Revisión sistemática de literatura: Visualización de Seguridad

Michael Phillips Pereira, Steven Mena Matarrita y Patricia Barrantes Chavarría

Escuela de Ingeniería,
Universidad Latinoamericana de Ciencia y Tecnología,
ULACIT, Urbanización Tournón, 10235-1000
San José, Costa Rica
[mphillipsp441,jmenam102,pbarrantesc813]@ulacit.ed.cr
http://www.ulacit.ac.cr

Abstract. La analítica visual aplicada a la seguridad informática, es un arte que está tomando fuerza debido a la necesidad de sintetizar grandes volúmenes de datos; estos son generados por los componentes de la infraestructura tecnológica de una organización. Mediante su uso, es posible combinar las capacidades visuales humanas con la potencia de procesamiento de las computadoras, esto para crear un entorno de conocimiento. Se realizó un análisis basado en una revisión sistemática de literatura relacionada con analítica visual, destacando elementos esenciales para esta técnica, como son: paradigmas de visualización, fuentes de datos, lenguajes de programación y técnicas de interacción. Se identificaron relaciones entre dichos elementos y los ataques o problemas de seguridad de los sistemas de información, con el fin de guiar a futuros desarrolladores o analistas sobre como se utiliza la visualización de datos para apoyar la compresión de la información que generan los procesos de seguridad informática. La analítica visual es un gran aporte en el mantenimiento y fortalecimiento de la seguridad tecnológica, reducen los esfuerzos requeridos para tener el panorama completo del estado de la red, amenazas y vulnerabilidades, así mismo colabora con la administración de las relaciones entre dispositivos, usuarios y aplicaciones en un entorno seguro, dotando herramientas para la toma de decisiones.

Keywords: Visualización, Seguridad, Datos, Información, Bitácoras

1 Introducción

Las organizaciones (e.g. gubernamentales, militares, servicios, entretenimiento o industriales) generan grandes volúmenes de distintos tipos de datos a velocidades y calidad variables. Esos volúmenes de datos están relacionados con el tamaño y estructura de la organización (i.e. número de empleados, productos o servicios que proveen y la distribución de operaciones en distintas regiones y países) y su seguridad es afectada por la necesidad de compartirlos entre las diferentes

localidades, a la vez que se tienen en cuenta los principios de confidencialidad, integridad y disponibilidad de la información.

Lo anterior genera la necesidad de diseñar e implementar sistemas para procesar, encriptar, mover y acceder la información de forma segura, así como también se requiere controlar las operaciones y cambios que se llevan a cabo mediante su registro en bitácoras [1] con el fin de monitorear y evaluar posibles vulnerabilidades que comprometan dicha información.

La complejidad para analizar fuentes de datos de distinto tipo y origen se incrementa cuando las cantidades son exorbitantes y aunado a esto se agrega el factor de que está información no es estática y cambia de forma constante; esto convierte el proceso de análisis en una tarea que requiere mucha experiencia por parte de sus ejecutores, debido a esto es crucial que visualizar estos grandes volúmenes de datos sea lo menos complicado posible, para obtener una visión general y de manera dinámica del estado de la infraestructura tecnológica de la organización, observando desde un nivel macro a uno detallado todos los aspectos que rodean la seguridad de la información.

2 Metodología de Investigación

El desarrollo de esta investigación, lleva a cabo una revisión sistemática de literatura, (Kitchenham et al., 2009) relacionada con analítica visual aplicada a la seguridad informática, tomando como referencia los artículos publicados durante los años del 2007 al 2012 en los Proceedings del International Symposium on Visualization for Cyber Security.

Con el fin de determinar como se aplican la analítica visual en los procesos de seguridad informática, se estudiaron 24 artículos completos de los cuales se recopiló información, esta se sintetizó mediante la utilización de una plantilla, la cual contempla varios aspectos importantes citados a continuación:

1. resumen,
2. problema que se busca resolver,
3. resultados,
4. conclusiones,
5. tipos de visualizaciones utilizadas,
6. fuente de datos requeridas,
7. lenguajes de programación,
8. técnicas de interacción

La información obtenida, así como las características pertenecientes a la herramienta o técnica investigada, además de ser analizada y procesada; fue tabulada para generar gráficos que mostraran de forma más contundente y precisa relaciones entre lenguajes, orígenes de datos, técnicas de interacción y paradigmas de visualización, todos estos puntos son el insumo para tener el conocimiento

[1] Logs

necesario para dictar un criterio relacionado a analítica visual aplicada a seguridad tecnológica.

Como base de la investigación se busca responder las siguientes preguntas:

1. ¿ Cómo se utiliza la visualización de datos para apoyar la comprensión de la información que se genera durante los procesos de seguridad informática?
2. ¿ Cómo llevar a cabo un análisis para relacionar tipos de visualizaciones, con tipos de ataques, fuentes de datos de distintos orígenes, lenguajes de programación y técnicas de interacción; que han sido utilizadas por las publicaciones bajo estudio?

De acuerdo con las preguntas anteriores, el objetivo establecido es: determinar cómo se utiliza la visualización de datos para apoyar la comprensión de la información que se genera durante los procesos de seguridad informática, mediante una revisión sistemática de literatura y un análisis para relacionar los tipos de visualizaciones, con tipos de ataques, fuentes de datos, lenguajes y técnicas de interacción que han sido utilizados por las publicaciones bajo estudio.

3 Resultados y Discusión

3.1 Visualizaciones

Frecuencia de uso Para poder responder la pregunta de investigación sobre ¿Cómo se utiliza la visualización de datos para apoyar la comprensión de la información que se genera durante los procesos de seguridad? se realizó un análisis con el fin de encontrar patrones en el uso de las técnicas de visualización, de acuerdo con las lecturas analizadas.

