

ESCUELA COLOMBIANA DE INGENIERÍA JULIO GARAVITO

SIEM: Plataforma para la detección de ataques informáticos en tiempo real a partir de eventos de seguridad

Realizado por

María Alejandra Blanco Uribe, Nicolás Gómez Solano y Edwin
Alexander Cerón Sánchez

Directores

Daniel Orlando Díaz López y Claudia Patricia Santiago Cely

Programa de Ingeniería de Sistemas

13 de diciembre de 2017

ESCUELA COLOMBIANA DE INGENIERA JULIO GARAVITO

Resumen

Programa de Ingeniería de Sistemas

Realizado por María Alejandra Blanco Uribe, Nicolás Gómez Solano y Edwin Alexander Cerón Sánchez

Este libro presenta el desarrollo del proyecto de grado "*Plataforma para la detección de ataques informáticos en tiempo real a partir de eventos de seguridad*". El libro tiene dos capítulos principales correspondientes a las dos fases desarrolladas durante la asignatura Proyecto de Grado, el primer capítulo Event Management describe el proceso de implementación de una herramienta SIEM y su configuración, el segundo capítulo muestra el proceso de aplicar ciencia de datos para la generación de modelos matemáticos de predicción sobre los eventos recolectados por herramientas SIEM.

Índice general

Lista de Tablas	V
Lista de Figuras	VI
Abreviaciones	XIV
Introducción	XV
1. Definición del proyecto.....	1
1.1 Objetivo General	1
1.2 Objetivos específicos.....	1
1.3 Logros	2
1.4 Productos y resultados	2
2. Desarrollo del proyecto	3
2.1 Event Management.....	4
2.1.1 Instalación OSSIM.....	5
2.1.2 Configuración OSSIM.....	13
2.1.3 Caracterización de activos	45
2.1.4 Definición de alertas y reglas de correlación.....	59
2.1.5 Definición de respuestas(10)	72
2.1.6 Escenarios aplicados	80
2.2 Data science applied in Cybersecurity	91
2.2.1 Conceptos clave	92
2.2.2 Caso 1: Sistema de Información del Laboratorio de Informática	98

2.2.3 Caso 2: Comando Conjunto Cibernético de las Fuerzas Militares.....	126
Conclusiones	141
Bibliografía	143

Lista de Tablas

Tabla 1 Requerimientos técnicos de OSSIM 5.3.4	5
Tabla 2 Tipos de reportes en OSSIM	26
Tabla 3 Nivel de severidad para tickets.....	31
Tabla 4 Clasificación de Directivas de correlación en OSSIM	60
Tabla 5 Propiedades Globales de las directivas de correlación en OSSIM.....	62
Tabla 6 Propiedades Globales de las reglas de correlación en OSSIM	64
Tabla 7 Tipos de eventos que son procesados en OSSIM	74
Tabla 8 Atributos de una política en OSSIM- Condiciones.....	75
Tabla 9 Atributos de una política en OSSIM- Consecuencias.....	76
Tabla 10 Atributos de eventos recibidos	127

Lista de Figuras

Ilustración 1 Cuadrante mágico de Gartner para tecnologías SIEM.	4
Ilustración 2 Máquina virtual OSSIM con especificaciones mínimas,	6
Ilustración 3 Pantalla de inicio de instalación de AlienVault OSSIM 5.3.4.....	7
Ilustración 4 Pantalla de configuración de lenguaje.....	8
Ilustración 5 Pantalla de selección de ubicación.....	9
Ilustración 6 Pantalla de configuración de teclado.....	9
Ilustración 7 Pantalla de configuración de red.	10
Ilustración 8 Pantalla de configuración de máscara de red.	10
Ilustración 9 Pantalla de configuración de Gateway	11
Ilustración 10 Pantalla de configuración de DNS.....	11
Ilustración 11 Pantalla de configuración de red	12
Ilustración 12 Pantalla de configuración de contraseñas	13
Ilustración 13 Interfaz web, mensaje de sitio inseguro.	14
Ilustración 14 Interfaz web. Formulario de creación de usuario administrador.	15
Ilustración 15 Interfaz web. Pantalla de inicio de sesión.	15
Ilustración 16 Interfaz web. Asistente de configuración inicial.	16
Ilustración 17 Interfaz web. Pantalla de configuración de interfaces de red.....	17
Ilustración 18 Interfaz web. Pantalla de descubrimiento de activos.....	17
Ilustración 19 Interfaz web. Despliegue de HIDS.....	18
Ilustración 20 Interfaz web. Configuración de dispositivos de red.	19
Ilustración 21 Interfaz web. Integración de OTX.	19
Ilustración 22 Pop up OTX.	20
Ilustración 23 Pantalla final del asistente de configuración inicial.....	20
Ilustración 24 Ventana de Dashboards OSSIM	21

Ilustración 25 Vista de alarmas OSSIM.....	22
Ilustración 26 Ventana de eventos de seguridad SIEM OSSIM	22
Ilustración 27 Ventana de tickets OSSIM.....	23
Ilustración 28 Vista de Activos y Grupos OSSIM.....	24
Ilustración 29 Vista de Vulnerabilidades OSSIM.....	24
Ilustración 30 Vista de Flujo de red OSSIM	24
Ilustración 31 Vista de Captura de tráfico OSSIM.....	25
Ilustración 32 Vista de Disponibilidad OSSIM	25
Ilustración 33 Vista de Detección OSSIM	25
Ilustración 34 Vista de Reportes OSSIM	27
Ilustración 35 Vista de Administración OSSIM.....	28
Ilustración 36 Vista de Despliegue OSSIM	28
Ilustración 37 Vista de Inteligencia de Amenazas OSSIM.....	28
Ilustración 38 Vista de configuración de OTX OSSIM	29
Ilustración 39 Página de OTX. Sección de API.....	30
Ilustración 40 Ventana de configuración de OTX.....	30
Ilustración 41 Configuración de límite para tickets	31
Ilustración 42 Pantalla de administración de OSSIM.....	32
Ilustración 43 Consola de configuración OSSIM.....	32
Ilustración 44 Edición de archivo para activar SNMP.....	33
Ilustración 45 Ventana de reconfiguración de OSSIM.....	33
Ilustración 46 Menú de configuración de usuarios.....	34
Ilustración 47 Formulario de creación de usuarios.....	34
Ilustración 48 Vista de Activos y Grupos.....	35
Ilustración 49 Vista de Activos y Grupos.....	36
Ilustración 50 Vista de Activos y Grupos.....	36
Ilustración 51 Formulario para agregar un activo manualmente.....	37

Ilustración 52 Descarga de OSSEC 2.8 vía consola.	38
Ilustración 53 Archivos dentro de la carpeta descomprimida de OSSEC.	39
Ilustración 54 Comando de Instalación de OSSEC	39
Ilustración 55 Opciones de lenguaje en el menú de instalación de OSSEC	39
Ilustración 56 Información sobre la instalación de OSSEC	40
Ilustración 57 Opciones de instalación servidor, agente, local o híbrido.	40
Ilustración 58 Ruta de instalación de OSSEC	40
Ilustración 59 Configuración de redirección de logs hacia el SIEM.	41
Ilustración 60 Habilitación de servicios dentro del sistema.	41
Ilustración 61 Instalación exitosa de OSSEC	42
Ilustración 62 Vista de agentes conectados al sistema	43
Ilustración 63 Formulario para agregar un nuevo agente a OSSIM	43
Ilustración 64 Llave de conexión con la plataforma OSSIM	44
Ilustración 65 Menú del script de administración de agentes	44
Ilustración 66 Scripts para la administración de OSSEC.....	44
Ilustración 67 Opción de insertar llave generada desde OSSIM.....	44
Ilustración 68 Inicio del agente HIDS dentro del servidor.....	45
Ilustración 69 Vista detallada del activo.	58
Ilustración 70 Vista de activos conectados al sistema.....	58
Ilustración 71 Formulario de edición del activo.	59
Ilustración 72 Relación entre reglas de Correlación	60
Ilustración 73 Vista de directivas de Correlación	63
Ilustración 74 Formulario para creación de Directivas de Correlación.....	64
Ilustración 75 Formulario para creación de reglas de correlación.....	66
Ilustración 76 Selección de eventos por tipo.....	67
Ilustración 77 Selección de eventos por taxonomía	67
Ilustración 78 Selección de tipos de eventos para la regla de correlación creada.	68

Ilustración 79 Selección de aspectos relacionados a la red.....	68
Ilustración 80 Selección de aspectos relacionados a la reputación de la IP.	69
Ilustración 81 Ventana emergente que aparece cuando terminamos de definir la regla de correlación.....	69
Ilustración 82 Regla de correlación definida	69
Ilustración 83 Creación de regla de nivel 2	70
Ilustración 84 Creación de regla de nivel 2. Configuraciones de red.....	70
Ilustración 85 Creación de regla de nivel 2. Definición del nivel de confianza.	71
Ilustración 86 Creación de regla de nivel 2. Modificación en la ocurrencia.	71
Ilustración 87 Creación de regla de nivel 2. Campos adicionales en la regla de correlación.	71
Ilustración 88 Creación de reglas de correlación. Recargar directivas.	72
Ilustración 89 Vista de configuración	73
Ilustración 90 Vista de políticas de OSSIM	74
Ilustración 91 Visualización de una política.....	75
Ilustración 92 Vista de acciones	78
Ilustración 93 Vista de políticas señalando la acción 'New'	78
Ilustración 94 Vista de una nueva política.....	79
Ilustración 95 Vista de consecuencias	79
Ilustración 96 Estado de recarga de políticas OSSIM	80
Ilustración 97 Vista de políticas señalando la acción 'Reload Policies'.....	80
Ilustración 98 Comprobación de la creación de la política	80
Ilustración 99 Esquema básico de un ataque distribuido	81
Ilustración 100 Esquema de reglas de correlación anidadas para el escenario 1 ..	81
Ilustración 101 Esquema de reglas de correlación anidadas vistas en OSSIM	82
Ilustración 102 Ataque de autenticación SSH por fuerza bruta.....	82
Ilustración 103 Alertas creadas por OSSIM al detectar el ataque.....	83

Ilustración 104 Detalle de una alerta proveniente de las reglas anteriormente creadas	83
Ilustración 105 Definición de políticas para el escenario 1	84
Ilustración 106 Configuración de la respuesta para el escenario 1	85
Ilustración 107 Ataque de denegación de servicio distribuido.....	86
Ilustración 108 Ilustración de las reglas de correlación anidada para el escenario 2	86
Ilustración 109 Alerta visualizada en OSSIM	86
Ilustración 110 Detalle de la alerta del escenario 2	87
Ilustración 111 Detalle de la respuesta del escenario 2	88
Ilustración 112 Detección de inicio de sesión sospechoso	89
Ilustración 113 Intervención del código de la aplicación para enviar el evento de inicio de sesión.....	90
Ilustración 114 Evento detallado del inicio de sesión del escenario 3	90
Ilustración 115 Ciclo de vida de la ciencia de los datos	93
Ilustración 116 Vista de Exploración de Weka	95
Ilustración 117 Fórmula del error absoluto medio.....	96
Ilustración 118 Fórmula de la raíz del error cuadrático medio	96
Ilustración 119 Fórmula del error absoluto relativo	97
Ilustración 120 Fórmula del error relativo al cuadrado.....	97
Ilustración 121 Explicación gráfica del coeficiente de correlación	98
Ilustración 122 Escenario idóneo de acceso al Sistema.	99
Ilustración 123 Escenario real de acceso al Sistema	100
Ilustración 124 Vista de eventos de seguridad del SIEM	101
Ilustración 125 Menú de exportación de eventos. Ventana de espera.	102
Ilustración 126 Menú de exportación de eventos.....	102
Ilustración 127 Ventana de archivo del navegador.....	102

Ilustración 128 Cantidad de eventos por hora laboratorio de informática	103
Ilustración 129 Top 10 direcciones IP con más peticiones al servidor del laboratorio de informática	103
Ilustración 130 Cantidad de peticiones a cada recurso web de la página del laboratorio de informática	104
Ilustración 131 Top 10 países que realizan más peticiones al la página del laboratorio de informática	104
Ilustración 132 Registros depurados – peticiones legitimas	105
Ilustración 133 Registros depurados – cantidad de eventos por hora	106
Ilustración 134 – Traducción eventos	107
Ilustración 135 Etiquetado manual indicanto si el comportamiento es legitimo....	108
Ilustración 136 Atributos de registros de comportamiento legitimo	108
Ilustración 137 Traducción archivo comportamiento legitimo	109
Ilustración 138 Interfaz inicial WEKA.....	109
Ilustración 139 Interfaz de selección de archivo WEKA	110
Ilustración 140 Selección de archivo WEKA	110
Ilustración 141 Selección de modelo matemático WEKA.....	111
Ilustración 142 Interfaz de configuración de modelo matematico WEKA	112
Ilustración 143 Opciones del clasificador WEKA	113
Ilustración 144 Interfaz selección formato de salida WEKA	114
Ilustración 146 Pantalla de resultados WEKA.....	116
Ilustración 147 Resultado del modelo MP5 aplicado a los datos de eventos por hora	116
Ilustración 148 Resultado del modelo Bagging aplicado a los datos de eventos por hora.....	117
Ilustración 149 Resultado del modelo REPTree aplicado a los datos de eventos por hora.....	117

Ilustración 150 Resultado del modelo NaiveBayes aplicado a los datos de comportamiento anómalo	118
Ilustración 151 Resultado del modelo DecisionTable aplicado a los datos de comportamiento anómalo	118
Ilustración 152 Resultado del modelo Bagging aplicado a los datos de comportamiento anómalo	119
Ilustración 153 Resultado del modelo KStar aplicado a los datos de comportamiento anómalo	119
Ilustración 154 Vista de preprocesamiento de datos	120
Ilustración 155 Vista de clusterización de WEKA	121
Ilustración 156 Ventana de modificación de atributos	122
Ilustración 157 Resultado de clusterización de eventos SSH.	123
Ilustración 158 Menú de opciones sobre los modelos generados	124
Ilustración 159 Gráfica de distribución de clústeres según puerto y fecha.	124
Ilustración 160 Resultado clusterización con SimpleKMeans	125
Ilustración 161 Formula matemática para calcular la variable Relevance	130
Ilustración 162 Formula matemática para calcular la variable Priority	130
Ilustración 163 Resultado del modelo DesicionStump aplicado a los datos de comportamiento sospechoso	131
Ilustración 164 Resultado del modelo J48 aplicado a los datos de comportamiento sospechoso	132
Ilustración 165 Resultado del modelo NaiveBayes aplicado a los datos de comportamiento sospechoso	132
Ilustración 166 Resultado del modelo SMO aplicado a los datos de comportamiento sospechoso	132
Ilustración 167 Resultado del modelo Kstar aplicado a los datos de comportamiento sospechoso	133

Ilustración 168 Pantalla principal de la aplicación.....	133
Ilustración 169 Diagrama de Componentes de la aplicación desarrollada	135
Ilustración 170 Diálogo para escoger archivo para realizar el análisis.	136
Ilustración 171 Diálogo de delimitador del programa.....	137
Ilustración 172 Ventana de análisis	137
Ilustración 173 Finalización del análisis	138
Ilustración 174 Vista de reentrenamiento del modelo.....	139
Ilustración 175 Salida de reentrenamiento del modelo.....	140

Abreviaciones

ECI Escuela Colombiana de Ingeniería Julio Garavito

SIEM Security Information and Event Management

HIDS Host Intrusion Detection System

MINTIC Ministerio de Tecnologías de la Información y la
Comunicación

OTX Open Threat Exchange

IoC Indicators of Compromise

TDS Team Data Science Process

Introducción

Este libro tiene como propósito mostrar el paso a paso del trabajo realizado en el proyecto de grado *Plataforma para la detección de ataques informáticos en tiempo real a partir de eventos de seguridad*, el cual tuvo origen en la idea de brindar una herramienta SIEM de bajo costo para la administración y manejo de eventos de seguridad, que apoye el cumplimiento de la norma MSPI (Modelo de seguridad y privacidad de la información) del MINTIC, la cual a su vez se basó en la norma ISO 27001, esta última específicamente en el dominio A. 12 nos habla de la seguridad en las operaciones, este dominio tiene tres subdominios en los que ésta enfocado nuestro proyecto los cuales son registro de eventos, protección de la información de registro, registros del administrador y el operador.

El proyecto durante su realización tuvo dos grandes fases.

En la primera fase del proyecto se realizó la implementación de la plataforma SIEM llamada OSSIM, la cual es un proyecto open source provisto por Alien Vault. En esta fase se abordaron temas técnicos como la instalación y configuración de la plataforma, caracterización de los activos, la definición de las diferentes reglas de correlación para la detección de ataques informáticos y por último la definición de respuestas para cada una de las reglas de correlación creadas.

En la segunda fase del proyecto se tomó un rumbo más investigativo producto de los retos encontrados en la primera fase del proyecto, como el gran volumen de eventos de seguridad que llegaban a la plataforma, los atributos y propiedades de cada evento de seguridad, la diversidad de los eventos, entre otros. Es por esta razón que se decidió realizar Data Science sobre los eventos de seguridad que arrojaba la plataforma, siguiendo la metodología TDS (Team Data Science Lifecycle Process) dividida en 4 fases entendimiento del negocio, adquisición y entendimiento de los datos, modelado y despliegue; este proceso se siguió en dos entornos uno académico en el Laboratorio de Informática y otro en un entorno real con el Comando Conjunto Cibernético de las Fuerzas Militares de Colombia.

Estos dos procesos se encuentran documentados a continuación y se muestran de la siguiente manera:

- En el primer capítulo se define el proyecto presentando la justificación, objetivos, logros, productos y resultados.

