Tilsor S.A.

Universidad de la República

Facultad de Ingeniería

Tesis de grado

Advanced Threats Information Sharing and Collaboration

Autor:
Julio Saráchaga

Supervisor:
Dr. Gustavo Betarte

Contraparte del cliente:
Ing. Fernanda Molina

Supervisor alterno:
Ing. Marcelo Rodríguez

20 de mayo de 2013

Capítulo 1

RTIR

1.1. Que es RTIR

RTIR es un sistema de manejo de incidentes diseñado para ser utilizado por los equipos de seguridad de sistemas. A sido creado en conjunto con equipos de CERT y CSIRT para manejar el creciente número de incidentes reportados. Presenta la ventaja de ser opensource, contener una API completa y una comunidad de usuarios grande y experta. Además es simple de integrar con otras herramientas existentes. Está implementado por medio de módulos PERL y las herramientas que provee RT. Se puede pensar en RTIR como una extensión de RT para ser utilizada por CERT's y CSIRT's.

Existen algunas alternativas a RTIR como lo son AIRT (Application for Incident Response Teams) cuya última versión data de Julio de 2009. AIRT es una aplicación web desarrollada para los equipos de respuesta a incidentes. Busca proveer facilidad ante los reportes de incidentes de seguridad así como un seguimiento simple de estos. Este sistema no cuenta con una comunidad comparable a la de RTIR así como con documentación tan extensa como la de RTIR.

Otra de las opciones existentes es OTRS (Open Technology Real Services), así como los anteriores también es open soruce. Presenta las ventajas de tener una comunidad más numerosa que AIRT y que el código esta siendo desarrollado continuamente. Así como RTIR esta desarrollado por medio de Perl y permite conectarse a varias bases de datos. Presenta documentación extensa para implementadores.

Welcome to this LaTeX Thesis Template, a beautiful and easy to use template for writing a thesis using the LaTeX typesetting system.

If you are writing a thesis (or will be in the future) and its subject is technical or mathematical (though it doesn't have to be), then creating it in LaTeX is highly recommended

as a way to make sure you can just get down to the essential writing without having to worry over formatting or wasting time arguing with your word processor.

LaTeX is easily able to professionally typeset documents that run to hundreds or thousands of pages long. With simple mark-up commands, it automatically sets out the table of contents, margins, page headers and footers and keeps the formatting consistent and beautiful. One of its main strengths is the way it can easily typeset mathematics, even *heavy* mathematics. Even if those equations are the most horribly twisted and most difficult mathematical problems that can only be solved on a super-computer, you can at least count on LaTeX to make them look stunning.

1.2. Learning LaTeX

LaTeX is not a WYSIWYG (What You See is What You Get) program, unlike word processors such as Microsoft Word or Apple's Pages. Instead, a document written for LaTeX is actually a simple, plain text file that contains *no formatting*. You tell LaTeX how you want the formatting in the finished document by writing in simple commands amongst the text, for example, if I want to use *italic text for emphasis*, I write the '`\textit{}`' command and put the text I want in italics in between the curly braces. This means that LaTeX is a "mark-up" language, very much like HTML.

1.2.1. A (not so short) Introduction to LaTeX

If you are new to LaTeX, there is a very good eBook – freely available online as a PDF file – called, "The Not So Short Introduction to LaTeX". The book's title is typically shortened to just "lshort". You can download the latest version (as it is occasionally updated) from here:
http://www.ctan.org/tex-archive/info/lshort/english/lshort.pdf

It is also available in several other languages. Find yours from the list on this page:
http://www.ctan.org/tex-archive/info/lshort/

It is recommended to take a little time out to learn how to use LaTeX by creating several, small 'test' documents. Making the effort now means you're not stuck learning the system when what you *really* need to be doing is writing your thesis.

1.2.2. A Short Math Guide for LaTeX

If you are writing a technical or mathematical thesis, then you may want to read the document by the AMS (American Mathematical Society) called, "A Short Math Guide for LaTeX". It can be found online here:
http://www.ams.org/tex/amslatex.html
under the "Additional Documentation" section towards the bottom of the page.

1.2.3. Common LaTeX Math Symbols

There are a multitude of mathematical symbols available for LaTeX and it would take a great effort to learn the commands for them all. The most common ones you are likely to use are shown on this page:
http://www.sunilpatel.co.uk/latexsymbols.html

You can use this page as a reference or crib sheet, the symbols are rendered as large, high quality images so you can quickly find the LaTeX command for the symbol you need.

1.2.4. LaTeX on a Mac

The LaTeX package is available for many systems including Windows, Linux and Mac OS X. The package for OS X is called MacTeX and it contains all the applications you need – bundled together and pre-customised – for a fully working LaTeX environment and workflow.

MacTeX includes a dedicated LaTeX IDE (Integrated Development Environment) called "TeXShop" for writing your '`.tex`' files and "BibDesk": a program to manage your references and create your bibliography section just as easily as managing songs and creating playlists in iTunes.

1.3. Getting Started with this Template

If you are familiar with LaTeX, then you can familiarise yourself with the contents of the Zip file and the directory structure and then place your own information into the '`Thesis.cls`' file. Section 1.5 on page 7 tells you how to do this. Make sure you read section 1.7 about thesis conventions to get the most out of this template and then get started with the '`Thesis.tex`' file straightaway.

If you are new to LaTeX it is recommended that you carry on reading through the rest of the information in this document.

1.3.1. About this Template

This LaTeX Thesis Template is originally based and created around a LaTeX style file created by Steve R. Gunn from the University of Southampton (UK), department of Electronics and Computer Science. You can find his original thesis style file at his site, here:
http://www.ecs.soton.ac.uk/~srg/softwaretools/document/templates/

My thesis originally used the '`ecsthesis.cls`' from his list of styles. However, I knew LaTeX could still format better. To get the look I wanted, I modified his style and also created a skeleton framework and folder structure to place the thesis files in.

This Thesis Template consists of that modified style, the framework and the folder structure. All the work that has gone into the preparation and groundwork means that all you have to bother about is the writing.

Before you begin using this template you should ensure that its style complies with the thesis style guidelines imposed by your institution. In most cases this template style and layout will be suitable. If it is not, it may only require a small change to bring the template in line with your institution's recommendations.

1.4. What this Template Includes

1.4.1. Folders

This template comes as a single Zip file that expands out to many files and folders. The folder names are mostly self-explanatory:

Appendices – this is the folder where you put the appendices. Each appendix should go into its own separate '`.tex`' file. A template is included in the directory.

Chapters – this is the folder where you put the thesis chapters. A thesis usually has about seven chapters, though there is no hard rule on this. Each chapter should go in its own separate '`.tex`' file and they usually are split as:

- Chapter 1: Introduction to the thesis topic
- Chapter 2: Background information and theory

- Chapter 3: (Laboratory) experimental setup
- Chapter 4: Details of experiment 1
- Chapter 5: Details of experiment 2
- Chapter 6: Discussion of the experimental results
- Chapter 7: Conclusion and future directions

This chapter layout is specialised for the experimental sciences.

Figures – this folder contains all figures for the thesis. These are the final images that will go into the thesis document.

Primitives – this is the folder that contains scraps, particularly because one final image in the 'Figures' folder may be made from many separate images and photos, these source images go here. This keeps the intermediate files separate from the final thesis figures.

1.4.2. Files

Included are also several files, most of them are plain text and you can see their contents in a text editor. Luckily, many of them are auxiliary files created by LaTeX or BibTeX and which you don't need to bother about:

Bibliography.bib – this is an important file that contains all the bibliographic information and references that you will be citing in the thesis for use with BibTeX. You can write it manually, but there are reference manager programs available that will create and manage it for you. Bibliographies in LaTeX are a large subject and you may need to read about BibTeX before starting with this.

Thesis.cls – this is an important file. It is the style file that tells LaTeX how to format the thesis. You will also need to open this file in a text editor and fill in your own information (such as name, department, institution). Luckily, this is not too difficult and is explained in section 1.5 on page 7.

Thesis.pdf – this is your beautifully typeset thesis (in the PDF file format) created by LaTeX.

Thesis.tex – this is an important file. This is the file that you tell LaTeX to compile to produce your thesis as a PDF file. It contains the framework and constructs that tell LaTeX how to layout the thesis. It is heavily commented so you can read exactly what each line of code does and why it is there. After you put your own information into the '`Thesis.cls`' file, go to this file and begin filling it in – you have now started your thesis!

vector.sty – this is a LaTeX package, it tells LaTeX how to typeset mathematical vectors. Using this package is very easy and you can read the documentation on the site (you just need to look at the '`vector.pdf`' file):
http://www.ctan.org/tex-archive/macros/latex/contrib/vector/

lstpatch.sty – this is a LaTeX package required by this LaTeX template and is included as not all TeX distributions have it installed by default. You do not need to modify this file.

Files that are *not* included, but are created by LaTeX as auxiliary files include:

Thesis.aux – this is an auxiliary file generated by LaTeX, if it is deleted LaTeX simply regenerates it when you run the main '`.tex`' file.

