Monitorización y selección de incidentes en seguridad de redes mediante EDA

J. Camacho, G. Maciá-Fernández, J. Díaz-Verdejo, P. García-Teodoro
Departamento de Teoría de la Señal, Telemática y Comunicaciones - CITIC
Universidad de Granada
Email: {josecamacho, jedv, gmacia, pgteodor}@ugr.es

***Resumen*—Uno de los mayores retos a los que se enfrentan los sistemas de monitorización de seguridad en redes es el gran volumen de datos de diversa naturaleza y relevancia que deben procesar para su presentación adecuada al equipo administrador del sistema, tratando de incorporar la información semántica más relevante. En este artículo se propone la aplicación de herramientas derivadas de técnicas de análisis exploratorio de datos para la selección de los eventos críticos en los que el administrador debe focalizar su atención. Adicionalmente, estas herramientas son capaces de proporcionar información semántica en relación a los elementos involucrados y su grado de implicación en los eventos seleccionados. La propuesta se presenta y evalúa utilizando el desafío VAST 2012 como caso de estudio, obteniéndose resultados altamente satisfactorios.**

***Palabras clave*—análisis exploratorio de datos (*exploratory data analysis*), *big data*, visualización de datos (*data visualization*), seguridad en redes (*network security*), sistemas de monitorización de seguridad en redes (*network security monitoring systems*)**

I. Introducción

Los sistemas de monitorización de la seguridad en redes (NSM, del inglés *Network Security Monitoring*) [1] tienen como finalidad la agregación y análisis de los datos procedentes de los diversos mecanismos y sensores desplegados en el entorno de red, a fin de validar y, en su caso, responder a incidentes de seguridad. Aunque suelen incorporar datos procedentes de sistemas de detección de intrusiones (IDS, del inglés *Intrusion Detection Systems*) [2] como elemento relevante, no son, en sí mismos, sistemas IDS. Por el contrario, su operación está orientada a seleccionar, priorizar y validar las alertas generadas por otros sistemas de monitorización y trazado de eventos.

Entre las limitaciones que deben afrontar los NSM podemos mencionar el gran volumen de datos que deben manejar, ya que integran información de múltiples fuentes, muchas de ellas generando un elevado número de registros (p.e., trazas de cortafuegos, de sesiones, alertas de IDS, etc.). Adicionalmente, los datos deben ser preprocesados, agregados y presentados al administrador de forma que éste pueda comprenderlos y gestionarlos fácilmente. En consecuencia, dos son los retos más relevantes para el diseño de NSM: el análisis de los datos y la presentación/visualización de los resultados.

La mayoría de los NMS existentes (p.e., Sguil[1] o Snorby[2]) se limitan básicamente a recopilar e interrelacionar los datos procedentes de los sensores con la finalidad de facilitar su análisis y consulta por parte del administrador, mostrándolos en base a secuencias temporales y/o priorizándolos a partir de esquemas simples. En algunos casos se incluyen algunas heurísticas y estadísticas simples (p.e., Pravail Security Analytics[3]), pero es evidente que se requieren métodos y técnicas más potentes y de mayores prestaciones para el análisis y visualización de los datos. En este contexto, las técnicas de análisis exploratorio de datos (EDA) [3] pueden resultar extremadamente útiles tanto para establecer los eventos relevantes en los que el administrador debería centrar su atención, como para mostrar las propiedades o características implicadas en cada evento.

En este trabajo proponemos y evaluamos una metodología basada en EDA que proporciona medidas y gráficas para conseguir el objetivo antes mencionado. Para ello se realiza una elección de herramientas que, secuenciadas adecuadamente, permiten, en primer lugar, determinar los eventos potencialmente relevantes sin intervención del administrador. A partir de estos, mediante la obtención e interpretación de algunas gráficas, el administrador puede recabar información semántica respecto de dichos eventos que puede serle de utilidad para la posterior comprobación o supervisión de los mismos.

