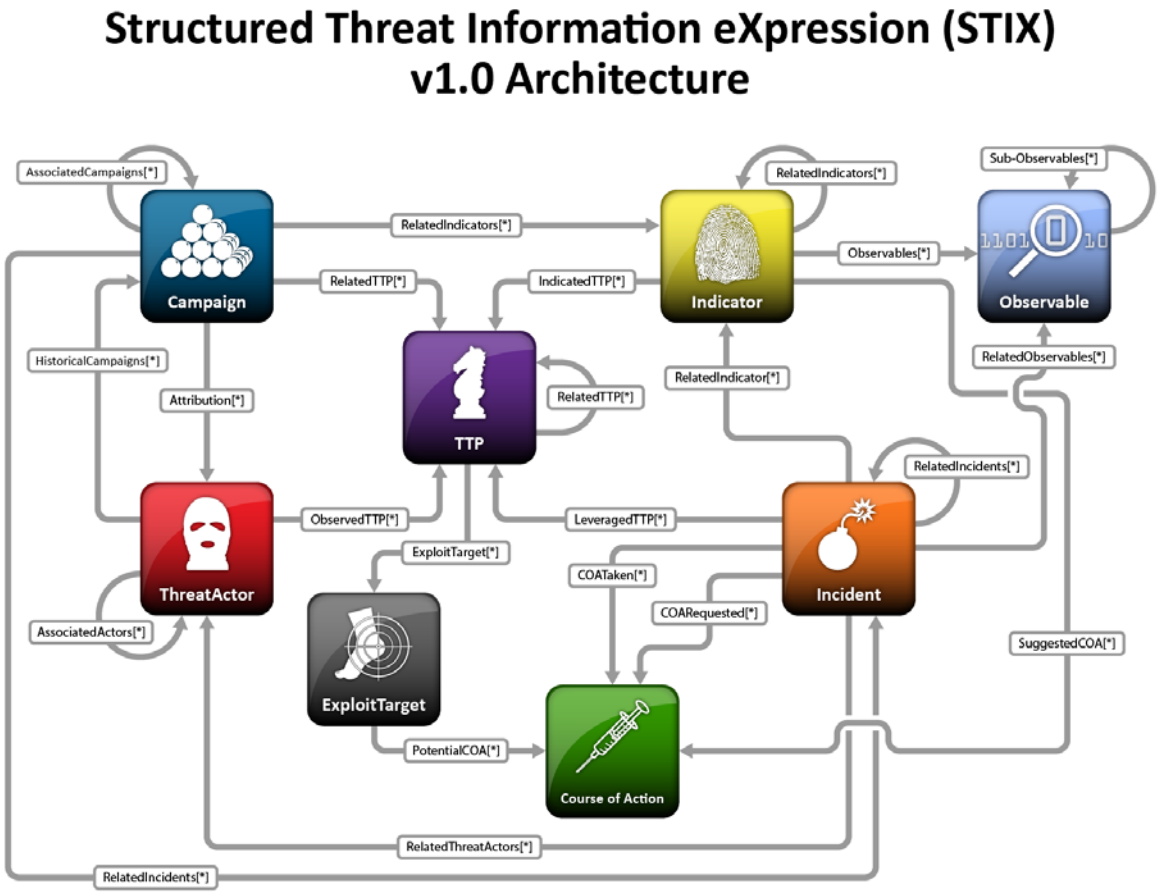
The STIX design approach intentionally seeks to maximize structure and consistency to support machine-processable automation.

Readability

The STIX design approach intentionally seeks for content structures to not only be machine-consumable and processable but also to, as much as possible, be human-readable. This human readability is necessary for clarity and comprehensibility during the early stages of development and adoption, and for sustained use in diverse environments going forward.