El resultado de ese análisis se muestra en la gráfica 1, donde se determinó 4 tipos de visualizaciones como los más utilizadas. Es importante destacar que estas no siempre son usadas en solitario, sino que en muchas ocasiones los autores hacen combinaciones de estas y otros paradigmas disponibles, todo con el fin de dar al analista de seguridad una herramienta, que de forma rápida, clara y concisa le permita observar los eventos de seguridad que ocurren dentro de la infraestructura informática, el estado o salud de su red y las vulnerabilidades que requieren atención inmediata.

La analítica visual aporta insumos valiosos en tiempo oportuno al proceso de toma de decisiones y administración de la seguridad en materia de TI [2]. Otro dato importante que podemos observar en la misma gráfica es que el valor más alto está en la categoría otros, lo que evidencia que las visualizaciones son tan versátiles y diversas como creativo sea su diseñador y al ser una técnica relativamente nueva, es atractiva para muchos informáticos que desean realizar su aporte a la rama desde su perspectiva. Debido a que la categoría otros incluye una variedad de visualizaciones amplia, se analizaran únicamente las técnicas con mayor incidencia, omitiendo dicha categoría.

[2] Acrónimo para Tecnologías de Información.

Comentaremos cada una de ellas en orden descendente, para determinar cuáles son las posibles razones, desde el punto de vista de sus características, fortalezas y tipos de ataque que sus autores investigaron. Los paradigmas de visualización son:

1. Grafos:Los grafos son representaciones de nodos casi siempre circulares unidos por líneas, estos sirven para representar relaciones o interacciones entre elementos que pueden ser nodos de una red, dispositivos de interconexión, flujos de red, usuarios, aplicaciones o bien la traza de un malware Ṡu forma sencilla y fácil de interpretar son las características más importantes, las cuales los hacen ser una de los paradigmas de visualización más utilizados por los autores. Prueba de esto es lo mostrado en los artículos de (Fang, Miller, & Kupsch, 2012) y (Tsigkas, Thonnard, & Tzovaras, 2012). Los primeros lo utilizan para su análisis profundo de seguridad, donde se visualiza la estructura de una solución, con relaciones entre procesos y diferentes host, con el fin de exponer posibles vulnerabilidades. A Tsigkas et.al le fue muy útil para la investigación de campañas de spam.
2. Lineas de tiempo: La característica principal de las visualizaciones de linea de tiempo es el mostrar el comportamiento de la red en periodos de específicos según los periodos de monitoreo y con ello observar los cambios presentados y determinar si corresponden a comportamientos anormales o de sintomatología de trazas de malware o bien de otro tipo de ataque.
3. Histogramas: Un histograma es un tipo de gráfica de barras para el análisis de frecuencias, le permite a los analistas determinar qué tan habitual es la ocurrencia de eventos de seguridad y con ello apoyar los procesos de toma de decisiones y administración de la red. La alta incidencia en uso de histogramas, demuestra el valor que tienen en el análisis comparativo, el conteo de ocurrencias de eventos o bien vulnerabilidades. Herramientas como Nessus, propuesta por (Harrison, Spahn, Iannacone, Downing, & Goodall, 2012), hacen un amplio uso de este paradigma, esta visualización le permite al administrador de red ver los niveles de severidad así como los tipos y puntuaciones de las vulnerabilidades de su red; mediante escaneos de sus conexiones o bien otras fuentes de datos que les sirvan como base.

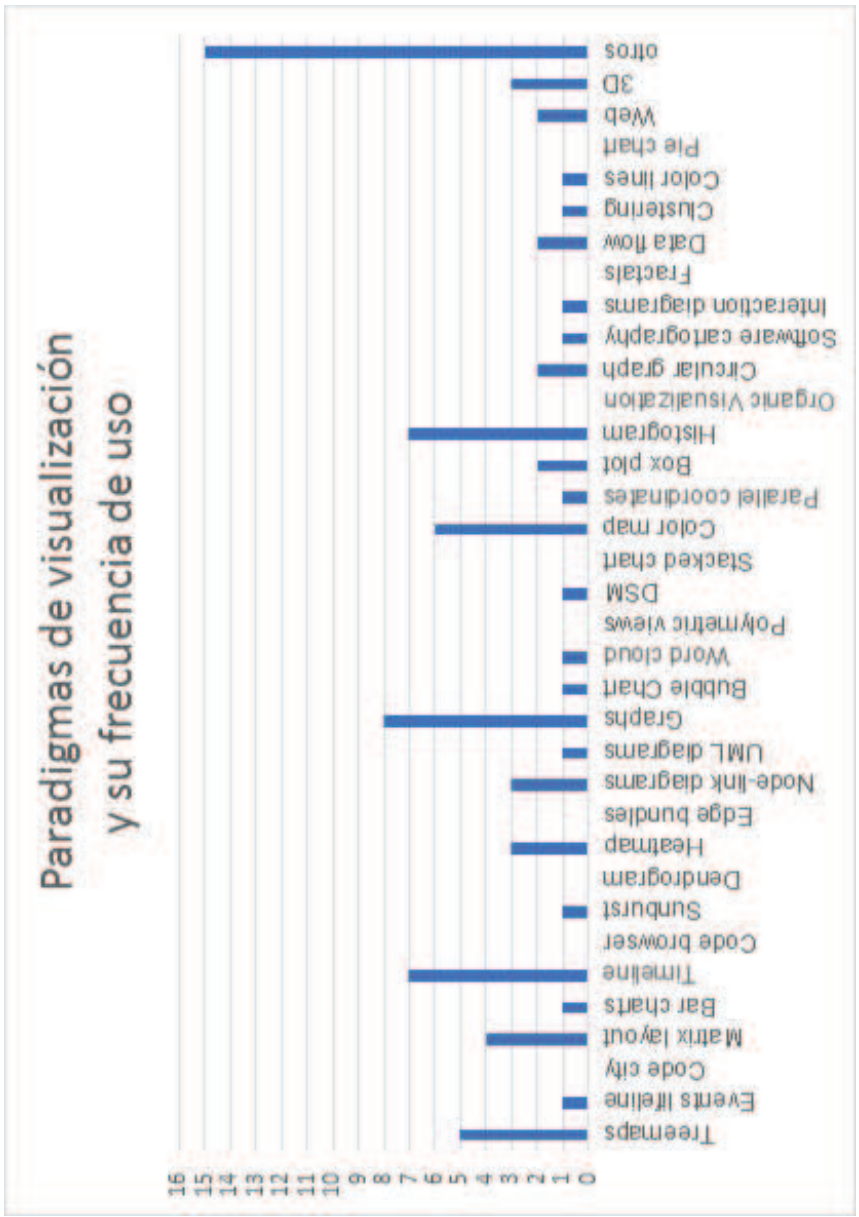

Fig. 1. Histograma: Analiza cuales paradigmas de visualización fueron más usados por los autores de los artículos estudiados

