Lenguajes de intercambio de inteligencia

Ronald Cordero Araya, Jerry Quesada Altamirano, and Profesor tutor: Randall Barneett Villalobos

[1] Escuela de Ingeniería,
Universidad Latinoamericana de Ciencia y Tecnología,
ULACIT, Urbanización Tournón, 10235-1000
San José, Costa Rica
Ronald Cordero,rcorderoa284@ulacit.ac.cr
[2] Jerry Quesada, jquesadaa651@ulacit.ac.cr
http://www.ulacit.ac.cr

Resumen Los lenguajes de intercambio de inteligencia son un tema relativamente nuevo el cual viene a dar un impulso a lo que es el intercambio de datos o información, de una forma estandarizada y segura entre dos o más puntos. Como bien se sabe, la información es un elemento de las empresas en que estas gastan más dinero para protegerse, ya que es su activo más preciado, sea cual sea su naturaleza. Hasta hoy, se han realizado lenguajes de intercambio de inteligencia aplicados a lo que es malware únicamente, por lo cual este artículo va a tomar en cuenta solamente estos. Se les investigará: su uso, como se relacionan entre ellos mismos, qué aplicaciones se les pueden dar y sus características principales; entre otros aspectos que puedan ser útiles para llegar a conocer o tener una visión general de estos lenguajes de intercambio de inteligencia.

Keywords: lenguajes, intercambio, inteligencia, intel, sharing, lenguajes, datos, información, protección, malware, seguridad

1. Introducción

El intercambio de información o datos de cualquier índole, o como se le conoce en inglés: "intel sharing" es una práctica que las empresas u oficinas de gobierno; entre otros, realizan con mucha regularidad. Esto se puede dar por varias razones como: que las empresas tengan varias sucursales instaladas en varios sitios geográficamente separados o por la comunicación que deben tener estas con entidades del gobierno o con sus proveedores de servicios o materiales necesarios para el quehacer de la empresa. En estos intercambios de información, suelen ir datos sensibles, por lo cual se debe tener muy en cuenta que la arquitectura que se utiliza para la transmisión de la información sea lo más segura y estable posible. Además de contar con aplicaciones que tengan alejados de ataques a la seguridad que puedan poner en peligro y en cuestión lo que es la integridad y consistencia de los datos.

Por un lado, los lenguajes de intercambio de inteligencia orientados a malware, le da a las empresas una herramienta con la cual pueden enviar y recibir información de manera estandarizada y segura, mientras que por el otro, ayuda en lo que es el estar al tanto de ataques a la seguridad, ya que varios lenguajes de intercambio de inteligencia otorgan datos estandarizados de vulnerabilidades y hallazgos de malware que pueden ser utilizados para tomar decisiones y así configurar los equipos de la infraestructura de red que posee la empresa para mitigar estos ataques del todo; además de otorgar a otras personas dicha información de un ataque que haya ocurrido para que ellos puedan tomar acciones y repeler el ataque.

Ahora, en lo que concierne a este artículo, se va a otorgar una visión general de lo que son los lenguajes de intercambio de inteligencia para la transmisión de información acerca de malware y el estado en que se encuentra actualmente. Además, se realizará un caso de uso en el cual se explique cómo es que estos lenguajes pueden trabajar en conjunto y que aplicación se les pueden dar. Como se ha mencionado anteriormente, este tema es muy nuevo, por lo que la investigación se va a realizar por medio de internet, en las páginas oficiales de cada lenguaje, las cuales dan información confiable y honesta, ya que no hay existencia de libros o artículos arbitrados que mencionen este tema.

De acuerdo con lo dicho anteriormente, los objetivos de la investigación, son los siguientes:

- Definir y analizar los diferentes tipos de lenguajes de intercambio de inteligencia.
- ¿Cuáles las relaciones entre los distintos lenguajes?
- Definir la utilidad de cada lenguaje.
- Preparar caso de uso de una amenaza de un tipo de ataque de malware en una Campaña.

Y las preguntas que se desean responder en la investigación, son las siguientes:

- ¿En qué consiste un lenguaje de intercambio de inteligencia?
- ¿Cuál es la importancia de utilizar lenguajes de intercambio de inteligencia?
- ¿Cómo se aplicaría este tipo de lenguajes de intercambio de inteligencia en nuestro entorno?