Architecture

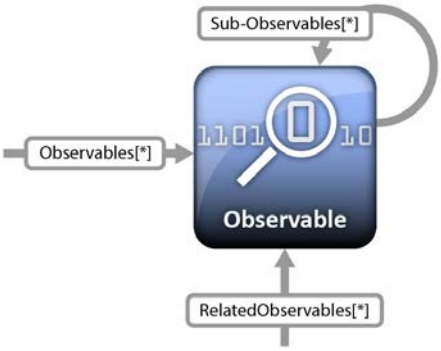
The STIX architecture diagram below identifies the core cyber threat concepts as independent and reusable constructs and characterizes all of their interrelationships based on the inherent meaning and content of each. Connecting arrows between construct icons indicate relationships in the form of content elements within the construct at the root of the connecting arrow, that is of the conceptual type of the construct at the head of the connecting arrow and is suggested but not required to utilize the specific STIX implementation of that construct. The bracketed asterisk on each of the arrow labels implies that each relationship may exist zero to many times. The structured content of each construct is fleshed out in detail within the language implementation (currently in the form of an XML Schema). The eight core constructs—Observable, Indicator, Incident, TTP, ExploitTarget, CourseOfAction, Campaign and ThreatActor— along with a cross-cutting Data Markings construct are briefly characterized below.



STIX Structure

Observables

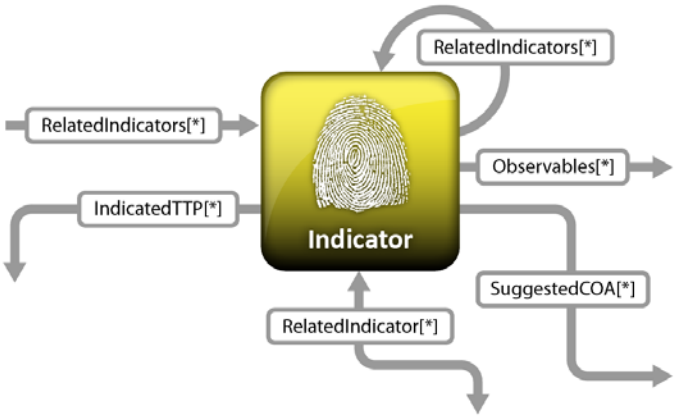
Observables are the "base" construct within the STIX architecture. Observables are stateful properties or measurable events pertinent to the operation of computers and networks. Information about a file (name, hash, size, etc.), a registry key value, a service being started, or an HTTP request being sent are all simple examples of observables. STIX leverages CybOX for its representation of Observables.



CybOX is a language for encoding and communicating standardized high-fidelity information about cyber observables, whether dynamic events or stateful measures properties that are observable in the operational cyber domain. CybOX, like STIX, is not targeted at a single cyber security use case but rather is intended to be flexible enough to offer a common solution for all cyber security use cases requiring the ability to deal with cyber observables. It is also intended to be flexible enough to allow both the high-fidelity description of instances of cyber observables that have been measured in an operational context as well as more abstract patterns for potential observables that may be targets for observation and analysis apriori.

Indicators

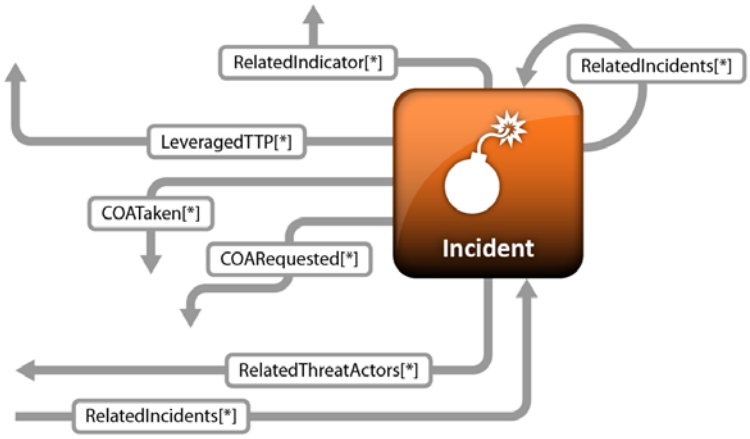
Indicators convey specific Observable patterns combined with contextual information intended to represent artifacts and/or behaviors of interest within a cyber security context. They consist of one or more Observable patterns potentially mapped to a related TTP context and adorned with other relevant metadata on things like confidence in the indicator's assertion, handling restrictions, valid time windows, likely impact, sightings of the indicator, structured test mechanisms for detection, suggested courses of action, related indicators, the source of the Indicator, etc.