4. Treemaps: Estos sirven para representar grandes cantidades de datos como jerarquías, donde la forma visual de la misma es una tabla con color en el que un rectángulo a nivel de columna contiene a otros rectángulos que se consideran de un nivel inferior, dependiente o perteneciente a su padre. (Mansmann, Göbel, & Cheswick, 2012) la utilizan como base para el análisis de políticas de seguridad para cortafuegos, desglosando por grupos y secciones las listas de control de acceso. Para el analista es la parte inicial a revisar y con ello guiarse sobre las relaciones y secuencias de las políticas aplicadas por el cortafuegos a grupos de objetos como ICMP (Protocolo de mensajes de control de Internet), elementos de red, protocolos y servicios. (Harrison et al., 2012),basan su análisis de vulnerabilidades y riesgos de la red en los treemaps, considerándola como la herramienta ideal para el fácil entendimiento por parte de los analistas. Los colores de los treemaps dan la apariencia de una matriz de calor para enfatizar los niveles de riesgo o bien daño que puede o pudo ocasionar el ataque o una ocurrencia de alguna amenaza.
5. Mapas de color: Los mapas de color explotan la predisposición humana a posar la mirada sobre los colores más brillantes o llamativos y con esto lograr de un solo vistazo, enfocar o resaltar esas zonas geográficas como por ejemplo identificar áreas de la red con altas vulnerabilidades, las cuales son el objetivo de los malware. El uso de los colores dentro del mapa también deja ver patrones de comportamiento, dan una visión de la presencia, importancia, o impacto que tiene el elemento analizado dentro del entorno. Tal como lo hacen (Fischer, Fuchs, Vervier, Mansmann, & Thonnard, 2012), para el análisis de anomalías en el protocolo BGP[3] y anomalías en el enrutamiento. A si mismo, (Veras, Thorpe, & Collins, 2012), lo utilizan para el estudio de la semántica en las contraseñas, y (Yu, Lippmann, Riordan, & Boyer, 2010) para el mapeo de direcciones IP [4] con posibles comportamientos maliciosos. También, en el articulo de (Leschke & Sherman, 2012) observamos que se pueden utilizar para análisis forenses, ellos logran mostrar información relativa al cambio de estructuras de una jerarquía de archivos, con el fin de comprender el comportamiento de algún objeto, directorios que no estén presentes o que cambien mucho en periodos de tiempo establecidos.

Se encontró que el uso de estás técnicas de visualización crece debido a lo versátiles que son ya que pueden desplegar mucha información en poco espacio y esto es un factor crucial para que el analista de seguridad, debido a que puede captar y aprovechar al máximo lo que se muestra en pantalla. Cabe recalcar que el uso de estos paradigmas es dependiente de los orígenes de datos y estos se relacionan muy estrechamente con los lenguajes de programación usados, lo cual se abordará más adelante.

Cantidad de Visualizaciones usadas por articulo Las técnicas de visualización tienen cada una sus fortalezas y debilidades; para realizar una propuesta

[3] Por sus siglas en Inglés Border Gateway Protocol
[4] Por sus siglas en Inglés Internet Protocol

hay que ser consciente de ellas, esto para lograr los objetivos establecidos y dar solución a la problemática bajo estudio. Cada articulo da muestra de ello, pues en cada uno intentan sacar provecho del paradigma de visualización elegido.

Ese conocimiento es observable en la forma en que autores como (Liao, Striegel, & Chawla, 2010) hacen uso no de uno, sino de varios paradigmas de visualización, que al complementarse ofrecen al analista de seguridad una solución robusta y completa. El treemap en la figura 2 nos muestra que como valor máximo se combinaron 6 tipos de visualizaciones y también, nos dice que en la mayoría de los artículos en estudio, se combinaron al menos dos tipos de visualizaciones, esto porque, con una técnica se da la visión general o completa y con la segunda se profundiza en detalles. Es común que sea está segunda, tercera o cuarta visualización la que presente los resultados mediante el uso de técnicas de interacción como el filtrado, la búsqueda o bien las animaciones.

Fig. 2. Técnicas de Visualización utilizadas por articulo

Uso de visualizaciones por Ataque La técnicas de visualización son versátiles y pueden usarse para distintos análisis , por ende los autores experimenten al menos una vez con ellas en los aspectos más críticos de la seguridad; con esto se evalúan sus aportes y determina su funcionalidad. Es de esperar que conforme se consolide el uso de las visualizaciones, vayan a ser más frecuentes unas y otras caigan en desuso para un tipo de ataque en concreto.

Tomando como base lo mostrado por la tabla de análisis de visualizaciones por tipo de ataque, en la gráfica 3 se observa que: en la columna de Malware, uno de los hitos de la seguridad, se requiere la integración de estrategias para su detección, estudio de comportamiento y evaluaciones de impacto. De ahí que para su análisis los autores utilizaran 15 de los 26 tipos de visualizaciones detectadas.