2. ¿Qué es un lenguaje de intercambio de inteligencia?

Antes de responder esta pregunta primero se debe de saber claramente qué es "intel sharing" o intercambio de inteligencia. Como bien lo dicen las palabras, esto es intercambiar datos entre dos o más puntos, los cuales pueden ser de cualquier índole, como:

- Seguridad.
- Datos de empresas.
- Clientes.
- Promociones.
- Entre muchos otros más.

Ahora, ¿qué es un lenguaje de intercambio de inteligencia?. Hoy, no existe una definición única o universal para este tema, por lo cual se ofrece el siguiente: "Un lenguaje de intercambio de inteligencia es un metalenguaje para describir y caracterizar de forma estándar los elementos relacionados con los ataques de seguridad". Se propone lo anterior de acuerdo con las investigaciones y reuniones con profesores que se realizaron en el transcurso de la investigación de este artículo.

3. Lenguajes de intercambio de inteligencia existentes

Los lenguajes de intercambio de inteligencia de los que se tiene conocimiento hasta el día de hoy y que se vienen desarrollando, son los detallados en la siguiente tabla.

Dentro de la misma los lenguajes que están resaltados con el simbolo (*) delante del nombre, son a los que se le van a dar más énfasis y se explicarán conceptos, usos, propiedades; entre otras características de cada uno de estos en este artículo.

LENGUAJE	DEFINICIÓN
*CybOX	CybOX es un esquema estandarizado para la especificación, la captura, la caracterización y comunicación de eventos o propiedades que son observables en el ámbito operativo. incluyendo: gestión de eventos / registro, caracterización malware, detección de intrusiones, respuesta a incidentes / gestión, caracterización del patrón de ataque , entre otros.(MITRE, 2015b)
*MAEC	MAEC es un lenguaje estandarizado para la codificación y la comunicación de información de alta fidelidad sobre el malware basado en atributos tales como comportamientos, artefactos y patrones de ataque.(MITRE, 2015c)
*STIX	El lenguaje STIX pretende transmitir toda la gama de información potencial de una amenaza cibernética y se esfuerza por ser totalmente expresiva, flexible, extensible, automatizable y tan legible como sea posible.(MITRE, 2015e)
*TAXII	TAXII, a través de sus especificaciones define conceptos, protocolos y los intercambios de mensajes para intercambiar información sobre la amenaza cibernética para la detección, prevención y mitigación de las amenazas cibernéticas. (MITRE, 2015f)
CAPEC	The Common Attack Pattern Enumeration and Classification (CAPEC™) Un patrón de ataque es un mecanismo de abstracción para ayudar a describir cómo se ejecuta un ataque contra los sistemas o redes vulnerables, La Lista CAPEC ofrece una lista formal de los patrones de ataque conocidos. (MITRE, 2015a)
OVAL	Oval es un esfuerzo de la comunidad de seguridad de información para estandarizar la forma de evaluar e informar sobre el estado de la máquina de los sistemas informáticos. OVAL incluye un lenguaje para codificar los detalles del sistema, y una variedad de repositorios de contenido celebradas en toda la comunidad. (MITRE, 2015d)

4. Lenguaje TAXII

4.1. ¿Qué es?

TAXII (Trusted Automated eXchange of Indicator Information), es un mecanismo de trasporte que automatiza y estandariza el intercambio de información de la cyber seguridad.

4.2. ¿Para qué sirve?

TAXII define un conjunto de servicios e intercambio de mensajes a las organizaciones que permite compartir dicha información de amenazas, para de igual manera advertir a las diferentes organizaciones.

A su vez este lenguaje define conceptos, protocolos y mensajes para intercambiar información sobre una amenaza cibernética para su respectiva detección, prevención y mitigación de las mismas.

4.3. ¿Como funciona?

TAXXI funciona por medio de modelos, los cuales se van a explicar a continuación.

Hub and Spoke: Este modelo se basa en que una organización (representada en el gráfico con la leyenda de fuente) tiene la funcionabilidad de actuar como base o centro de actividad de información, la cual coordina y trasmite datos a sus receptores (el resto de rectángulos gráfico, productores, consumidores). En este caso se podría llamar otras empresas, productores, consumidores, entre otros. Es importante rescatar que en este modelo las diferentes entidades pueden detectar los malware, pero es la principal (fuente) la que se encarga de enviar la información a las diferentes entidades para su respectiva protección.