Recognizing limitations in current standardized approaches of representation, STIX leverages community knowledge and best practices to define a new Indicator structure for representing Indicator information.

Incidents

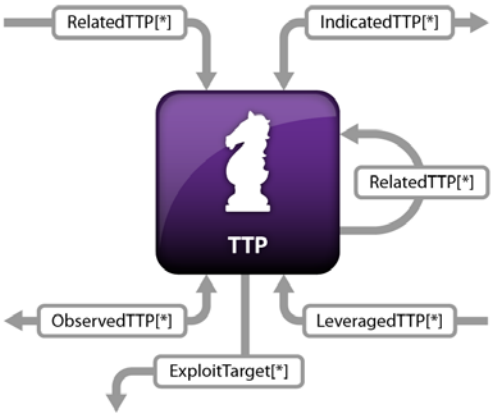
Incidents are discrete instances of Indicators affecting an organization along with information discovered or decided during an incident response investigation. They consist of data such as time-related information, parties involved, assets affected, impact assessment, related Indicators, related Observables, leveraged TTP, attributed Threat Actors, intended effects, nature of compromise, response Course of Action requested, response Course of Action taken, confidence in characterization, handling guidance, source of the Incident information, log of actions taken, etc.



Recognizing limitations in current standardized approaches of representation, STIX leverages community knowledge and best practices to define a new Incident structure for representing Incident information.

Tactics, Techniques and Procedures (TTP)

TTPs are representations of the behavior or *modus operandi* of cyber adversaries. It is a term taken from the traditional military sphere and is used to characterize what an adversary does and how they do it in increasing levels of detail. For instance, to give a simple example, a tactic may be to use malware to steal credit card credentials. A related technique (at a lower level of detail) may be to send targeted emails to potential victims, which have documents attached containing malicious code which executes upon opening, captures credit card information from keystrokes, and uses http to communicate with a command and control server to transfer information. A related procedure (at a lower level of detail) may be to perform open source research to identify potentially gullible individuals, craft a convincing socially engineered email and document, create malware/exploit that will bypass current antivirus detection, establish a command and control server by registering a domain called mychasebank.org, and send mail to victims from a Gmail account called accounts-mychasebank@gmail.com.

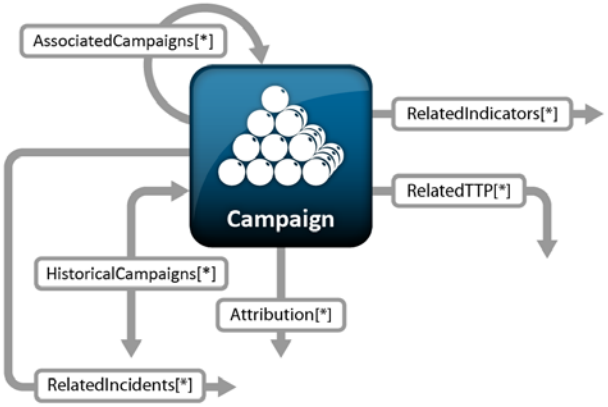


TTPs consist of the specific adversary behavior (attack patterns, malware, exploits) exhibited, resources leveraged (tools, infrastructure), information on the victims targeted (who, what or where), relevant ExploitTargets being targeted, intended effects, relevant kill chain phases, handling guidance, source of the TTP information, etc.

TTPs play a central role in cyber threat information and cyber threat intelligence. They are relevant for Indicators, Incidents, Campaigns, and ThreatActors. In addition, they hold a close relationship with ExploitTargets that characterize the specific targets that the TTPs seek to exploit.