- En el segundo capítulo se presenta la instalación y configuración de la herramienta OSSIM de Alien Vault, subdividido en tres componentes, sistemas operativos, software base y aplicaciones.
- En el tercer capítulo se presenta el proceso de implementación de data science aplicado a eventos de seguridad, el modelo de implementación que se desarrolló sus diferentes etapas y la aplicación en el laboratorio y en el CCOC.
- Por último, se dan las conclusiones del trabajo realizado y se exponen los trabajos futuros.

1. Definición del proyecto.

1.1 Objetivo General

Generar una propuesta de sistema de gestión de eventos de seguridad de la información que permita detectar oportunamente la existencia de incidentes de seguridad asociados a activos de información de alta criticidad de una empresa, y permita de esta forma mitigar los riesgos asociados y reducir el impacto operativo, reputacional o legal.

1.2 Objetivos específicos

1. Desplegar una solución de gestión de eventos que permita recolectar eventos de seguridad de diferentes tipos de activos tecnológicos sobre la cual se puedan realizar las pruebas de las mejoras propuestas en la investigación.
2. Generar habilidades técnicas en la configuración y administración de los componentes de una solución típica de gestión de eventos de seguridad que permitan realizar una efectiva gestión de riesgos en un ambiente real.
3. Caracterizar los activos de información críticos comunes en una empresa a nivel de atributos de seguridad (confidencialidad, integridad y disponibilidad) y contexto de riesgos sobre los mismos.
4. Proponer un conjunto de alertas y reglas de correlación que identifiquen de una forma oportuna la existencia de riesgos de seguridad en un contexto empresarial con múltiples activos de información y vectores de ataque.
5. Aplicar los conceptos de ciencias de los datos para apoyar la gestión de eventos de seguridad a través de modelos predictivos y descriptivos.

1.3 Logros

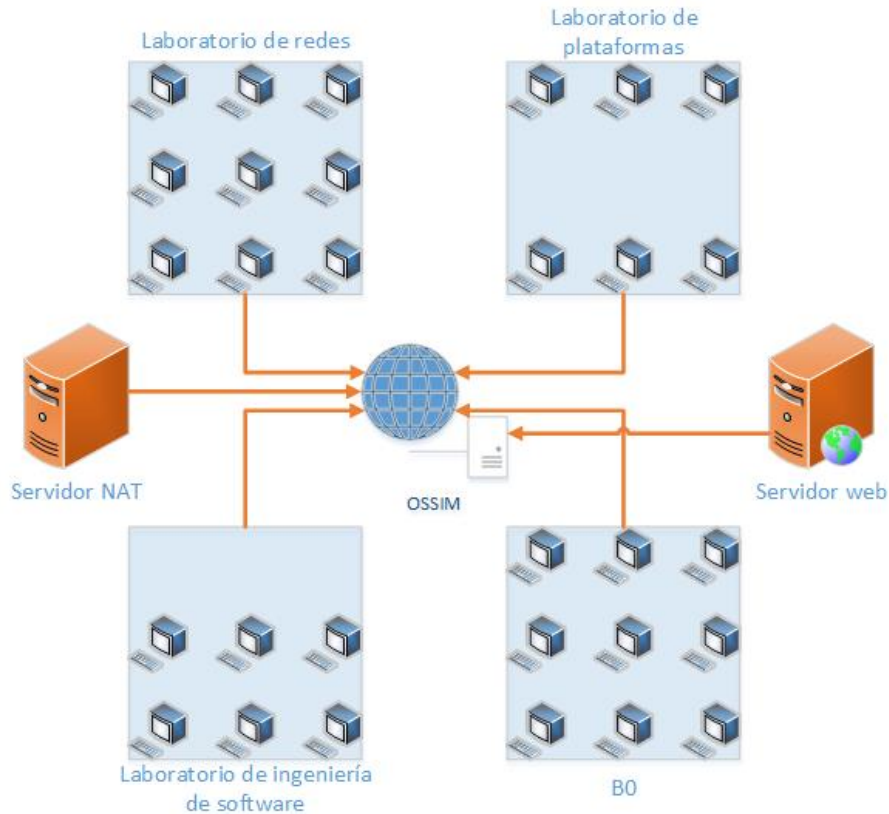
1. Entendimiento del contexto de negocio del Laboratorio de Informática con el fin de identificar problemáticas a resolver con Data Science.
2. Obtención y análisis de eventos de seguridad obtenidos de una solución SIEM (Security Information and Event Management) para el sistema de información del Laboratorio de Informática.
3. Desarrollo de un modelo predictivo y descriptivo para el sistema de información del Laboratorio de Informática con el fin de resolver las problemáticas identificadas previamente
4. Aplicación del ciclo de vida de Data Science para la identificación de incidentes en sistemas de información críticos protegidos por el CCOC (Conjunto Comando Cibernético) de las Fuerzas Armadas de Colombia.

1.4 Productos y resultados

1. 2 artículos de investigación:
 - “Event Management and Information Security: State of Art and Challenges”*
 - “Data Science applied in Cybersecurity”*
2. 2 Implementaciones realizadas
 - a. Entorno académico: Laboratorio de Informática
 - b. Entorno real: Comando Conjunto Cibernético de las Fuerzas Militares
3. Presentación y validación con expertos en seguridad informática en el Security Barcamp 2017. “Data Science Applied to Cybersecurity”.

2. Desarrollo del proyecto

Antes que nada, empecemos entendiendo la arquitectura final de la solución SIEM implementada en el laboratorio de informática la cual recibía eventos de seguridad por medio de plugins o HIDS que serán explicados en el desarrollo del documento, la arquitectura consiste en cuatro laboratorios de informática y dos servidores uno web (laboratoriois) y uno NAT (coral).



Comentado [CPSC1]: Existe un punto 1, pero no existe, 2 o más.

Comentado [CPSC2]: Creo que antes de Event Management debería hablar de la infraestructura que se debe proteger, algo como un diagrama, con su respectiva explicación, en donde se incluyan redes y equipos, la generación de logs que se pueden tener y la manera como OSSIM apoyo.

Comentado [CPSC3]: No vi en el desarrollo del documento la manera como se dividieron sistemas operativos, aplicación y plataforma. Y lo que hicieron en las tres y las herramientas que usaron.

2.1 Event Management

Alienvault OSSIM (Open Source Security Information and Event Management) es una herramienta SIEM de código abierto que permite recolectar, normalizar y correlacionar eventos además de realizar descubrimiento de activos, escaneo de vulnerabilidades, detección de intrusos y monitoreo de comportamiento. La empresa Alienvault se encuentra dentro del cuadrante de jugadores de nicho, acercándose a la zona de retadores, en la primera fase de nuestro proyecto de grado se realizó la instalación y configuración de una herramienta SIEM para el Laboratorio de Informática de la Decanatura de Ingeniería de Sistemas de la ECI, el proceso se encuentra documentado a continuación.



Ilustración 1 Cuadrante mágico de Gartner para tecnologías SIEM.

2.1.1 Instalación OSSIM

En esta sección se presenta el procedimiento para la instalación de OSSIM, como primera medida se incluirán los requerimientos técnicos y posteriormente el paso a paso de la instalación.

2.1.1.1 Requerimientos técnicos para instalación en máquina virtual:

Tabla 1 Requerimientos técnicos de OSSIM 5.3.4

Requerimientos técnicos mínimos para la instalación	
Versión de OSSIM utilizada	5.3.4
Número de núcleos	8
RAM(GB)	16 u 8 (pero tendrá un rendimiento muy bajo)
Almacenamiento	500GB
Entorno de virtualización	VMware ESXi 4.x, 5.x, 6.x Hyper-V v3.0+ (Windows Server 2008 SP2 y superior)

2.1.1.2 Proceso de Instalación

1. Crear una máquina virtual utilizando cualquier herramienta de virtualización (Vmware Workstation Player 12.5.8 en mi caso) con los requerimientos mínimos presentados anteriormente.

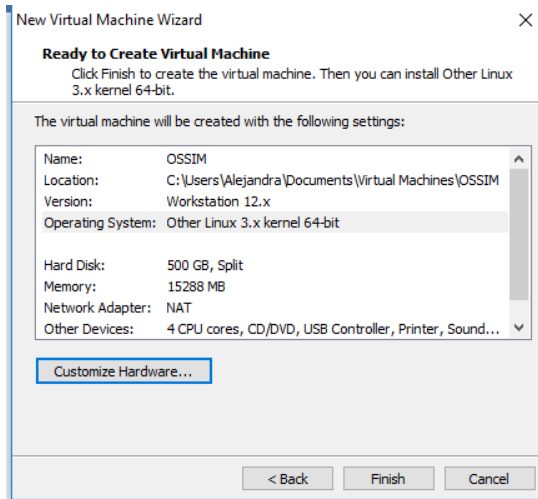


Ilustración 2 Máquina virtual OSSIM con especificaciones mínimas,

2. Al iniciar la máquina virtual nos aparece la pantalla de la ilustración 3, donde deberemos escoger la primera opción "Install AlienVault OSSIM 5.3.4 (64



Ilustración 3 Pantalla de inicio de instalación de AlienVault OSSIM 5.3.4

- bit)".
3. Escogemos el idioma utilizado en la instalación, en nuestro caso escogemos idioma inglés.

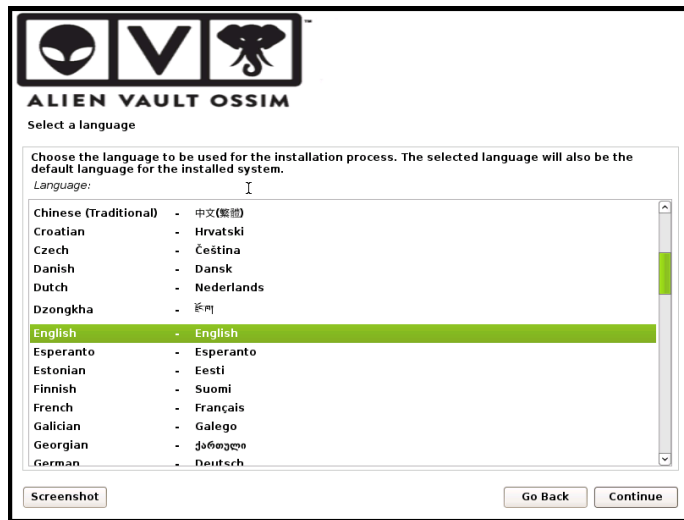


Ilustración 4 Pantalla de configuración de lenguaje

4. En la siguiente pantalla escogemos la ubicación en nuestro caso será: "other->South America->Colombia".

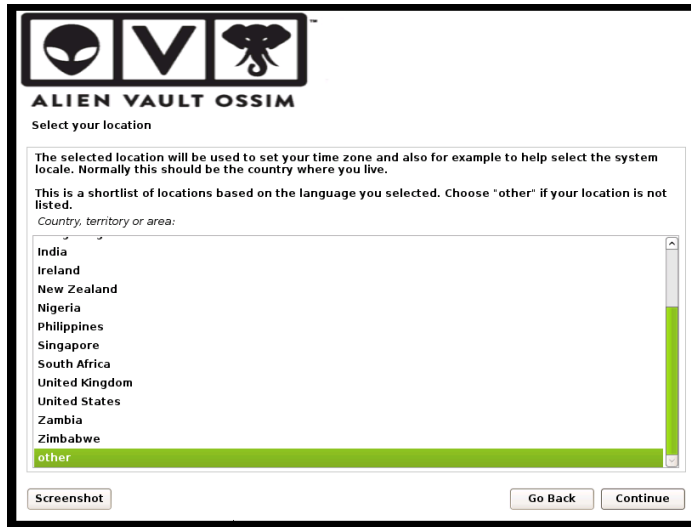


Ilustración 5 Pantalla de selección de ubicación

5. La siguiente pantalla nos pedirá la configuración de teclado, en nuestro caso escogeremos "Latin American".

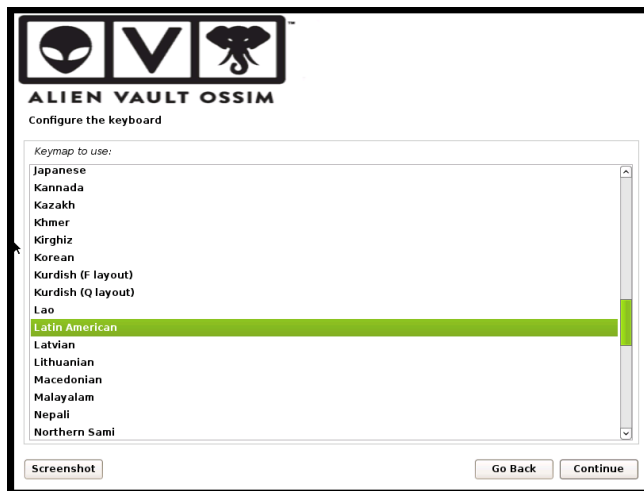


Ilustración 6 Pantalla de configuración de teclado

- Después del paso anterior, comenzará la configuración para el proceso de instalación, y posteriormente tendremos que realizar la configuración de la red. Nos solicita la IP, que en el caso del Laboratorio de Informática es 10.2.78.8 (Revisar con los administradores del Laboratorio).

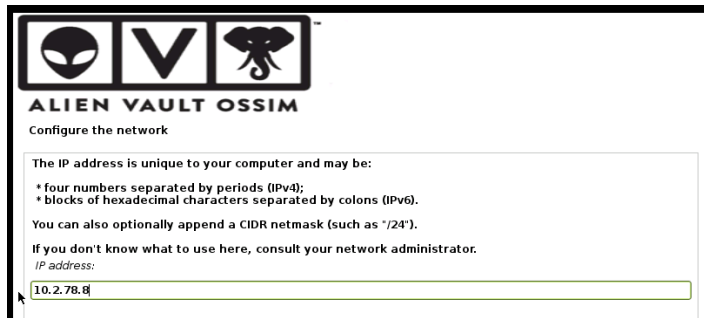


Ilustración 7 Pantalla de configuración de red.

- Después de ingresada la IP nos solicitará la máscara de red, que en el caso del Laboratorio de Informática es 255.255.0.0

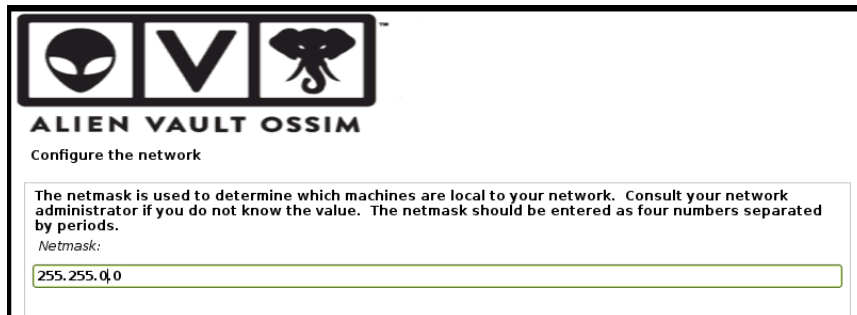


Ilustración 8 Pantalla de configuración de máscara de red.

- Después de ingresar la máscara de red, ingresamos el Gateway, que en nuestro caso es el 10.2.65.1

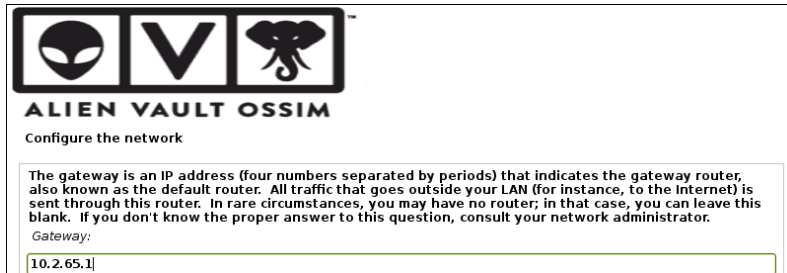


Ilustración 9 Pantalla de configuración de Gateway

9. En la siguiente pantalla nos pedirá el DNS, en nuestro caso es el 10.2.65.2, cabe anotar que todas las configuraciones de red se deben validar con los administrativos del Laboratorio de Informática.