El resto del artículo se estructura como sigue. En la Sección II se presentan brevemente las herramientas y técnicas en las que se basan los análisis de datos subsiguientes. En la Sección III se describe el funcionamiento del sistema propuesto, explicitándose la secuenciación de las técnicas y procedimientos a aplicar. En la Sección IV se describe la aplicación del sistema desarrollado al reto VAST 2012 [4], para lo que se describirá previamente dicho reto así como la parametrización realizada, aspecto clave del análisis. Finalmente, en la Sección V se presentan las contribuciones más relevantes del trabajo y se apuntan brevemente algunos trabajos de futuro.

II. Herramientas de análisis exploratorio de datos

El análisis exploratorio de datos (EDA) tiene como objetivo facilitar el conocimiento y visualización de la estructura que

[1] http://sguil.sourceforge.net
[2] https://snorby.org
[3] http://www.arbornetworks.com/products/pravail/securityanalytics

ISBN: 978-84-9717-323-0

presenta un conjunto de datos. Para ello utiliza una serie de técnicas y herramientas que permiten analizar su propiedades relevantes y presentarlas de forma adecuada para facilitar su interpretación. Entre las técnicas empleadas se encuentran algunas bien conocidas como PCA (*Principal Component Analysis*) [6], así como otras más novedosas, propuestas recientemente por parte de los autores, como son MEDA [7] y oMEDA [8].

A continuación, se describen brevemente las técnicas utilizadas:

- PCA: El análisis de componentes principales permite transformar un conjunto de N observaciones, cada una de ellas con M variables o componentes que pueden estar correlacionadas entre sí, a un nuevo espacio de características decorrelacionadas denominadas componentes principales (*Principal Components* o *PCs*). Sin entrar en detalles, que pueden consultarse en [5], si $\mathbf{X}$ es la matriz de datos, de dimensión $N \times M$, el análisis PCA permite expresar estas observaciones de acuerdo a:

$$\mathbf{X} = \mathbf{T}_A \cdot \mathbf{P}_A^t + \mathbf{E}_A, \quad (1)$$

  donde A es el número de PCs incluidas en el modelo, $\mathbf{T}_A$ es la matriz $N \times A$ de puntuaciones (*scores*), $\mathbf{P}_A$ la matriz $M \times A$ de cargas (*loadings*), compuesta por los A autovectores de $\mathbf{XX} := \mathbf{X}' \cdot \mathbf{X}$ con los mayores autovalores asociados, y $\mathbf{E}_A$ la matriz $N \times M$ de residuos. En el contexto del análisis de datos podemos decir, de forma coloquial, que el objetivo de este análisis es retener la mayor información posible sobre los datos con el menor número posible de parámetros. El procedimiento para la selección adecuada de A depende de la aplicación concreta considerada [9]. El resto de herramientas se basan en el modelo PCA.
- Gráficos de evolución: Los gráficos de evolución, utilizados ampliamente en el entorno industrial, permiten visualizar de forma simple la parte del modelo y de los residuos obtenida en la ec. (1) para el conjunto de observaciones. Para ello, se obtienen una pareja de gráficos a partir del *leverage* o estadístico T^2 de Hotelling, que comprime la información en el modelo, y la estadística Q [10], que comprime la información en el residuo. En el contexto de la seguridad, ambas gráficas permiten identificar con sencillez cualquier evento anómalo.
- MEDA: Los gráficos MEDA son mapas de color de tamaño $M \times M$ en los que se representa la relación (positiva o negativa) existente entre las parejas de variables de un conjunto de datos. Los coeficientes de MEDA son una variante de la correlación menos sensible al ruido y, por tanto, con mejores cualidades para detectar la estructura en los datos. Para facilitar la visualización de los gráficos MEDA, se suele usar un método de serialización [11] que reordena las variables de acuerdo a un criterio de similitud. De esta forma es más fácil identificar grupos de variables, ya que tienden a formar cuadrados en el gráfico. En el contexto de la seguridad, los grupos de variables nos permiten identificar los tipos de tráfico o incidentes que tienen lugar en la red.
- oMEDA: Los gráficos oMEDA permiten comparar valores de variables en dos grupos de observaciones a partir de un diagrama de barras. Así, un valor positivo para una variable en oMEDA significa que el primer grupo de observaciones presenta un valor mayor para dicha variable que el segundo grupo, mientras que un valor negativo representa lo contrario. En el contexto de la seguridad, oMEDA se utiliza para identificar las variables relacionadas con un evento anómalo, comparando dicho evento con la tendencia genérica en la red. El resultado nos permite determinar características del evento anómalo, que potencialmente nos pueden permitir identificar las causas de dicho evento y, en su caso, proponer medidas paliativas o de respuesta de forma veloz y eficaz.

