Lenguajes de Intercambio de Inteligencia

Adrián Cambronero[1], Alberto Ugalde[1], Io Meng Wong[1], and Profesor Antonio González[1]

[1] Escuela de Ingeniería,
Universidad Latinoamericana de Ciencia y Tecnología,
ULACIT, Urbanización Tournón, 10235-1000
San José, Costa Rica
alumno1,alumno2@ulacit.ac.cr
http://www.ulacit.ac.cr
[2] Departamento,
Institución,
Siglas, Dirección, Apartado postal
San José, Costa Rica
alumno1@institucion.cr
http://www.institucion.cr

Abstract. Los lenguajes de intercambio de inteligencia fueron definidos mediante la Ley de Seguridad Nacional de la Agencia Central de Inteligencia del Gobierno de los Estados Unidos, con fines militares para prevenir el robo de información y los ataques de enemigos. Sin embargo, los lenguajes de intercambio de inteligencia se utilizan hoy en día para proteger la información de las empresas, la cual es considerada como su principal activo. Lo anterior, teniendo en cuenta que el desarrollo de un gran número de productos y negocios está basado en el conocimiento la propiedad intelectual. En consecuencia, el presente trabajo tiene como fin realizar una investigación acerca de los estándares que definen los lenguajes de intercambio de inteligencia más conocidos (MAEC, STIX, CybOX y TAXII), tomando en cuenta sus principales características y aplicaciones.

Keywords: intercambio, inteligencia, lenguajes, ataques, protección

1 Introducción

En el contexto actual el principal capital de muchas empresas está basado en la información y la propiedad intelectual. Las condiciones económicas de un mercado agresivo y competitivo obligan a las empresas a establecerse en diferentes países localizados en diversas regiones, con el fin de reducir costos. Eso obliga a que las empresas tengan que llevar a cabo un intenso intercambio de información entre sus diferentes oficinas, socios de negocio y autoridades gubernamentales.

Conviene tomar en consideración que los socios de negocios también pueden distribuir sus operaciones en diferentes países, y que es necesario cumplir con las leyes y la normativa vigente en cada uno de ellos. Es necesario tener en

cuenta que las empresas generan, procesan e intercambian en tiempo real grandes volúmenes de información. Por lo cual conviene recordar que existen personas y entidades que, por razones económicas, de competencia o incluso por diversión, se dedican a encontrar y explotar las vulnerabilidades en las infraestructuras y sistemas de las compañías. Esto hace que la responsabilidad de las empresas por salvaguardar su información, así como la recibida de otras empresas y entidades estatales las obligue a tomar las medidas necesarias para proteger su operación de penalizaciones legales y económicas.

En este escenario han surgido diversas estrategias, métodos y técnicas que buscan salvaguardar la información y los sistemas, o al menos minimizar el impacto de los ataques de individuos y organizaciones, tanto internas como externas. Entre los esfuerzos que se han realizado en esa dirección surgieron los lenguajes de intercambio de información o inteligencia, conocido en inglés como "Intel Sharing".

"Intel sharing" corresponde a una política establecida para intercambiar información de forma segura, donde las empresas o entidades lo utilizan para prevenir ataques de personas u organizaciones que se dedican al robo de datos. Cabe tener en cuenta que en el año 1947 se creó la Ley de Seguridad Nacional de la Agencia Central de Inteligencia en los Estados Unidos, y que después de los ataques del 11 de setiembre de 2001 se trabajó en mejorar dicha ley. Como resultado, en el 2002 se hizo una modificación estructural, se creó una red en todo el sistema federal estadounidense para mitigar cualquier amenaza; después, en 2004 se creó la ley de Reforma de Inteligencia y Prevención del Terrorismo que es una ley modificada de la Ley de Seguridad Nacional .(Harknett & Stever, 2011). Dicha ley es la base que originó el desarrollo de lenguajes para intercambio de inteligencia.

En este contexto, para intercambiar información de forma segura se ha definido el uso de lenguajes estandarizados con el fin de que personas o entidades externas no comprometan información sensible de carácter confidencial.