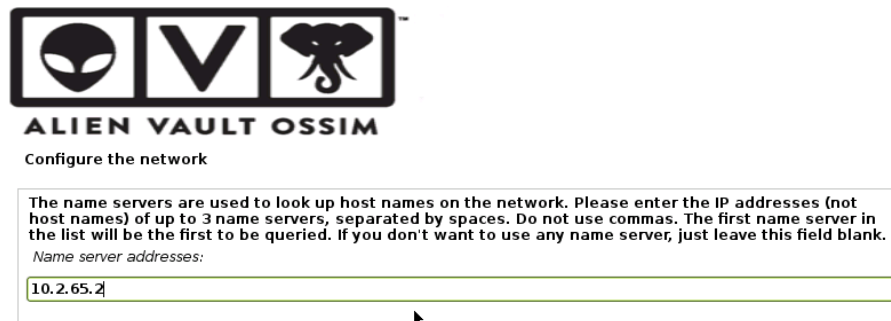


Ilustración 10 Pantalla de configuración de DNS

10. A continuación, nos aparecerá la ventana de configuración de red, debe estar atento a cualquier error que surja.



Ilustración 11 Pantalla de configuración de red

11. Después de configurada la red debemos ingresar la contraseña de root del sistema.

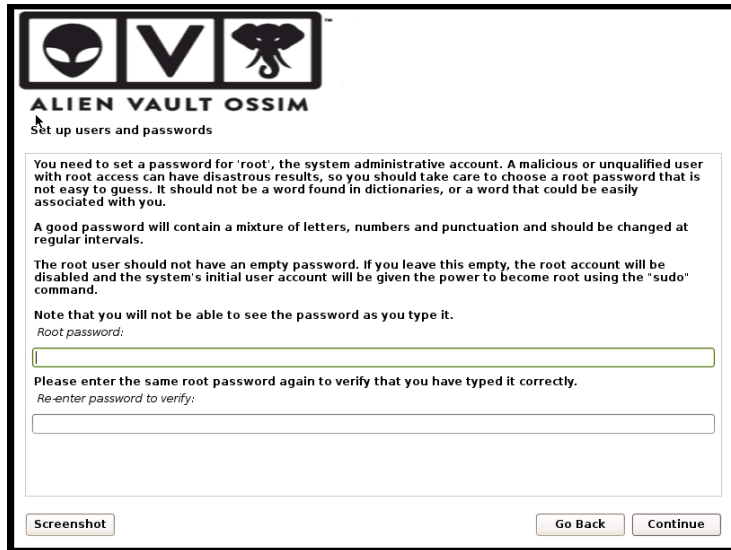


Ilustración 12 Pantalla de configuración de contraseñas

12. Después de ingresada la contraseña comenzará la instalación del sistema, es importante vigilar la instalación ya que puede tener errores asociados a versiones o requerimientos de memoria. Después de instalado el sistema podremos acceder a la interfaz ingresando a la url: <https://10.2.78.8> y en la consola aparecerá el login.

Y con esto finaliza la instalación del sistema.

2.1.2 Configuración OSSIM

Una vez instalado el sistema nos dirigimos a la interfaz web para realizar la configuración inicial de OSSIM. También realizamos algunas configuraciones al sistema como agregar usuarios, entre otras, que se mostrarán en el desarrollo de esta sección.

2.1.2.1 Configuración inicial de OSSIM

1. Al ingresar a la interfaz web de configuración nos aparecerá un mensaje de que nos dirigimos a un sitio web inseguro, en la misma página nos aparece un link de “opciones avanzadas” damos clic allí y elegimos la opción “Continuar a sitio no seguro”:

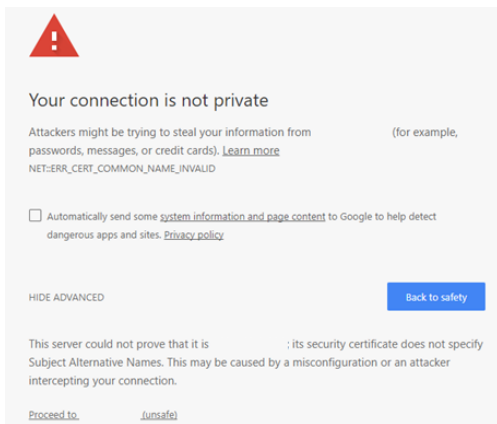


Ilustración 13 Interfaz web, mensaje de sitio inseguro.

2. Al proceder a la interfaz web de OSSIM nos aparecerá un formulario para registrar la instalación realizada y configurar el usuario administrador de la interfaz web, cabe anotar que este usuario es diferente al usuario root del sistema. Llenamos cada uno de los campos y hacemos clic en el botón “Start using AlienVault”.

The screenshot shows a web interface for creating an administrator account. At the top, it says "Welcome" and provides a congratulatory message. Below this is the "Administrator Account Creation" section, which includes a heading "Create an account to access your AlienVault product." and a note that asterisks indicate required fields. The form contains several input fields: "FULL NAME", "USERNAME" (with "admin" pre-filled), "PASSWORD", "CONFIRM PASSWORD", "E-MAIL", "COMPANY NAME", and "LOCATION" (with a "View Map" link). A checkbox at the bottom is checked, indicating consent to share usage statistics. A blue "START USING ALIENVAULT" button is at the bottom center.

Ilustración 14 Interfaz web. Formulario de creación de usuario administrador.

3. Al dar clic al botón nos aparecerá la pantalla de login, ingresamos con las credenciales que configuramos anteriormente y el usuario "admin".

The screenshot shows the login screen for AlienVault OSSIM. It features the AlienVault logo at the top, followed by the text "ALIEN VAULT OSSIM" and "alienvault". Below this are input fields for "USERNAME" (with "admin" pre-filled) and "PASSWORD". A "Forgot Password?" link is located below the password field. A blue "LOGIN" button is positioned at the bottom center of the form.

Ilustración 15 Interfaz web. Pantalla de inicio de sesión.

4. Al ingresar a la interfaz de administración por primera vez, deberemos realizar la configuración inicial del sistema siguiendo el asistente de configuración inicial, allí revisaremos las interfaces de red, los activos

conectados, desplegaremos HIDS y configuraremos los logs en dispositivos de red.

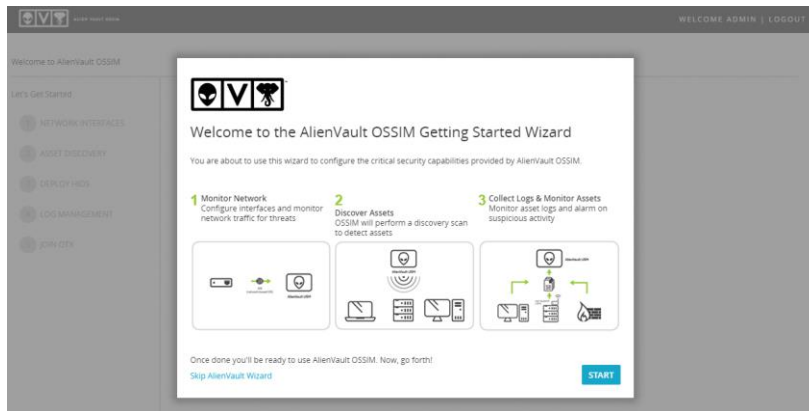


Ilustración 16 Interfaz web. Asistente de configuración inicial.

5. Primero se deben configurar las interfaces de red, que aparecen listadas en la pantalla, cuando tenemos más de una interfaz podemos configurarlas para la administración (acceso a interfaz web) y recibir logs. En este caso se configurará para administración, cuando se configuren todas las interfaces hacemos clic en el botón siguiente.

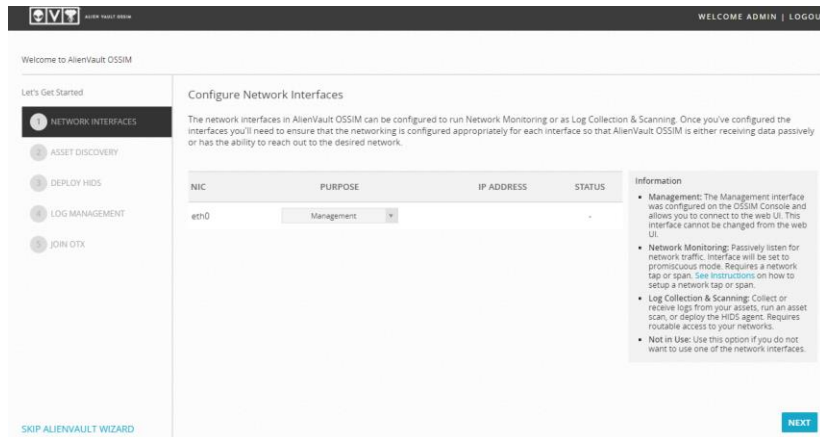


Ilustración 17 Interfaz web. Pantalla de configuración de interfaces de red.

6. En la siguiente pantalla nos mostrará los activos que descubrió en la red, si queremos que realice un nuevo descubrimiento hacemos clic en el botón “Scan networks”, y si el activo que estamos tratando de agregar no se encuentra de forma automática, se debe ingresar en el formulario “Add Asset Manually” hostname (nombre del activo), IP y tipo de activo.

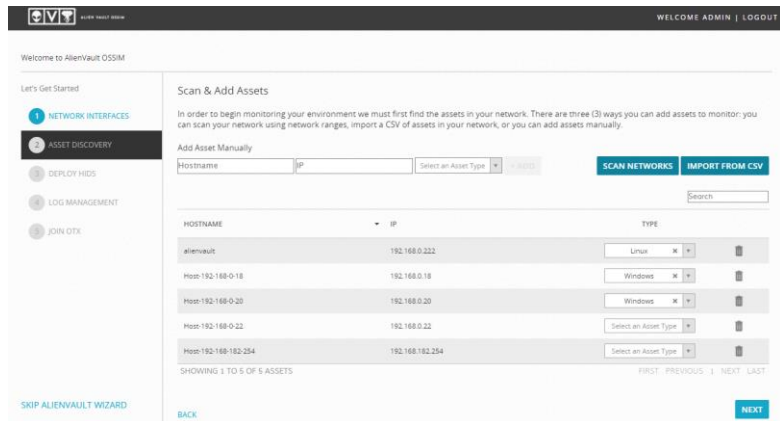


Ilustración 18 Interfaz web. Pantalla de descubrimiento de activos.

7. En la pantalla de despliegue de HIDS, podremos instalar OSSEC en los servidores Windows que escojamos, en este caso no se desplegarán automáticamente, sino que los instalaremos manualmente más adelante.

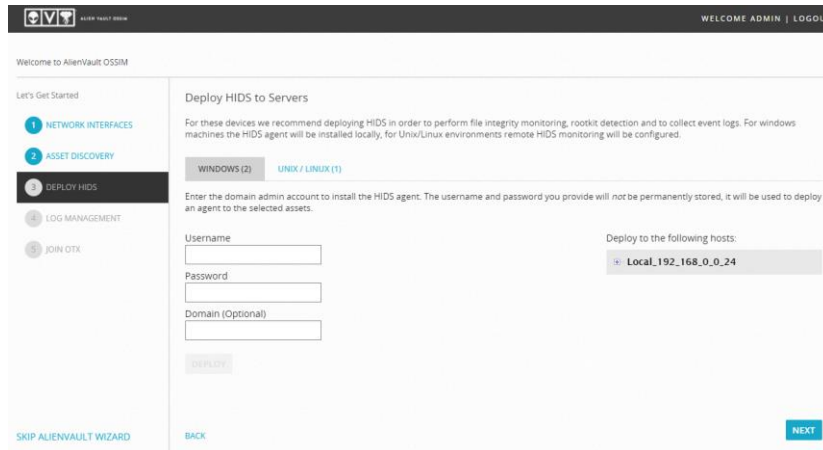


Ilustración 19 Interfaz web. Despliegue de HIDS.

8. En la siguiente pantalla podremos escoger las opciones para cada uno de los dispositivos de red (routers, switches y demás), como no encontré dichos dispositivos en nuestro caso (al realizar la instalación en una red local sin acceso a la red del laboratorio) continuaremos al paso final.

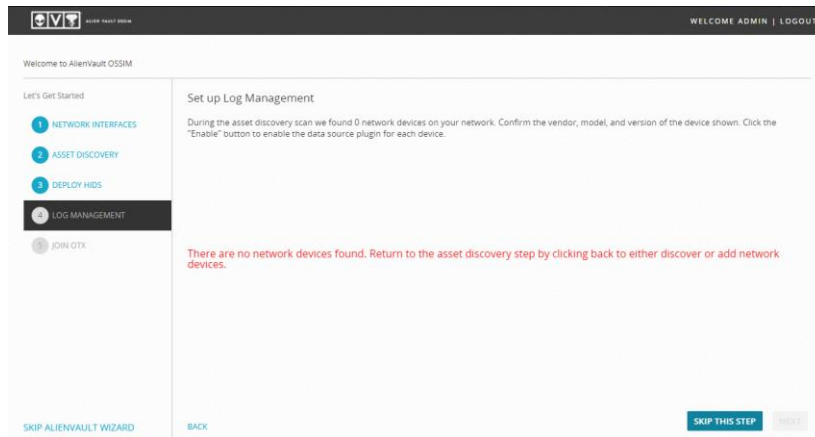


Ilustración 20 Interfaz web. Configuración de dispositivos de red.

- En el último paso de la configuración inicial deberemos conectar la instalación a la plataforma OTX que nos permite compartir datos de inteligencia de amenazas en forma de IoC's entre usuarios de OSSIM y otros productos de AlienVault.

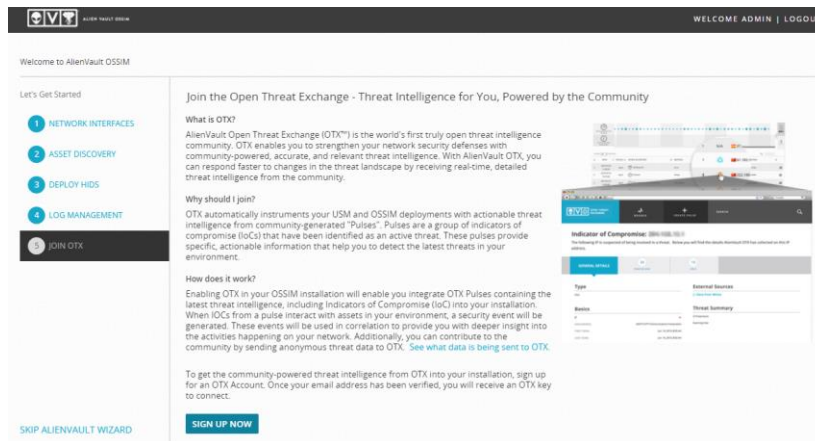


Ilustración 21 Interfaz web. Integración de OTX.

Al hacer clic en el botón “SIGN UP NOW” aparecerá la pantalla de crear usuario o iniciar sesión, la conexión a OTX se mostrará más adelante. Si no queremos realizar la conexión con OTX, saltamos este paso.

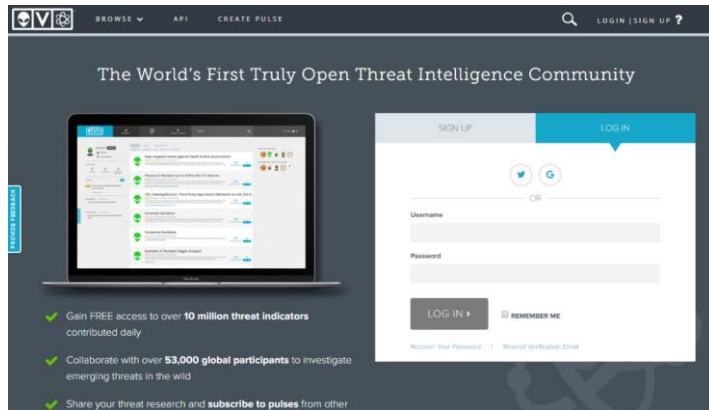


Ilustración 22 Pop up OTX.

10. Al finalizar nos mostrará un mensaje de que la configuración inicial ha sido completada, y procederemos a dar clic en la opción “Explore AlienVault OSSIM”.

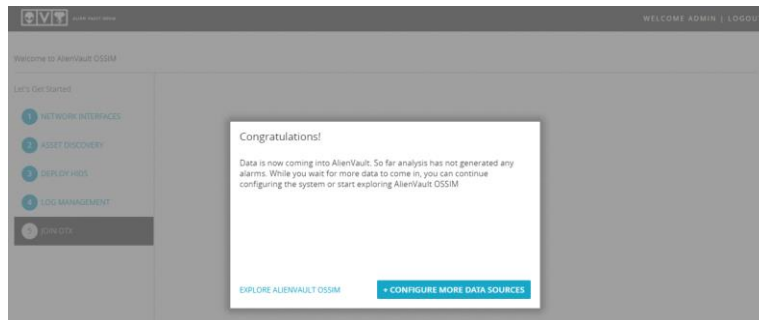


Ilustración 23 Pantalla final del asistente de configuración inicial.

2.1.2.2 VISTA GENERAL DE OSSIM

La interfaz web de OSSIM se compone de 5 menús principales Dashboards, Analysis, Environment, Reports y Configuration, los cuales iremos explorando a medida que desarrollamos cada una de las secciones del libro. En esta sección se mostrará una vista general de cada uno de estos menús.

Dashboards:

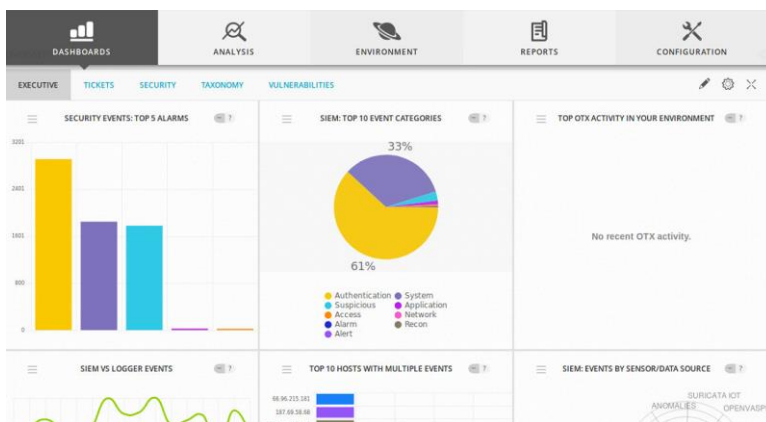


Ilustración 24 Ventana de Dashboards OSSIM

En esta ventana podremos encontrar diferentes gráficas asociadas a eventos, alarmas, actividad en OTX, tickets, vulnerabilidades, seguridad del sistema a nivel de hosts, alarmas, tendencias y demás.

Analysis:

En este menú encontramos algunas opciones, alarmas, donde podremos revisar las alarmas generadas por el sistema; Security events (SIEM) donde encontramos todos los eventos que han llegado al sistema, podemos hacer filtros sobre ellos y exportarlos en formato csv o pdf, o bien ver cómo llegan en tiempo real; y en la

opción de tickets podremos encontrar los que se han generado en el sistema automáticamente, su encargado y su estado actual.