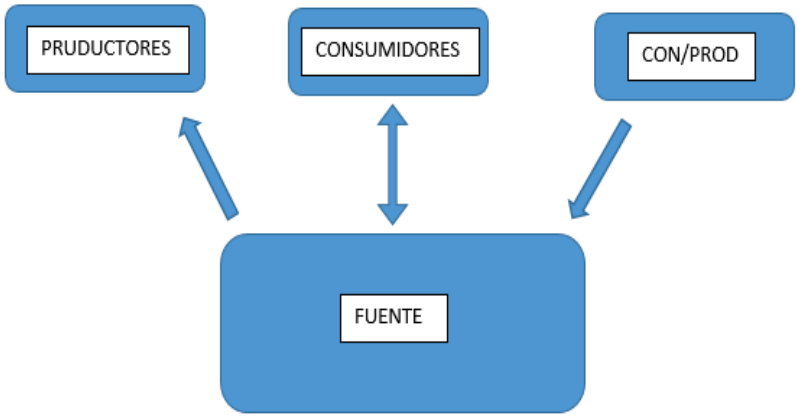

Figura 1. Hub and Spoke

Source/Subcriber: En este otro modelo existe una organización (la que se representa con el rectángulo con la leyenda que indica fuente) que de igual manera tiene la funcionabilidad de ser la única fuente de información de malware hacia las demás entidades para así protegerlas del ataque ya detectado.

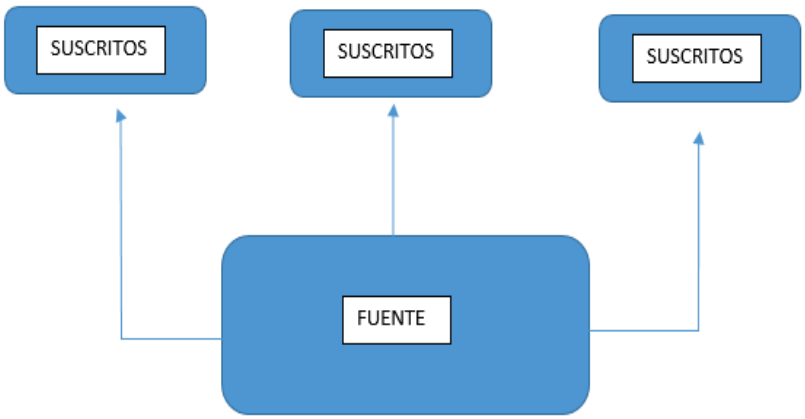

Figura 2. Source/Subcriber

Peer to Peer: En este último modelo existen dos o más organizaciones (representada con los rectángulos del gráfico) que comparten la información recopilada, con las demás entidades, para su respectiva protección. En este modelo no existen fuentes únicas para difundir la información, sino que todas se encargan de pasarla y de universalizarla, para que las entidades se puedan proteger ante cualquier ataque de cyberseguridad.

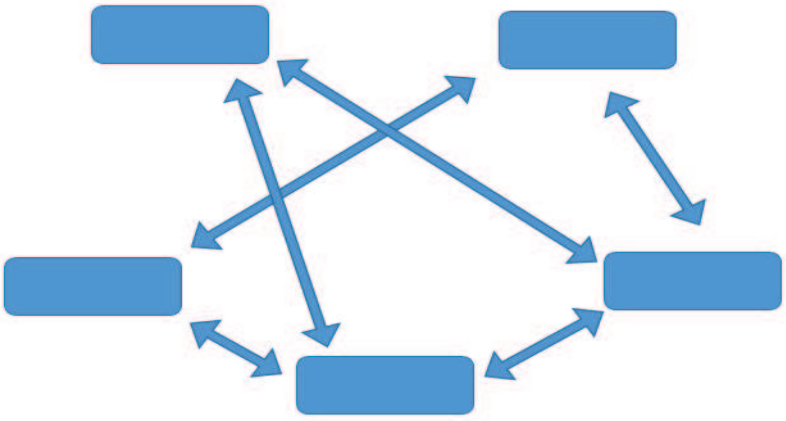

Figura 3. Peer to Peer

4.4. Servicios.

TAXII, a su vez, define cuatro servicios importantes:

- Servicios de entrada.
- Servicio de encuesta.
- Servicio de gestión o cobranza.
- Descubrimiento del servicio.