Recognizing a lack of current standardized approaches, STIX leverages community knowledge and best practices to define a new TTP structure for representing TTP information. However, portions of the TTP construct utilize defined extension points to enable leveraging of other existing standardized approaches for detailed characterization of things like behaviors in the form of attack patterns and malware and resources in the form of tools and infrastructure. CAPEC may be utilized for structured characterization of TTP attack patterns. MAEC may be utilized for structured characterization of TTP malware. CybOX is utilized for characterization of tools and infrastructure.

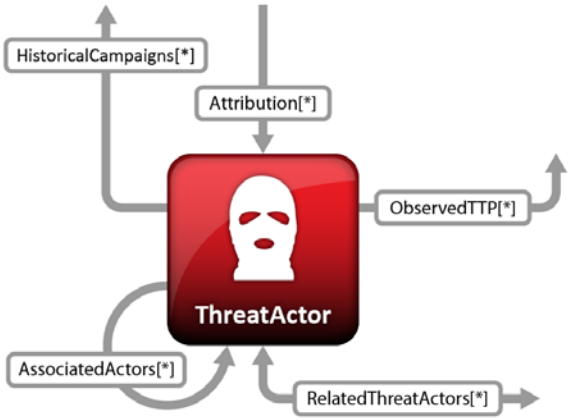
Campaigns



Campaigns are instances of ThreatActors pursuing an intent, as observed through sets of Incidents and/or TTP, potentially across organizations. In a structured sense, Campaigns may consist of the suspected intended effect of the adversary, the related TTP leveraged within the Campaign, the related Incidents believed to be part of the Campaign, related Indicators for behavior associated with the Campaign, attribution to the ThreatActors believed responsible for the Campaign, other Campaigns believed related to the Campaign, confidence in the assertion of aggregated intent and characterization of the Campaign, activity taken in response to the Campaign, source of the Campaign information, handling guidance, etc.

Recognizing a lack of current standardized approaches, STIX leverages community knowledge and best practices to define a new Campaign structure for representing Campaign information.

ThreatActors



ThreatActors are characterizations of malicious actors (or adversaries) representing a cyber attack threat including presumed intent and historically observed behavior. In a structured sense, ThreatActors consist of a characterization of identity, suspected motivation, suspected intended effect, historically observed TTP used by the ThreatActor, historical Campaigns believed associated with the ThreatActor, other ThreatActors believed associated with the ThreatActor, handling guidance, confidence in the asserted characterization of

the ThreatActor, source of the ThreatActor information, etc.

Recognizing a lack of current standardized approaches, STIX leverages community knowledge and best practices to define a new ThreatActor structure for representing ThreatActor information.

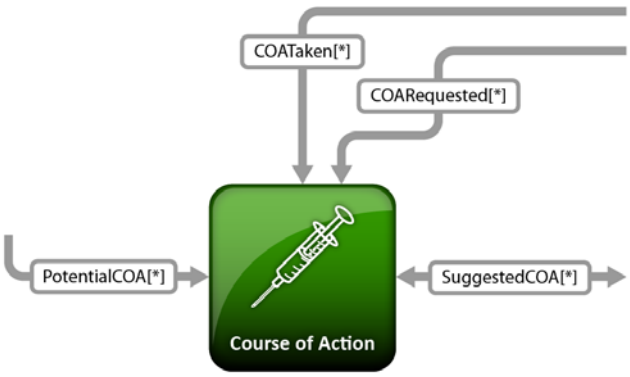
ExploitTargets

ExploitTargets are vulnerabilities or weaknesses in software, systems, networks or configurations that are targeted for exploitation by the TTP of a ThreatActor. In a structured sense, ExploitTargets consist of vulnerability identifications or characterizations, weakness identifications or characterizations, configuration identifications or characterizations, potential Courses of Action, source of the ExploitTarget information, handling guidance, etc.