Un lenguaje de intercambio de inteligencia, al igual que al de programación corresponde a que codifica la información o instrucciones utilizando un conjunto de símbolos, así como reglas sintácticas y semánticas (Barber & Ferrís, 2004). Entonces una vez que la información ha sido codificada, es enviada las distintas oficinas, socios de negocio y autoridades gubernamentales. De acuerdo con lo anterior, en el presente trabajo de investigación se realiza una revisión sobre qué es y cómo funciona Intel Sharing, así como posibles usos que se le están dando tanto en el país como de forma internacional. De acuerdo con lo mencionado anteriormente, los objetivos de esta investigación son:

1. Definir y analizar los diferentes tipos de lenguajes de intercambio de inteligencia.
2. Determinar la utilidad e importancia de cada lenguaje de "Intel Sharing".
3. Identificar las relaciones entre los distintos lenguajes de intercambio de inteligencia.
4. Determinar la forma como se puede aplicar este tipo de lenguajes de intercambio de inteligencia en nuestro entorno.

2 Antecedentes

En la actualidad existen muchos tipos de “malware”, es decir software malicioso. Estos han ganado terreno desde su aparición en 1971 y hasta la fecha (Durán, 2010-2011). Otro tipo de amenazas son la distribución masiva de correos conocida como “SPAM” y la ingeniería social (Castellanos, 2011).

De acuerdo con diversos informes, estas amenazas causan grandes perjuicios las empresas mediante el robo de su propiedad intelectual, lo cual ocasiona grandes pérdidas económicas. Además, según muchos gerentes de empresas y Ceos en general, el 2013 fue uno de los años con más violaciones de seguridad sus sistemas (Paganini, 2013). Esto sucede como consecuencia de la disminución en los niveles de seguridad de las empresas y organizaciones

En este orden de ideas, un factor que es considerado como una de las principales causas de los problemas de seguridad es la falta de cultura sobre seguridad de los colaboradores de las compañías. Debido a este y otros aspectos, que en su mayoría son factores internos de las organizaciones, estas han puesto en marcha planes de evaluación de los riesgos de seguridad, incluyendo el conocimiento del personal, para poder efectuar las correcciones necesarias y aplicar las buenas practicas que dictan los estándares y organizaciones especializadas.

Cabe indicar que con el paso del tiempo, la detección y prevención de malware se ha vuelto más eficaz y eficiente. Sin embargo, para mitigar esta amenaza se requiere de una gran variedad de métodos, técnicas y herramientas de análisis y detección. Entre las herramientas que se han implementado con ese objetivo, por ejemplo, un gran número son los antivirus .(SecuritySupervisor.com., 2014).

Asimismo, con el avance en la detección de malware se ha mejorado el intercambio de información entre organismos y entidades. En contextos de este intercambio de información de carácter sensible entre entidades financieras, gubernamentales y de inteligencia, los lenguajes de intercambio de inteligencia son utilizados de forma común. Una de las organizaciones encargadas de la investigación e implementación de estos lenguajes es MITRE, cuyos aportes son significativos y relevantes en esta área.

Al respecto esta empresa, arriba mencionada, utiliza varios tipos de lenguajes que son aptos para compartir información sobre vulnerabilidades y hallazgos, y hasta lograr que organizaciones puedan compartir información segura sin que su información sea capturada por terceros por medio de “sniffers” o métodos de detección (Symantec, 2010).

3 Dessarrollo

Entre algunos de los leguajes utilizados para compartir información podemos encontrar MAEC, STIX, TAXII y CybOX. Cada uno de estos lenguajes tiene diferentes funciones, pero cumpliendo siempre el mismo objetivo, mantener la información sensible segura y lejos de los ataques.