5. Lenguaje CYBOX.

5.1. ¿Qué es?

Es un esquema estandarizado hecho en XML para la especificación, captura, caracterización y comunicación de eventos o propiedades de los observables cibernéticos.

Estos observables cibernéticos son, como por ejemplo: las especificaciones de un archivo (el nombre, tamaño, tipo de archivo), un registro del sistema o una petición HTTP que se envía; entre otros.

5.2. ¿Para qué sirve?

Este lenguaje de intercambio de inteligencia, ayuda a estandarizar la información que es capturada acerca de los observables cibernéticos para ser utilizada, según convenga.

5.3. ¿Cómo funciona?

Como se explicó anteriormente, este lenguaje es un esquema, o sea, un formulario, el cual se va a llenar de información únicamente de observables cibernéticos. Este esquema está hecho en XML, por lo cual tiene su sintaxis y forma única de llamar sus datos u objetos.

5.4. Ejemplo

Figura 4. Ejemplo de CYBOX

6. Lenguaje STIX.

6.1. ¿Qué es?

Es un lenguaje estandarizado creado en XML para representar información estructurada de una amenaza cibernética. Esta información debe tener las siguientes características: debe ser expresiva, flexible, automatizable y que las personas lo puedan entender.

6.2. ¿Para qué sirve?

Este lenguaje se utiliza para:

- Analizar amenazas.
- Especificar patrones indicadores de amenazas cibernéticas.
- Administrar actividades de respuesta a estas amenazas.
- Intercambiar información acerca de amenazas.

6.3. ¿Cómo funciona?

Este lenguaje permite mostrar información estandarizada y estructurada acerca de amenazas, dentro de la información, que se puede obtener. Al utilizar este lenguaje, se encuentra la siguiente:

- Observables cibernéticos: explicados anteriormente en el lenguaje Cybox.
- Indicadores: Son patrones en los observables cibernéticos que pueden ser de importancia para la seguridad cibernética.
- Incidentes: Es como afecta una amenaza a una organización.
- Tácticas adversarias, técnicas y procedimientos (TTP): es el modo operanti de los actores de amenazas.
- Exploits: Son vulnerabilidades que hay en los sistemas o dispositivos de red que pueden ser explotados.
- Cursos de acción: Es el paso a paso de que hacer para hacerle frente a una amenaza cibernética.
- Campañas de ataques: Son el propósito que tienen a futuro los actores de amenazas.
- Actores de amenazas cibernéticas: son caracterizaciones que se le dan a los actores maliciosos que representan una amenaza cibernética, incluyendo el comportamiento observado históricamente.

6.4. Ejemplo

Figura 5. Ejemplo de STIX

7. Lenguaje MAEC.

7.1. ¿Qué es?

Este lenguaje de intercambio de inteligencia consiste en la detección y prevención de ataques de malware. Se basa en tomar al malware y determinar información de los mismos, como lo son los patrones de ataque, su comportamiento y sus características; entre otros. Esta información que MAEC recopila normalmente lo hace en un archivo plano, o XML, en los cuales se detalla toda la información recopilada anteriormente para así hacerlo vulnerable ante cualquier otro ataque.

7.2. ¿Para qué sirve?

Su objetivo principal es la detección y prevención temprana de malware. Existen muchos repositorios que MAEC posee, estos con el fin de tener como un listado de los malware ya recopilados para que en un futuro no afecten a la organización, y su ataque no afecte en lo más mínimo a las mismas.

7.3. ¿Cómo funciona?

Para la recopilación de los datos, el lenguaje MAEC lo almacena en el siguiente modelo de datos:

- Nivel 1: MAEC Bundle.
- Nivel 2: MAEC Package.
- Nivel 3: MAEC Conteiner.