Thesis.bbl – this is an auxiliary file generated by BibTeX, if it is deleted, BibTeX simply regenerates it when you run the main tex file. Whereas the '`.bib`' file contains all the references you have, this '`.bbl`' file contains the references you have actually cited in the thesis and is used to build the bibliography section of the thesis.

Thesis.blg – this is an auxiliary file generated by BibTeX, if it is deleted BibTeX simply regenerates it when you run the main '`.tex`' file.

Thesis.lof – this is an auxiliary file generated by LaTeX, if it is deleted LaTeX simply regenerates it when you run the main '`.tex`' file. It tells LaTeX how to build the 'List of Figures' section.

Thesis.log – this is an auxiliary file generated by LaTeX, if it is deleted LaTeX simply regenerates it when you run the main '`.tex`' file. It contains messages from LaTeX, if you receive errors and warnings from LaTeX, they will be in this '`.log`' file.

Thesis.lot – this is an auxiliary file generated by LaTeX, if it is deleted LaTeX simply regenerates it when you run the main '`.tex`' file. It tells LaTeX how to build the 'List of Tables' section.

Thesis.out – this is an auxiliary file generated by LaTeX, if it is deleted LaTeX simply regenerates it when you run the main '`.tex`' file.

So from this long list, only the files with the '`.sty`', '`.bib`', '`.cls`' and '`.tex`' extensions are the most important ones. The other auxiliary files can be ignored or deleted as LaTeX and BibTeX will regenerate them.

1.5. Filling in the 'Thesis.cls' File

You will need to personalise the thesis template and make it your own by filling in your own information. This is done by editing the 'Thesis.cls' file in a text editor.

Open the file and scroll down, past all the '\newcommand...' items until you see the entries for 'University Name', 'Department Name', etc....

Fill out the information about your group and institution and ensure you keep to block capitals where it asks you to. You can also insert web links, if you do, make sure you use the full URL, including the 'http://' for this.

The last item you should need to fill in is the Faculty Name (in block capitals). When you have done this, save the file and recompile 'Thesis.tex'. All the information you filled in should now be in the PDF, complete with web links. You can now begin your thesis proper!

1.6. The 'Thesis.tex' File Explained

The Thesis.tex file contains the structure of the thesis. There are plenty of written comments that explain what pages, sections and formatting the LaTeX code is creating. Initially there seems to be a lot of LaTeX code, but this is all formatting, and it has all been taken care of so you don't have to do it.

Begin by checking that your information on the title page is correct. For the thesis declaration, your institution may insist on something different than the text given. If this is the case, just replace what you see with what is required.

Then comes a page which contains a funny quote. You can put your own, or quote your favourite scientist, author, person, etc... Make sure to put the name of the person who you took the quote from.

Next comes the acknowledgements. On this page, write about all the people who you wish to thank (not forgetting parents, partners and your advisor/supervisor).

The contents pages, list of figures and tables are all taken care of for you and do not need to be manually created or edited. The next set of pages are optional and can be deleted since they are for a more technical thesis: insert a list of abbreviations you have used in the thesis, then a list of the physical constants and numbers you refer to and finally, a list of mathematical symbols used in any formulae. Making the effort to fill these tables means the reader has a one-stop place to refer to instead of searching the

internet and references to try and find out what you meant by certain abbreviations or symbols.

The list of symbols is split into the Roman and Greek alphabets. Whereas the abbreviations and symbols ought to be listed in alphabetical order (and this is *not* done automatically for you) the list of physical constants should be grouped into similar themes.

The next page contains a one line dedication. Who will you dedicate your thesis to?

Finally, there is the section where the chapters are included. Uncomment the lines (delete the '`%`' character) as you write the chapters. Each chapter should be written in its own file and put into the 'Chapters' folder and named '`Chapter1`', '`Chapter2`, etc... Similarly for the appendices, uncomment the lines as you need them. Each appendix should go into its own file and placed in the 'Appendices' folder.

After the preamble, chapters and appendices finally comes the bibliography. The bibliography style (called '`unsrtnat`') is used for the bibliography and is a fully featured style that will even include links to where the referenced paper can be found online. Do not under estimate how grateful you reader will be to find that a reference to a paper is just a click away. Of course, this relies on you putting the URL information into the BibTeX file in the first place.

1.7. Thesis Features and Conventions

To get the best out of this template, there are a few conventions that you may want to follow.

One of the most important (and most difficult) things to keep track of in such a long document as a thesis is consistency. Using certain conventions and ways of doing things (such as using a Todo list) makes the job easier. Of course, all of these are optional and you can adopt your own method.

1.7.1. Printing Format

This thesis template is designed for single sided printing as most theses are printed and bound this way. This means that the left margin is always wider than the right (for binding). Four out of five people will now judge the margins by eye and think, "I never noticed that before.".

The headers for the pages contain the page number on the right side (so it is easy to flick through to the page you want) and the chapter name on the left side.

The text is set to 11 point and a line spacing of 1.3. Generally, it is much more readable to have a smaller text size and wider gap between the lines than it is to have a larger text size and smaller gap. Again, you can tune the text size and spacing should you want or need to. The text size can be set in the options for the '`\documentclass`' command at the top of the '`Thesis.tex`' file and the spacing can be changed by setting a different value in the '`\setstretch`' commands (scattered throughout the '`Thesis.tex`' file).

1.7.2. Using US Letter Paper

The paper size used in the template is A4, which is a common – if not standard – size in Europe. If you are using this thesis template elsewhere and particularly in the United States, then you may have to change the A4 paper size to the US Letter size. Unfortunately, this is not as simple as replacing instances of '`a4paper`' with '`letterpaper`'.

This is because the final PDF file is created directly from the LaTeX source using a program called '`pdfTeX`' and in certain conditions, paper size commands are ignored and all documents are created with the paper size set to the size stated in the configuration file for pdfTeX (called '`pdftex.cfg`').

What needs to be done is to change the paper size in the configuration file for `pdfTeX` to reflect the letter size. There is an excellent tutorial on how to do this here:
http://www.physics.wm.edu/~norman/latexhints/pdf_papersize.html

It may be sufficient just to replace the dimensions of the A4 paper size with the US Letter size in the `pdftex.cfg` file. Due to the differences in the paper size, the resulting margins may be different to what you like or require (as it is common for Institutions to dictate certain margin sizes). If this is the case, then the margin sizes can be tweaked by opening up the `Thesis.cls` file and searching for the line beginning with, '`\setmarginsrb`' (not very far down from the top), there you will see the margins specified. Simply change those values to what you need (or what looks good) and save. Now your document should be set up for US Letter paper size with suitable margins.

1.7.3. References

The '`natbib`' package is used to format the bibliography and inserts references such as this one [**?**]. The options used in the '`Thesis.tex`' file mean that the references are listed in numerical order as they appear in the text. Multiple references are rearranged

in numerical order (e.g. [1**?**]) and multiple, sequential references become reformatted to a reference range (e.g. [1**? ?**]). This is done automatically for you. To see how you use references, have a look at the '`Chapter1.tex`' source file. Many reference managers allow you to simply drag the reference into the document as you type.

Scientific references should come *before* the punctuation mark if there is one (such as a comma or period). The same goes for footnotes[1]. You can change this but the most important thing is to keep the convention consistent throughout the thesis. Footnotes themselves should be full, descriptive sentences (beginning with a capital letter and ending with a full stop).

To see how LaTeX typesets the bibliography, have a look at the very end of this document (or just click on the reference number links).

1.7.4. Figures

There will hopefully be many figures in your thesis (that should be placed in the 'Figures' folder). The way to insert figures into your thesis is to use a code template like this:

```
\begin{figure}[htbp]
  \centering
    \includegraphics{./Figures/Electron.pdf}
    \rule{35em}{0.5pt}
  \caption[An Electron]{An electron (artist's impression).}
  \label{fig:Electron}
\end{figure}
```

Also look in the source file. Putting this code into the source file produces the picture of the electron that you can see in the figure below.

Sometimes figures don't always appear where you write them in the source. The placement depends on how much space there is on the page for the figure. Sometimes there is not enough room to fit a figure directly where it should go (in relation to the text) and so LaTeX puts it at the top of the next page. Positioning figures is the job of LaTeX and so you should only worry about making them look good!

Figures usually should have labels just in case you need to refer to them (such as in Figure 1.1). The '`\caption`' command contains two parts, the first part, inside the square brackets is the title that will appear in the 'List of Figures', and so should be short. The

[1] Such as this footnote, here down at the bottom of the page.

FIGURA 1.1: An electron (artist's impression).

second part in the curly brackets should contain the longer and more descriptive caption text.

The '`\rule`' command is optional and simply puts an aesthetic horizontal line below the image. If you do this for one image, do it for all of them.

The LaTeX Thesis Template is able to use figures that are either in the PDF or JPEG file format.