Malware Attribute Enumeration and Characterization (MAEC^{TM}): Es un lenguaje estructurado para la decodificación y transmisión de información. Diseñado

por analistas de malware, con el fin de dar soluciones de captura de información completas, detalladas y eficientes. Tiene también entre sus objetivos eliminar la ambigüedad y la imprecisión que existe actualmente en las descripciones de malware, logrando así reducir la dependencia de las firmas de autenticación . (Martin, 2014b). Entre las mayores virtudes de MAEC podemos citar:

- Eliminación de la imprecisión en las descripciones de malware
- Reducción de la duplicación de los esfuerzos de análisis de malware
- Mejora de la conciencia general de malware
- Disminución del tiempo de respuesta global a las amenazas de malware

The Structured Threat Information eXpression (STIX™): Es un lenguaje que describe o detalla de una manera estandarizada y estructurada, sobre la amenaza cibernética. Stix incluye indicadores de actividad maliciosa (por ejemplo, las direcciones IP), así como información contextual adicional de las amenazas (por ejemplo, tácticas, técnicas y procedimientos [TTP]; objetivos de explotación; campañas y líneas de acción [COA]) que en conjunto caracterizan por completo las motivaciones, capacidades y actividades del ataque; y por lo tanto, como defenderse de mejor forma contra ellos. El lenguaje de STIX está diseñado para poder capturar una amplia gama de datos de información de ciberseguridad relacionada con la amenaza a la información básica del Malware, alguna de la información capturada con este lenguaje es:

- Indicadores: es un patrón observable de la actividad del adversario, relevante en el dominio cibernético operativo junto con información contextual acerca de la interpretación; por ejemplo, este dominio ha sido comprometido, este correo electrónico es falso, este archivo hash se asocia con un troyano, etc.
- TTP (técnica, tácticas y procedimientos): es una representación de la conducta o modus operandi de un atacante, que incluye el uso de determinados patrones de ataque, malware, exploits, herramientas, infraestructura, o la focalización de las víctimas en particular.

Trusted Automated eXchange of Indicator Information (TAXII™): Define un conjunto de servicios e intercambios de mensajes que permiten la compartición de información sobre las amenazas cibernéticas accionables a través dela organización y los límites del producto o servicio. Taxii, a través de sus especificaciones, define conceptos, protocolos, y los intercambios de mensajes para la detección, prevención y mitigación de las amenazas cibernéticas. Taxii no es una iniciativa o aplicación específica para compartir o solicitar información; por el contrario, permite a las organizaciones lograr un mejor conocimiento dela situación sobre las amenazas emergentes, permitiéndole a estas compartir la información que elijan con sus socios.

Cyber Observable eXpression (CybOX™): Es un lenguaje estandarizado, sin embargo este no está dirigido a casos de ciberseguridad, sino más bien está orientado a ofrecer soluciones cibernéticas y que sean lo suficientemente flexibles para los usuarios y permitir que estas se puedan compartir. Entre algunos de los tipos de información que es capturada por este lenguaje se encuentra:

- Características de malware
- Gestión de eventos operacionales
- Ejemplos cibernéticos
- Claves de registros

Puede observarse, entonces, cómo los lenguajes de intercambio de inteligencia juegan un papel determinante, al convertirse en herramientas que no solo sirven como el medio para proteger la información y hacer de su transferencia un método seguro y confiable, sino que también interesan para establecer una estructura y es fiable para el intercambio de información como tal. Otro punto importante a acotar sobre los lenguajes de intercambio, es que entre ellos interactúan y pueden estar presentes de forma embebida, lo cual permite que se complementen uno del otro, con lo cual se dan mejores resultados y brindan una mejor y más detallada información sobre software maliciosos.

Seguidamente se mencionan algunos tipos de información que es capturada por estos lenguajes:

Información sobre el Malware:

- Capacidades
- Comportamientos
- Acciones
- Relaciones
- Tipo nombre
- Descripción

Dichos lenguajes son muy aplicados en el entorno cotidiano; un ejemplo claro, en Costa Rica ha logrado tener éxito de intercambio de inteligencia multilateral con las autoridades de México y Colombia, para descubrir una red de tráfico de drogas. Las autoridades confirman que la cooperación regional está dando frutos, la detención del capo mexicano Edgar Valdez Villareal el 30 de agosto de2010, luego se capturaron a 11 personas más por parte de las autoridades colombianas, por sospecha de conexiones con la Fuerzas Armadas Revolucionarias de Colombia (FARC). (“Intelligence-Sharing Success To be Limited.”, 2010).