Recognizing a lack of current standardized approaches for generalized characterizations, STIX leverages community knowledge and best practices to define a new ExploitTarget structure for representing ExploitTarget information. However, portions of the ExploitTarget structure utilize defined extension points to enable leveraging of other existing standardized approaches for characterizing things like vulnerabilities, weaknesses, and configurations. The identifier constructs from the Common Vulnerabilities and Exposures (CVE®) and the Open Source Vulnerability Database (OSVDB) are utilized for identification of publicly disclosed vulnerabilities. The Common Vulnerability Reporting Framework (CVRF) format may be utilized for detailed structured characterization of vulnerabilities not identified in CVE or OSVDB including the potential for characterizing 0-day vulnerabilities. The identifier construct from the Common Weakness Enumeration (CWE™) is utilized for identification of weaknesses. The identifier construct from the Common Configuration Enumeration (CCE™) is utilized for identification of configuration issues.

CoursesOfAction (COA)

CoursesOfAction are specific measures to be taken to address threat whether they are corrective or preventative to address ExploitTargets, or responsive to counter or mitigate the potential impacts of Incidents. In a structured sense, COA consist of their relevant stage in cyber threat management (e.g., remedy of an ExploitTarget or response to an Incident), type of COA, description of COA, objective of the COA, structured representation of the COA (e.g., IPS rule or automated patch/remediation), the likely impact of the COA, the likely cost of the COA, the estimated efficacy of the COA, handling guidance, etc.



Recognizing a lack of current standardized approaches for generalized characterizations, STIX leverages community knowledge and best practices to define a new COA structure for representing Courses of Action information.

Data Markings

A consistent requirement across many of the core STIX constructs is the ability to represent markings of the data to specify things like handling restrictions given the potentially sensitive nature of many forms of cyber threat information. This construct is not conveyed as a separate entity in the STIX architecture diagram but exists as a cross-cutting structure within all of the top-level constructs described above. There currently exists no broad consensus standardized approach for such markings but rather various approaches within differing communities and driven by different motivations and usage contexts. Rather than adopting a single marking approach (e.g., Traffic Light Protocol (TLP)[3]) and attempting to force everyone else to accept it, STIX has taken a flexible and generic approach.

The Data Markings structure defined for STIX is flexible in two ways. First, instances are specified independent of the structures being tagged (pointing to their locations) rather than embedded everywhere which is typically less efficient and more difficult to update and refine. Second, it allows the definition and use of any data marking structure simply as an abstraction from a base type structure. This allows varying marking schemes to be leveraged and when combined with the independent specification described above it can also easily enable marking any given data structure with multiple different marking schemes for leveraging within different use cases or by different communities. The initial implementation of this Data Marking structure has been created utilizing XML Schema (XSD) and has the potential to be leveraged not only throughout an XML implementation of STIX but also for any other XML-based structured representation.

Implementations

The initial implementation for STIX is utilizing XML Schema as a ubiquitous, portable and structured mechanism for detailed discussion, collaboration and refinement among the communities involved. It also provides a practical structure for real-world prototyping and proof-of-concept implementations in structured threat information management and sharing. This sort of real-world usage of the language will be encouraged and supported through the development of various supporting utilities such as programmatic language bindings, representation translation transforms, APIs, etc. Only through appropriate levels of collaborative iteration among a relevant community of experts and vetted through real-world data and use cases can a practical and effective solution evolve.

The structure for the language is also planned to be abstracted into an implementation-independent specification. This will then enable other potential implementations to be derived including possibilities such as semantic web (RDF/OWL), JSON-centric, ProtocolBuffers, etc.