Las herramientas descritas se encuentran implementadas en un módulo para ©Matlab desarrollado por uno de los autores [12].

III. Monitorización y visualización: arquitectura y metodología

La metodología de monitorización y visualización de incidentes de seguridad propuesta se basa en la detección e interpretación de anomalías a partir del análisis PCA, para lo que se usan las gráficas de *Hotelling* T^2 *y* Q, junto con MEDA y oMEDA para determinar las variables y relaciones entre ellas asociadas a dichas anomalías. Esta combinación de herramientas resulta extremadamente útil para los fines mencionados en escenarios caracterizados por un elevado número de datos y parámetros.

En la Figura 1 se muestra un diagrama de bloques del sistema planteado para NSM. Como puede observarse, se consideran dos bloques diferenciados que se discuten a continuación.

III-A. Preprocesado

En este bloque se preparan los datos procedentes de las fuentes para su análisis. Las secuencias de datos de entrada son preprocesadas y parametrizadas de acuerdo a un conjunto de características/variables seleccionadas.

Cada variable contabiliza el número de veces que, durante un cierto intervalo de tiempo w, aparece cierto valor o valores en los registros (*logs*) del dispositivo fuente. A modo de ejemplo, una variable podría contabilizar el número de veces que un puerto determinado, p.e., el 21 (ftp), aparece en las trazas durante un periodo de 1 minuto. La motivacion para esta elección es que el número de entradas en una traza en las que aparece un puerto concreto puede proporcionar información para detectar eventos asociados a un protocolo.

Aunque el sistema de parametrización puede diseñarse de forma específica para los dispositivos de monitorización y detección disponibles en la red, se pueden definir ciertas buenas prácticas de diseño:

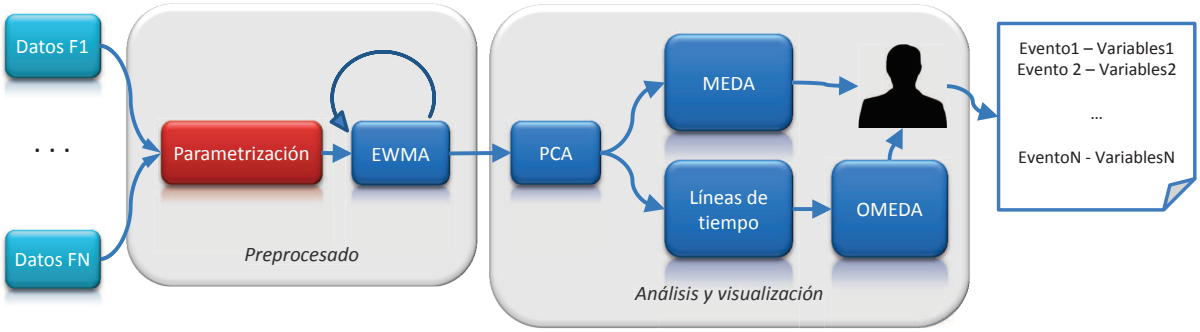

Figura 1. Diagrama de bloques del sistema.

- Seleccionar una variable por nivel de seguridad, prioridad o alarma en los registros del dispositivo (por ejemplo, variable "*warning*z variable "*critical*").
- Seleccionar una variable por código en los registros del dispositivo (por ejemplo, variable contador del código ASA-4).
- Seleccionar una variable por dirección IP o grupo de direcciones IP con sentido topológico o funcional en la red (por ejemplo, variable contador de la IP del DNS local o de las IPs de un departamento concreto).
- Seleccionar una variable por grupo de puertos relacionados con un protocolo de interés (por ejemplo, variable contador del puerto 80 y el 8080).