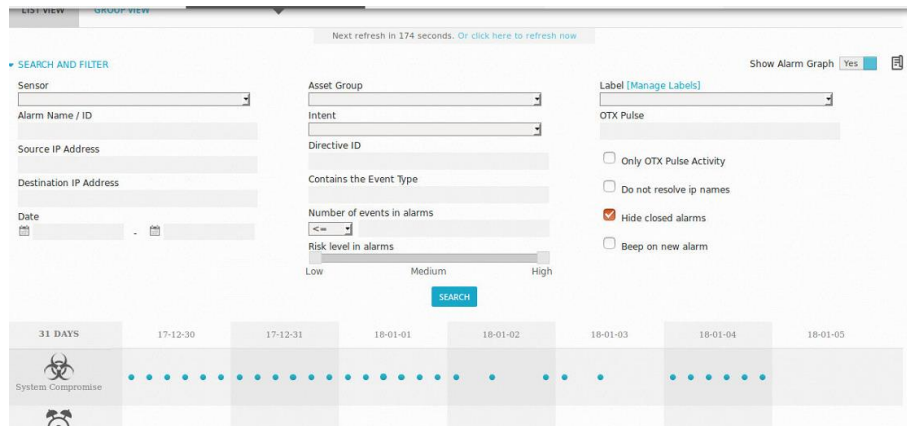


Ilustración 25 Vista de alarmas OSSIM

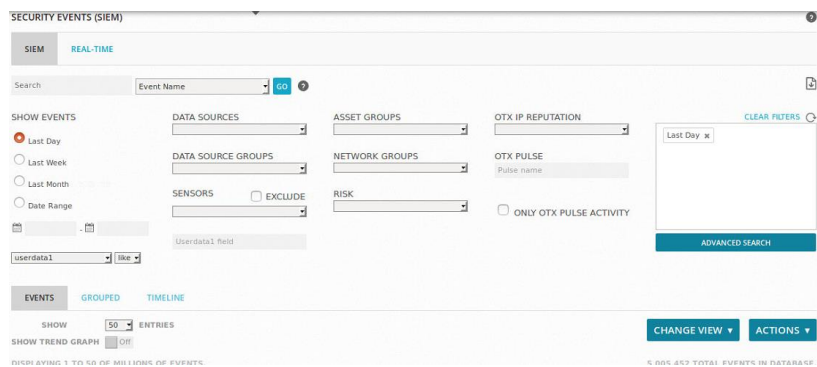


Ilustración 26 Ventana de eventos de seguridad SIEM OSSIM

TICKETS

SIMPLE FILTERS [SWITCH TO ADVANCED]

Class: ALL | Type: ALL | Search text: | Assignee: | Status: Open | Priority: ALL | SEARCH

TICKET	TITLE	PRIORITY	CREATED	LIFE TIME	ASSIGNEE	SUBMITTER	TYPE	STATUS	LABELS
<input type="checkbox"/>	EVE25592	Automatic Incident Ticket	2017-12-31 16:06:48	4 Days 10:23	Escuela Colombiana de Ingenieria	admin	Generic	Open	
<input type="checkbox"/>	EVE25593	Automatic Incident Ticket	2017-12-31 16:06:48	4 Days 10:23	Escuela Colombiana de Ingenieria	admin	Generic	Open	
<input type="checkbox"/>	EVE25590	Automatic Incident Ticket	2017-12-31 16:06:45	4 Days 10:23	Escuela Colombiana de Ingenieria	admin	Generic	Open	

Ilustración 27 Ventana de tickets OSSIM

Environment

En el menú de Environment encontramos diferentes opciones, la primera de ellas Assets & Groups, nos permite explorar los diferentes activos conectados al sistema, añadirlos, filtrarlos por sus características y asignarlos a grupos; en la opción de Vulnerabilities, podemos encontrar las vulnerabilidades que se han encontrado en los diferentes hosts conectados al sistema, realizar nuevos escaneos en la opción Scan Jobs y revisar la base de datos de amenazas; en la opción Netflow, se encuentran los gráficos de flujo de los diferentes protocolos que OSSIM monitorea; en Traffic Capture, podemos tomar capturas del tráfico en la red como lo haríamos en cualquier otro Sniffer; la opción de Availability; nos muestra el estado de los servidores conectados a OSSIM; y por último la opción de Detection, nos permite controlar los diferentes IDS conectados a OSSIM.

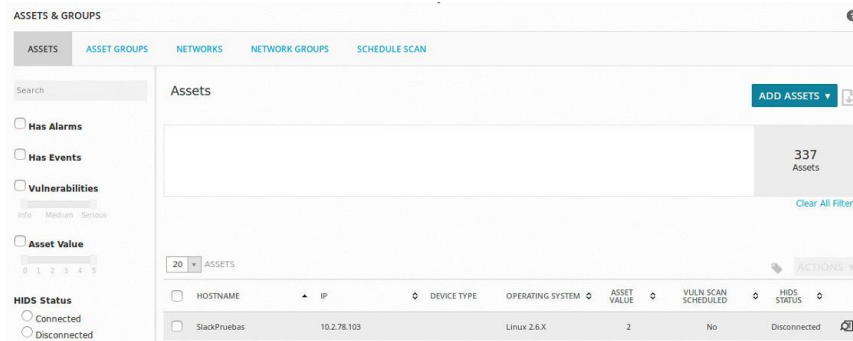


Ilustración 28 Vista de Activos y Grupos OSSIM

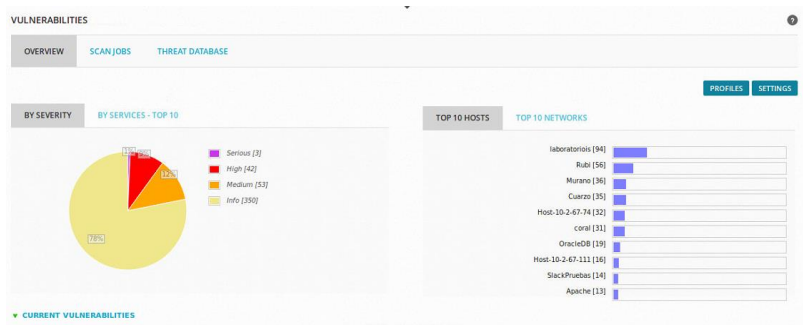


Ilustración 29 Vista de Vulnerabilidades OSSIM



Ilustración 30 Vista de Flujo de red OSSIM

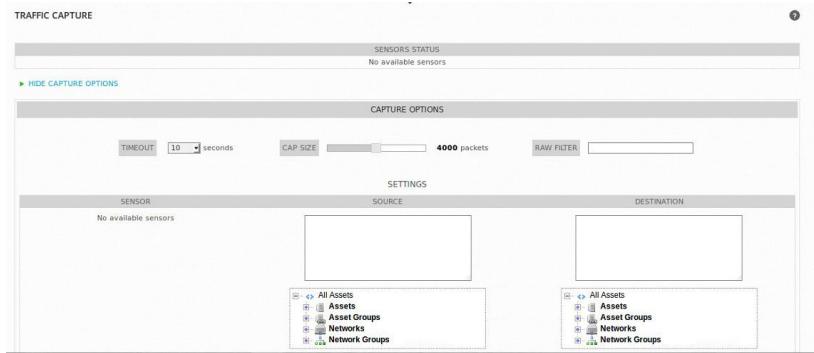


Ilustración 31 Vista de Captura de tráfico OSSIM

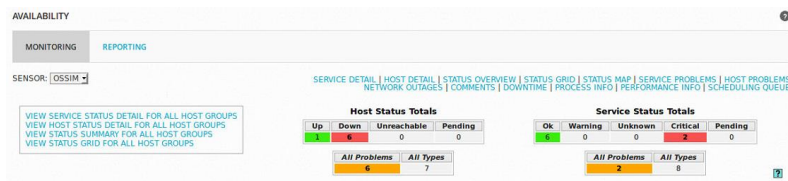


Ilustración 32 Vista de Disponibilidad OSSIM

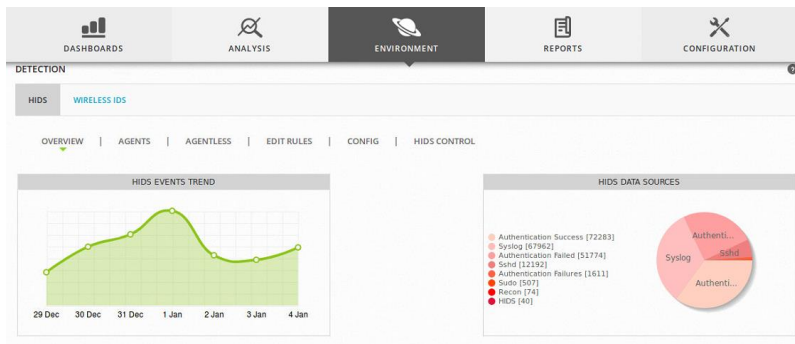


Ilustración 33 Vista de Detección OSSIM

Reports

En este menú podremos generar reportes sobre el estado del sistema, algunos de los reportes que se pueden generar son (1):

Tabla 2 Tipos de reportes en OSSIM

Categoría	Explicación
Alarms Report	Reporte sobre alarmas, fuentes y destinos de ataque, y puertos más afectados.
Assets Details	Este reporte contiene propiedades de los activos, vulnerabilidades, eventos, alarmas y logs para los activos seleccionados.
Business & Compliance ISO PCI Report	Reporta sobre el cumplimiento de regulaciones, incluyendo FISMA, HIPAA, ISO 27001, PCI 2.0, PCI 3.0, PCI DSS 3.1, and SOX. Muestra información en el formato requerido por cada estándar.
Tickets	Reporta sobre tickets abiertos, alarmas, vulnerabilidades y anomalías.
SIEM events	Reporta sobre eventos de seguridad de varias fuentes.
Tickets Status	Reporta sobre operaciones de seguridad como tickets, y eventos y alarmas más significativos.
User Activity	Reporta la actividad registrada de los usuarios de la plataforma.
Geographic	Muestra información geográfica sobre los eventos.
Threats & Vulnerabilities Database	Muestra el estado actual de la base de datos de vulnerabilidades y amenazas.

Vulnerabilities Report

Realiza un reporte sobre las vulnerabilidades teniendo en cuenta su criticidad y activos asociados.

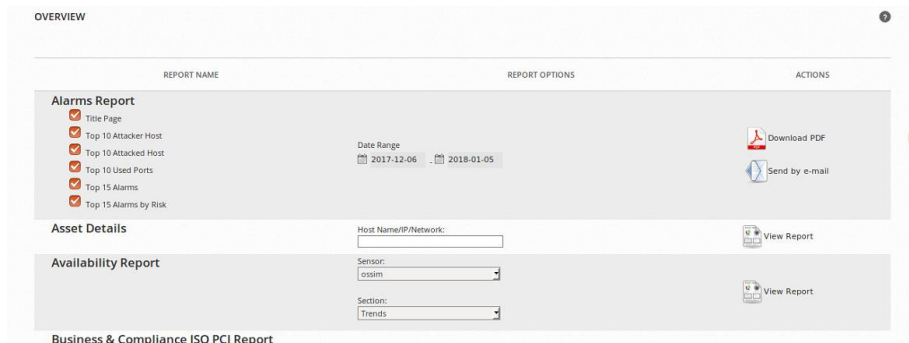


Ilustración 34 Vista de Reportes OSSIM

Configuración

Este menú se compone de 4 opciones, la opción Administration que permite realizar configuraciones al sistema, como lo son usuarios, configuraciones generales y backups; la opción Deployment nos permite revisar los componentes del sistema, adicionar plugins y configurar la ubicación de los componentes de OSSIM; en la sección Threat Intelligence, podemos crear políticas, configurar acciones, directivas de correlación y de correlación cruzada, ver el mapa de cumplimiento de normas como ISO 27001 o PCI DSS 2.0 y 3.0; y la sección de OTX nos permite configurar la conexión con la plataforma.

ADMINISTRATION

USERS MAIN BACKUPS

USER INFORMATION | ACTIVITY

SHOW 20 ENTRIES NEW MODIFY DELETE SELECTED DUPLICATE SELECTED

LOGIN	NAME	EMAIL	VISIBILITY	STATUS	LANGUAGE	CREATION DATE	LAST LOGIN DATE
claudia.santiago	Claudia Patricia Santiago Cely	claudia.santiago@escuelaing.edu.co	ECI	✓	English	2017-01-25 18:29:45	2017-01-31 07:37:04
daniel.diaz	Daniel Orlando Diaz Lopez	daniel.diaz@escuelaing.edu.co	ECI	✓	English	2017-01-25 18:26:48	2017-10-09 16:44:31
admin	Escuela Colombiana de Inger	maria.blanco@mail.escuelaing.edu.co	ECI	✓	English	2017-01-25 12:15:25	2018-01-04 20:19:41
gerardo.ospina	Gerardo Ospina	gerardo.ospina@escuelaing.edu.co	ECI	✓	English	2017-01-25 18:31:21	-
pg1	Proyecto SIEM Engine	nicolas.gomez-s@mail.escuelaing.edu.co	ECI	✓	English	2017-01-25 18:33:42	2017-12-30 07:27:23
pg2	Proyecto SIEM for IoT	nicolas.moreno-g@mail.escuelaing.edu.co	ECI	✓	English	2017-01-25 18:34:54	2017-11-27 16:20:03

Ilustración 35 Vista de Administración OSSIM

DEPLOYMENT

COMPONENTS PLUGIN BUILDER LOCATIONS

ALIENVAULT CENTER | SENSORS | SERVERS

ALIENVAULT CENTER

ALIENVAULT COMPONENTS INFORMATION

NAME	STATUS	RAM USAGE	SWAP USAGE	CPU USAGE	NEW UPDATES
ossim [10.2.78.8] Server Sensor Web interface Database	UP	90.20 %	7.50 %	39.05 %	Patch 5.5

SHOWING 1 TO 1 OF 1 ENTRIES FIRST PREVIOUS 1 NEXT LAST

Ilustración 36 Vista de Despliegue OSSIM

THREAT INTELLIGENCE

POLICY ACTIONS PORTS DIRECTIVES COMPLIANCE MAPPING CROSS CORRELATION DATA SOURCE TAXONOMY KNOWLEDGE BASE

EDIT POLICY GROUPS

Default policy group: Default group policy objects

New Modify Delete selected Duplicate selected Reload Policies Enable/Disable policy

STATUS	ORD	NAME	SOURCE	DESTINATION	SOURCE PORT	DEST PORT	EVENT TYPES	SENSORS	TIME RANGE	TARGETS	SIEM
✓	1	Prueba Apache BruteForce	ANY	OracleDB	ANY	ANY	DS Groups: Apache BruteForce	ANY	America/Bogota (0h : 00m) : 23h : 59min	ossim	✓
✓	2	DoS Policy Test	ANY	ANY	ANY	ANY	DS Groups: DoS DoS Failures	ANY	America/Bogota (0h : 00m) : 23h : 59min	ossim	✓
✗	3	IoT GeoFencing	ANY	ANY	ANY	ANY	DS Groups: IoT	ANY	America/Bogota (0h : 00m) : 23h : 59min	ossim	✗

AV default policies: Filter events from AlienVault avapi user

Ilustración 37 Vista de Inteligencia de Amenazas OSSIM

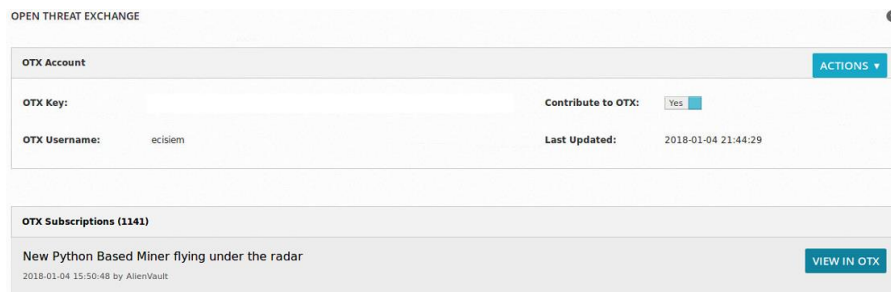


Ilustración 38 Vista de configuración de OTX OSSIM

2.1.2.3 OTX

Es una comunidad global que permite el intercambio de IoC's permitiendo discusiones, investigación y validación de datos asociados a amenazas, tendencias y técnicas, fortaleciendo las defensas de la organización y permitiendo a los demás miembros de la comunidad hacer lo mismo (2). Esta información se comparte como pulsos de OTX, donde son revisados por la comunidad para revisar que sean legítimos y luego de esta revisión son publicados. Para conectarnos a OTX seguimos los pasos presentados a continuación:

1. Crear una cuenta de OTX: Nos dirigimos a la página <https://otx.alienvault.com> (Ilustración 22 Pop up OTX.) y entramos a la opción "SIGN UP" y llenamos el formulario.
2. En OSSIM para poder realizar la conexión nos pedirá una clave de OTX (OTX Key), para obtener esta clave iniciamos sesión en <https://otx.alienvault.com> y nos dirigimos a la zona de API, allí en la zona lateral derecha encontraremos nuestra clave de OTX.