Otro dato importante que podemos extraer de la gráfica es que el monitoreo del tráfico es el segundo con mayor número de visualizaciones utilizadas, esto porque en el área de seguridad esta técnica es una de las que más genera datos para analizar, ya que el registro de eventos es continuo y sus cambios a través del tiempo son de gran importancia para el analista; esto debido a que le permiten detectar anomalías y debilidades de la red.

Según el nivel de criticidad que tenga el tipo de ataque o problema a analizar dentro de la seguridad, así es el número de visualizaciones utilizadas, basta con ver los totales de temas como: patrones de tráfico, códigos fuente, archivos de sistema, entre otros.

Visualization	Attack										
	FileSystem	Firewall	Intrusion ..	Malware	Others	Passwords	Source C..	SPAM	Traffic Mo..	Traffic Pa..	Total
3D				1					1	1	3
Bar charts							1				1
Box plot				1							1
Bubble treemap				1					1		2
Circular graph				1					1	1	3
Clustering					1						1
Color lines									1	1	2
Color map	1			1	1	1	1				5
Data flow				1			1				2
DSM					1						1
Events lifeline					1						1
Graphs	1			1	1		1	1	1	1	7
Heatmap				1			1		1	1	4
Histogram	1			1	1		1		1	1	6
Interaction diagrams				1							1
Matrix layout	1			1					1	1	4
Node-link diagrams			1		1		1				3
Others	1		1	2	1	1	1	1	2	1	11
Parallel coordinates									1	1	2
Software cartography					1						1
Sunburst		1									1
Timeline			1	1			1		1	1	5
Treemaps		1		1	1		1		1	1	6
UML diagrams							1				1
Word cloud						1					1
Total	5	2	3	15	10	3	11	2	13	11	75

Visualization Count: 1 – 75

Fig. 3. El gráfico muestra cantidad de paradigmas de visualización utilizados desglosado por tipo de ataque la intensidad de color se relaciona a la cantidad de incidencias

3.2 Orígenes de datos y Lenguajes de Programación

Las visualizaciones se construyen a partir de grandes fuentes de datos, que se recopilan diariamente en los departamentos de seguridad del área de tecnologías

de información, en esta sección se analizará cuáles son esos tipos que se requieren para evaluar aspectos de la seguridad informática.

De la gráfica 4 se puede analizar que desde el punto de vista de los ataques, se evidencia que las visualizaciones que usan más tipos de datos son: malware, monitoreo de tráfico patrones de tráfico.Otro enfoque se basa en el punto de vista de los datos; ya que para malware, monitoreo de tráfico y patrones de tráfico, se usan principalmente archivos PCAP [5] que contienen capturas de estados de la red y también Netflows que mantienen información de tráfico IP, recolectada mediante este protocolo que fue creado por CISCO System.

Otro tipo de dato que se usa con frecuencia son los archivos de sistema (i.e como se le conoce en ingles File System). Estos aparecen en análisis de malware, monitoreo de tráfico y código fuente.En menor grado los logs o registros de eventos de seguridad, que pueden estar asociados a usuarios, aplicaciones o dispositivos dentro de la red. Hay datos como los IDS [6] que por su naturaleza solo tienen una aplicabilidad que para este caso es la detección de intrusos.

Quien diseñe una visualización tomará un único tipo de dato o hará uso de combinaciones de ellos según los recursos con los que cuente el analista, tal como aplicaciones, o configuraciones que algunos dispositivos proveen como parte de su diseño y que son parte de la seguridad de la red.

Cabe recalcar que siempre existirá una dependencia entre las fuentes datos y los lenguajes de programación, a causa de que los datos necesitan ser formateados, manipulados y preparados debido a que estos son insumos que pueden estar de forma estructurada o no y sus orígenes suelen ser muy variados. Para poder ser interpretados y generar a partir de estos información útil; los lenguajes de programación los procesan y proveen a las aplicaciones de materia prima para cumplir objetivos.

En el contexto actual la gran cantidad de lenguajes de programación dotan a los desarrolladores de distintas posibilidades de elección, buscando que se ajusten de la mejor manera para solucionar un problema concreto. En el proceso de análisis se determinaron algunos hechos relevantes entre las fuentes de datos y los lenguajes como se muestra en la gráfica 5, se vislumbro que los archivos PCAP y NetFlows son las fuentes de datos que más se utilizan ya que de los 8 lenguajes que se listaron (omitiendo la categoría otros), en 7 ocasiones fueron la base para visualizar información.

Con respecto a lo anterior se deduce que: impera el uso de archivos generados por los dispositivos activos de la red sean enrutadores o concentrados, así como también la importancia de los analizadores de protocolos los cuales brindan en gran medida fuentes de datos de gran valor para ser analizadas. Esto da por sentado la gran valía que tienen las diferentes bitácoras a las que debe tener acceso un analista de seguridad dentro de las infraestructuras.

Por otra parte los lenguajes de programación que usaron mayor variedad de orígenes de datos fueron Java y SQL [7]; esto demuestra su versatilidad para

[5] Por sus siglas en Inglés Border Gateway Protocol
[6] Sistema de detección de Intrusos
[7] Acrónimo para Structure Query Languaje en ingles.