En cada intervalo de tiempo w, se combinan las características evaluadas para cada fuente en un único vector, $\overrightarrow{x_t}$, que será la observación correspondiente al instante t sobre la que se realizarán los análisis posteriores. A continuación, esta observación se utiliza para actualizar, siguiendo una estrategia de media móvil de peso exponencial o EWMA (*Exponentially Weighted Moving Average*), la matriz $\mathbf{XX}$ que representa el estado actual de la red monitorizada. La matriz se actualiza con la entrada de nuevos datos de la forma $\mathbf{XX}_t = \lambda \cdot \mathbf{XX}_t + \overrightarrow{x_t}' \cdot \overrightarrow{x_t}$, donde $0 \leq \lambda \leq 1$ es un factor de olvido que permite descartar información pasada.

III-B. Análisis y visualización

En el segundo bloque se realizan todas las operaciones necesarias para el análisis de los datos, visualización y posterior interpretación. Se inicia el procesamiento realizando un análisis PCA de $\mathbf{XX}$. La aproximación utilizada para la parametrización, a diferencia de las habitualmente utilizadas en los NSM, puede generar un elevado número de parámetros y, consecuentemente, una alta dimensionalidad de las observaciones. Sin embargo, esto no supone un problema dado el análisis PCA que se realiza a continuación. Usando PCA, el sistema permite identificar eventos donde se correlacionan las variables contador antes mencionadas, permitiendo de forma sencilla establecer puertos, segmentos de red, niveles de seguridad vulnerados en firewall o IDS, etc., asociados a cada evento anómalo.

A partir del modelo PCA se obtiene la estructura de las variables con los gráficos MEDA y la evolución temporal (líneas de tiempo) de los estadísticos *Hotelling* T^2 y Q, que serán utilizados para detectar anomalías, las cuales se reflejan en estos gráficos por picos en la evolución. Para cada anomalía o conjunto de anomalías próximas en el tiempo se obtienen gráficos oMEDA para determinar cuáles son las variables relacionadas con dicha anomalía.

De la metodología propuesta resulta relevante su capacidad, no sólo para manejar grandes volúmenes de datos, sino también para gestionar una alta dimensionalidad. Es decir, los eventos u observaciones del sistema pueden ser representados con tantos parámetros como se estime oportuno, no siendo problemática la introducción de información redundante o relacionada, que será adecuadamente procesada por los esquemas PCA subyacentes. Por el contrario, cuantos más parámetros se incluyan, mayor sera la información que podrá extraerse. Esta es una característica diferencial de la propuesta, ya que la mayoría de las herramientas de análisis de redes operan sobre series de datos unidimensionales o de reducida dimensionalidad [13].

IV. Caso de estudio: aplicación a VAST 2012

La mejor forma de explicitar y mostrar las potencialidades de la metodología propuesta en la sección anterior es aplicarla y explicarla en un escenario concreto. Para ello consideraremos el segundo reto del VAST 2012 [4].

Este reto considera un escenario correspondiente a una red corporativa bancaria con varias sedes y acceso a Internet (Figura 2) en la que ocurren incidentes de seguridad durante dos días. El desafío consiste en determinar los eventos más relevantes, sus causas y las posibles soluciones.

Los datos proporcionados consisten en una traza de un cortafuegos Cisco ASA, conteniendo 23.711.341 registros, y la salida generada por un IDS, que incluye 35.948 registros. Los conjuntos de datos, su descripción y los detalles sobre el reto se encuentran disponibles en [4].

IV-A. Parametrización y preprocesado

De acuerdo a la metodología propuesta, los datos procedentes de las trazas del IDS y del cortafuegos se han