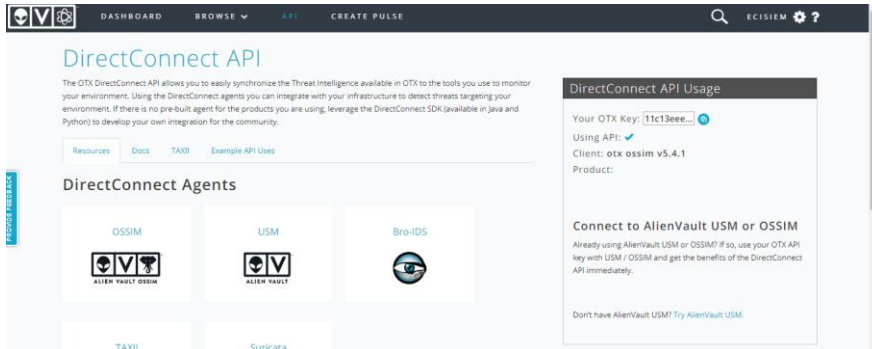


Ilustración 39 Página de OTX. Sección de API.

3. Copiamos la llave haciendo clic en el botón azul, y nos dirigimos a la interfaz web de OSSIM e iniciamos sesión. En la ventana principal nos dirigimos al menú CONFIGURATION > OPEN THREAT EXCHANGE donde nos aparece la ventana de configuración:

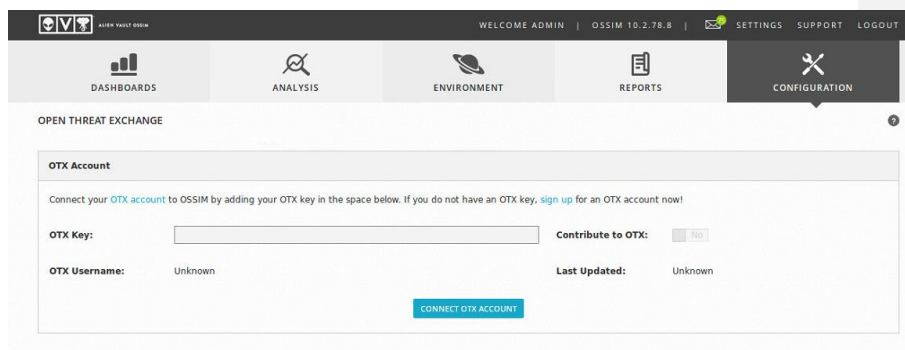


Ilustración 40 Ventana de configuración de OTX.

En el input de OTX Key pegamos la llave que anteriormente habíamos copiado, y hacemos clic en el botón “CONNECT OTX ACCOUNT”. Aparecerá un mensaje de que la cuenta se ha conectado correctamente.

2.1.2.4 Configuración del vulnerability ticket threshold

Este límite mide el nivel de riesgo necesario como para generar un ticket en el sistema. Se mide del 1 al 10 siendo 1 una situación crítica, y 10 una situación no crítica. Una vulnerabilidad con un riesgo mayor al valor establecido abre un ticket automáticamente.

Los valores de severidad se muestran a continuación:

Tabla 3 Nivel de severidad para tickets

Severidad	Valor Interno
Serio	1
Alto	2
Medio	3
Bajo	6
Informativo	7

Para cambiar la configuración, nos dirigimos al menú CONFIGURATION > ADMINISTRATION > MAIN > VULNERABILITY TICKET THRESHOLD, en nuestro caso cambiaremos la configuración a medio. Y hacemos clic en el botón “Update configuration” (3).

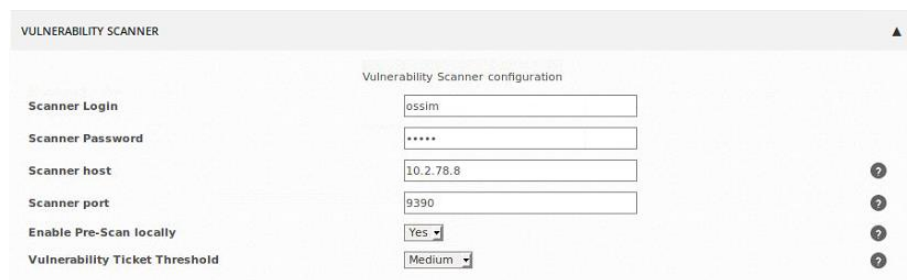


Ilustración 41 Configuración de límite para tickets



Ilustración 42 Pantalla de administración de OSSIM.

2.1.2.5 Configuración de SNMP y OSSIM

Las SNMP traps son utilizadas para realizar monitoreo de la red, son mensajes de alerta de un dispositivo que se envían a un recolector central. Una trampa SNMP puede alertar sobre el sobrecalentamiento de un dispositivo. En OSSIM se reciben mensajes de dichas trampas que pueden ser insumo para las directivas de correlación. El SNMP no está activado por defecto y OSSIM no soporta por el momento SNMPv3(4).

Para activar el SNMP:

1. Ingresar a la consola de OSSIM por la opción "Jailbreak System".

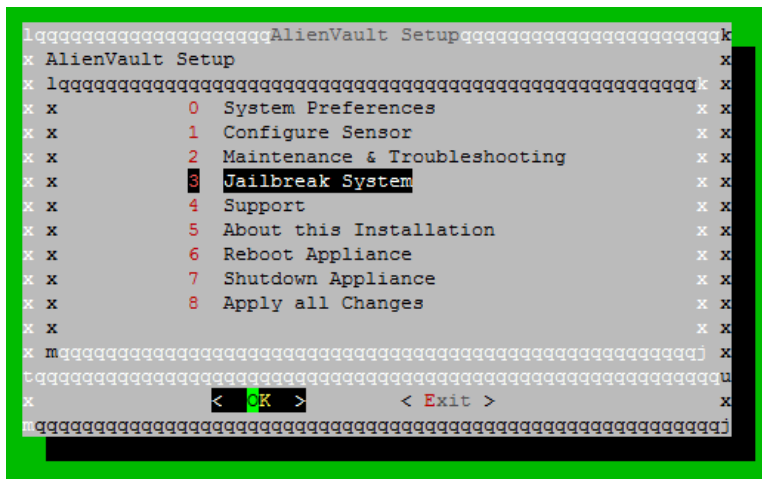


Ilustración 43 Consola de configuración OSSIM.

2. Editar el archivo `/etc/ossim/ossim_setup.conf` cambiando `snmpd` y `snmptrap` a "yes".

```
[snmp]
community=public
snmpd=yes
snmptrap=yes
```

Ilustración 44 Edición de archivo para activar SNMP.

3. Por último, ejecutamos el comando desde consola, `alienvault-reconfig`. Y aparecerá la siguiente pantalla.

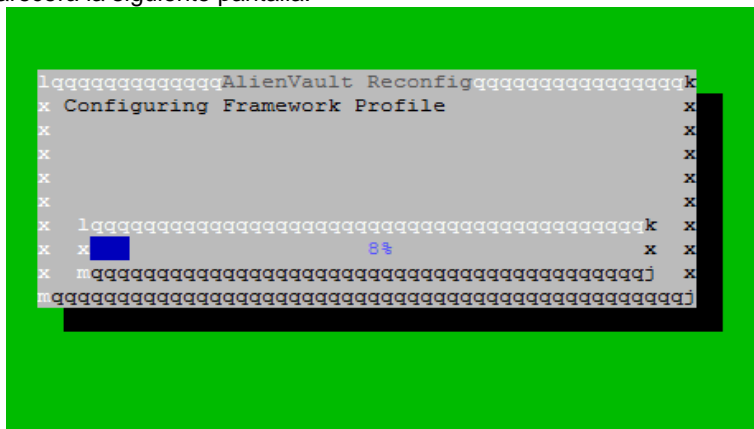


Ilustración 45 Ventana de reconfiguración de OSSIM.

2.1.2.6 Agregar usuarios al sistema

Para agregar usuarios al sistema seguimos estos pasos:

1. Entrar al menú `CONFIGURATION> USERS`

WELCOME ADMIN | OSSIM 10.2.78.8 | SETTINGS SUPPORT LOGOUT

DASHBOARDS ANALYSIS ENVIRONMENT REPORTS CONFIGURATION

ADMINISTRATION

USERS MAIN BACKUPS

USER INFORMATION | ACTIVITY

SHOW 25 ENTRIES

NEW MODIFY DELETE SELECTED DUPLICATE SELECTED

LOGIN	NÂME	EMAIL	VISIBILITY	STATUS	LANGUAGE	CREATION DATE	LAST LOGIN DATE
claudia.santiago	Claudia Patricia Santiago Cely	claudia.santiago@escuelaing.edu.co	ECI	✓	English	2017-01-25 18:29:45	2017-01-31 07:37:04
daniel.diaz	Daniel Orlando Diaz Lopez	daniel.diaz@escuelaing.edu.co	ECI	✓	English	2017-01-25 18:26:48	2017-10-09 16:44:31
admin	Escuela Colombiana de Inger	maria.bianco@mail.escuelaing.edu.co	ECI	✓	English	2017-01-25 12:15:25	2018-01-04 20:19:41
gerardo.ospina	Gerardo Ospina	gerardo.ospina@escuelaing.edu.co	ECI	✓	English	2017-01-25 18:31:21	-
pp1	Proyecto SIEM Engine	nicolas.gomez-sil@mail.escuelaing.ec	ECI	✓	English	2017-01-25 18:33:42	2017-12-30 07:27:23
pp2	Proyecto SIEM for IoT	nicolas.moreno@mail.escuelaing.edu.co	ECI	✓	English	2017-01-25 18:34:54	2017-11-27 16:20:03

Ilustración 46 Menú de configuración de usuarios.

- Hacer clic en el link “new” donde aparecerá un formulario para crear el nuevo usuario.

USER LOGIN *

USER NAME *

USER EMAIL *

USER LANGUAGE * English

TIMEZONE * UTC

COMPANY

DEPARTMENT

ENTER YOUR CURRENT PASSWORD *

ENTER USER PASSWORD *

RE-ENTER USER PASSWORD *

ASK TO CHANGE PASSWORD AT NEXT LOGIN Yes No

MAKE THIS USER A GLOBAL ADMIN Yes No

▶ ALLOWED MENUS

▶ ASSET FILTERS

SAVE

Ilustración 47 Formulario de creación de usuarios.

Llenamos los campos y marcamos la opción de “ASK TO CHANGE PASSWORD AT NEXT LOGIN” si queremos que el usuario cambie su contraseña la próxima vez que inicie sesión, podemos hacerlo un usuario administrador si marcamos la opción de “MAKE THIS USER A GLOBAL ADMIN”.

Añadir un nuevo activo

En OSSIM se pueden añadir activos de diferentes formas, por medio del descubrimiento de activos, importándolos desde otro SIEM o archivo y añadiéndolo manualmente.

- Descubrimiento de activos:
 - Dirigirnos al menú Environment>Assets&Groups:

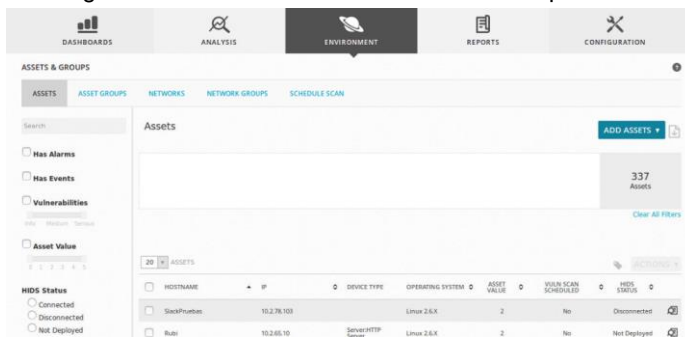
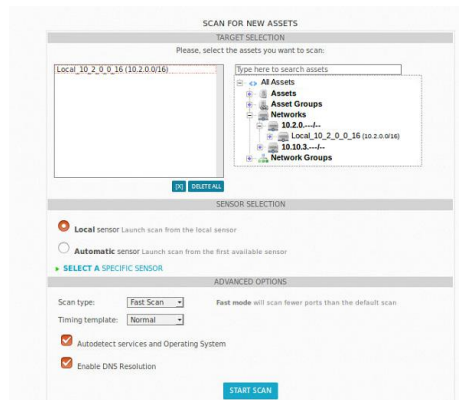


Ilustración 48 Vista de Activos y Grupos

- Hacemos clic en el botón “Add Assets” y seleccionamos la opción “scan for new assets”, nos dirigirá a otra ventana donde podremos escoger la red sobre la cual se va a hacer al descubrimiento de



activos, grupos y demás, en este caso se realizará sobre la red de estudiantes del Laboratorio, se escoge el tipo de escaneo y se hace clic en el botón “Start Scan”.

Ilustración 49 Vista de Activos y Grupos

- Nos aparecerá una ventana de progreso sobre el escaneo de la red, esperamos y nos aparecerán los nuevos activos agregados al SIEM.

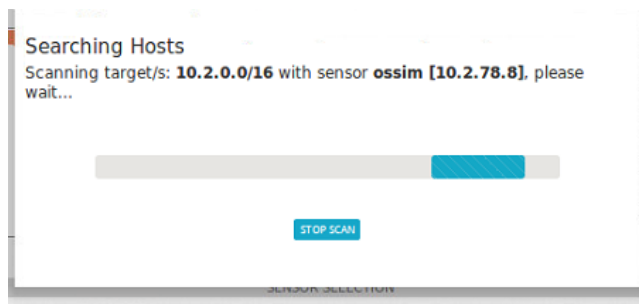


Ilustración 50 Vista de Activos y Grupos

- Añadir activos de forma manual
 - En la vista de activos (Figura 48) hacemos clic al botón “Add Assets” y escogemos la opción “Add Host”.
 - Nos aparecerá un formulario donde deberemos llenar los siguientes datos obligatorios:
 - Nombre: Nombre del activo.
 - IP Address: Dirección IP del activo.
 - Asset value: Nivel de importancia de 1 a 5 del activo para la organización.

- External asset: Si el activo es externo o no

The screenshot shows a 'NEW ASSET' form with the following fields and options:

- Name ***: Input field containing 'Activo de prueba'.
- IP Address ***: Input field containing '10.2.65.200'.
- FQDN/Aliases**: Empty input field.
- Location**: Input field containing 'Undetermined location'.
- Asset Value ***: Dropdown menu with '2' selected.
- External Asset ***: Radio buttons for 'Yes' (unselected) and 'No' (selected).
- Sensors ***: Checkboxes for '10.2.78.8 (ossim)' (checked) and another sensor.
- ICON**: 'Choose icon ...' button.
- Map**: A Google Map showing the Gulf of Guinea, Equatorial Guinea, and Gabon.
- Latitude/Longitude**: Two empty input fields.

Ilustración 51 Formulario para agregar un activo manualmente.

- Después de llenar el formulario hacemos clic en el botón “Save” ubicado al final, y ya quedará añadido nuestro activo.

Conexión de activos

En nuestro caso la conexión de activos se hizo a diferentes niveles, a nivel de sistema operativo con OSSEC, de software base por medio de plugins y en aplicaciones utilizando AppSensor, que actúa como plugin para la conexión al SIEM. A continuación, se muestra el proceso a seguir para realizar la conexión de los activos en los tres niveles:

OSSEC

Es un proyecto open source, que permite monitorear y controlar los sistemas. Combina aspectos de HIDS (Host-Based intrusion detection), monitoreo de logs y manejo de incidentes.(5)

En este caso se realiza la instalación de OSSEC sobre Slackware 13.37, los paquetes requeridos que se muestran a continuación cambian según la versión:

- a/
 - glibc-solibs-2.13-i486-4.txz
 - kernel-modules-2.6.37.6-i486-2.txz
- d/
 - make-3.82-i486-2.txz
 - gcc-4.5.2-i486-2.txz
 - gcc-g++-4.5.2-i486-2.txz
 - binutils-2.21.51.0.6-i486-1.txz
 - kernel-headers-2.6.37.6_smp-x86-2.txz
- k/
 - kernel-source-2.6.37.6_smp-noarch-2.txz
- l/
 - mpfr-3.0.1-i486-1.txz
 - glibc-2.13-i486-4.txz
 - glib2-2.28.6-i486-1.txz
 - glib-1.2.10-i486-3.txz
 - libmpc-0.8.2-i486-2.txz
 - libmcrypt-2.5.8-i486-1.txz
- n/
 - wget-1.12-i486-1.txz

PROCESO DE INSTALACIÓN:

1. Descargar la última versión de OSSEC, se ejecutando el siguiente comando:

wget -U ossec -O ossec.tar.gz <https://bintray.com/artifact/download/ossec/ossec-hids/ossec-hids-2.8.3.tar.gz> --no-check-certificate

```

root@alackware:~# wget -U ossec -O ossec.tar.gz https://bintray.com/artifact/download/ossec/ossec-hids/ossec-hids-2.8.3.tar.gz --no-check-certificate
--2017-02-28 13:16:11-- https://bintray.com/artifact/download/ossec/ossec-hids/ossec-hids-2.8.3.tar.gz
Resolving bintray.com (bintray.com)... 108.168.194.93
Connecting to bintray.com (bintray.com)|108.168.194.93|:443... connected.
WARNING: cannot verify bintray.com's certificate, issued by '/C=US/O=GoDaddy Inc./CN=GoTrust SSL CA - G3':
Unable to locally verify the issuer's authority.
HTTP request sent, awaiting response... 302 found
Location: https://dl.bintray.com/ossec/ossec-hids/ossec-hids-2.8.3.tar.gz [following]
--2017-02-28 13:16:15-- https://dl.bintray.com/ossec/ossec-hids/ossec-hids-2.8.3.tar.gz
Resolving dl.bintray.com (dl.bintray.com)... 108.168.243.150
Connecting to dl.bintray.com (dl.bintray.com)|108.168.243.150|:443... connected.
WARNING: cannot verify dl.bintray.com's certificate, issued by '/C=US/O=GoTrust Inc./CN=GoTrust SSL CA - G3':
Unable to locally verify the issuer's authority.
HTTP request sent, awaiting response... 302
Location: https://akamai.bintray.com/87/87/ta1804d508c7c7f3e42bd470055b14b0f8a7_gda__resp=48832803-bmao=3ddc5078b9468891762aef4b18183c27718a5bd364f8b14oub89ab0b3deceresponse-content-dispo
sitionattachment;filename=ossec-hids-2.8.3.tar.gz&response-content-type=application&2Fgip4requestInfo=02Fad07XK18CW8d2XVDTpA3j1YzYQ8tjWfW6-6FFPb6vIM6Opaillyc0HleRgRW5p8Rgo1bFS106
HTDQvqLl7VwB13z89yE1a1e8_vfVr-csmMbl19T [following]
--2017-02-28 13:16:15-- https://akamai.bintray.com/87/87/ta1804d508c7c7f3e42bd470055b14b0f8a7_gda__resp=48832803-bmao=3ddc5078b9468891762aef4b18183c27718a5bd364f8b14oub89ab0b3deceresponse
-content-disposition=attachment;filename=ossec-hids-2.8.3.tar.gz&2response-content-type=application&2Fgip4requestInfo=02Fad07XK18CW8d2XVDTpA3j1YzYQ8tjWfW6-6FFPb6vIM6Opaillyc0HleRgRW5
Wp8eap0bRf08t7Dm0q1VwB13z89yE1a1e8_vfVr-csmMbl19T
Resolving akamai.bintray.com (akamai.bintray.com)... 23.32.201.90
Connecting to akamai.bintray.com (akamai.bintray.com)|23.32.201.90|:443... connected.
WARNING: cannot verify akamai.bintray.com's certificate, issued by '/C=NL/I=Meurescom/O=Verizon Enterprise Solutions/GN=Cybertrust/CN=Verizon Akamai SecureServer CA 014-SHA2':
Unable to locally verify the issuer's authority.
HTTP request sent, awaiting response... 200 OK
Length: 1442095 (1.4M) [application/gzip]
Saving to: 'ossec.tar.gz'

100%[=====] 1,442,095  2.09MB/s  in 0.6s

2017-02-28 13:16:18 (2.03 MB/s) - 'ossec.tar.gz' saved [1442095/1442095]

```

Ilustración 52 Descarga de OSSEC 2.8 vía consola.