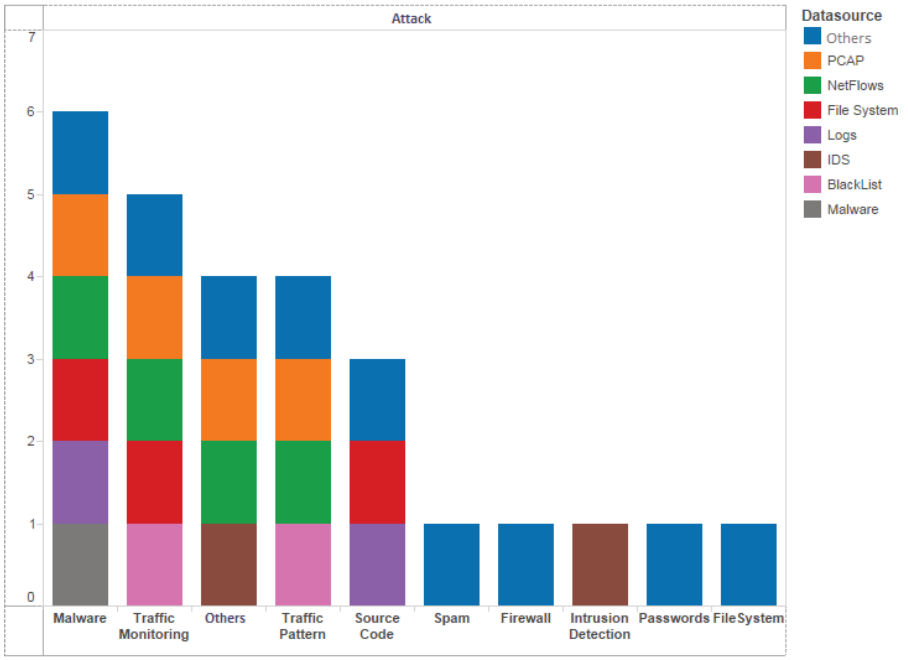

Fig. 4. El gráfico muestra la cantidad de orígenes de datos diferentes por tipo de Ataque, el color se relaciona a cada origen de datos

trabajar y manipular datos con el fin de obtener información que pueda ser visualizada y analizada de una forma menos complicada.

De igual forma otro de los hallazgos que se descubrió durante el proceso de investigación, fue la relación entre los lenguajes de programación y los ataques, por ejemplo los ataques que más lenguajes utilizaron son aquellos que monitorean tráfico y patrones de la red; esto debido a que es una técnica habitual socavar información de lo que ocurre en nuestra infraestructura de red para tomar decisiones.

Así pues otra de los resultados obtenidos como se muestra en la gráfica, 6 es que el lenguaje más utilizado en este caso fue java, esto se puede deber a su polimorfismo, sus características de multiplataforma y el gran número de adeptos que posee, en contra posición .Net de Microsoft solo se utilizó en una ocasión con lo cual se evidencia que mucho del desarrollo se lleva a cabo en plataformas que son de Código Abierto.

Del mismo modo se obtuvo información importante como lo es la dependencia de los lenguajes de consultas estructuradas, como es el caso de SQL , el cual debido a la necesidad de manipular volúmenes de datos cuantiosos; se requiere para trabajar de forma eficiente con las bases de datos, en donde se fusionan los insumos. Este se encontró en 5 de los diferentes ataques que se investigaron. Además se puede visualizar en la gráfica 7 que SQL es utilizado en varios de los

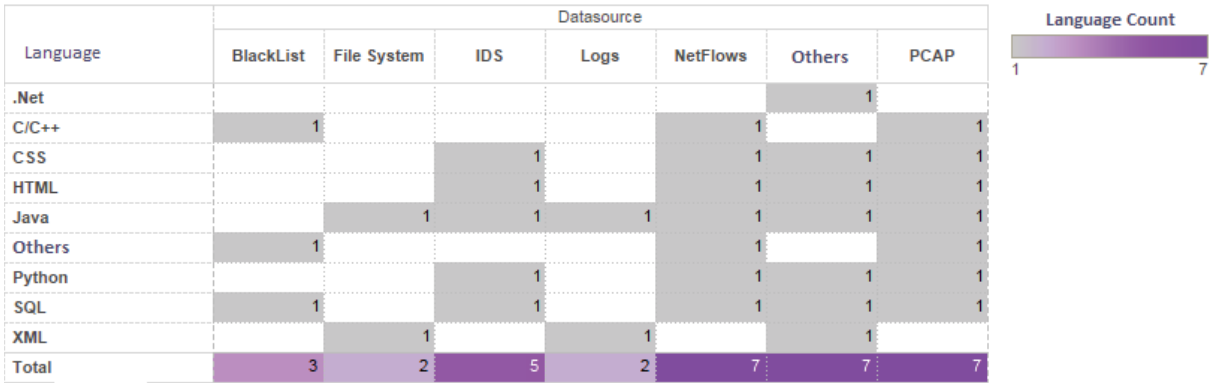

Fig. 5. El gráfico muestra el conteo definido de cantidad de lenguajes desglosado por origen de datos, el color muestra la cantidad filtrada excluyendo valores nulos

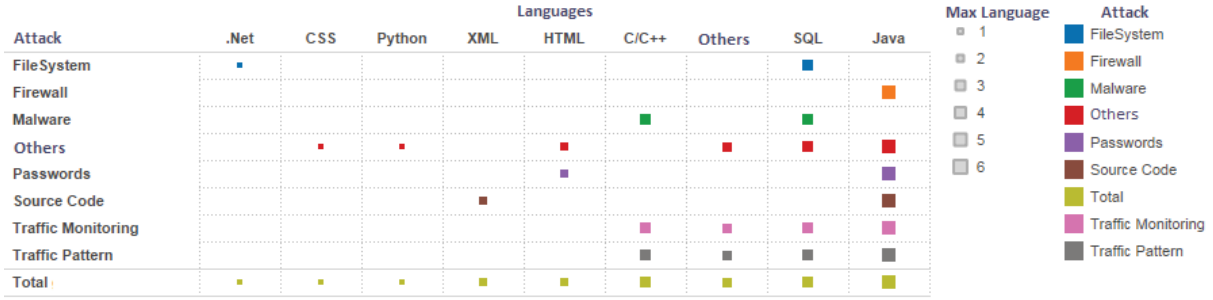

Fig. 6. El gráfico muestra el conteo definido de cantidad de visualizaciones desglosado por tipos de lenguaje, el color muestra la cantidad filtrada excluyendo valores nulos

tipos de visualizaciones identificados, específicamente en 11, esto se debe a que es un lenguaje ideal puesto que permite efectuar consultas con el fin de recuperar de forma sencilla, información de bases de datos, así como hacer cambios en ellas.