Seguidamente se explicará cada uno de estos niveles en el modelo de datos de MAEC:

- Nivel 1: En este nivel se realiza la recolección de toda la información del malware, características, comportamiento o patrón de ataque; entre otros. Además, es un contenedor independiente para los datos de salida que fueron recopilados.
- Nivel 2: En este nivel se hace una recopilación y agrupación que abarca toda la información que se recolectó en el nivel 1 el MAEC Bundle. En determinado caso de que encuentren dos o más malware similares en este nivel, se agrupan para un mejor análisis, prevención y protección.
- Nivel 3: En el último nivel es simplemente un contenedor de datos MAEC. En este se almacenan los archivos planos con la información del malware, además; funciona como mecanismo de transporte de los paquetes de datos.

8. Sinergia entre lenguajes.

Estos lenguajes de intercambio de inteligencia se complementan uno con otro para trabajar en equipo y así dar información acerca de malware más detallada y estandarizada, a como lo hicieran cada uno de estos por separado.

STIX encapsula información detallada acerca de amenazas cibernéticas, mientras que MAEC proporciona una manera estructurada de capturar información detallada sobre muestras de malware.

Por un lado, estos dos lenguajes se pueden complementar. El MAEC por ser parte de STIX para cuando se quiera hacer un análisis más detallado de un malware para hacerle frente antes de que ataque.

Por otro lado, STIX y MAEC utilizan dentro de estos, el lenguaje CYBOX, el cual les da la información detallada de los diferentes observables cibernéticos que utilizan cada uno de estos para realizar sus esquemas.

Por último, si se desea transportar la información que se crea a partir de estos lenguajes, se utiliza TAXII, el cual brinda servicios para el intercambio de información.

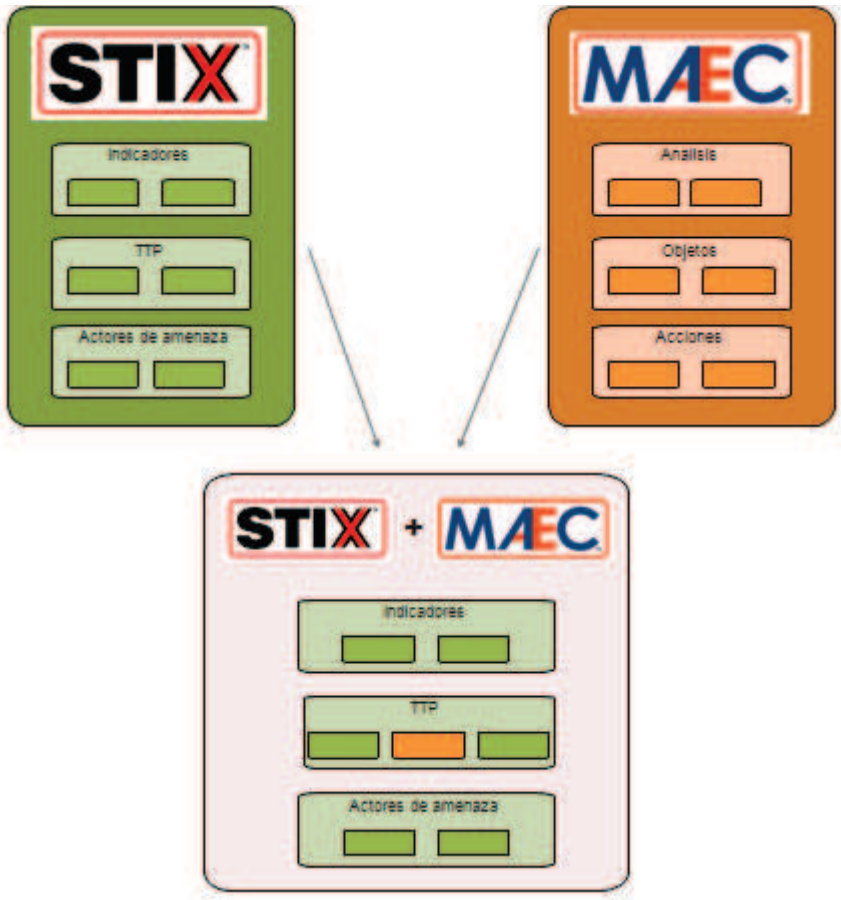

Figura 6. STIX y MAEC

9. ¿Cómo son utilizados?