1.7.5. Typesetting mathematics

If your thesis is going to contain heavy mathematical content, be sure that LaTeX will make it look beautiful, even though it won't be able to solve the equations for you.

The "Not So Short Introduction to LaTeX" (available here) should tell you everything you need to know for most cases of typesetting mathematics. If you need more information, a much more thorough mathematical guide is available from the AMS called, "A Short Math Guide to LaTeX" and can be downloaded from:
`ftp://ftp.ams.org/pub/tex/doc/amsmath/short-math-guide.pdf`

There are many different LaTeX symbols to remember, luckily you can find the most common symbols here. You can use the web page as a quick reference or crib sheet and because the symbols are grouped and rendered as high quality images (each with a downloadable PDF), finding the symbol you need is quick and easy.

You can write an equation, which is automatically given an equation number by LaTeX like this:

```
\begin{equation}
E = mc^{2}
  \label{eqn:Einstein}
\end{equation}
```

This will produce Einstein's famous energy-matter equivalence equation:

$$E = mc^2 \tag{1.1}$$

All equations you write (which are not in the middle of paragraph text) are automatically given equation numbers by LaTeX. If you don't want a particular equation numbered, just put the command, '`\nonumber`' immediately after the equation.

1.8. Sectioning and Subsectioning

You should break your thesis up into nice, bite-sized sections and subsections. LaTeX automatically builds a table of Contents by looking at all the '`\chapter{}`', '`\section{}`' and '`\subsection{}`' commands you write in the source.

The table of Contents should only list the sections to three (3) levels. A '`\chapter{}`' is level one (1). A '`\section{}`' is level two (2) and so a '`\subsection{}`' is level three (3). In your thesis it is likely that you will even use a '`\subsubsection{}`', which is level four (4). Adding all these will create an unnecessarily cluttered table of Contents and so you should use the '`\subsubsection*{}`' command instead (note the asterisk). The asterisk (*) tells LaTeX to omit listing the subsubsection in the Contents, keeping it clean and tidy.

1.9. In Closing

You have reached the end of this mini-guide. You can now rename or overwrite this pdf file and begin writing your own '`Chapter1.tex`' and the rest of your thesis. The easy

work of setting up the structure and framework has been taken care of for you. It's now your job to fill it out!

Good luck and have lots of fun!

Guide written by —
Sunil Patel: www.sunilpatel.co.uk

Capítulo 2

TAXII

2.1. Introducción

A lo largo de la historia de la computación se ha visto un aumento en la sofisticación, velocidad e impacto de los ataques informáticos. Se ha vuelto necesario que las estrategias de defensa se adapten a los nuevos actores y ataques. El intercambio de información se ha tornado critico contra los adversarios del presente. Actualmente un número creciente de organizaciones comparten información de amenazas con el fin de tener un visión más amplia de las actividades de los adversarios y así ayudar a administrar los recursos de las organizaciones para obtener el mejor resultado posible de sus defensas. El problema que se presenta hoy en día es que compartir información es una tarea manual que lleva mucho tiempo ó un proceso automatizado con grandes limitaciones el cual sea realiza dentro de una comunidad. La capacidad de compartir información de forma automática con un gran número de comunidades y que dicha información sea completa no existe en la actualidad. El objetivo que plantea TAXII es extender la capacidad de compartir indicadores permitiendo intercambios robustos, seguros y de gran volumen de datos los cuales deben poseer información mas expresiva de amenazas informáticas. TAXII es un conjunto de especificaciones técnicas y de documentación para permitir el intercambio de información procesable entre organizaciones. Para ello se definen protocolos y formatos de datos para intercambiar información de forma segura la cual ayude a detectar, prevenir y mitigar amenazas informáticas en tiempo real. No se buscan definir acuerdos para el intercambio de información, gobierno o aspectos técnicos del intercambio de información. En su lugar, permite a las organizaciones alcanzar una mejora en su situación respecto a las nuevas amenazas y además compartir información que ellos elijan con quien elijan de forma simple y rápida aprovechando las relaciones y sistemas existentes. En el desarrollo de TAXII se busco consenso y participación de la comunidad. TAXII

permite intercambio de información sobre amenazas de forma eficiente y comprensiva por medio de *automatización* y *articulación* de un modelo detallado de información. Para lograr esto, se utiliza una representación estándar de información de amenazas y un framework para soportar el intercambio de datos. El modelo permite el envío y recepción de un conjunto amplio de información de seguridad para soportar un amplio número de necesidades referentes al intercambio de información. TAXII cubre un amplio número de casos de uso, tecnologías, especificaciones e implementaciones. Los casos de uso son desarrollados de manera secuencial, permitiendo un conjunto inicial de casos de usos que permite el intercambio e información. TAXII utiliza protocolos y especificaciones existentes siempre que es posible y se integra con mecanismos de intercambio de información existentes para reducir los costos de implementación y permitir la adopción rápida por parte de organizaciones ya establecidas que ya intercambian información.

2.1.1. Fundamentos

Las estrategias de defensa se deben adaptar al creciente número, frecuencia y complejidad de los ataques que se llevan a cabo. Actualmente la estrategia predominante se refiere a bloquear los ataques y arreglar las vulnerabilidades, esta estrategia se basa en alertas. Si bien puede ser efectiva contra algunas amenazas no logra detener ataques avanzados o proveer información sobre las actividades de un atacante luego de que la red fue penetrada. Una estrategia más adecuada es *çyber kill-chain"* la cual se representa a continuación.

ACA va la figura 1 de White paper de taxii

La estrategia presentada busca descomponer las fases de un ataque con la finalidad de obtener una amplia comprensión del ataque y el atacante así como mejorar las posibilidades de defensa.

Aca pongo la tabla 1 del White paper

Los primeros pasos de esta estrategia representan una oportunidad para detectar y mitigar las amenazas de forma proactiva antes de que el adversario realice un acceso no autorizado en los sistemas de la organización. En los pasos posteriores es donde se realiza la detección, respuesta y aseguramiento de los activos más importantes. Al entender al adversario los defensores tienen un mejor oportunidad para descubrir y responder al ataque. Actualmente se busca que las defensas anticipen y mitiguen las amenazas antes de que sean mas difíciles de encontrar y erradicar utilizando los métodos tradicionales de detección y respuesta. Para poder realizar esto es necesario que se realice una

actividad de cyber inteligencia, recolecte información referente a ataques, con esta información analistas pueden agrupar patrones de actividades similares, atribuir actividades a ciertos actores, identificar e implementar estrategias de mitigación de forma rápida y anticiparse al lanzamiento de ataques similares en el futuro. Para aprovechar de forma más adecuada los beneficios de la cyber inteligencia, las organizaciones deben compartir la información recolectada (incluyendo las estrategias de defensa entre otras) con socios de su confianza. De esta forma se obtiene una imagen mas completa de las actividades del adversario y de las acciones defensivas que se deben realizar. [1] Por medio del análisis del comportamiento de los adversarios en distintos objetivos y en un periodo de tiempo, los defensores son capaces de identificar un conjunto importante de indicadores y tácticas, técnicas y procedimientos (TTPs). De esta forma se obtiene información de los objetivos y las estrategias lo cual permite al defensor predecir el comportamiento del ataque y generar defensas dinámicas.

2.1.2. Comunidades

Hoy en día un número creciente de organizaciones buscan compartir información de las amenazas. Con esto ha crecido el número y tipo de las comunidades que buscan compartir información. Se pueden encontrar tres tipos de comunidades :

- Peer
- Comerciales
- Gobierno

La necesidad primordial entre dichas organizaciones es la confianza dado que compartir información sensible podría exponer a una organización a un daño en su reputación, demandas o advertir a un adversario con lo cual el trabajo realizado fuera inútil. Se deben definir medidas para la protección de los datos como restricciones en el manejo de los datos, sanitización de los datos y el establecimiento de confianza entre las dos partes. Esto es particularmente importante cuando las organizaciones forman parte de varias comunidades para el intercambio de amenazas de seguridad. Se puede ver que lo que es compartido con una comunidad no necesariamente debería ser compartido con otra. Las comunidades entre peers son las más comunes, en estas organizaciones o individuos con un propósito común se unen para mejorar las defensas colectivas contra adversarios comunes o conjuntos de adversarios. Las comunidades comerciales se basan en membresias por parte de los miembros y son altamente anónimas. La organización comercial maneja de forma centralizada la información y la distribuye entre los miembros de la organización. Estas organizaciones proveen una forma rápida de obtener información,

además puede ser más amplia que la información especializada que es compartida por pares y puede que no siempre sea aplicable a las necesidades de una organización. Las comunidades gubernamentales son establecidas y manejadas por el gobierno, son voluntarias u obligatorias e incluyen participantes tanto del gobierno como de la industria privada. En ellas el gobierno controla la información y la diseminación de esta, cabe señalar que así como en las comunidades comerciales la información y los participantes son altamente confidenciales.