Lo anterior se puede evidenciar en lo propuesto por (Zhuo & Nadjin, 2012), que consiste en una herramienta para visualizar e identificar patrones de comportamiento de posibles ataques de malware en una infraestructura de red, tomando como base el análisis de paquetes y tráfico en tiempo real; para ello utiliza archivos de tipo PCAP, estos luego de ser recolectados mediante analizadores de protocolos, generan una colección de datos que contienen mucha información, la cual se analiza gramaticalmente para obtener los insumos que serán visualizados, es aquí donde entra en juego el desarrollo mediante SQL.

Por otro lado se puede observar que XML[8] es otro de los lenguajes más utilizados para diversos tipos de visualizaciones, 11 ocasiones en total, lo anterior debido a que su función es almacenar datos en forma legible y permite definir la gramática de lenguajes específicos para estructurar documentos grandes, lo que lo hace ideal para estructurar los orígenes de datos para las diferentes herramientas de visualizaciones.

Como complemento de lo anterior se puede observar en la gráfica 7, que el lenguaje de mayor uso para representar diversos tipos de visualizaciones fue Java, el mismo fue utilizado en 14 de un total de 17 paradigmas identificados, incluyendo visualizaciones propias, o sea, desarrolladas por los autores de los artículos; como lo propuesto por (Veras et al., 2012), donde describe una herramienta que combina visualizaciones cronológicas, de texto y un gráfico de azulejos, la cual posee muchas vistas coordinadas. Lo anterior pone en evidencia las bondades del lenguaje Java (e.g. portabilidad, dinámico, multihilo, robusto, independiente a la arquitectura) que permiten desarrollar aunque lo objetivos sean complejos.

3.3 Técnicas de interacción

La analítica visual busca la asimilación rápida de información o el monitoreo de grandes cantidades de datos mediante el uso interactivo de representaciones gráficas, es por ello que dicho análisis no solo se beneficia de métodos de representación, sino también de ser capaz de integrarse con una adecuada técnica de interacción.

De acuerdo a lo anterior se analizaron las diferentes técnicas de interacción utilizadas para cada uno de los tipos de visualización identificados.En la gráfica 8 se puede observar que para todos los tipos de visualizaciones excepto para el "Box Plot" se utilizaron dos o más técnicas de interacción, esto porque la analítica visual no es solo una representación gráfica estática, sino que requiere de visualizaciones dinámicas que muestren la información desde diversas perspectivas y ese dinamismo lo proporciona las distintas técnicas de interacción.

Debido a lo anterior es que la técnica más utilizada son los filtros, utilizados en 21 tipos de visualización, lo que se busca con esta técnica es enfatizar obje-

[8] Por sus siglas en Inglés eXtensible Markup Language

l

Fig. 7. El gráfico muestra la cantidad de lenguajes desglosado por tipo de visualización, la saturación de color muestra la cantidad de lenguajes filtrada excluyendo valores nulos

tos individuales o grupos de ellos, de manera que se pueda obtener información específica y detallada de acuerdo a algún parámetro.Como se demuestra en el articulo desarrollado por (Inoue, Eto, Suzuki, Suzuki, & Nakao, 2012) cuya herramienta , DAEDALUS-VIZ, ofrece una función de filtrado de gran flexibilidad para el tráfico darknet [9], donde se pueden mostrar parámetros y combinaciones de acuerdo a esta técnica ya que se puede hacer mediante direcciones IP de origen y destino, protocolos (e.g. TCP, UDP, ICMP), números de puerto de origen y destino, ID [10] del sensor, etc.

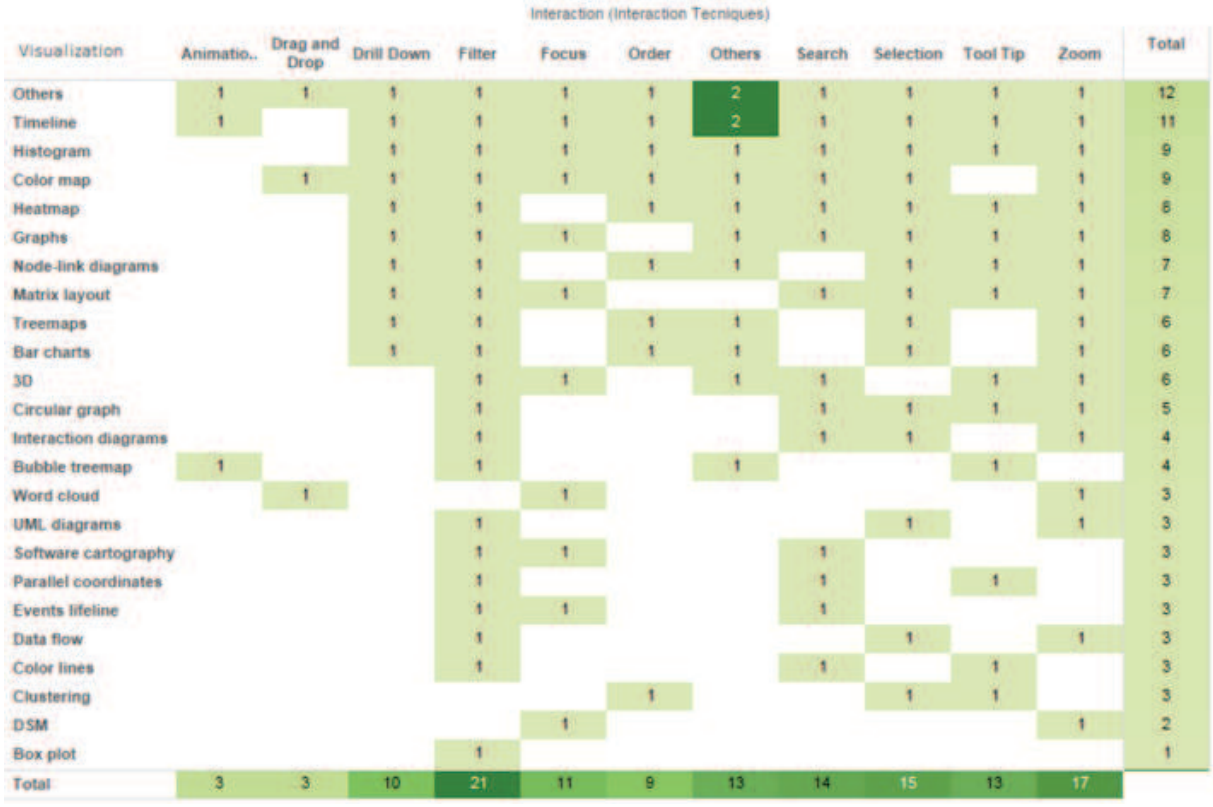

Fig. 8. El gráfico muestra las técnicas de interacción utilizadas desglosado por tipo de visualización