2. Descomprimir el archivo y entrar a la carpeta OSSEC-HIDS-2.8.3 con los comandos:

```
tar -xvf ossec.tar.gz
cd ossec-hids-2.8.3/
```

Deben aparecer los archivos que se muestran en la imagen a continuación:

```
root@slackware:~/ossec-hids-2.8.3# ls
BUGS CONFIG CONTRIBUTORS INSTALL LICENSE README.md active-response/ contrib/ doc/ etc/ install.sh* src/
root@slackware:~/ossec-hids-2.8.3#
```

Ilustración 53 Archivos dentro de la carpeta descomprimida de OSSEC.

3. Ejecutamos el archivo install.sh con el comando que aparece en la imagen.

```
root@slackware:~/ossec-hids-2.8.3# ./install.sh
```

Ilustración 54 Comando de Instalación de OSSEC

4. Cuando ejecutemos el archivo, se iniciará el proceso de instalación, como queremos instalar un agente HIDS debemos escoger las siguientes opciones:

```
root@slackware:~/ossec-hids-2.8.3# ./install.sh
which: no host in (/usr/local/sbin:/usr/sbin:/sbin:/usr/local/bin:/usr/bin:/bin:/usr/games)

** Para instalação em português, escolha [br].
** 要使用中文进行安装, 请选择 [cn].
** Für eine deutsche Installation wählen Sie [de].
** Για εγκατάσταση στα Ελληνικά, επιλέξτε [el].
** For installation in English, choose [en].
** Para instalar en Español , eliga [es].
** Pour une installation en français, choisissez [fr]
** A Magyar nyelvű telepítéshez válassza [hu].
** Per l'installazione in Italiano, scegli [it].
** 日本語でインストールします。選択して下さい。 [jp].
** Voor installatie in het Nederlands, kies [nl].
** Aby instalować w języku Polskim, wybierz [pl].
** Для инструкций по установке на русском , введите [ru].
** Za instalaciju na srpskom, izaberi [sr].
** Türkçe kurulum için seçin [tr].
(en/br/cn/de/el/es/fr/hu/it/jp/nl/pl/ru/sr/tr) [en]: en
```

Ilustración 55 Opciones de lenguaje en el menú de instalación de OSSEC


```
which: no host in (/usr/local/sbin:/usr/sbin:/sbin:/usr/local/bin:/usr/bin:/bin:/usr/games)
./install.sh: line 967: clear: command not found
OSSEC HIDS v2.8.3 Installation Script - http://www.ossec.net

You are about to start the installation process of the OSSEC HIDS.
You must have a C compiler pre-installed in your system.
If you have any questions or comments, please send an e-mail
to dcid@ossec.net (or daniel.cid@gmail.com).

- System: Linux slackware 3.10.17
- User: root
- Host: slackware

-- Press ENTER to continue or Ctrl-C to abort. --
```

Ilustración 56 Información sobre la instalación de OSSEC

Presionamos Enter.

```
1- What kind of installation do you want (server, agent, local, hybrid or help)? agent
```

Ilustración 57 Opciones de instalación servidor, agente, local o híbrido.

En el siguiente paso nos preguntará la ruta en la que queremos instalar el agente, la ruta estándar en el Laboratorio es /usr/local/ossec

```
2- Setting up the installation environment.
- Choose where to install the OSSEC HIDS [/usr/local/ossec]: /usr/local/ossec
```

Ilustración 58 Ruta de instalación de OSSEC

Aquí debemos ingresar la IP del servidor que va a recolectar los eventos, en este caso OSSIM:

```
3- Configuring the OSSEC HIDS.
  3.1- What's the IP Address or hostname of the OSSEC HIDS server?: 10.2.78.8
    - Adding Server IP 10.2.78.8
  3.2- Do you want to run the integrity check daemon? (y/n) [y]:
    - Running syscheck (integrity check daemon).
  3.3- Do you want to run the rootkit detection engine? (y/n) [y]:
    - Running rootcheck (rootkit detection).
  3.4 - Do you want to enable active response? (y/n) [y]:
```

Ilustración 59 Configuración de redirección de logs hacia el SIEM.

Para las siguientes 3 opciones, damos enter para que habilite los servicios que

```
3.5- Setting the configuration to analyze the following logs:
  -- /var/log/messages
  -- /var/log/secure
  -- /var/log/syslog
  -- /var/adm/syslog
  -- /var/adm/messages
  -- /var/log/maillog

- If you want to monitor any other file, just change
  the ossec.conf and add a new localfile entry.
  Any questions about the configuration can be answered
  by visiting us online at http://www.ossec.net .

--- Press ENTER to continue ---
```

Ilustración 60 Habilitación de servicios dentro del sistema.

necesita.

Presionamos Enter y comenzará el proceso de instalación, en caso de generarse un error se debe revisar que todos los paquetes listados

```
Started ossec-agentd...
Started ossec-logcollector...
Started ossec-syscheckd...
Completed.

- Configuration finished properly.

- To start OSSEC HIDS:
  /var/ossec/bin/ossec-control start

- To stop OSSEC HIDS:
  /var/ossec/bin/ossec-control stop

- The configuration can be viewed or modified at /var/ossec/etc/ossec.conf

Thanks for using the OSSEC HIDS.
If you have any question, suggestion or if you find any bug,
contact us at contact@ossec.net or using our public maillist at
ossec-list@ossec.net
( http://www.ossec.net/main/support/ ).

More information can be found at http://www.ossec.net

--- Press ENTER to finish (maybe more information below). ---
```

Ilustración 61 Instalación exitosa de OSSEC

anteriormente se encuentren instalados.

Si la instalación se completó con éxito, se debe mostrar el mensaje anterior.

CONEXIÓN DEL SERVIDOR CON LA PLATAFORMA OSSIM:

1. Nos dirigimos al menú Environment->Detection->Agents

The screenshot shows a web-based form titled "NEW HIDS AGENT". At the top, it says "Values marked with (*) are mandatory". Below that, there is a instruction: "Select an asset to connect to HIDS agent. This will associate the agent with the asset so that you can see the status of the agent from the asset views. *". There is a search bar labeled "Search by IP address or name". Below the search bar is a tree view of assets with the following structure:


- All Assets
 - Assets
 - Asset Groups
 - Networks
 - Network Groups

Below the tree view, there are two input fields: "Agent Name *" and "IP/CIDR *". There is a checkbox labeled "This is a dynamic IP address (DHCP)". At the bottom of the form is a "SAVE" button.

Ilustración 63 Formulario para agregar un nuevo agente a OSSIM

2. Damos clic en el botón ADD AGENT y aparecerá la siguiente ventana:

En esta ventana buscamos la IP del activo que queremos agregar, y automáticamente pondrá los campos correspondientes a la IP y el nombre del agente. Damos clic en el botón Save y se guardará la configuración del agente HIDS.

3. Copiamos la llave generada haciendo clic en el botón  aparecerá un diálogo con la llave que se debe utilizar para conectar el servidor con la plataforma:

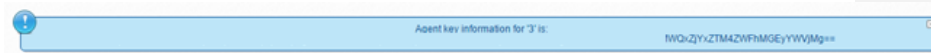


Ilustración 64 Llave de conexión con la plataforma OSSIM

CONFIGURAR LLAVE DE LA PLATAFORMA OSSIM EN EL SERVIDOR CON SISTEMA OPERATIVO LINUX, DISTRIBUCIÓN SLACKWARE:

1. Nos dirigimos a la ruta donde está instalado OSSEC (/usr/local/ossec/bin):

```
root@slackware:/usr/local/ossec/bin# ls
agent-auth*  ossec-agentd*  ossec-execd*  ossec-lua*  ossec-syscheckd*
manage_agents*  ossec-control*  ossec-logcollector*  ossec-luac*  util.sh*
```

Ilustración 66 Scripts para la administración de OSSEC

2. Ejecutamos el script manage_agents:

```
root@slackware:/var/ossec/bin# ./manage_agents

*****
* OSSEC HIDS v2.8.3 Agent manager.      *
* The following options are available: *
*****
(I) Import key from the server (I).
(Q) Quit.
Choose your action: I or Q: _
```

Ilustración 65 Menú del script de administración de agentes

3. Seleccionamos la opción "I" y escribimos la llave generada:

```
* Provide the Key generated by the server.
* The best approach is to cut and paste it.
*** OBS: Do not include spaces or new lines.

Paste it here (or '\q' to quit):
```

Ilustración 67 Opción de insertar llave generada desde OSSIM

Nos aparece el menú anterior y seleccionamos la opción "Q".

4. Corremos el script de inicio del agente:

```
/ossec-control start
```

```
root@slackware:/usr/local/ossec/bin# ./ossec-control start
Starting OSSEC HIDS v2.8.3 (by Trend Micro Inc.)...
Started ossec-execd...
2017/03/09 13:23:20 ossec-agentd: INFO: Using notify time: 600 and max time to reconnect: 1800
Started ossec-agentd...
Started ossec-logcollector...
Started ossec-syscheckd...
Completed.
root@slackware:/usr/local/ossec/bin# █
```

Ilustración 68 Inicio del agente HIDS dentro del servidor

Con estos pasos ya quedará conectado el servidor a la plataforma OSSIM.

Plugin Apache HTTPD y Plugin Oracle Database

La configuración de estos dos plugins se encuentra especificada en los archivos Plugin Apache HTTP On Ubuntu Installation Manual y Plugin Oracle Database On Ubuntu Installation Manual, los cuales se encuentran anexos a este documento.

2.1.3 Caracterización de activos

Como se había mencionado anteriormente el nivel de criticidad de los activos de información es heredado de la información que soporta o maneja, es por ello por lo que las empresas se preocupan cada vez más por asegurar dichos activos, y para poder realizar controles efectivos sobre los activos primero deben ser identificados y clasificados.

En esta sección se tratarán de exponer las características de seguridad que diferencian a cada nivel de activo, sistemas operativos, software base y aplicaciones, para ello se expondrán sus principales vulnerabilidades y amenazas, los tipos de eventos de seguridad asociados a cada nivel de activo y los desafíos relacionados a la gestión de eventos.

2.1.3.1 Sistemas Operativos

Un sistema operativo debe proteger su Trusted Computing Base (TCB) con el fin de tener un punto de referencia, esta base segura se podría ver afectada por la explotación de vulnerabilidades. Se dice que una vulnerabilidad está asociada al sistema operativo cuando afecta al kernel o a los componentes que son propios del sistema desarrollados bien sea por Microsoft, Apple o un proveedor Unix. Algunas de las vulnerabilidades más comunes dentro de los diferentes sistemas operativos se presentan a continuación:

- **Network-facing daemons:** Dentro del sistema operativo existen procesos que mantienen abiertos puertos usados en conexiones remotas (sshd, ftpd, etc) estos puertos son conocidos como “demonios de cara a la red”, se han reportado diferentes vulnerabilidades asociadas a dichos puertos particularmente buffer overflows, que permiten a un atacante comprometer el sistema.
- **Rootkits:** Conocidos por ser una variedad de malware, que permite la ejecución de código directamente desde el kernel, lo que implica privilegios sobre todo el sistema. Los rootkits se aprovechan del proceso carga del kernel para así permitir la implementación de funciones del atacante y evadir su detección.
- **Variables de entorno:** Los sistemas operativos soportan dichas variables, para que sean utilizadas por diferentes procesos. Un ataque común que explota dicha vulnerabilidad se ejecuta cuando un atacante cambia la variable de entorno LIBPATH (determina el orden de búsqueda sobre librerías dinámicas) para incluir un archivo como una librería dinámica. Si un proceso del sistema operativo no configura la variable de entorno por su cuenta, podría llegar a correr dicho código malicioso dentro del sistema.
- **Recursos compartidos:** Los procesos que hacen parte del trusted computing base (TCB) pueden llegar a compartir recursos con otros procesos, y esto los hace vulnerables a un ataque. Un problema común es compartir del

directorio temporal (/tmp UNIX – temp Windows), cualquier proceso puede crear archivos en este directorio y permitir que procesos TCB accedan a ellos.

- Time-of-Check-to-Time-of-Use (TOCTTOU): Un proceso no confiable puede llegar a cambiar el estado del sistema operativo y cambiar los tiempos en que una acción fue autorizada y realizada.
- Registro: El registro es una base de datos global que almacena información de todos los programas, cuando se carga una nueva aplicación actualiza el registro con información sensible como lo son paths a librerías y ejecutables, cada registro es asociado a un contexto de seguridad que limita el acceso, pero dichas limitaciones no siempre se configuran correctamente, convirtiéndolas en un posible vector de ataque.
- Usuarios administradores: La mayoría de usuarios utilizan cuentas administrativas para realizar sus tareas, el instalar programas o realizar configuraciones dentro del sistema se hace mucho más sencillo, pero se debe tener mucho cuidado ya que los procesos creados tienen los mismos privilegios que el usuario que los crea. Por lo tanto, un malware se ejecutaría con los mismos permisos que tiene el usuario.
- Configuraciones por defecto: Las configuraciones por defecto del sistema operativo, permiten a los atacantes explotar vulnerabilidades comunes dentro del sistema.
- Contraseñas inseguras: Es una vulnerabilidad muy común, no sólo a nivel de sistema operativo. Las contraseñas comunes o fáciles de adivinar les permiten a los atacantes adivinarlas con mayor facilidad poniendo en peligro la integridad del sistema.

Dichas vulnerabilidades se materializan en amenazas que pueden poner en riesgo el sistema operativo y cualquier aplicación que soporte. Con el fin de asegurar las diferentes plataformas computacionales, se han creado diferentes guías de endurecimiento (hardening) de sistemas operativos, donde se dictan pautas de lo

que debemos asegurar para garantizar un TCP, se coincide en que como administrador de un sistema se deben tener en cuenta al menos los siguientes aspectos:

- Actividades de login.
- Información de autorización.
- Información de autenticación.
- Comandos y aplicaciones usados por los usuarios.
- Cambios de estado del sistema.
- Transacciones a nivel de red.

En la sección de auditoría de sistemas operativos se trata de exponer la importancia de la revisión de los logs, que almacenan los eventos generados por el sistema operativo y sus aplicaciones, con el fin de detectar anomalías.

Cabe aclarar que dependiendo de la familia a la que pertenece el sistema operativo, ya sea Windows, Linux, Unix, BSD, etc. Se reportarán más o menos eventos, por ejemplo, el Sistema Operativo Windows, reporta eventos asociados a servicios como Kerberos, los cuales a su vez se dividen en autorización o autenticación.

Los eventos listados a continuación se han categorizado teniendo en cuenta su relación con los requerimientos de diseño seguro que buscan garantizar su Trusted Computing Base (TCB), dichas categorías no se encuentran relacionadas a algún tipo de sistemas, al contrario, se buscó que fueran lo más transversales posible:

- Inicios y cierres de sesión: Estos eventos se dividen a su vez en dos categorías, inicio y cierre. Dentro de esta categoría encontramos eventos asociados al tiempo en el que el usuario mantuvo la sesión abierta dentro del sistema, fallos y éxitos en el proceso de inicio de sesión, máquina desde la cual se realiza el inicio de sesión y demás.
- Administración de cuentas: Esta categoría encontramos eventos relacionados a la creación, actualización y eliminación de cuentas e

información relacionada a la administración de grupos del sistema, también podemos encontrar eventos relacionados a actualización de contraseñas, cuentas habilitadas o deshabilitadas, entre otros.