2.1.3. Modelos

Hay tres modelos principales para el intercambio de información entre organizaciones:

- hub and spoke
- peer to peer
- source/subscriber

En el método hub and spoke, una entidad controla la recepción y la diseminación de los datos. La entidad hub usualmente realiza una anonimiza los datos recolectado de las amenazas y provee una análisis adicional a los participantes. Este modelo es comúnmente visto en comunidades de gobierno o comerciales. En el modelo peer to peer, los participantes intercambian y reciben información directamente de los otros participantes. La información es compartida entre todos los miembros de la comunidad por igual y la fuente esta claramente identificada. El modelo faltante es el de source/suscriber, este modelo es utilizado por las comunidades comerciales que proveen de información. El proveedor de información envía regularmente información a todos los suscriptores y estos podrían eventualmente enviarle información a la fuente. Usualmente, en este modelo la información esta codificada de una manera propietaria y puede faltar información esencial sobre algunas intentos de irrupción. Presenta la ventaja de que se tiene acceso rápido a un conjunto de datos amplio y es útil para organizaciones con recursos limitados.

2.1.4. Métodos para el intercambio de información

Hay múltiples métodos para el intercambio de información. Usualmente, el método juega un rol significante en los tipos, volúmenes y naturaleza de la información compartida con la comunidad. Algunos medios de intercambio limitan el tipo de contenido que es compartido de forma sencilla mientras que otros promueven ciertos tipos de intercambio. Algunos métodos comunes son:

- Email lista de servidores
- Foros de discusión
- wikis
- repositorios de datos

La mayoría de estos métodos no permiten el consumo de información de amenazas de forma automática. La mayoría de los consumidores rutinariamente toman esta información y la sintetizan en sus bases de datos locales. Existen varios esfuerzos para generar arquitecturas abiertas, estándar basados en indicadores e información de incidentes. La realidad es que ninguno de estos a podido convertirse en un estándar para el intercambio entre comunidades.

2.1.5. Información compartida

Actualmente los indicadores que se comparten son sistemas malignos o actividades de red o cyber observables de interes como las direcciones de IP, nombres de dominio, nombres de archivos o direcciones de email. En algunos casos, la información compartida esta enfocada en una amenaza en particular, como las botnets. En otros casos se incluye malware en uso u otros TTP. La información compartida esta establecida generalmente por la comunidad o por el grado de confianza entre las partes.

Limitaciones Compartir información ha ayudado a mejorar las capacidades defensivas de numerosas organizaciones. Sin embargo, las aproximaciones actuales no han logrado que se llegue al máximo potencial. Los procesos para compartir información son manuales, llevan mucho tiempo, son repetitivos y en muchos casos requieren que las organizaciones re escriban o traduzcan la información a una amplia variedad de formatos. La información es además compartida por medios inseguros. Debido a la variedad de formatos y los protocolos en uso, así como los procesos manuales involucrados, esta técnica se lleva a cabo entre pocas organizaciones en las que se confía. Otro factor que hace que hace ineficiente el intercambio de información ineficiente, menos escalable y que consuma mucho tiempo es el uso de tecnologías y/o formatos propietarios, presentandose la necesidad de desarrollar un amplio número de scripts y módulos para permitir compartir información por fuera de las comunidades. Para aquellas comunidades con algún grado de automatización, sus modelos son generalmente bajos en prestaciones y usan soluciones propietarias, comerciales o adaptadas a su comunidad. Otra limitación se presenta en la naturaleza atómica de la mayoría de los indicadores. Por ejemplo, cuando una IP en particular es identificada como sospechosa el esfuerzo que debe realizar el adversario

para cambiar la IP es prácticamente cero. Confiar únicamente en indicadores atómicos sin contexto puede proveer un gran número de falsos positivos llevando a un desperdicio en tiempo de análisis.

Motivación Una mejor solución para compartir información es necesaria, una que sea utilizada por diferentes comunidades y modelos para compartir información, permitiendo diferentes métodos para compartir, y que soporte un rango amplio de datos. En particular, los objetivos de la solución ideal son:

- Permitir el poder compartir información de forma más rápida y precisa
- Reducir el análisis humano y liberar a los recursos humanos para realizar trabajo de análisis más valioso.
- Mover las amenazas más conocidas para que sean analizadas por computadoras.
- Permitir que se comparta de forma automática un gran rango de datos, siendo estos datos complejos y no los simples, atómicos indicadores. Esto debería permitir una defensa activa.
- Proteger la información intercambiada.
- Permitir que se agregue información a las bases locales pero contexto y discreción, pero con un menor número de analistas que vean las información.
- Permitir la colaboración de analistas de distintas organizaciones en los incidentes que sean un reto.

2.1.6. Que es TAXII

TAXII es un conjunto de especificaciones técnicas y documentación para el intercambio de información de alta fidelidad, dicho intercambio es independiente de la plataforma y realizado de forma segura. Esta diseñado de forma que permita la interoperabilidad de diferentes soluciones en lugar de ligarse a una tecnología o producto en particular. Además se busca incentivar a los proveedores de tecnología a incorporar soporte para las especificaciones de TAXII en sus productos. Ha sido desarrollado con consenso y participación de la comunidad, con la finalidad de permitir un intercambio de eficiente y comprensivo de la información detallada, de forma automática y articulada. Para lograr esto, TAXII utiliza una representación estándar de la información y define un framework para soportar el intercambio. TAXII ofrece una forma de describir e intercambiar los indicadores, dejando a los proveedores la libertad de determinar como sus productos

producen, consumen o toman ventaja de los flujos de información especificados por TAXII.

Imagen 2

2.1.7. Objetivos de TAXII

Los objetivos de TAXII son:

- Permitir el intercambio seguro y rápido de información referente a amenazas entre comunidades de defensores de seguridad.
- Lograr un standard para permitir compartir indicadores entre otros elementos entre organizaciones.
- Extender el intercambio de indicadores para permitir intercambios seguros, robustos y de gran volumen que tengan una expresividad mayor a la actual.
- Soportar un amplio número de casos de uso y practicas comunes a las comunidades.
- Tomar los estandares existentes que sean adecuados.
- Llegar a una adopción por parte de organizaciones internacionales de standars.

TAXII no ha creado una comunidad para compartir, sino que permite que las comunidades compartan. TAXII mejora las deficiencias existentes proveyendo especificaciones abiertas y comunes para transportar los mensajes con información con capacidades como encriptación, autenticación, direccionamiento, alertas y pedidos entre sistemas.

2.1.8. Representación estándar de la información

TAXII utiliza el lenguaje STIX para representar la información. STIX es un lenguaje desarrollado por la comunidad para la especificación, captura, caracterización y comunicación de información de amenazas cibernéticas de forma estandarizada. STIX provee una arquitecutra unificada que soporta varios tipos de información entre los cuales se incluyen Cyber Observables, Indicadores, Incidentes, tacticas, técnicas y procedimientos de los adversarios, etc.

Para maximizar la compatibilidad y facilidad de adopción, STIX utiliza varios standards como CybOX, Common vulnerabilities and exposures (CVE) y common platform enumeration (CPE).

2.1.9. Un framework de Intercambio

La segunda parte necesaria en automatizar el intercambio de información es especificar como esta es compartida. Para alcanzar esto, TAXII define especificaciones técnicas y documentación de soporte. En particular, las especificaciones de TAXII definen un conjunto de capacidades necesarias para el transporte exitoso de mensajes, o como los mensajes TAXII llegan del punto A al B. Los mensajes TAXII llevan datos de amenazas informáticas transformadas a formato STIX. El conjunto completo de los mensajes incluye mensajes con datos y de control. TAXII utiliza protocolos y especificaciones existentes siempre que es posible y los integra con los mecanismos actuales para reducir los costos de implementación y permitir una adopción rápida por parte de las organizaciones ya establecidas que ya comparten información. TAXII esta siendo desarrollado de forma modular para soportar una variedad de mecanismos y formatos de datos para ser intercambiados.

2.1.10. Casos de Uso

TAXII ha sido desarrollado para soportar casos de uso comunes para el intercambio de información.

Alertas o Advertencias publicas Estas son advertencias al publico en general o a varios asistentes de varios CSIRT, estas son enviadas a todos los suscriptores. Estas alertas son de una naturaleza tan amplia que no necesitas ser encriptadas o no se necesitan autorizaciones. Sin embargo es importante una firma digital para asegurar la autenticidad. Las entidades u organizaciones deben ser especificadas para identificar la fuente de la alerta.

Alertas y Reportes privados Las alertas privadas son similares a las publicas, exceptuando que la información compartida es sensible y restringida a los socios que comparten datos. Los mecanismos para enviar datos deberían ser similares a los de las alertas publicas. Como las alertas y reportes se suponen sensibles y no para el uso general, es importante que TAXII soporte formas adecuadas de encriptacion, autenticación, autorización y identificación de datos. Dependiendo en la naturaleza de las comunicaciones, manejo explícito de marcas o restricciones en datos compartidos debería ser necesario. Las alertas son generalmente cortas, mensajes estándar con indicadores muy específicos o acciones especificadas. Los reportes son mensajes mas largos, y pueden incluir reportes de incidentes, análisis de malware, análisis de amenazas u otras observaciones.