Para entender de una mejor manera lo que son estos lenguajes, se puede decir que son muy parecidos a un formulario que está hecho en XML, el cual se va a llenar de información estandarizada y necesaria para ser tomada en cuenta acerca de malware. Esta puede ser extraída de múltiples maneras. Una de ellas es por medio de "Network Programing", en español: "Programación de Red". Cuando es el caso de que se quiera extraer información de amenazas por medio de un programa alojado en el firmware de un dispositivo de red (comúnmente router) en la intranet; el cual va a capturar los paquetes de información enviados por la red y por medio de "firmas" de malware se detecten estos y se les extraiga la información para agregarla a algún lenguaje de intercambio de inteligencia.

Otra manera es por medio de "API"s los cuales son pequeños programas que capturan información en los que se pueden utilizar estos lenguajes de intercambio de inteligencia para almacenarla.

Por último, toda esta información que es capturada puede ser almacenada en repositorios de datos (MAEC), los cuales son consultados por aplicaciones para así mostrar la información, realizar acciones o tomar decisiones que ayuden a hacer frente a las diferentes amenazas cibernéticas que existen actualmente.

10. Tabla de casos de uso.

Los lenguajes de intercambio de inteligencia son unas herramientas muy versátiles, ya que pueden ser utilizadas según sea necesario, además se puede utilizar uno o varios lenguajes para dar solución al caso de uso, ya que estos pueden trabajar conjuntamente sin ningún impedimento.

Caso de uso	Proceso	Lenguaje
Detección de actividades maliciosas utilizando caracterizaciones del comportamiento del malware.	Detección de ataque	MAEC
Habilitar el intercambio de indicadores de ataque colectivos.	Intercambio de información	STIX, TAXII
Habilitar nuevos niveles de meta-análisis sobre,observables cibernéticos operacionales.	Estar al tanto de la situación cibernética	CYBOX, STIX
Habilitar controles de intercambio explícitas e implícitas para información de observables.	Intercambio de información	STIX, CYBOX, TAXII

En la tabla anterior se puede observar cuáles lenguajes de intercambio de inteligencia se pueden utilizar dependiendo del caso de uso que se tenga que resolver. Por ejemplo: cuando se necesita realizar nuevos meta-análisis a los observables cibernéticos, se puede utilizar CYBOX y STIX, ya que CYBOX es el encargado de caracterizar los estos observables, por lo tanto se realizan esquemas nuevos en los que se tomen en cuenta, a la hora de realizarlos, las necesidades que se tienen que resolver para los nuevos meta-análisis, mientras que STIX funciona para la hora de mostrar la información de dichos observables cibernéticos para realizar los meta-análisis.

11. Conclusiones

Como bien se probó en la investigación, los lenguajes de intercambio de inteligencia son un tema sumamente reciente, por ende la poca o escasa información hace el asunto sea más exhaustivo. Con la implementación de estos tipos de lenguajes, nos indican que en la actualidad los entes gubernamentales o bien las empresas, buscan como principal objetivo la seguridad de la información, dándole a éste un grado de importancia muy alto. En la investigación se pudieron definir los tipos más importantes de lenguajes de intercambio de inteligencia, al definir tanto su concepto como funcionalidad y sus diferentes aplicaciones. Con esto se realizó una comparación de los mismos, y además un análisis de cómo logran trabajar en equipo para lograr un resultado más satisfactorio al momento de otorgar la información que por ellos mismos capturan y caracterizan.

12. Bibliografia

Referencias

MITRE. (2015a). *Common attack pattern enumeration and classification.* `http://capec.mitre.org/data/index.html`. ([Online; accesado en febrero-2015]) pages 4

MITRE. (2015b). *Cyber observable expression.* `http://cybox.mitre.org/`. ([Online; accesado en febrero-2015]) pages 4

MITRE. (2015c). *Malware attribute enumeration and characterization.* `http://maec.mitre.org/`. ([Online; accesado en febrero-2015]) pages 4

MITRE. (2015d). *Open vulnerability and assessment language.* `http://oval.mitre.org/`. ([Online; accesado en febrero-2015]) pages 4

MITRE. (2015e). *Structured threat information expression.* `http://stix.mitre.org/`. ([Online; accesado en febrero-2015]) pages 4

MITRE. (2015f). *Trusted automated exchange of indicator information.* `http://taxii.mitre.org/`. ([Online; accesado en febrero-2015]) pages 4