En segundo lugar en cuanto a utilización se encuentra el Zoom [11], la cual tiene presencia en 17 de los 24 tipos de visualización identificados, esto demuestra la importancia de poder reducir el número de objetos visibles aumentando el nivel

[9] Red Oscura: Conjunto de direcciones IP no utilizadas anunciadas a nivel mundial
[10] Abreviatura de Identificación
[11] Conocido como técnica de acercamiento

de detalle; posiblemente mediante el incremento del número de variables que se muestran de cada objeto anterior, permite a los analistas observar toda la información detallada de el objeto en estudio.

Malware	(Erbacher, 2012; Inoue et al., 2012; Nataraj et al., 2011) (Tsigkas et al., 2012; Zhuo & Nadjin, 2012; Saxe et al., 2012) (Yu et al., 2010; Roveta et al., 2011; ?, ?)
Source Code	(Fang et al., 2012)
DDoS	(Goodall et al., 2010)
FileSystem	(Leschke & Sherman, 2012)
SPAM	(Tsigkas et al., 2012)
Traffic Monitoring	(Kintzel et al., 2011; Singh et al., 2011; Boschetti et al., 2011) (Roveta et al., 2011; Chu et al., 2010; Glatz, 2010)
Traffic Pattern	(Chu et al., 2010; Glatz, 2010) (Kintzel et al., 2011; Singh et al., 2011; Boschetti et al., 2011)
IDS	(Guenther et al., 2010)
Firewall	(Mansmann et al., 2012)
Passwords	(Veras et al., 2012)
Otros	(Horn & D'Amico, 2011; Liao et al., 2010) (Harrison et al., 2012; Fischer et al., 2012)

Table 1. Tabla de Referencias desglosado o subdivido por ataque

4 Conclusiones

La analítica visual desde el punto de vista del diseño e implementación requiere muchos elementos para producir buenos resultados.Durante la investigación se logró identificar relaciones de dependencia entre fuentes de datos y lenguajes de programación como se observo en la sección , así como técnicas de interacción que son parte fundamental de casi cualquier tipo de visualización, también se encontraron los valores máximos de utilización en lo que respecta a lenguajes, fuentes de datos y visualizaciones.

Con respecto a las técnicas de visualización se concluye que, entre más representativa sea de los elementos a analizar, mejores resultados se obtendrán y el nivel de experiencia requerido por los analistas será menor. Características como la sencillez, la claridad y la simplicidad de los elementos ayudan a la fácil interpretación de los resultados y se agilizan los procesos de toma de decisiones, administración de la infraestructura y el control de vulnerabilidades.

Otro deducción importante de recalcar es la utilización complementaria de técnicas de visualización, por ejemplo para que dentro de una misma pantalla se puedan representar varios valores a considerar durante el proceso de análisis de una infraestructura particular. De ahí que muchos autores utilizaron hasta 6 tipos de visualización para mostrar y navegar entre los datos, de esta forma se obtiene información desde niveles macro a micro en lo que respecta a detalle.

Aunado a esto se desprende que un dibujo estático no es suficiente para extraer información de los datos. La visualización debe ser construida y reconstruida, es decir manipulada hasta que todas las relaciones que subyacen de los insumos de la red hayan sido percibidas. Es necesario que los datos puedan ser explorados de forma interactiva y para ello se deben combinar buenos métodos de representación visual con buenas interacciones de estas representaciones, con el fin de facilitar o detallar la información a los usuarios para su entendimiento. Es por ello que se pudo observar que para todos los tipos de ataque existen más de dos técnicas de interacción, de forma que se puede mostrar la información desde distintas perspectivas.

References

Boschetti, A., Salgarelli, L., Muelder, C., & Ma, K.-L. (2011). Tvi: A visual querying system for network monitoring and anomaly detection. In *Proceedings of the 8th international symposium on visualization for cyber security* (pp. 1:1–1:10). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2016904.2016905` doi: 10.1145/2016904.2016905 pages 15

Chu, M., Ingols, K., Lippmann, R., Webster, S., & Boyer, S. (2010). Visualizing attack graphs, reachability, and trust relationships with navigator. In *Proceedings of the seventh international symposium on visualization for cyber security* (pp. 22–33). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/1850795.1850798` doi: 10.1145/1850795.1850798 pages 15

Erbacher, R. F. (2012). Visualization design for immediate high-level situational assessment. In *Proceedings of the ninth international symposium on visualization for cyber security* (pp. 17–24). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2379690.2379693` doi: 10.1145/2379690.2379693 pages 15

Fang, W., Miller, B. P., & Kupsch, J. A. (2012). Automated tracing and visualization of software security structure and properties. In *Proceedings of the ninth international symposium on visualization for cyber security* (pp. 9–16). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2379690.2379692` doi: 10.1145/2379690.2379692 pages 4, 15

Fischer, F., Fuchs, J., Vervier, P.-A., Mansmann, F., & Thonnard, O. (2012). Vistracer: A visual analytics tool to investigate routing anomalies in traceroutes. In *Proceedings of the ninth international symposium on visualization for cyber security* (pp. 80–87). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2379690.2379701` doi: 10.1145/2379690.2379701 pages 6, 15