Soportes para Queries Es común que los analistas de amenazas busquen información de otros entre o por fuera de sus comunidades.

- RFI (Request for Information): es un mensaje simple que se espera sea manejado de forma manual y que permite que se pida información.
- Repository Search: Para este tipo de query, es esperado que una organización ofrezca repositorios en los que buscar, los cuales podrían ser compatibles con TAXII o STIX.

Transferencia Varias organizaciones que transfieren información necesitan en algunas instancias agregar miembros. El nuevo miembro necesitas obtener los datos del repositorio de la organización. Por ello es esperada la transferencia de un gran volumen de datos.

2.2. Componentes de TAXII

- Especificación de TAXII: Define la especificación de los componentes y provee guía y requerimientos sobre como dichas especificaciones interoperan en TAXII.
- Especificación de servicios de TAXII: Define una serie de servicios que deben ser implementados para ser compatible con TAXII. Describe información intercambiada a un nivel alto y no se limita a ningún mecanismo de intercambio en especial.
- Implementación de Servicios: Se realiza una implementación de los servicios TAXII para un mecanismo de intercambio. Cada implementación de servicios provee guía técnica y requerimientos para implementar la especificación de los mecanismos de intercambio.
- Modelo de datos de mensajes: Se define una estructura para los mensajes TAXII, incluyendo header, payload, control y mensajes de datos. Los mensajes de datos utilizan STIX para el payload de los mensajes TAXII.
- Implementaciones de los mensajes de datos: Es una implementación del modelo de datos de mensajes, incluyendo el payload STIX, a un formato en particular. Cada implementación de mensajes define la guía técnica y requerimientos para utilizar un formato de mensajes particular para expresar el modelo de datos de mensajes.

2.2.1. TAXII Toolkit

Es provisto para soportar la adopción de TAXII y asistir en el desarrollo de capacidades compatibles. El toolkit provee una colección de implementaciones de referencia, un conjunto de herramientas y una colección de librerías e interfaces.

Especificaciones de TAXII

TAXII esta definido por múltiples especificaciones relacionadas. Esta sección describe las especificaciones definidas en TAXII.

- Especificación de Servicios (Service Specification): Provee los requerimientos por los cuales se definen los servicios e intercambios de TAXII. No provee detalles respecto al formato de los datos o como los mensajes TAXII son transportados por la red. Dichos detalles y requerimientos pueden ser encontrados en Protocol Binding Specification y Message Binding Specification.
- Especificación de protocolos de enlace (Protocol Binding Specification): Define los requerimientos para transportar mensajes TAXII por la red. Puede haber varias especificaciones creadas para TAXII. Cada especificación define requerimientos para transportar mensajes TAXII usando protocolos de red y se proveen requerimientos respecto a como los servicios TAXII son soportados por los protocolos de red.
- Especificación de Mensajes (Message Binding Specification): Se definen requerimientos para representar mensajes TAXII en un formato particular. Puede haber múltiples especificaciones para dichos mensajes. Se provee información detallada sobre como la información definida en los especificación de servicios es expresada en los mensajes.

ACA VA LA FIGURA 1 DEL DOCUMENTO DE SERVICE SPECIFICATION

Separación de la especificación de servicios, los protocolos de enlace y los mensajes existe para dar flexibilidad mientras TAXII evoluciona. Debido a que las organizaciones generalmente tienen restricciones respecto a los protocolos que soportan, TAXII busca no ligarse a un único protocolo que excluya a una parte de la comunidad. Cuando se ve que la comunidad expresa interés en un nuevo protocolo o tipo de mensaje, TAXII puede dar soporte para ellos sin cambiar los componentes centrales. Dos grupos que usen el mismo protocolo de red y formato de mensajes serán capaces de intercambios de información estructurada de forma automática. Las políticas de intercambio de los participantes puede limitar estos intercambios si es necesario, pero el uso de servicios compatibles con TAXII asegura que se puede intercambiar cualquier información con los mecanismos definidos por TAXII. Los grupos que usen diferentes protocolos o formatos de mensajes no serán capaces de comunicarse directamente, pero como están utilizando mensajes y servicios en el núcleo de las comunicaciones de sus comunidades significa que es posible establecer caminos para que ocurra la interacción.

2.2.2. Especificación de Servicios

Esta especificación provee normativas respecto a los servicios, mensajes e intercambio de mensajes en TAXII. No provee detalles respecto a como los mensajes son transportados, dejando eso a la especificación de los protocolos de enlace. Si bien se provee información relacionada a la información presente en los mensajes TAXII, no da información respecto a como los mensajes TAXII son expresados. *Versión de los servicios TAXII* En la especificación de los servicios se hace referencia a "version IDs", especificamente TAXII services version Id, TAXII protocol binding version ID y TAXII message binding version ID. Los protocolos de red que transportan mensajes TAXII así como los mensajes en si es necesario en algunos casos que indiquen el número de versión de TAXII así como de la especificación de mensajes o de protocolos de transporte. Los string de version id representan la versión de las especificaciones TAXII utilizadas en los intercambios TAXII. Cada especificación de TAXII está identificada por su propia version id. Diferentes versiones de cada especificación proveerán diferentes version ids. La especificación actual tiene un version id: $TAXII_1$,0

2.2.3. Terminios y definiciones

Conceptos utilizados en TAXII

- Cyber Threat information: Es cualquier información representable en STIX. Esto incluye, pero no esta limitado a, Observables, Indicadores, incidentes, TTPs (Tactics, Techniques y Procedures), exploit targets, campaigns, threat actors y courses of action.
- TAXII Data Feed: Es una colección de cyber threat information estructurada expresable en uno o mas documentos STIX que pueden ser intercambiados utilizando TAXII. Cada TAXII Data Feed *debe* tener un nombre que lo identifica de forma única entre el resto de los feeds de un productor dado. Cada elemento de un TAXII data feed debe ser etiquetado con un timestamp y puede tener otras etiquetas a discreción del productor.
- Mensae TAXII: Un bloque de información que es pasado de una entidad a la otra. Un mensaje TAXII representa un pedido o una respuesta.
- Intercambio de mensajes TAXII: Una secuencia definida de mensajes TAXII intercambiados entre dos entidades.
- Servicio TAXII: Son funcionalidades albergadas por algunas entidades y que es accedido o invocado usando uno o mas TAXII Message Exchange.

- TAXII Capability: Una actividad de alto nivel soportado por TAXII por medio del uso de uno o mas servicios TAXII.

Unidades funcionales de TAXII Las unidades funcionales de TAXII representan conjuntos discretos de actividades requeridas para soportar TAXII. Una unidad funcional representa algún componente con un rol bien definido en TAXII.

- TAXII Transfer Agent (TTA) : Es una unidad funcional conectada a la red que envía o recibe mensajes TAXII. Una TTA interactua con otras TTAs por medio de la red y maneja los detalles de los requerimientos del protocolo de uno o más TAXII Protocol Binding Specifications. Una TTA provee mensajes TAXII a un TAXII Message Handler permitiendo que este último sea independiente del protocolo de red utilizado. De la misma forma, el TTA puede ser independiente del contenido de los mensajes TAXII, dejando el manejo de la información al TAXII Message Handler.
- TAXII Message Handler (TMH): Es una unidad funcional que produce y consume mensajes TAXII. El TMH es responsable de parsear y construir mensajes con el formato especificado en uno o mas TAXII Message Binding Specifications. Un TMH interactua con un TTA, el cual maneja los detalles necesarios para transmitir mensajes por la red. El Back-end TAXII interactua con el TMH para convertir su contenido en mensajes TAXII, y llevar a cabo actividades basadas en los mensajes TAXII que son recibidos por el TMH.
- TAXII Back-end: Cubre todas las unidades funcionales distintas al TTA y al TMH. Las especificaciones de TAXII no proveen requerimientos sobre como son implementadas las capacidades en un back-end mas allá de como debe interactuar con el TMH. Las organizaciones o implementadores pueden decidir que capacidades implementar según los servicios TAXII que deseen soportar o según como quieren dar ese soporte.
- Arquitectura TAXII: Cubre los aspectos de las unidades funcionales de la infraestructura de productor o consumidor que prove o utiliza servicios TAXII. Una arquitectura TAXII incluye una TTA, un TMH y un back-end TAXII.

  IMAGEN DE LA ARQUITECTURA FIG 2 DE DOCS

Roles en TAXII Los roles en TAXII son utilizados de acuerdo al uso que se hace de los servicios especificados en TAXII.

- Productor es el rol de una entidad que es la fuente de información estructurada de amenazas.

- Consumidor es el rol de la entidad que recibe información estructurada sobre amenazas.

Componentes de red Los siguientes términos son utilizados para definir componentes de una implementación TAXII utilizando un modelo cliente-servidor.