“Intel Sharing” además de poder intercambiar información entre entidades, también puede utilizarse entre dispositivos como celulares, Tablet, etc., proporcionando un ambiente más seguro para transmitir datos, de esta forma se evita el robo de información personal. Sin embargo, “Intel Sharing” es aún un tema muy innovador para muchas personas y entidades, por lo que se debe de investigar a fondo y realizar pruebas para poder obtener el mejor resultado, y así poder salvaguardar nuestra información.

4 Discusión: comparación de los lenguajes de intercambio de inteligencia

Si bien es cierto los lenguajes de intercambio de inteligencia son distintos entre sí, estos tienen la facilidad de relacionarse uno con otro y según sean las necesidades.

Entre este tipo de relaciones se puede citar la que existe entre los lenguajes STIX y CybOX, este último utiliza al primero para poder describir e importar de Forma originaria las características del sistema y eventos, utilizando como propias sus caracterizaciones. A su vez, TAXII va a utilizar la información de STIX para representar la información estandarizada y estructurada de una amenaza cibernética.

Otro lenguaje que tiene la facilidad de relacionarse es MAEC, este va a fusionarse con CybOX, juntos van a poder obtener las operaciones, descripciones, características del sistema, eventos, conductas en el dominio y observables cibernéticos relacionados con el malware. MAEC al relacionarse con CybOX va hacer uso de campos de objetos y de acción de CybOX. Por su parte CybOX utiliza patrones de generalización de contenidos en los cuales va a permitir que los usuarios y desarrolladores puedan caracterizar una serie de atributos en conjunto, mediante los patrones observables.

	Definición	Objetivo	Enfoques	Funciones
CybOX (Corporation, 2014)	Lenguaje estructurado para la especificación, captura, caracterización y comunicación de eventos.	Se pretende que sea lo suficientemente flexible como para ofrecer una solución común para todos los casos de uso de seguridad cibernética.	Dirigido a apoyar una amplia gama de dominios de seguridad cibernética, incluyendo: — Evaluación de amenazas y caracterización — Caracterización de malware — Gestión de eventos Operacional — Registro — Ciber conocimiento de la situación — Respuesta de Incidentes	— Analizar datos de eventos de diverso conjunto de sensores de diferentes tipos y diferentes proveedores. — Detectar la actividad de la utilización de patrones de ataque maliciosos. — Detectar la actividad maliciosa utilizando caracterizaciones de comportamiento de malware. — Identificar nuevos patrones de ataque. — Potenciar y orientar la gestión de incidentes utilizando patrones de ataque y caracterizaciones de malware.

	Definición	Objetivo	Enfoques	Funciones
MAEC (Chase & Kirillov, 2013)	Lenguaje estandarizado para el intercambio de información sobre malwares. Se basa en atributos como comportamientos y patrones de ataque.	Su objetivo es eliminar la ambigüedad y la imprecisión en las descripciones de malware, reducir la duplicación de esfuerzos de análisis y mejorar el conocimiento general de malware.	Permitir descripciones de malware completas, a través de diversos componentes de un paquete, para ser utilizados como firmas e indicadores en la empresa.	— Intercambio de información estructurada sobre el malware • Gramática • Vocabulario • Formato — Enfoque en atributos y comportamientos. — Activación de correlación, integración y automatización — Operación — Análisis: • Ayuda proceso guía • Herramienta estandarizada • Malware Repositorios
STIX (Corporation, 2013)	Es un lenguaje para describir la información de la amenaza cibernética de una manera estandarizada y estructurada	Defensa de las redes o sistemas contra las amenazas cibernéticas. Además de proporcionar un lenguaje común para describir la información y de esta forma poder compartir, almacenar y utilizar el contenido adquirido.	Stix representa un esfuerzo de colaboración impulsado por la comunidad, para definir y desarrollar un lenguaje estructurado que permita representar la información sobre las amenazas cibernéticas.	— Análisis de amenazas cibernéticas. — Especificación de indicadores de patrones para las ciberamenazas. — Gestión de amenazas cibernética y actividades de prevención y respuesta. — Intercambio de información de la amenaza cibernética.