Glatz, E. (2010). Visualizing host traffic through graphs. In *Proceedings of the seventh international symposium on visualization for cyber security* (pp. 58–63). New York, NY, USA: ACM. Retrieved from `http://doi.acm`

.org/10.1145/1850795.1850802 doi: 10.1145/1850795.1850802 pages 15

Goodall, J. R., Radwan, H., & Halseth, L. (2010). Visual analysis of code security. In *Proceedings of the seventh international symposium on visualization for cyber security* (pp. 46–51). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/1850795.1850800 doi: 10.1145/1850795.1850800 pages 15

Guenther, J., Volk, F., & Shaneck, M. (2010). Proposing a multi-touch interface for intrusion detection environments. In *Proceedings of the seventh international symposium on visualization for cyber security* (pp. 13–21). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/1850795.1850797 doi: 10.1145/1850795.1850797 pages 15

Harrison, L., Spahn, R., Iannacone, M., Downing, E., & Goodall, J. R. (2012). Nv: Nessus vulnerability visualization for the web. In *Proceedings of the ninth international symposium on visualization for cyber security* (pp. 25–32). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2379690.2379694 doi: 10.1145/2379690.2379694 pages 4, 6, 15

Horn, C., & D'Amico, A. (2011). Visual analysis of goal-directed network defense decisions. In *Proceedings of the 8th international symposium on visualization for cyber security* (pp. 5:1–5:6). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2016904.2016909 doi: 10.1145/2016904.2016909 pages 15

Inoue, D., Eto, M., Suzuki, K., Suzuki, M., & Nakao, K. (2012). Daedalus-viz: Novel real-time 3d visualization for darknet monitoring-based alert system. In *Proceedings of the ninth international symposium on visualization for cyber security* (pp. 72–79). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2379690.2379700 doi: 10.1145/2379690.2379700 pages 14, 15

Kintzel, C., Fuchs, J., & Mansmann, F. (2011). Monitoring large ip spaces with clockview. In *Proceedings of the 8th international symposium on visualization for cyber security* (pp. 2:1–2:10). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2016904.2016906 doi: 10.1145/2016904.2016906 pages 15

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering â" a systematic literature review. *Information and Software Technology*, *51*(1), 7 - 15. Retrieved from http://www.sciencedirect.com/science/article/pii/S0950584908001390 (Special Section - Most Cited Articles in 2002 and Regular Research Papers) doi: http://dx.doi.org/10.1016/j.infsof.2008.09.009 pages 2

Leschke, T. R., & Sherman, A. T. (2012). Change-link: A digital forensic tool for visualizing changes to directory trees. In *Proceedings of the ninth international symposium on visualization for cyber security* (pp. 48–55). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/2379690.2379697 doi: 10.1145/2379690.2379697 pages 6, 15

Liao, Q., Striegel, A., & Chawla, N. (2010). Visualizing graph dynamics and similarity for enterprise network security and management. In *Proceedings of the seventh international symposium on visualization for cyber security* (pp. 34–45). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/1850795.1850799` doi: 10.1145/1850795.1850799 pages 7, 15

Mansmann, F., Göbel, T., & Cheswick, W. (2012). Visual analysis of complex firewall configurations. In *Proceedings of the ninth international symposium on visualization for cyber security* (pp. 1–8). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2379690.2379691` doi: 10.1145/2379690.2379691 pages 6, 15

Nataraj, L., Karthikeyan, S., Jacob, G., & Manjunath, B. S. (2011). Malware images: Visualization and automatic classification. In *Proceedings of the 8th international symposium on visualization for cyber security* (pp. 4:1–4:7). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2016904.2016908` doi: 10.1145/2016904.2016908 pages 15

Roveta, F., Caviglia, G., Di Mario, L., Zanero, S., Maggi, F., & Ciuccarelli, P. (2011). Burn: Baring unknown rogue networks. In *Proceedings of the 8th international symposium on visualization for cyber security* (pp. 6:1–6:10). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2016904.2016910` doi: 10.1145/2016904.2016910 pages 15

Saxe, J., Mentis, D., & Greamo, C. (2012). Visualization of shared system call sequence relationships in large malware corpora. In *Proceedings of the ninth international symposium on visualization for cyber security* (pp. 33–40). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2379690.2379695` doi: 10.1145/2379690.2379695 pages 15

Singh, A., Bradel, L., Endert, A., Kincaid, R., Andrews, C., & North, C. (2011). Supporting the cyber analytic process using visual history on large displays. In *Proceedings of the 8th international symposium on visualization for cyber security* (pp. 3:1–3:8). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2016904.2016907` doi: 10.1145/2016904.2016907 pages 15

Tsigkas, O., Thonnard, O., & Tzovaras, D. (2012). Visual spam campaigns analysis using abstract graphs representation. In *Proceedings of the ninth international symposium on visualization for cyber security* (pp. 64–71). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2379690.2379699` doi: 10.1145/2379690.2379699 pages 4, 15

Veras, R., Thorpe, J., & Collins, C. (2012). Visualizing semantics in passwords: the role of dates. In *Vizsec'12* (p. 88-95). pages 6, 12, 15

Yu, T., Lippmann, R., Riordan, J., & Boyer, S. (2010). Ember: A global perspective on extreme malicious behavior. In *Proceedings of the seventh international symposium on visualization for cyber security* (pp. 1–12). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/1850795.1850796` doi: 10.1145/1850795.1850796 pages 6, 15

Zhuo, W., & Nadjin, Y. (2012). Malwarevis: Entity-based visualization of mal-

ware network traces. In *Proceedings of the ninth international symposium on visualization for cyber security* (pp. 41–47). New York, NY, USA: ACM. Retrieved from `http://doi.acm.org/10.1145/2379690.2379696` doi: 10.1145/2379690.2379696 pages 12, 15