- Una implementación TAXII es un implementación especifica de una arquitectura TAXII.
- Servidor TAXII Es una implementación que provee uno o mas servicios TAXII. Para soportar esta funcionalidad, se asume que un servidor TAXII esta continuamente esperando trafico de red.
- Cliente TAXII Es una implementación TAXII que inicia intercambio con un servidor TAXII. Un cliente no necesita una conexión persistente con internet para operar pero puede abrir conexiones cuando desea interactuar con un servidor y desconectarse de internet cuando la conexión a terminado.
- TAXII endpoint denota una implementación TAXII que puede ser un servidor o un cliente

2.2.4. Capacidades

La existencia de TAXII provee capacidades especificas para aquellos que desean compartir información de amenazas cibernéticas. Las capacidades TAXII son el nivel mas alto en el cual se pueden expresar las acciones de TAXII. Hay tres capacidades que soporta la actual versión de TAXII: push messaging, pull messaging y discovery. *Push Messaging* La información puede ser enviada de un productor a un consumidor. Esto puede reflejar una relación pre-existente entre el productor y el consumidor en la que el consumidor a pedido que se le envíen datos desde el productor. También puede usarse en caso de que el consumidor desee aceptar contribuciones de cualquier productor, y estos le envien datos en cualquier momento.

Pull Messaging Un consumidor puede requerir información de un productor. Esto no solo le permite al consumidor el control sobre el momento en el que recibe los datos sino que también le permite hacerlo sin tener que aceptar conexiones entrantes. Así como en push messaging, el productor y consumidos pueden tener acuerdos pre-existentes para que el consumidor tenga acceso a los datos del productor. De forma alternativa, un productor puede hacer su información pública y cualquer consumidor puede requerir sus datos. La versión actual de pull messaging, limita a los consumidores a hacer pedidos por medio

de las organizaciones productoras de los datos en lugar de por los datos en si. Todos los datos provistos por un productor deben estar organizados en grupos llamados "TAXII Data Feeds". Piezas individuales de información en un TAXII Data Feed son etiquetadas utilizando timestamps. El productor tiene total discresión sobre como el contenido se mapea en TAXII Data Feeds y en el significado de los timestamps. La capacidad de pull messaging esta atada a entender el contenido del productor.

Discovery Para facilitar las comunicaciones automatizadas, TAXII soporta capacidades para descubrir los servicios específicos que ofrece un servidor o grupo de servidores, así como los protocolos o mensajes que este servidor ofrece. Esto no quita la necesidad del involucramiento humano para establecer acuerdos de cooperación lo cual esta por fuera del objetivo de TAXII. Sin embargo permite el intercambio de información respecto a las capacidades que un productor pudiera soportar y cuales son los mecanismos que utiliza para hacerlo.

2.2.5. Servicios TAXII

Los servicios TAXII representan un conjunto de mecanismos necesarios para soportar capacidades TAXII. Una implementación TAXII pudiera implementar alguno, todos o incluso ninguno de los servicios definidos. TAXII define los siguientes servicios:

- Servicio de descubrimiento: Es utilizado para recibir y responder a mensajes que requieren información sobre los servicios ofrecidos.
- Feed Managment Service: Es utilizado para recibir o responder a mensajes utilizados para el manejo de subscripciones a TAXII Data Feed.
- Inbox Service: Es utilizado para recibir información de amenazas cibernéticas por medio de intercambios iniciados por el productor en intervalos dictados por este.
- Poll Service: Es utilizado para recibir y responder a mensajes de pedido a el TAXII Data Feed iniciados por el consumidor.

En las siguientes subsecciones se describen los distintos servicios. *Discovery Service* Es un mecanismo para comunicar información referente al uso de servicios TAXII y a su disponibilidad. Para un pedido al servicio, se retorna una lista de los servicios TAXII y como estos pueden ser invocados. Un solo servicio de descubrimiento puede reportar servicios TAXII en diferentes equipos finales o incluso en múltiples organizaciones, los propietarios del servicio pueden definir su alcance a su gusto. Un servicio de descubrimiento puede utilizar varios factores para determinar cuales servicios revelar ante

una petición, incluyendo pero no limitado a la entidad del cliente TAXII. El servicio de descubrimiento debe soportar "Discovery Message Exchange".

Feed Managment Service Es el mecanismo con el cual un consumidor pide información referente a TAXII Data Feeds, pidiendo subscripciones a estos, o modificando las existentes. Este servicio facilita el intercambio de mensajes para manejar las subscripciones. Este servicio no entrega contenido de los TAXII Data Feed, en su lugar se envia contenido del TAXII Data Feed al servicio de Inbox de un consumidor en intercambios iniciados por un productor o en respuesta directa a un pedido del consumidor al servicio de poll. Dicho servicio debe implementar soporte para subscription managment exchange. Dicho servicio podría implementar soporte de feed information exchange.

Inbox service Este servicio es el mecanismo con el cual un consumidor acepta los mensajes en un intercambio iniciado por el productor. Un consumidor puede implementar este servicio para recibir datos del TAXII Data Feed. El servicio de inbox debe implementar soporte para Data Push Exchange.

Poll service Es provisto por un productor para permitir pedidos al TAXII Data Feed iniciados por el consumidor. Un consumidor contacta a este servicio explícitamente pidiendo el contenido del TAXII Data Feed. Los productores podrían ofrecer Data Feeds combinando envios al Inbox service del consumidor o por medio de pedidos al servicio de poll de productor. Un implementación de este servicio debe dar soporte a Data Poll Exchange.

Capítulo 3

STIX

Chapter 3. *STIX*

3.1. Background

Las aproximaciones tradicionales para la seguridad, se focalizan en entender y registrar las vulnerabildes, debilidades y configuraciones necesarias pero insuficientes. Las defensas efectivas contra las amenazas actuales y futuras también requiere la adición sobre el comportamiento, capacidades e intenciones del atacante. Entendiendo al adversario y a nosotros podemos entender lo suficiente respecto a la naturaleza de las amenazas a las que se enfrenta la organización para tener decisiones para una defensa efectiva. El comportamiento de los adversarios no esta solamente focalizado de forma extendida en actividades destructivas, sino que también se focaliza en objetivos de mas bajo nivel que apuntan a conseguir objetivos tacticos y establecer puntos de entrada a las organizaciones. //ACA SE PUEDE UNIR CON KILL CHAIN QUE ESTA EN TAXII La cyber inteligencia busca entendenr y caracterizas las acciones que un atacante puede o podria realizar, como estas pueden ser detectadas y reconocidas, como pueden ser mitigadas y cuales son los actores relevantes, etc. Un entendimiento de la amenaza que implica el adversario permite la realización de decisiones mas efectivas, la priorización de recursos y llegar a tener una oportunidad de tomar la ventaja ante el adversario. El efecto de las defensas en base a inteligencia es una mejor postura respecto a las respuestas dado que los atacantes ajustan sus operaciones basandose en el exito o falla de sus intentos. En un modelo de kill chain los intentos que realiza un adversario pueden ser reconocidos y logrando asi que los defensores tengan la posiblidad de ajustar sus tacticas para una mejor respuesta. Esto hace que al adversario le sea mas dificil alcanzar sus objetivos. En esto se presenta un desafio dado que ninguna organizacion por si sola tiene acceso a un

conjunto de información relevante para tener una identificación precisa de una amenaza. La forma de sobrepasar esta limitación es por medio del intercambio de información relevante sobre las amenazas con socios y comunidades de confianza. Por medio del intercambio de información, cada ente puede alcanzar un nivel mas alto de entendimiento del panorama de las amenazas, no solamente de forma abstracta sino que también que cosas especificas indican la presencia de una atacante. Dada la forma en la complejidad con la que evoluciona el panorama de amenazas, la velocidad a la cual ocurren los eventos, y la basta cantidad de datos que se deberian intercambiar, es necesario establecer una forma automatica para ayudar a los humanos o para ejecutar acciones defensivas para que esta aproximación sea efectiva. La automatizacion requiere de información de calidad y las mayoria de las capacidades defensivas son construidas con arquitecturas heterogeneas. La combinación de estos factores requiere estandarización, representaciones estructuradas de información y un aprovechamiento de la información sin saber de antemano quien va a proveer que información. Un desafio que tienen las organizaciones a la hora de realizar intercambio de información es el de tener la habilidad de estructurar la información sin perder el juicio y control que puede tener un humano. La información intercambiada tiene que ser leible por un humano y parseable por una maquina. Este requerimiento es utilizado por porgramas de intercambio de información en los cuales las organizaciones no solo consumen los datos sino que también los evaluan como parte del proceso de inteligencia. Este proceso es llevado a cabo por analistas que se focalizan en tipos de analisis que son inapropiados para ser automatizados o focalizados en tomas de decisiones por parte de humanos en donde el analista lee la información para obtener información de la situación. Además, el analista esta constantemente evaluando la fidelidad de las fuentes y los métodos utilizados para producir la información. Dados estos factores, es necesaria la existencia de representaciones estructuradas de la información y que esta información sea expresiva, flexible, extensible, automatizable y que se pueda leer. En este punto es donde surge STIX como una solución al problema presentado.