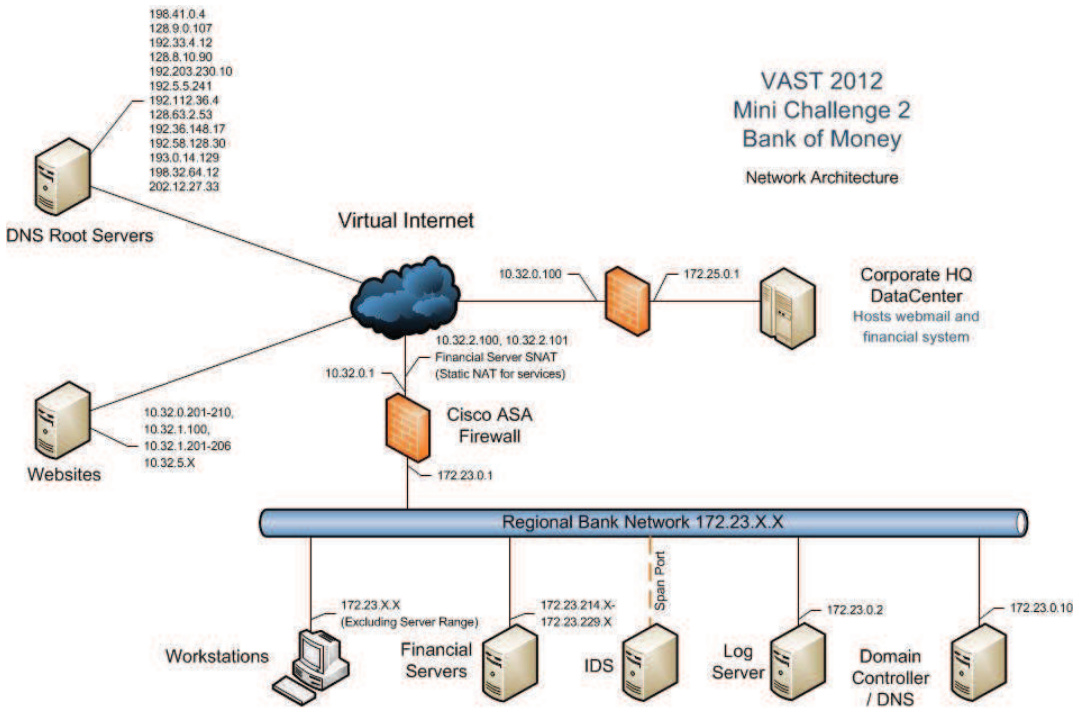

Figura 2. Red utilizada en el reto VAST 2012.

parametrizado agregando los datos durante periodos de $w = 1$ minuto. La elección de la ventana temporal viene determinada por la resolución temporal de las trazas del IDS, que impiden el uso de ventanas de menor tamaño. Se obtienen así 2.350 observaciones, ordenadas temporalmente.

Cada observación del sistema corresponde a un vector de 112 características o variables que representa la información procedente de ambas fuentes de datos. En particular, se han asignado 69 variables para las trazas del cortafuegos y las restantes 43 para las trazas del IDS. En la Tabla I se muestran los números de parámetros establecidos para cada campo presente en las trazas. Así, se consideran 17 parámetros asociados a cada uno de los puertos correspondientes a servicios estándar (número de puerto inferior a 1024) que aparecen en las trazas. En el caso de las direcciones IP, se han establecido 9 parámetros a partir de la topología de la red y de los rangos de direcciones existentes. En la Tabla II se muestran algunos de los parámetros seleccionados. La notación utilizada hace referencia a la fuente de los datos (*fw* o *ids*) y al significado o *flag* asociado a cada uno.

Tabla I
NÚMERO DE PARÁMETROS DEFINIDOS PARA CADA TIPO DE CARACTERÍSTICA.

	Campo	#parámetros
Trazas cortafuegos	Prioridad syslog	5
	Operación	6
	Código del mensaje	25
	Protocolo	3
	Dirección IP	9
	Puerto	17
	Dirección	2
	Tiempos conexión	2
	Subtotal	**69**
Trazas IDS	Dirección IP	9
	Puerto	17
	Tipo alerta	5
	Prioridad	3
	Etiqueta	9
	Subtotal	**43**

IV-B. Análisis y visualización

Una vez realizada la parametrización de los datos de entrada se procede a realizar un análisis PCA que será la base para el resto del estudio. A partir del modelo PCA, se obtiene un gráfico MEDA (Figura 3) que muestra la existencia de agrupaciones de variables en el conjunto de datos (cuadrados rojos). MEDA nos permite establecer, a nivel general, las relaciones comunes entre variables en nuestra red. A modo de ejemplo, uno de los cuadrados rojos relaciona *logs* de prioridad media en el IDS (`ids_prio2`) reportando intentos de robo de información (`ids_leak`) en el firewall (`ids_ipfwhq`) utilizando el protocolo VNC (`ids_lvnc`).