	Definición	Objetivo	Enfoques	Funciones
TAXII (Mark Davidson, 2013)	Es un esfuerzo impulsado por la comunidad para estandarizar el intercambio confiable y automatizado de la información.	Define conceptos, protocolos y mensajes para intercambiar información sobre la amenaza cibernética para la detección, prevención y mitigación de las amenazas cibernéticas.	TAXII permite a las organizaciones tener un mejor panorama de la situación de las amenazas emergentes, y permite a estos entes compartir fácilmente la información elegida con sus socios, y con esto, aprovechar las relaciones y los sistemas existentes.	— Hub and spoke: una organización actúa como centro de intercambio (hub) para todos los participantes (spokes). — Source/subscriber: una organización actúa como única fuente de información para todos los suscriptores. — Peer to peer: cualquier organización actúa como productor y consumidor de información.

Table 1: Tabla de Caracteristicas.

5 Caso de uso: análisis de malware usando MAEC

A continuación se ilustrará cómo se hace un análisis utilizando el lenguaje MAEC. Como se ha mencionado anteriormente, MAEC se utiliza para codificar los datos de malware. Con estos análisis se pretende brindar informes, ya que son útiles para poder determinar la naturaleza de una instancia de malware.

Los procedimientos de TRIAGE brindarán información detallada sobre los malware, tal como lo son encabezados de correos electrónicos, phishing o URL. Pero con los informes de análisis de malware se pueden obtener hasta nombres y direcciones IP ya que estos hacen una análisis más profundo.

Los análisis de malware utilizan dos métodos, los estáticos y los dinámicos. Ambos métodos son fundamentales para poder entender cómo funcionan los malware internamente. Los análisis estáticos van a usar las capturas particulares que se extraen de las instancias de malware. Mientras que los análisis dinámicos hacen capturas de conductas particulares como código binario malicioso. Este proceso se hace mediante niveles de abstracción, comenzando desde el más bajo hasta llegar al más alto.

Es importante recalcar que MAEC puede capturar la información de los análisis estáticos y dinámicos por separado. En el ejemplo siguiente se muestra un diagrama de análisis, en este escenario se utilizaron paquetes y el esquema MAEC. (Martin, 2014c)

5.1 MAEC Esquema

- Nivel 3 - MAEC Container: es un contenedor de datos MAEC, de alto nivel que funciona como transporte de uno o varios paquetes.
- Nivel 2 - MAEC Package: hace una agrupación abarcando toda la información de análisis que debe estar relacionada a un malware, como los capturados en el nivel 1 (MAEC Bundle), si se encuentra más de un malware en el paquete que se está examinando, este nivel los agrupa.
- Nivel 1- MAEC Bundle: captura todas las características de un malware, como acciones y objetos relacionados. Es un contenedor independiente para los datos de salida que fueron capturados.

Ejemplo: Para empezar vamos a detallar el primer paso que es el MAEC Bundle.

1. Se crea el MAEC Bundle

   En esta parte es donde se van a capturar todos los datos del malware, y estos van a ser ubicados en cada uno de los atributos correspondientes. En esta primera etapa de creación se le asigna un identificador y nombre único al a cada malware para identificarlos.

2. Static Analysis Output

   Cuando el Malware ingresa al lenguaje, este pasa por el "Static Aalysis Output" donde se van a hacer una extracción automática de los archivos binarios del malware. Seguidamente, se continúa con el paso 2 donde se va a crear un MAEC Object donde las características obtenidas en el paso 1 van a ser capturadas en uno o más objetos. Luego se concluye con el paso 3 donde los objetos creados van a ser agregados al MAEC Bundle.