3.2. Aproximaciones de la actualidad

La información que se intercambia y utiliza hoy en día es atómica, inconsistente y muy limitada en sofisticación y expresividad. En donde se utilizan estructuras estandarizadas, estas son tipicamente focalizadas en una porción del problema. Además no se integran de forma adecuada entre si o carecen de flexibildiad. Las actividades de intercambio de indicadores son entre humanos, siendo los indicadores desestructurados o semi-estructurados intercambiados por portales web o encriptados via email. Más recientemente se ha visto el surgimiento de transferencias entre maquinas de conjutnso de indicadores simples de modelos de ataques bien conocidos. STIX busca extender los indicadores para permitir

el manejo e intercambio de indicadores de forma más expresiva así como de un espectro más amplio de información. Actualmente, el intercambio y manejo de información de forma automatica de información es visto tipicamente en lineas de productos, servicios ofrecidos o soluciones especificas de una comunidad. STIX busca permitir el intercambio de información de forma comprensiva, rica, de alta fidelidad entre organizaciones, comunidades, productos y servicios ofrecidos. STIX es un lenguaje para la especificación, captura, caracterización y comunicación de información de seguridad estandar. Lo hace de forma estructurada para soportar un mejor manejo de la información y la aplicación de automatización. Una variedad de casos de uso para dicha información son realizados.

- Analisis de las amenazas
- Especificación de patrones de indicadores para la amenaza.
- Manejo de las actividades de respuesta
- Intercambiar información

STIX provee un mecanismo común para hacer frente a estos casos de uso dando consistencia, eficiencia, interoperabildiad y conciencia global de la situación.

STIX provee una arquitectura unificada juntando un conjunto amplio de información incluyendo:

- Cyber Observables
- Indicadores
- Incidentes
- Tacticas, tecnicas y procedimientos de los adversarios
- Objetivos de exploits
- Cursos de acción
- Capañas de ataques
- Actores de ataques

Para permitir que dicha solución sea practica para cualquier caso de uso, lenguajes existentes y estandarizados son utilizados.

3.3. Casos de uso

3.3.1. Analisis de amenazas

Un analista revisa información estructurada y no estructurada respecto a actividades de amenazas de una variedad de fuentes de entrada manuales o automaticas. El analista buscan entender la naturaleza de las amenazas relevantes, identificarlas y caracterizarlas totalmente para que todo el conocimiento relevante de la amenaza sea totalmente expresado y evolucionado a traves del tiempo. Este conocimiento relevante incluye acciones relacionadas con la amenaza, comportamientos, capacidades, intenciones, actores atribuidos, etc. Por medio del entendimiento y la caracterización, el analista puede especificar patrones de indicadores relevantes, sugerir acciones para las actividades de respuesta y compartir la información con miembros de la comunidad.

3.3.2. Especificación de patrones de indicadores

Un analista especifica patrones medibles representando las características observables de una amenaza junto con su contexto y metadata para ser interpretadas, manejadas y aplicar el patrón y sus resultados. Esto puede ser realizado de forma manual o con la asistencia de una herramienta automatizada.

3.3.3. Manejo de las actividades de respuesta

Los tomadores de decision y el personal de operaciones trabajan en conjunto para prevenir o detectar actividades amenasantes y responder a los incidentes detectados que sean realizados por dichas amenazas. Las acciones preventivas pueden mitigar vulnerabilidades, debilidades, o malas configuraciones que sean que sean objetivos de los exploits. Luego de la detección e investigación de incidentes específicos, acciones reactivas pueden ser realizadas. *Prevención de amenazas* //REPASAR ESTO Los tomadores de decisiones evaluan acciones preventivas para amenazas relevantes que sean identificadas y seleccionan acciones apropiadas para su implementación. El personal de operaciones implementa las acciones seleccionadas para prevenir la ocurrencia de amenazas especificas aplicando mitigaciones amplias o con objetivos especificos iniciados por intepretación de los indicadores. *Detección de amenazas* //REPASAR ESTO El personal de operaciones aplica mecanismos (automaticos y manuales) para monitorear y asistir las operaciones con la finalidad de detectar la ocurrencia de amenazas especificas basandose en la evidencia historica, analisis del contexto actual, interpretación de los indicadores que se presentan. Esta detección se realiza generalmente por medio de patrones de indicadores.

Respuesta a incidentes El personal de operaciones responde a amenazas detectadas, investiga que ha ocurrido o esta ocurriendo, trata de identificar y caracterizar la naturaleza de las amenazas y lleva a cabo la mitigación o lleva a cabo cursos de acción preventivos. Una vez que los efectos son entendidos, el personal de operaciones puede implementar mitigaciones acordes o llevar a cabo cursos de acción correctivos.

3.3.4. Compartir información de amenazas

Los tomadores de decisión establecen políticas respecto a que tipo de información será compartida, además toman decisiones respecto a con quienes se comparte y como debería ser manejada basandose en frameworks de confianza de forma de mantener niveles adecuados de consistencia, contexto y control. Esta politica es luego implementada para compartir los indicadores de amenazas apropiados y otra información de amenazas.

3.4. Principios de guia

En el enfoque dado a STIX se ha buscado implementar un conjunto de principios con el consenso de la comunidad. Esos principios son los siguientes:

3.4.1. Expresividad

Con el fin de soportar la diversidad de casos de uso relevantes, STIX apunta a proveer una cobertura expresiva en todos sus casos de uso especificos en lugar de dirigirse especificamente a alguno de ellos.

3.4.2. Integración en lugar de duplicación

Cuando STIX abarca conceptos de información estructurada para los cuales ya existen representaciones estandarizadas con el consenso adecuado y que se encuentran disponibles, se busca integrar estas representaciones a la arquitectura STIX en lugar de duplicar esta información innecesariamente.

3.4.3. Flexibildiad

Con el fin de soportar un amplio rango de casos e información variable con varios niveles de fidelidad, STIX esta diseñado para ofrecer tanta flexibilidad como sea posible.

STIX se adhiere a una politica de permitir a los usuarios utilizar cualquier porcion de representaciones estandar que sean relevantes para un contexto dado y evita elementos obligatorios siempre que sea posible.

3.4.4. Extensibildiad

//REVISAR Con la finalidad de soportar un amplio rango de casos de uso con un potencial diferente de representación y para asegurar la facilitar el perfeccionamiento impulsado por la comunidad, se ha diseñado STIX para construir mecanismos de extensión para uso específicos, para usos localizados, para refinamientos del usuario y evolución y para facilidad de refinamiento y evolución centralizada.

3.4.5. Automatización

El diseño de STIX busca estructuras del contenido para que nos ea únicamente consumible o procesable por maquinas sino que también sea leíble por humanos. Esto es necesario para claridad y comprensibilidad durante las primeras etapas de desarrollo y adopción y para el uso sostenido en diversos ambientes,

3.5. Arquitectura

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

3.5.1. Subsection 1

Nunc posuere quam at lectus tristique eu ultrices augue venenatis. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Aliquam erat volutpat. Vivamus sodales tortor eget quam adipiscing in vulputate ante ullamcorper. Sed eros ante, lacinia et sollicitudin et, aliquam sit amet augue. In hac habitasse platea dictumst.

3.5.2. Subsection 2

Morbi rutrum odio eget arcu adipiscing sodales. Aenean et purus a est pulvinar pellentesque. Cras in elit neque, quis varius elit. Phasellus fringilla, nibh eu tempus venenatis, dolor elit posuere quam, quis adipiscing urna leo nec orci. Sed nec nulla auctor odio aliquet consequat. Ut nec nulla in ante ullamcorper aliquam at sed dolor. Phasellus fermentum magna in augue gravida cursus. Cras sed pretium lorem. Pellentesque eget ornare odio. Proin accumsan, massa viverra cursus pharetra, ipsum nisi lobortis velit, a malesuada dolor lorem eu neque.

3.6. Main Section 2

Sed ullamcorper quam eu nisl interdum at interdum enim egestas. Aliquam placerat justo sed lectus lobortis ut porta nisl porttitor. Vestibulum mi dolor, lacinia molestie gravida at, tempus vitae ligula. Donec eget quam sapien, in viverra eros. Donec pellentesque justo a massa fringilla non vestibulum metus vestibulum. Vestibulum in orci quis felis tempor lacinia. Vivamus ornare ultrices facilisis. Ut hendrerit volutpat vulputate. Morbi condimentum venenatis augue, id porta ipsum vulputate in. Curabitur luctus tempus justo. Vestibulum risus lectus, adipiscing nec condimentum quis, condimentum nec nisl. Aliquam dictum sagittis velit sed iaculis. Morbi tristique augue sit amet nulla pulvinar id facilisis ligula mollis. Nam elit libero, tincidunt ut aliquam at, molestie in quam. Aenean rhoncus vehicula hendrerit.