- **Acceso a objetos:** Estos eventos buscan llevar un registro de las acciones que realizan los usuarios sobre los objetos que residen en el sistema operativo como lo son aplicaciones, archivos, directorios, servicios o registro. Encontramos eventos relacionados a inicialización de aplicaciones, cambios sobre objetos del sistema de archivos como creación de archivos, cambios en los permisos y demás, peticiones de acceso a objetos, entre otros.
- **Privilegios:** Los eventos de esta categoría buscan alertar sobre posibles accesos o acciones no autorizadas de los usuarios ya sea sobre objetos o servicios del sistema.
- **Trazado de procesos:** Buscan llevar un registro de los procesos que se están llevando a cabo dentro del sistema con el fin de detectar procesos anormales, registra creación de eventos y terminación de los mismos, además de registrar eventos de procesos generados por dispositivos de entrada y salida o dispositivos externos al sistema.
- **Sistema:** Estos eventos se encuentran relacionados a garantizar la integridad del sistema operativo, podemos encontrar registros de instalación de aplicaciones, servicios, encendido o apagado del sistema, fallos y excepciones, dispositivos de entrada y salida, dispositivos externos del sistema, configuración del sistema operativo, operaciones criptográficas, entre otros.
- **Servicios:** Estos eventos registran las acciones que se realizan sobre los servicios que brinda el sistema operativo, aunque muchos de estos eventos se encuentran asociados al software base existen algunas categorías que son propias del sistema operativo, como el manejo de memoria, puertos

utilizados, accesos no autorizados y relaciones con otros procesos del sistema operativo.

2.1.3.2 Software base

El software base, como ya se mencionó anteriormente, comprende varias categorías de software como lo son las bases de datos, servidores web y servidores de aplicaciones cada una con vulnerabilidades, desafíos y tipos de eventos que podemos almacenar de cada uno.

En cuanto a las bases de datos podemos ver que las vulnerabilidades más comunes según un estudio de la empresa Imperva son(6):

- Privilegios excesivos, inapropiados o no usados: Esta vulnerabilidad se evidencia cuando se asignan más privilegios de los justamente necesarios para que las personas realicen su labor. También se ve presente cuando los usuarios cambian de roles dentro de las compañías y mantienen sus privilegios anteriores.
- Abuso de privilegios: Esta vulnerabilidad hace referencia a los usuarios que poseen gran cantidad de privilegios y por esta razón realizan acciones ilícitas sobre las bases de datos. Los usuarios más comunes que poseen este tipo de privilegios son los administradores de las bases de datos y los desarrolladores.
- Seguridad insuficiente en las aplicaciones web: Esta vulnerabilidad parte de que las aplicaciones son vulnerables y por medio de éstas se viola la seguridad de las bases de datos, los principales ejemplos de este tipo de vulnerabilidad son la inyección SQL y consolas de administración remotas.
- Rastro de auditorías débiles: Esta vulnerabilidad hace referencia a el registro que se debe tener de todas las operaciones que se realicen sobre la base de datos, para así prevenir posibles fraudes, como cambios no autorizados.
- Medios inseguros de almacenamiento: Esta vulnerabilidad hace referencia a la forma de almacenar los backups de las bases de datos, a la eliminación

de los dispositivos de almacenamiento y la protección física de los dispositivos de almacenamiento.

La auditoría en las bases de datos en la actualidad es un proceso que se puede realizar en tres diferentes lugares:

- Controles en las aplicaciones: Este tipo de monitoreo implica que cada aplicación que tenga acceso a la base de datos debe adicionar controles en el código para asegurar la base de datos. Este tipo de controles suele ser muy vulnerable.
- Controles en la red: Esta alternativa consiste poner el control en medio de la aplicación y la base de datos. Este tipo de control se realiza con un aplicativo que capture todas las comunicaciones que se realicen entre el aplicativo y la base de datos. Esta forma de auditoría no permite conocer el valor que se encontraba almacenado en las bases de datos, solo las sentencias de las transacciones y las respuestas a ellas.
- Control en la fuente: Esta alternativa consiste en poner los controles en la base de datos como tal, permitiendo tener un único control y permitiéndonos saber los datos que se encontraban almacenados antes de una transacción. Se considera que es la mejor alternativa.

Los tipos de auditoría que se pueden realizar sobre las bases de datos son:

- Auditoría de actividades: Consiste en monitorear y controlar las actividades que realizan los usuarios sobre los objetos de la base de datos, entiéndase por objetos, las tablas, vistas, etc.
- Auditoría de transacciones: Consiste en implementar una serie de controles que permiten llevar un registro de las transacciones que realiza un usuario con alto grado de detalle. Esto no solo nos permite saber que un usuario

interactuó con un objeto, sino que nos permite saber todas las operaciones que realizó con o sobre él.

En cuanto a servidores web podemos encontrar que existen varias vulnerabilidades, dentro de las que se encuentran:

- Verificación de versiones: Esta vulnerabilidad hace referencia a mostrar la versión del servidor web o del sistema operativo que lo está soportando, debido a que esta información puede ser usada para explorar formas de atacar basados en la falta de actualizaciones.
- Actualizaciones: Esta vulnerabilidad va ligada con la anterior ya que, si nuestro sistema se encuentra desactualizado (y además los atacantes lo pueden visualizar) la probabilidad de que un ataque suceda aumentará; es vital siempre mantener actualizado el servidor web.
- Uso inapropiado de módulos: Esta vulnerabilidad hace referencia a activar módulos en los servidores web que no se utilizan o dejar los que se encuentran por defecto. El problema radica en que algunos módulos poseen vulnerabilidades o configuraciones por defecto que pueden ser utilizadas como vector de ataque.
- Métodos HTTP innecesarios: Esta vulnerabilidad hace referencia a utilizar un conjunto por defecto de métodos HTTP que nuestro servidor web pueda aceptar. La mejor práctica para un servidor convencional es utilizar únicamente los métodos necesarios que en usualmente suelen ser: GET, POST y HEAD.
- Tamaño de peticiones no restringido: Esta vulnerabilidad parte del hecho de que si no existe una restricción al tamaño de las peticiones HTTP que recibe nuestro servidor estaremos expuestos a ataques de denegación de servicio.

En cuanto a la auditoría de servicios web, se realiza a través de los logs que estos generan, estos logs proporcionan información sobre:

- Alertas sobre actividades sospechosas

- Seguimiento de las actividades de un atacante
- Asistencia en la recuperación del sistema
- Asistencia en la investigación posterior a un evento
- Información requerida en procesos legales

El sistema de registro de eventos es diferente en cada servidor web. Las formas en las que pueden registrar los diferentes eventos son:

- Transfer Log: Cada transacción es almacenada, con toda la información relacionada a ella.
- Error Log: Cada error es registrado, incluyendo una explicación del por qué sucedió.
- Agent Log: Este tipo de registro contiene la información sobre el software que utilizó el cliente para acceder al contenido web.
- Referrer Log: Este tipo de registro recopila información sobre la petición HTTP.
- Common Log Format (CLF): Este formato guarda información sobre las transferencias en el siguiente orden:
 - Host remoto
 - Identidad del usuario remoto
 - Usuario autenticado
 - Fecha
 - URL pedida
 - Estado de la petición
 - Número de bytes transferidos
- Combined Log Format: Este formato contiene los mismos siete campos que el anterior y también provee información normalmente guardada en el Agent Log y el Referrer Log.
- Extended Log Format: Este formato proporciona una manera de describir todos los elementos que se pueden recolectar en un archivo de registro.

Algunos servidores proporcionan otros formatos diferentes como formatos de base de datos, con los que se pueden ampliar la información que se puede auditar.

La mayoría de servidores web aceptan el transfer log, que generalmente es considerado el más importante.

2.1.3.3 Aplicaciones

Las vulnerabilidades asociadas a las aplicaciones han sido de interés por diversas organizaciones, y por ello están estudiadas y clasificadas. Para este caso tomaremos como referente el OWASP Top Ten Project, un proyecto liderado por OWASP que busca generar conciencia acerca de la seguridad en las aplicaciones, y presenta una lista con las vulnerabilidades (y posibles riesgos que pueden ocurrir si se explotan) que se muestran a continuación(7):

1. Inyección (Injection): Ocurre cuando datos no confiables son enviados a un intérprete como parte de un comando o consulta, buscando engañar al mismo y ejecutar o acceder a datos no autorizados.
2. Pérdida de Autenticación y Gestión de Sesiones (Broken Authentication and Session Management): La mala implementación de funciones relacionadas a la autenticación y gestión de sesiones puede terminar en que los atacantes puedan suplantar la identidad de los usuarios.
3. Secuencia de Comandos en Sitios Cruzados (Cross-Site Scripting) (XSS): Ocurre cuando una aplicación toma datos no confiables y los envía al navegador web sin ninguna validación y/o codificación adecuada. El atacante a través de estos datos puede secuestrar la sesión del usuario, modificar sitios web o redireccionar al usuario a sitios maliciosos, entre muchos otros.
4. Referencia Directa Insegura a Objetos (Insecure Direct Object References): Ocurre cuando la aplicación hace vulnerable una referencia a un objeto interno (bien puede ser un archivo, un directorio, una base de datos, etc.).

Si no se verifica que el usuario tenga acceso a la referencia que está expuesta, ésta puede ser cambiada para acceder a datos no autorizados.

5. Configuración de Seguridad Incorrecta (Security Misconfiguration): Todas las configuraciones relacionadas con la seguridad de la aplicación deben ser definidas, implementadas y mantenidas ya que por lo general no son seguras por defecto. Si, por ejemplo, hay configuraciones por defecto en una librería utilizada, un atacante que conozca la librería podrá acceder a ella y podría afectar directamente el funcionamiento de la aplicación.
6. Exposición de datos sensibles (Sensitive Data Exposure): La mala protección de datos sensibles (como los números de tarjetas de crédito, credenciales, datos médicos, etc.) puede conllevar a la realización de fraudes, robos de identidad u otros delitos.
7. Ausencia de Control de Acceso a Funciones (Missing Function Level Access Control): Las aplicaciones deberían verificar el control de acceso a una función en el servidor, evitando que los atacantes puedan realizar peticiones sin la autorización apropiada.
8. Falsificación de Peticiones en Sitios Cruzados (Cross-Site Request Forgery) (CSRF): Un ataque CSRF obliga al navegador de una víctima autenticada a enviar una petición HTTP falsificada, con todos los datos de autenticación incluidos en el paquete, a una aplicación vulnerable. Esto permite al atacante forzar la sesión de la víctima y generar acciones que la aplicación vulnerable piensa que son peticiones legítimas.
9. Utilización de componentes con vulnerabilidades conocidas (Using Components with Known Vulnerabilities): Algunos componentes tales como las librerías, frameworks y otros módulos de software casi siempre funcionan con todos los privilegios. Si se ataca un componente vulnerable esto podría facilitar la intrusión en el servidor o una pérdida seria de datos. Las aplicaciones que utilicen componentes con vulnerabilidades conocidas debilitan las defensas de la misma y permiten ampliar la superficie de ataque.

10.Redirecciones y reenvíos no validados (Unvalidated Redirects and Forwards): La mala validación de los datos utilizados por la aplicación para realizar una redirección puede facilitar un ataque de phishing o malware por un atacante.

Ante esto, se han buscado maneras de prevenir y detectar a tiempo el momento en el que un atacante busque explotar alguna de las vulnerabilidades (incluyendo, pero no limitándose a las mencionadas anteriormente). Uno de los proyectos más importantes que se alinea a este propósito es el OWASP AppSensor que ya se había definido anteriormente, y que provee:

- Recomendaciones acerca de qué acciones de la aplicación deben ser detectadas como maliciosas junto con las respuestas sugeridas.
- Una guía para diseñar e implementar un sistema de detección y respuesta de ataques dentro de una aplicación.
- Una implementación de referencia Java que puede integrar en su aplicación como base para su mecanismo de Detección de Intrusos y Respuesta de la capa de aplicación (Application Layer Intrusion Detection and Response mechanism)

OWASP AppSensor, además, provee una lista con sondas (llamados puntos de detección) que deben ser colocadas en partes específicas del código donde se crea necesario, para identificar posibles atacantes que estén intentando probar vulnerabilidades o debilidades de la aplicación. Éstas se dividen en 12 categorías, la mayoría relacionadas con el OWASP Top Ten (por ejemplo, AccessControlException o CommandInjectionException).

También, al momento de que una sonda sea invocada, ésta puede generar una o varias respuestas automáticas previamente configuradas (avisar al administrador, enviar un evento a un SIEM, etc.). Dichas respuestas también han sido definidas por OWASP AppSensor como "Acciones de Respuesta" (Response Actions), y las categorizan de la siguiente manera:

- Respuesta silenciosa: El usuario no percibe ningún cambio.
- Respuesta pasiva: El proceso que está llevando a cabo el usuario se altera, pero puede continuar hasta terminarlo.
- Respuesta activa: Funcionalidades reducidas o desactivadas.
- Respuesta intrusiva: Acción no maliciosa sobre la información del usuario.

Para el propósito del trabajo, se entiende que cada evento será reportado vía respuesta silenciosa hacia el SIEM.

2.1.3.4 Caracterización de Activos en OSSIM

Para realizar la caracterización de un activo en OSSIM debemos seguir estos pasos:

1. Dirigirnos al menú Environment > Assets & Groups

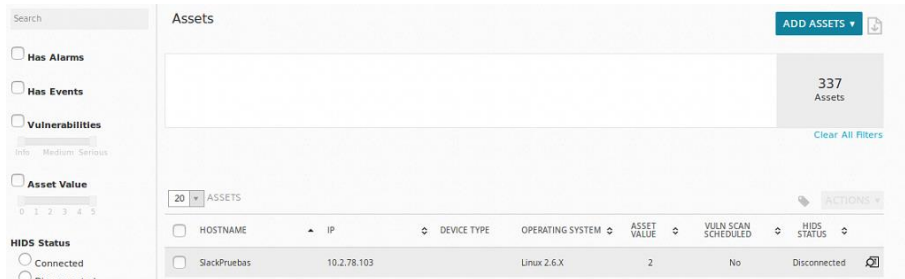


Ilustración 70 Vista de activos conectados al sistema.

2. En la lista de activos buscamos el que queremos caracterizar y damos click en el botón de lupa ubicado a la derecha. Nos aparecerá la vista detallada

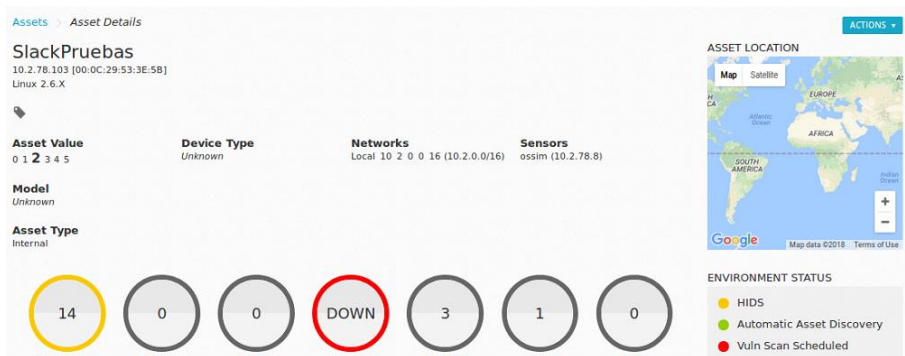


Ilustración 69 Vista detallada del activo.

del activo:

3. Abrimos el menú ACTIONS y seleccionamos la opción EDIT, donde nos aparecerá un formulario donde podremos modificar las características del activo como:
 - a. Nombre: Nombre del activo
 - b. IP: Dirección IP del activo.
 - c. Ubicación: Ubicación geográfica.
 - d. Alias: Si el activo tiene otro identificador.
 - e. Valor del activo: Importancia del activo para la organización. (1-5)

- f. Si es activo externo o no.
- g. Sistema operativo: Sistema operativo instalado.
- h. Descripción: Breve descripción del activo o sus funciones.
- i. Modelo: Modelo dado por el fabricante.
- j. Tipo de activo: Determinar si es un servidor, dispositivo de red, host, dispositivo médico, entre otros.

The screenshot shows a web form titled "EDIT ASSET". It has several sections:

- IP Address ***: Text input with "10.2.78.103".
- Location**: Text input with "AK 45 (Autonorte) #205-59, Bogotá, Cundinamarca, Colombia".
- FQDN/Aliases**: Empty text input.
- Asset Value ***: Dropdown menu with "0".
- External Asset ***: Radio buttons for "Yes" (unselected) and "No" (selected).
- Sensors ***: A list with a checked checkbox for "10.2.78.8 (ossim)".
- Operating System**: Text input with "Linux 2.6.X".
- Model**: Empty text input.
- Description**: Empty text input.
- Devices Types**: A dropdown menu with "Server" selected, and a sub-menu with "HTTP Server" selected. An "ADD" button is next to it.
- Map**: A Google Map showing Bogotá, Colombia, with a red pin. Below the map, "Latitude/Longitude" is "4.7827" and "-74.0426".
- SAVE**: A blue button at the bottom center.

Ilustración 71 Formulario de edición del activo.

4. Al finalizar damos clic al botón SAVE y quedará caracterizado el activo.

2.1.4 Definición de alertas y reglas de correlación

La correlación de eventos identifica amenazas potenciales por medio de la detección de patrones de comportamiento a través de diferentes tipos de activos. La correlación enlaza diferentes eventos, convirtiendo los datos en información

más valiosa. En OSSIM la correlación se realiza por medio de directivas de correlación que a su vez se componen de reglas, la relación entre reglas se determina según su indentación, si se encuentran en el mismo nivel se toma como OR y si están en un nivel inferior se toma como AND(8).