3. Dynamic Analysis Output

   En este paso el MAEC se propone llenar el espacio "Action", donde el malware va a ser ejecutado en un entorno limitado "sandbox". Luego de eso, se va a generar un informe de ejecución de todos los análisis sometidos y resultados obtenidos. Después de haber obtenido esos datos, se procede a crear cada acción, las cuales van a tener archivos creados y claves de registros, todo esto lo hace la herramienta "sandbox". Por último, se procede a buscar algún tipo de asociación de los "objects" con las "actions", si es así, se añade al MAEC Bundle; hay algunas acciones que no se puede descifrar.

4. Análisis Manual

   En esta parte interactúa un analista humano, quien va a profundizar más en el código del malware y poder descifrar las acciones que no se pudieron determinar mediante la herramienta. Una vez descifradas las acciones, se extraen y se añaden al MAEC Bundle en el espacio de comportamiento (behaviors), luego definirá a qué tipo de acción corresponde cada comportamiento; de esta forma se crearán los comportamientos que definen cada acción y se añade a los objetos relacionados.(Martin, 2014a)

En la Figura 1 muestra un framework...

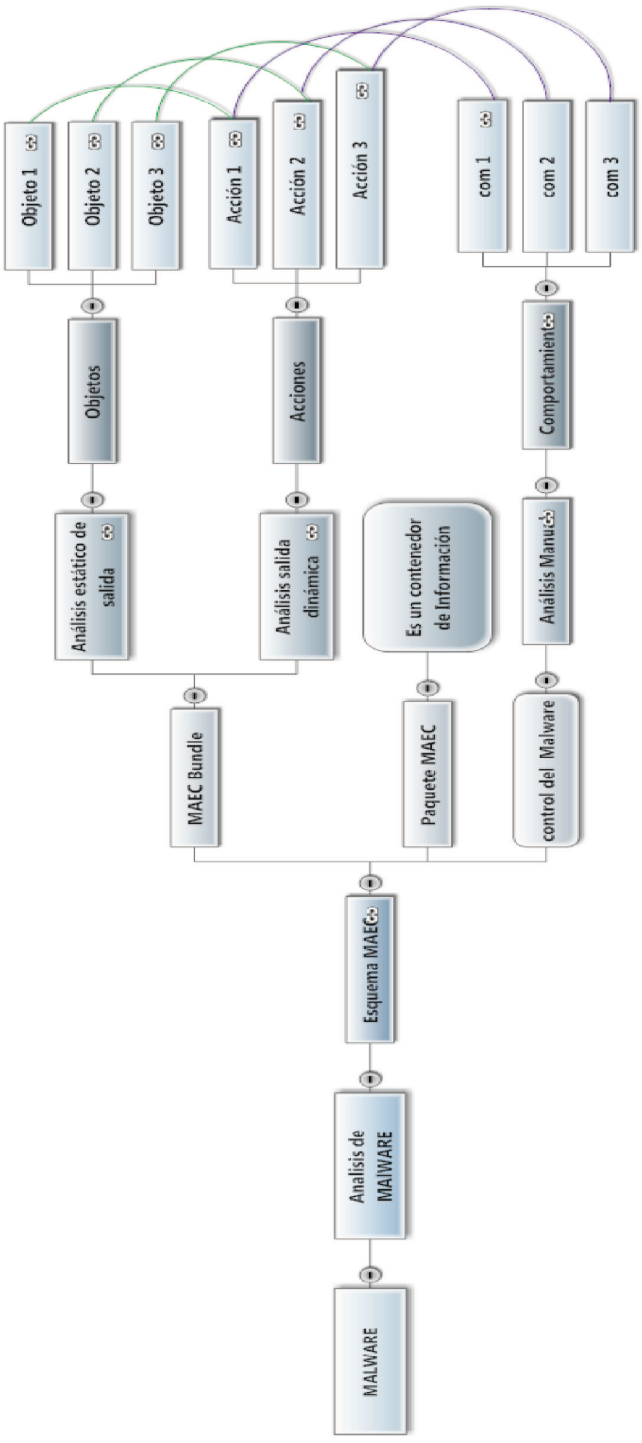

Fig. 1. Lenguajes de intercambio de inteligencia.