MEDA nos da una idea de eventos de seguridad comunes en nuestra red, pero no incorpora información temporal. La evolución temporal se analiza con los gráficos de evolución, Figura 4, donde las posibles anomalías se identifican como los

Tabla II
EJEMPLOS DE PARÁMETROS UTILIZADOS Y VALORES ASOCIADOS.

Parámetro	Campo	Valor(es) asociado(s)
fw_syscritical	Prioridad syslog	Critical
fw_syserror	Prioridad syslog	Error
fw_as37	Código mensaje	asa-3-710003
fw_pshell	Puerto	514
ids_ipfwr	Dirección IP	10.32.0.100 o 172.25.0.1
ids_iplog	Dirección IP	172.23.0.2
ids_misc	Clasificación	Misc. activity

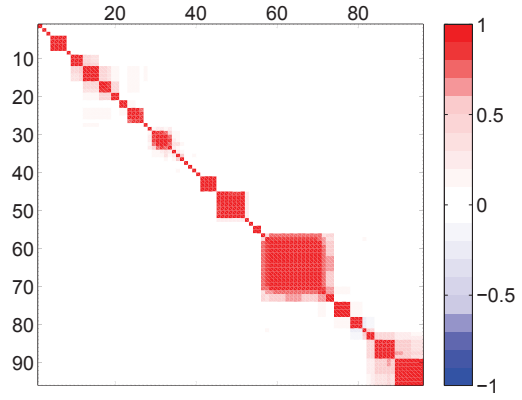

Figura 3. MEDA para todo el conjunto de datos.

valores más altos.

Las observaciones seleccionadas como anomalías son analizadas, bien en solitario, bien en grupos de observaciones consecutivas, para determinar las variables que hacen de ellas valores anómalos. A este fin, como se indicó en la Sección III, se obtienen gráficos oMEDA.

Para ilustrar el procedimiento, consideraremos el gráfico oMEDA de las observaciones {1,11,13,15} (Figura 5). A partir de esta gráfica se identifican dos parámetros con valores muy altos: *fw_iplog* (variable 54) y fw_syslog (variable 55). De acuerdo a esta información, el administrador puede concluir que las anomalías se encuentran relacionadas con el puerto *syslog* en el servidor de trazas *fw_iplog*. En el caso de las observaciones {374, 375}, dos de las que mayores valores proporcionan en las líneas de tiempo, se identifican de forma análoga las variables *ids_lssh, ids_pssh, fw_ptelnet, ids_limap, ids_lpop3, ids_leak, ids_ipfwhq* e *ids_prio2* como asociadas a la anomalía. Esto apunta a la existencia de problemas relacionados con intentos de acceso o fuga de información en los servicios SSH, IMAP y POP. El análisis manual de las trazas para el periodo de tiempo asociado a las observaciones nos lleva a la conclusión de que en este periodo se producen escaneos e intentos reiterados de acceso en los puertos correspondientes, lo que resulta coherente con la información proporcionada por el sistema.

IV-C. Resultados

A partir de la información obtenida en los pasos previos, tanto a nivel de observaciones a supervisar como de las variables implicadas en cada caso, se ha procedido a la