Capítulo 4

STIX

Chapter 4. *IODEF*

4.1. IODEF

Las organizaciones deben colaborar entre ellas para mitigar las actividades maliciosas que atacan sus redes así como para ganar conocimiento de posibles amenazas. Esta coordinación puede requerir coordinación con ISPs, sitios remotos o intercambiar datos con socios. IODEF son las siglas para Incident Object Description Exchange Format, este define una representación de datos que provee un framework para el intercambio de información entre CSIRTs, dicho tipo de información es intercambiado comúnmente por CSIRTs y es referente a incidentes de seguridad. IODEF provee una representación en XML para transportar información de incidentes entre dominios administrativos entre pares que tienen una responsabilidad operacional de remediar o de analizar y establecer advertencias en un dominio definido. El modelo de datos provisto en IODEF codifica la información referente a hosts, redes y servicios corriendo en estos sistemas; metedologia de ataques y evidencia forense asociada; impacto de la actividad realizada; aproximaciones limitadas para documentar el flujo. El objetivo primordial de IODEF es mejorar las capacidades operacionales de los CSIRTs. La adopción por parte de la comunidad provee una habilidad mejorada para resolver los incidentes y transmitir el contexto simplificando la colaboración e intercambio de datos. El formato estructurado provisto por IODEF permite:

- Incrementar la automatización en el procesamiento de los datos, ya que los recursos de los analistas de seguridad de parsear documentos se verá reducido.
- Bajar el esfuerzo necesario para normalizar datos similares de diferentes fuentes.

- Un formato común con el cual construir herramientas interoperables para el manejo de incidentes y análisis subsecuente, específicamente cuando los datos provienen de dominios distintos.

La coordinación entre CSIRTs no es un problema estrictamente técnico. Hay muchas consideraciones: procedimientos, confianza, legales, las cuales pueden ocasionar que las organizaciones intercambien información. IODEF no busca evadir dichas consideraciones, sin embargo, las implementaciones operacionales de IODEF deben considerar este contexto.

4.1.1. Modelo de daos de IODEF

A la hora de diseñar IODEF se realizaron ciertas consideraciones de diseño

- El modelo de datos sirve como formato de transporte. Por ello, su representación especifica no es optima para almacenamiento en disco, archivamiento a largo plazo o procesamiento en memoria.
- Por medio de la implementación no se busca establecer un consenso respecto a la definición de un incidente. Se busca en su lugar un entendimiento amplio que sea lo suficientemente flexible como para abarcar la mayoría de las operaciones.
- Describir un incidente para todas las definiciones requeriria un modelo de datos extremadamente complejo. Por ello, IODEF solo busca dar un marco para transmitir información de incidentes comúnmente intercambiada. Se asegura un mecanismo amplio para ser extendido para soportar información propia de la organización, y técnicas para referenciar información mantenida por fuera del modelo de datos.
- El dominio del análisis de seguridad no esta totalmente estandarizado y debe basarse en descripciones textuales libres. IODEF busca conseguir un balance entre el contenido libre, pero permitiendo el procesamiento automático de la información de los incidentes.
- IODEF es solo una de las representaciones que han sido estandarizadas. El modelo de datos de IDMEF influencio el diseño de IODEF.

4.1.2. Implementaciones de IDMEF

Las implementaciones del protocolo se especifican como un esquema XML. Implementar IODEF en XML provee varias ventajas. Que sea extensible lo hace ideal para especificar

codificación de datos que soporta varias codificaciones de caracteres. Es mas simple la manipulación debido a la presencia de varias tecnologías para ello. Aunque fundamentalmente XML es una representación de texto, lo cual lo hace ineficiente cuando se deben embeber datos binarios o grandes volúmenes de datos deben ser intercambiados.

4.1.3. Internacionalización

Internacionalización y localización son de interés para IODEF, dado que solo por medio de colaboración (a menudo con barreras idiomáticas) son resueltos ciertos incidentes. IODEF soporta esto dependiendo de las construcciones de XML, y por medio de disen explicito en el modelo de datos. Como IODEF es implementado como un esquema XML, este soporta las diferentes codificaciones de caracteres. Además, los documentos IODEF deben especificar el lenguaje en el cual sus contenidos son codificados.

4.1.4. Consideraciones de Seguridad

El modelo de datos de IODEF no introduce problemas de seguridad. Este solo define una representación simple para información de incidentes. Como los datos codificados por IODEF pueden ser considerados sensibles por las partes que los intercambian o por los descritos por los datos, se deben tomar precauciones para asegurar la confidencialidad durante el intercambio y el subsecuente procesamiento. El primero debe ser resguardado por un formato de mensaje, pero luego se deben tener consideraciones de seguridad por el sistema que procesa los datos, los almacena y archiva la documentación y la información derivada de estos. El contenido de un documento IODEF puede incluir un pedido de acción o un parser puede independientemente tener lógica para realizar cierta acción basandose en información que encuentra. Por esta razón, se debe tener cuidado de autenticar apropiadamente el beneficiario del documento y atribuir un nivel de confidencialidad apropiado a los datos antes de realizar la acción. El formato de mensaje subyacente y el protocolo utilizado para intercambiar datos provee una garantía de confidencialidad, integridad y autenticidad. El uso de protocolos de seguridad estandarizados es recomendado. Por ejemplo el uso de IODEF/RID. Con el fin de sugerir buenas practicas en el procesamiento y manejo de los datos codificados, IODEF permite a un emisor de documentos transmitir una política de privacidad utilizando un atributo de restricción. Las distintas instancias de este atributo permite a los distintos elementos del documento tener las mismas políticas. Esto sirve como una guía para el receptor, peor este podría decidir ignorarlo, este problema no es un reto técnico.

4.2. IDMEF

IDMEF es un protocolo experimental que no especifica ningún estándar. Sin embargo IODEF se basa en algunas de las cosas definidas por este para su definición. El propósito de IDMEF es definir formatos de datos y procedimientos de intercambio para compartir información de interés con sistemas de detección de intrusos y sistemas de respuesta con los sistemas de administración que podrían necesitar interactuar con estos. El rfc de IDMEF describe el modelo de datos para representar información tomada de los sistemas de intrusión y explica la razón de usar este modelo. IDMEF busca ser un formato de datos estándar el cual pueda ser utilizado por los IDSs para reportar alertas sobre eventos que parecen sospechosos. Se puede proveer interoperabildiad entre sistemas comerciales, open source y sistemas de investigación, permitir a los usuarios esta mezcla de sistemas ayuda a obtener una implementación óptima de sus sistemas.

4.2.1. Modelo de datos de IDMEF

El modelo de datos de IDMEF es una representación orientada a objetos de los datos del alerta enviados a los administradores de los sistemas de intrusión. El modelo de datos presenta varios problemas referentes a la representación de los datos:

- La información de alertas es mayormente heterogénea. Algunas alertas son definidas con muy poca información, como orígenes, destinos, nombre u hora del evento. Otras alertas proveen mucha mas información, como puertos de servicios, procesos, información de usuarios, etc. El modelo de datos que debe representar dicha información debe ser flexible para adaptarse a las distintas necesidades. Un modelo orientado a objetos es extensible por medio de agregación y sub clases. Si una implementación del modelo de datos extiende con una nueva clase, por medio de agregación o subclases, una implementación que no entienda estas extensiones podrá seguir entendiendo el subconjunto de información que esta definido en el modelo. Estas dos formas de extender el modelo permiten que se mantenga la consistencia del modelo.
- Los IDS son diferentes, algunos analizadores detectan ataques analizando el trafico en la red, otros utilizando los logs de sistemas operativos o aplicaciones que auditan información. Alertas para el mismo ataque enviadas por analizadores con diferentes fuentes de información, no contendran los mismos datos. El modelo de datos define clases que soportan las diferencias en las fuentes de los datos. En particular, las nociones de fuente y objetivo para el alerta son representadas por una combinación de nodo, procesos, servicio y clase de usuario.

- Las capacidades de los analizadores son diferentes. Por ello el modelo de datos debe permitir la conversión de formatos utilizados por herramientas distintas a IDSs con el propósito de mayor procesamiento de la información.
- Los ambientes operacionales son diferentes. Dependiendo en el tipo de red o sistemas operativos utilizados, los ataques serán observados y reportados con diferentes características. El modelo de datos se adapta a estas diferencias.
- Los instrumentos comerciales persiguen objetivos diferentes. Por varias razones, estos desean entregar mas o menos información sobre ciertos tipos de ataques. El modelo orientado a objetos permite la flexibilidad necesaria preservando la integridad del modelo.

4.2.2. El diseño del modelo de datos

Representación de eventos El objetivo del modelo de datos es proveer una representación estándar de la información que el analizador de un sistema de intrusión reporta cuando detecta la ocurrencia de eventos inusuales. Estas alertas pueden ser simples o complejas dependiendo de las capacidades del analizador que la creo.

Bibliografía

[1] D. Rolsky D. Chamberlain J. Vincent, R. Spier and R. Foley. *RT Essentials*. O'Reilly Media, Inc, 105 Gravensein Highway North, Sebastopol, CA 95472, 2005.