[3] http://www.us-cert.gov/tlp/

Usage

STIX has garnered significant interest from a broad range of organizations and communities facing the challenges of undertaking or supporting cyber threat intelligence and information sharing. Examples of organizations that have already chosen to begin leveraging STIX (or its constituent components) to convey cyber threat information include:

- The U.S. Department of Homeland Security (DHS) is leveraging STIX in a number of critical areas including the Trusted Automated eXchange of Indicator Information (TAXII)[4] effort which allows the Office of Cybersecurity and Communications (CS&C) and its partners in both government and the private sector to exchange data elements and relationships defined by STIX using secure automated mechanisms. Through the use of STIX, they seek to enable the rapid detection, prevention and mitigation of cyber threats and where possible, automate key elements of this process. Initial proof-of-concept efforts for TAXII are currently underway.
- The DHS Cyber Information Sharing and Collaboration Program (CISCP) is currently utilizing STIX for publication of all of its operational threat information to program partners.
- US-CERT is working on integration of STIX as an enabler for its internal incident response and incident management processes.
- In Q2 2012, the Financial Services Information Sharing and Analysis Center (FS-ISAC) agreed to implement the STIX architecture (including CybOX) for cyber threat information sharing amongst its constituent members in the financial services arena. As of May 2013, FS-ISAC stood up an operational threat information repository utilizing STIX and available to all of its ~4200 member organizations.
- The MITRE Corporation is currently in the process of integrating STIX and the underlying components into its cyber threat intelligence platform known as CRITs (Collaborative Research Into Threats) [14].
- The Japanese IPA (Information-technology Promotion Agency, Japan) is currently undertaking an active feasibility study of applying elements of the STIX architecture (CybOX, MAEC, etc.) as an international exchange format for cyber observables and threat information.

In addition, STIX (and its constituent components) is under active consideration for use and initial prototyping among a large variety of different U.S. and international public-public, public-private and private-private cyber threat information sharing communities and by several vendors supporting the domain.

Conclusion and Future Work

There is an urgent need for new and more outward-looking collaborative approaches to cyber security defense. Cyber threat intelligence and cyber threat information sharing are on the leading edge of novel approaches with a high potential for shifting the balance of power between the attacker and the defender. A core requirement for maturing effective cyber threat intelligence and cyber threat information sharing is the availability of an open-standardized structured representation for cyber

[4] http://taxii.mitre.org

threat information. STIX is a community-driven effort to provide such a representation adhering to guiding principles to maximize expressivity, flexibility, extensibility, automatability, and readability. STIX provides expressive coverage of the full-spectrum of cyber threat information—observables, indicators, incidents, TTP, exploit targets, courses of action, threat actors and campaigns—to provide support for a broad set of cyber security defense use cases. Though relatively new and still evolving, STIX has already generated a great deal of interest from a wide range of stakeholders and communities, both public and private. Evidence so far indicates that STIX has a high potential to be a solution to the cyber threat information problem that is effective, practical, and acceptable to the community of practitioners and supporting entities.

The STIX effort foresees increasing levels of community involvement as it continues to gain awareness and initial real-world use. Community effort will be focused on evaluation and refinement of the schematic implementation of the language (eventually leading to the abstraction to an implementation-independent language specification), on continued prototyping and use, and on development of supporting utilities. In addition, effort will be applied to the establishment of a branding and adoption program to facilitate and support the emerging ecosystem of products, services, and information sources that can be leveraged together to address cyber threat information sharing.

All parties interested in becoming part of the collaborative community discussing, developing, refining, using and supporting STIX are welcome and invited to join the effort. More information is available on the STIX website (http://stix.mitre.org/). Questions or comments may be sent to the STIX team at stix@mitre.org or to the STIX community on the STIX discussion list (http://stix.mitre.org/community/registration.html). Access to full schema implementations, utilities and issue trackers is available from the STIX Project GitHub site (https://github.com/STIXProject).

Acknowledgments

MITRE's services as community coordinator of the STIX effort and editor of this paper are provided under the sponsorship of the U.S. Department of Homeland Security. While the summary work contained in this paper is based on the efforts, comments, and conversations on these topics by a large number of individuals from many organizations, special thanks is made for those individuals that took time to review and comment on the specific text in this document.