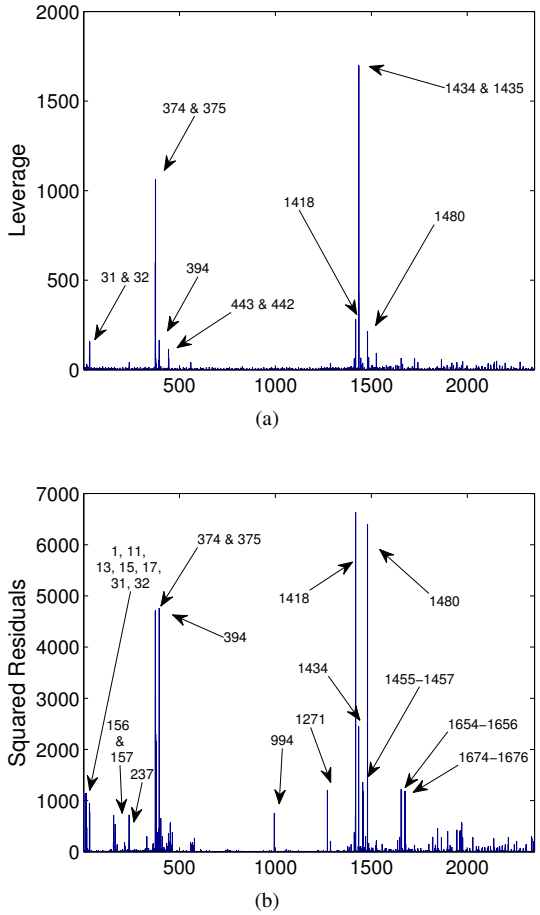

Figura 4. Evolución temporal de *leverage* (a) y residuo (b).

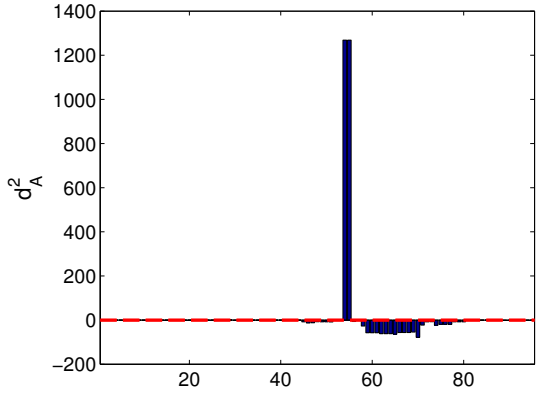

Figura 5. oMEDA para las observaciones {1,11,13,15}

inspección e intrepretación de los registros asociados a dichas observaciones, tanto en las trazas de IDS como de cortafuegos. Los detalles de dicho análisis exceden los objetivos del presente artículo, por lo que a continuación nos limitaremos a relacionar los incidentes encontrados y a comparar nuestros hallazgos con los de otros autores participantes en el reto, incluyendo los ganadores [14] (Tabla III).

Como resultado del análisis realizado, se han identificado

Tabla III
INCIDENTES DE SEGURIDAD REPORTADOS POR NUESTRO SISTEMA Y POR LOS DE OTROS AUTORES.

Anomalía	Propuesta	[14]	[15]	[16]
Ataques DNS/Controlador	X			
Intentos de intrusión al cortafuegos	X	X	X	
Tráfico FTP hacia nodos externos	X			X
Actividad IRC	X	X	X	
Errores en trazas	X			

las siguientes actividades sospechosas, para las que también se han obtenido los intervalos de actividad:

- Ejecución remota no interactiva.
- *DNS spoofing* hacia el servidor DNS y el controlador de dominio. Esta actividad se circunscribe a la red interna.
- Escaneo de puertos en el cortafuegos.
- Ataques de *buffer overflow* y denegación de servicio hacia el servidor DNS y el controlador de dominio.
- Actividad IRC continuada.

Adicionalmente, algunas de las anomalías encontradas han resultado en la constatación de errores de formato en los archivos de traza en un volumen no despreciable. Estos errores no habían sido informados por ninguno de los autores participantes en el reto ni en publicaciones posteriores.

Se puede comprobar (Tabla III) que el sistema propuesto ha permitido identificar no sólo los eventos previamente hallados por otros autores, sino algunos nuevos (ataques DNS y errores en las trazas) a partir del análisis de un reducido número de observaciones seleccionadas por el mismo. Adicionalmente, dicho análisis ha sido realizado de forma dirigida, focalizando la atención en las variables sugeridas a partir de la metodología propuesta.