6 Conclusiones

La creación de las políticas para intercambio de inteligencia ha sido un importante avance en materia de seguridad cibernética, al establecer una estructura para permitir el intercambio información de forma segura y confidencial; asimismo, contribuyó con el surgimiento de importantes herramientas, los lenguajes de intercambio de inteligencia, los cuales permiten la prevención de ataques provenientes personas u organizaciones dedicadas a la delincuencia cibernética.

En esta investigación se definieron y analizaron los diferentes lenguajes de intercambio involucrados en este tema, se determinaron sus funciones y utilidades, sus distintos enfoques y servicios; de esta forma, se identificaron las relaciones entre ellos. Como resultado no se puede concluir cuál es el mejor o el más robusto, pero es posible indicar y destacar sus características y funcionalidades específicas, de forma que se puedan conocer las herramientas para salvaguardar la información. Aunque no existe un elemento que permita determinar cuál lenguaje es mejor, en términos comparativos debido a la diversidad de sus funciones, sí se puede afirmar con certeza que todos tienen y cumplen un objetivo similar: mantener la información sensible segura y lejos de los ataques.

References

Barber, F., & Ferrís, R. (2004, 05). Tema 2: Lenguajes de programación. *Universitat de València*, 9-19. Retrieved from http://informatica.uv.es/iiguia/AED/oldwww/2004_05/AED.Tema.02.pdf pages 2

Castellanos, E. J. S. (2011). ingeniería social: Corrompiendo la mente humana. *seguridad cultura de prevencion para TI 2011*. pages 3

Chase, P., & Kirillov, I. (2013). Malware attribute enumeration and characterization (maec™). *MITRE*. pages 8

Corporation, T. M. (2013). Structured threat information expression — stix™ a structured language for cyber threat intelligence information. *MITRE*. pages 8

Corporation, T. M. (2014). Cyber observable expression — cybox™. *A Structured Language for Cyber Observables*. pages 7

Durán, F. (2010-2011). Seguridad informática en la empresa. *seguridadinformatica*. Retrieved from http://seguridadinformati.ca/articulos/malware pages 3

Harknett, R. J., & Stever, J. A. (2011). The struggle to reform intelligence after 9/11. *Public Administration Review*, *71*(5), 700–706. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6210.2011.02409.x/full pages 2

Intelligence-sharing success to be limited. (2010). *Latin America Monitor: Central America Monitor*, *27*(11), 8. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=54354452&lang=es&site=ehost-live pages 5

Mark Davidson, C. S. (2013). Taxii overview. *MITRE*. pages 9

Martin, I. K. P. C. R. (2014a, febrero). *The maec™ language v4.0.1 detailed examples.* 202 Burlington Road Bedford, MA 01730. Retrieved from http://maec.mitre.org/language/MAEC_Detailed_Examples_v4.0.1.pdf pages 11

Martin, I. K. P. C. R. (2014b, enero). *Malware attribute enumeration and characterization.* MITRE corporation. 202 Burlington Road Bedford, MA 01730. Retrieved from http://maec.mitre.org/about/inuse.htmlhttps://maec.mitre.org/about/docs/Introduction_to_MAEC_white_paper.pdf pages 4

Martin, I. K. P. C. R. (2014c, abril 18). *Use cases maec.* 202 Burlington Road Bedford, MA 01730. Retrieved from http://maec.mitre.org/language/usecases.html pages 10

Paganini, P. (2013). Data breaches: All you need to know. *violaciones de datos*. Retrieved from http://resources.infosecinstitute.com/2013-data-breaches-need-know/ pages 3

SecuritySupervisor.com. (2014). Top 10 antivirus review. *Antivirus*. Retrieved from http://www.securitysupervisor.com/security-articles/antivirus-reviews/52-top-10-antivirus-review pages 3

Symantec. (2010). Sniffers: qué son y cómo protegerse. *symantec*. Retrieved from `http://www.symantec.com/connect/articles/sniffers-what-they-are-and-how-protect-yourself` pages 3