A sampling of some of the organizations contributing to the STIX discussion includes:

- MITRE Corporation
- United States Computer Emergency Readiness Team (US-CERT)
- DHS Cyber Information Sharing and Collaboration Program (CISCP)
- DHS National Cybersecurity and Communications Integration Center (NCCIC)
- National Institute of Standards and Technology (NIST)
- Financial Services Information Sharing and Analysis Center (FS-ISAC)
- CERT Coordination Center (CERT/CC)

- Research and Engineering Networking Information Sharing and Analysis Center (REN-ISAC)
- Industrial Control System Information Sharing and Analysis Center (ICS-ISAC)
- Electricity Sector Information Sharing Analysis Center (ES-ISAC)
- Department of Defense Cyber Crime Center (DC3)
- National Cyber Investigative Joint Task Force Analytical Group (NCIJTF-AG)
- Depository Trust & Clearing Corporation (DTCC)
- USAA
- GE
- Siemens
- Visa
- CrowdStrike, Inc.
- Mandiant
- CyberIQ
- LookingGlass
- iSIGHT Partners
- Red Sky Alliance
- Tripwire
- Verisign
- Verizon
- Lockheed Martin
- General Dynamics
- Northrop Grumman
- Booz Allen Hamilton
- Threat Stream
- ThreatGRID
- Bromium
- Reversing Labs
- Cyber2
- Internet Identity
- StegoSystems
- G2, Inc.
- Cyveillance
- Critical Intelligence
- Incident Logic
- Punch Cyber Analytics Group
- NCI Security LLC
- Foreground Security
- SHARKSEER Program
- Public Regional Information Security Event Management (PRISEM)
- NATO
- World Bank
- CERT-EU
- Gendarmerie Nationale (France)
- Trustworthy Software Initiative (TSI) (UK Software Security)

References

[1] Cyber Observable eXpression (CybOX). URL https://cybox.mitre.org

[2] Common Attack Pattern Enumeration and Classification (CAPEC). URL https://capec.mitre.org

[3] Malware Attribute Enumeration and Characterization (MAEC). URL https://maec.mitre.org

[4] Common Vulnerabilities and Exposures (CVE). URL https://cve.mitre.org

[5] Common Weakness Enumeration (CWE). URL https://cwe.mitre.org

[6] Common Configuration Enumeration (CCE). URL https://cce.mitre.org

[7] Open Source Vulnerability Database (OSVDB). URL http://www.osvdb.org

[8] Common Vulnerability Reporting Framework (CVRF). http://www.icasi.org/cvrf/

[9] E.M. Hutchins, M.J. Cloppert and R.M Amin PH.D., "Intelligence-Driven Computer Network Defense Informed by Analysis of Adversary Campaigns and Intrusion Kill Chains," *Proc. 6th Int'l Conf. Information Warfare and Security* (ICIW 11), Academic Conferences Ltd., 2010, pp. 113–125; URL http://www.lockheedmartin.com/content/dam/lockheed/data/corporate/documents/LM-White-Paper-Intel-Driven-Defense.pdf

[10] Luc Dandurand, Emmanuel Bouillon and Philippe Lagadec, “High-level Requirements for a Cyber Defence Data Collaboration and Exchange Infrastructure,” Reference Document 3175, NATO Consultation, Command and Control Agency, April 2011.

[11] “TAXII: The Trusted Automated eXchange of Indicator Information,” U.S. Department of Homeland Security

[12] ”Cyber Information-Sharing Models: An Overview,” MITRE Corporation, May 2012.

[13] ”Active Defense Strategy for Cyber,” MITRE Corporation, July 2012.

[14] ”A New Cyber Defense Playbook,” MITRE Corporation, July 2012.

[15] IDXWG@NICWG.ORG, Oct 2011 - May 2013

[16] Traffic Light Protocol (TLP). URL http://www.us-cert.gov/tlp/

[17] Julie Connolly, Mark Davidson, Matt Richard, Clem Skorupka, “The Trusted Automated eXchange of Indicator Information (TAXII™),” November 2012. http://taxii.mitre.org/about/documents/Introduction_to_TAXII_White_Paper_November_2012.pdf