V. CONCLUSIÓN

En este trabajo se ha propuesto un sistema para la mejora de las prestaciones de los NSM existentes en tres aspectos clave: la integración y parametrización de la información procedente de diversas fuentes heterogéneas, la selección automática de los incidentes más relevantes y la incorporación de información semántica al proceso de análisis. Cada una de estas contribuciones resulta relevante, ya que facilitan la tarea de los administradores de seguridad durante el proceso de monitorización y verificación de las alertas generadas por los sistemas automáticos, que pueden resultar muy numerosas y, consecuentemente, inmanejables.

La metodología propuesta ha mostrado una gran capacidad para dirigir al administrador hacia los incidentes relevantes y su interpretación. La evaluación realizada sobre el reto VAST12 ha permitido identificar todos los incidentes reportados hasta la actualidad en dicho reto, así como algunos que no habían sido detectados.

El sistema se encuentra actualmente implementado en laboratorio a nivel de realización de los análisis PCA y la obtención de las diferentes gráficas de forma no integrada, esto es, se requiere de la intervención del administrador en cada paso para ejecutar y proporcionar las entradas a cada módulo. Consecuentemente, una de las líneas de trabajo futuro debe centrarse en la integración de todas las herramientas en un NMS de fácil uso, automatizando el sistema y posibilitando el acceso a los datos originales a partir de los hallazgos del mismo para su inspección por parte del administrador.

AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por el MICINN a través del proyecto TEC2011-22579.

REFERENCIAS

[1] R. Bejtlich, "The Tao of Network Security Monitoring", *Addison-Wesley*, 2004.
[2] P. García-Teodoro, J. Díaz-Verdejo, G. Maciá-Fernández, E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges", *Computers & Security*, 28(1):18–28, 2009.
[3] G. Keren, C. Lewis, "A Handbook for data analysis in the behavioral sciences: statistical issues,", *L. Erlbaum*, 1993.
[4] "Vast challenge 2012", http://www.vacommunity.org/vast+challenge+2012.
[5] I.T. Jollifee, "Principal component analysis", *Springer-Verlag*, 2002.
[6] P. Geladi, B.R. Kowalski, "Partial Least-Squares Regression: a tutorial", *Analytica Chimica Acta*, 185:1–17, 1986.
[7] J. Camacho, "Missing-data theory in the context of exploratory data analysis," *Chemometrics and Intelligent Laboratory Systems,* 103:8–18, 2010.
[8] J. Camacho, "Observation-based missing data methods for exploratory data analysis to unveil the connection between observations and variables in latent subspace models," *Journal of Chemometrics,* 25(11):592–600, 2011.
[9] J. Camacho, A. Ferrer,"Cross-validation in {PCA} models with the element-wise k-fold (ekf) algorithm: Practical aspects," *Chemometrics and Intelligent Laboratory Systems,* 131:37–50, 2014.
[10] J.E. Jackson, "A User's Guide to Principal Components," *Wiley*, 2003.
[11] G. Caraux and S. Pinloche, "Permutmatrix: a graphical environment to arrange gene expression profiles in optimal linear order," *Bioinformatics*, 21(7):1280–1, 2005.
[12] J. Camacho, "EDA toolbox", disponible en http://wdb.ugr.es/~josecamacho/downloads.php, 2013.
[13] R. Marty, "Applied Security Visualization," *Pearson Education*, 2008.
[14] F. Fischer, J. Fuchs, F. Mansmann, D. A. Keim, "Banksafe: A visual situational awareness tool for large-scale computer networks: Vast 2012 challenge award: Outstanding comprehensive submission, including multiple vizes," en *Proc. IEEE VAST*, pp. 257–258, IEEE Computer Society, 2012.
[15] Y. Cao, R. Moore, P. Mi, A. Endert, C. North, R. C. Marchany, "Dynamic analysis of large datasets with animated and correlated views: Vast 2012 mini challenge 2 award: Honorable mention for good use of coordinated displays.'", en *Proc. IEEE VAST*, pp. 283–284, IEEE Computer Society, 2012.
[16] L. Shi, Q. Liao, C. Yang, "Investigating network traffic through compressed graph visualization: Vast 2012 mini challenge 2 award: good adaptation of graph analysis techniques," en *Proc. IEEE VAST*, pp. 279–280, IEEE Computer Society, 2012.