	NAME	RELIABILITY	TIMEOUT	OCCURRENCE	FROM	TO	DATA SOURCE	EVENT TYPE	STICKY DIF	
	Firewall dropped packet	2	None	1	ANY	ANY	cisco-pix (1514)	SIDs: 106011	None	More
AND	Firewall dropped packets	4	10	3	ANY	1.DST_IP	cisco-pix (1514)	SIDs: 106011	None	More
AND	Firewall dropped packets	6	20	5	ANY	1.DST_IP	cisco-pix (1514)	SIDs: 106011	None	More
OR	Firewall dropped packets	8	30	10	ANY	1.DST_IP	cisco-pix (1514)	SIDs: 106011	None	More

Ilustración 72 Relación entre reglas de Correlación

En OSSIM existen diferentes tipos de directivas que se presentan en la siguiente tabla:

Tabla 4 Clasificación de Directivas de correlación en OSSIM

Categoría	Explicación	Ejemplo
AlienVault Network	Directivas para detectar anomalías en la red y ataques.	AV Network attack, too many dropped inbound packets from DST_IP
AlienVault Policy	Directivas para detectar violación de políticas.	AV Policy violation, vulnerable Java version detected on SRC_IP
AlienVault Scada	Directivas para detectar ataques en sistemas SCADA.	AV SCADA attack, Modbus scanning or fingerprinting against DST_IP

AlienVault DoS	Directivas que detectan ataques de denegación de servicio sobre diferentes aplicaciones y servicios.	AV Service attack, successful denial of service against IIS web server on DST_IP (MS07-041)
AlienVault Malware	Directivas para detectar malware	AV Malware, botnet Koobface activity detected on SRC_IP (source IP)
AlienVault Misc	Directivas para detectar actividades que no se relacionan a ninguna categoría.	AV Misc, suspicious executable download from a dynamic domain on SRC_IP
AlienVault Scan	Directivas para detectar actividades de escaneo.	AV Network scan, Nmap scan against DST_IP
User Contributed	Directivas creadas o modificadas por el usuario. Por defecto esta categoría está vacía.	
Ataques Alienvault	Directivas para detectar ataques sobre servicios y	AV Attacks, Successful OpenSSL HeartBeat attack

	aplicaciones vulnerables.	
AlienVault BruteForce	Directivas para detectar ataques de fuerza bruta que requieren autenticación.	AV Bruteforce attack, SSH authentication attack against DST_IP (destination IP)

Directivas de Correlación

Cada directiva de correlación tiene las siguientes propiedades globales:

Tabla 5 Propiedades Globales de las directivas de correlación en OSSIM

Propiedad Global	Descripción
ID	Es un identificador único para la directiva, se convierte en el id del tipo de evento cuando un nuevo evento es creado. Nota: El ID de la directiva no se muestra en la interfaz web.
Name	Es el nombre significativo que se le da a la directiva. Se convierte en el nombre del evento o alarma.
Intent, Strategy, Method	Describe lo que la directiva de correlación trata de detectar, estas categorías ayudan a clasificar la directiva dentro del USM Appliance Taxonomy.
Priority	Define el impacto del ataque. Es usado para el cálculo el riesgo de un evento.

Para crear una Directiva de Correlación se deben seguir estos pasos:

1. Dirigirse al menú Configuration > Threat Intelligence > Directives.
2. Click en el botón New directive

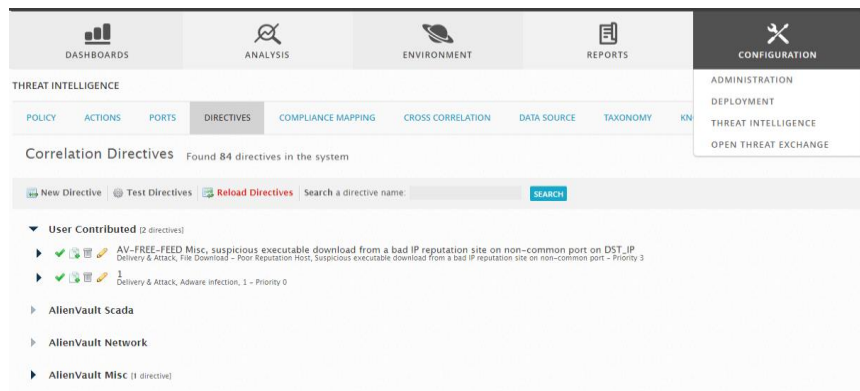


Ilustración 73 Vista de directivas de Correlación

3. Aparecerá la siguiente ventana donde debemos llenar los campos correspondientes, en este caso se creará una directiva de ataques de denegación de servicio. Al finalizar hacemos clic en el botón “Next” y la directiva será agregada.

Ilustración 74 Formulario para creación de Directivas de Correlación

Reglas de correlación

Una regla de correlación define una condición para detectar anomalías o ataques de los que llegan al SIEM, cada regla de correlación tiene los siguientes atributos(9):

Tabla 6 Propiedades Globales de las reglas de correlación en OSSIM

Atributo	Descripción
Nombre	Nombre de la regla, cada regla tiene su propio nombre dentro de la directiva.
Confiabilidad	El nivel de confiabilidad que se le da al evento. En una escala del 1 al 10.

Timeout	La cantidad de tiempo que toma que la regla expire y el proceso de correlación asociado a esa regla aborte. Se mide en segundos. Nota: El timeout por defecto es None, es decir que la regla no expira.
Ocurrencia	Número de veces que el evento debe ocurrir.
Desde	IP y puertos de origen.
A	IP y puertos de destino.
Data Source	Nombre e ID del plugin que la regla trata de emparejar.
Event Type	Tipo de evento (SID). Note: Cuando se incluyen diferentes SID's la regla trata de detectar alguno de ellos.
Sensor	El Sensor que envía los eventos.
Protocol	Protocolo especificado en un evento, se aceptan valores como cualquiera, TCP, UDP e ICMP.
Sticky Dif	Por defecto los atributos en una directiva se definen como "sticky".
Username	Usuario definido en el evento
Pass	Contraseña definida para el evento
Userdata1- Userdata9	Los datos de usuario definidos para el evento

Para crear una regla de correlación se deben seguir los pasos que se mencionan a continuación: En este caso se agregará una regla de nivel 1 para detectar accesos a un activo Cisco ASA por el puerto 139.

1. Establecemos el nombre de la regla:



Ilustración 75 Formulario para creación de reglas de correlación

2. Seleccionamos los tipos de eventos que serán correlacionados, se pueden agregar de dos formas por plugin (eventos asociados a un tipo de plugin) o por taxonomía (según tipo de activo). Al seleccionar por plugin se podrán elegir los tipos de eventos que serán emparejados por la regla.

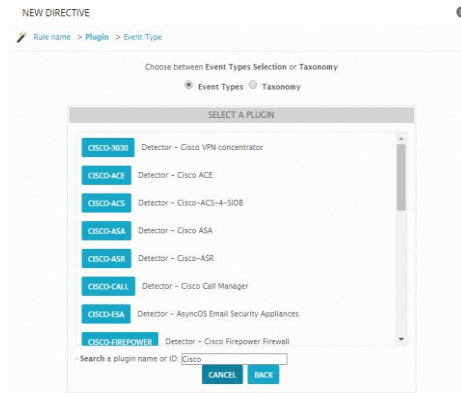


Ilustración 76 Selección de eventos por tipo

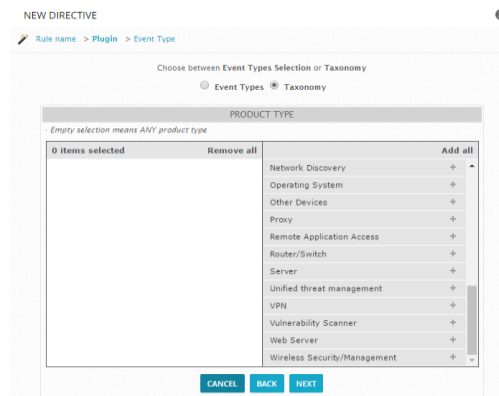


Ilustración 77 Selección de eventos por taxonomía

Seleccionamos Tipos de eventos para plugin Cisco ASA, para agregar damos clic en el botón “+”, si queremos añadirlos todos hacemos clic en el label “add all”, también podemos hacer búsquedas específicas, después de agregados los tipos de eventos hacemos clic en el botón next:

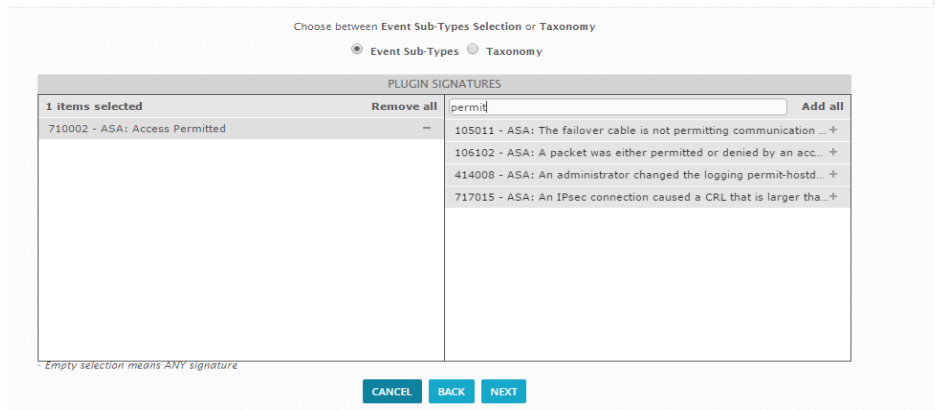


Ilustración 78 Selección de tipos de eventos para la regla de correlación creada.

- Determinamos los aspectos relacionados a la red, aparecerá la ventana de network, donde podremos establecer si serán eventos de un activo en específico, a un activo en específico o un grupo de activos. Si lo dejamos vacío significa que será de cualquier fuente y a cualquier activo agregado a OSSIM.

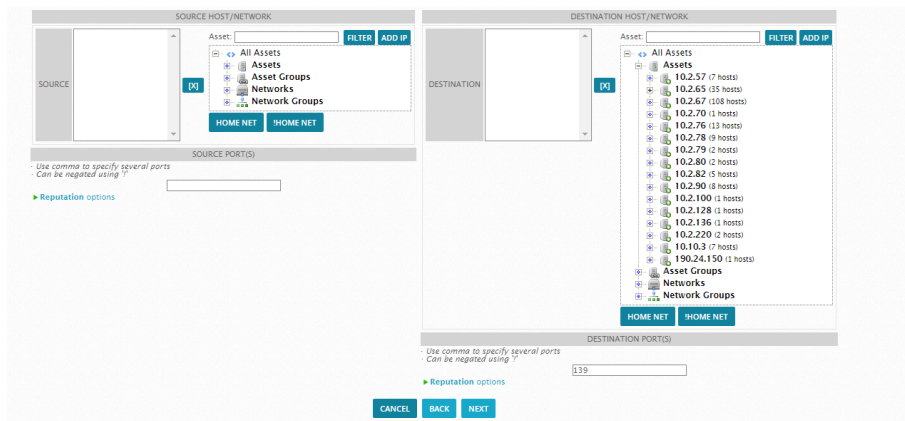


Ilustración 79 Selección de aspectos relacionados a la red.

- Si queremos que también se tengan en cuenta aspectos de reputación de la IP, seleccionamos Reputation options -> Yes, deberemos especificar la

The screenshot shows the 'NETWORK' configuration interface. It is divided into two main sections: 'SOURCE HOST/NETWORK' and 'DESTINATION HOST/NETWORK'. Each section has an 'Asset' dropdown menu with options like 'All Assets', 'Assets', 'Asset Groups', 'Networks', and 'Network Groups'. Below these are 'SOURCE PORT(S)' and 'DESTINATION PORT(S)' input fields. A 'Reputation options' section is expanded, showing 'Reputation from: No', 'Min Priority: -', and 'Min Reliability: -'. The 'DESTINATION' side has 'Reputation to: Yes', 'Min Priority: 1', and 'Min Reliability: 1'. At the bottom, there are 'CANCEL', 'BACK', and 'NEXT' buttons.

prioridad y la confiabilidad. Hacemos click en el botón "Next".

Ilustración 80 Selección de aspectos relacionados a la reputación de la IP.

- Nos aparecerá la siguiente pantalla donde podremos terminar de definir la regla (clic en "Finish") o especificar parámetros asociados a protocolos, sensor, y demás campos (clic en "Next"). (Para más información revise la tabla 5)

The screenshot shows a dialog box titled 'RULE DEFINED'. It contains the text 'Would you like to specify any other condition for this rule (Protocol, Sensor, Special fields...)?' and three buttons: 'BACK', 'FINISH', and 'NEXT'.

Ilustración 81 Ventana emergente que aparece cuando terminamos de definir la regla de correlación.

The screenshot shows a table of rules. The table has columns: NAME, RELIABILITY, TIMEOUT, OCCURRENCE, FROM, TO, DATA SOURCE, EVENT TYPE, and ACTION. The rule shown is 'Established connections' with a reliability of 10, no timeout, and an occurrence of 1. The data source is 'cisco-asa (1636)' and the event type is 'SIDs: 710002'. There is a 'More' link and a plus sign in the action column.

NAME	RELIABILITY	TIMEOUT	OCCURRENCE	FROM	TO	DATA SOURCE	EVENT TYPE	ACTION
Established connections	10	None	1	ANY	ANY	cisco-asa (1636)	SIDs: 710002	More +

Ilustración 82 Regla de correlación definida

Reglas de correlación de nivel 2 o más:

Para este ejemplo de creación de reglas de nivel 2 o más se debe tener en cuenta que para que se haga la correlación de forma correcta se debe tener el mismo tipo de evento, los mismos activos de fuente y destino y el mismo puerto. La diferencia se verá en la ocurrencia:

1. Hacemos clic en el botón “+” bajo la opción de “Action”, nos aparecerá el mismo menú de agregar una nueva regla.

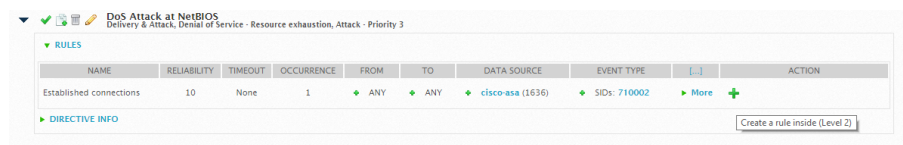


Ilustración 83 Creación de regla de nivel 2

2. En el menú de “Network” elegiremos que el destino y la fuente sean los mismos que la regla padre:

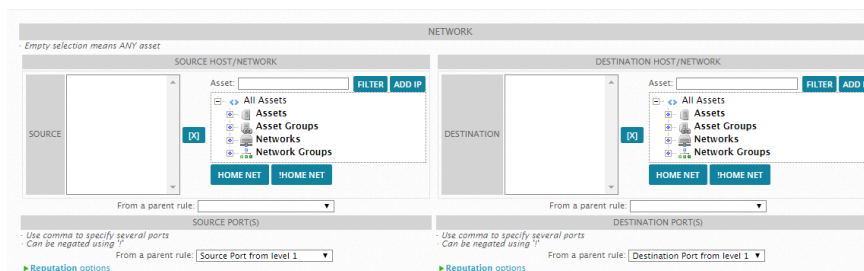


Ilustración 84 Creación de regla de nivel 2. Configuraciones de red.

3. Nos aparecerá la definición de confianza, las opciones con valor “=” son absolutas y las que tienen valor “+” son relativas, en este ejemplo usamos “+2”. Hacemos clic en el botón “Finish”.



Ilustración 85 Creación de regla de nivel 2. Definición del nivel de confianza.

- Para cambiar la ocurrencia hacemos clic sobre la opción “Ocurrence” y la editamos a 100.

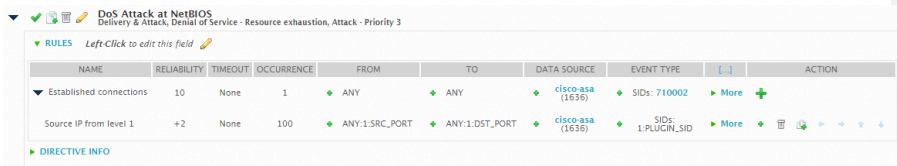


Ilustración 86 Creación de regla de nivel 2. Modificación en la ocurrencia.

Para editar las reglas simplemente hacemos clic sobre el campo que queremos modificar, en la opción “more” podremos editar los campos más específicos:

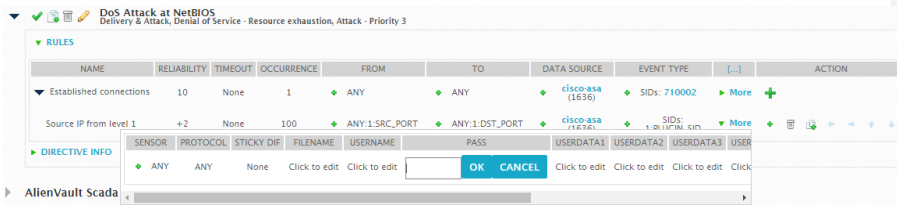


Ilustración 87 Creación de regla de nivel 2. Campos adicionales en la regla de correlación.

- Para que se carguen las reglas que acabamos de definir, se debe hacer click en el botón “Reload directives”, y seleccionamos la opción “Yes” en el menú emergente, que refrescará los datos y añadirá las directivas y reglas añadidas, se debe tener en cuenta que la nueva directiva se añadirá en la sección “User contributed”.

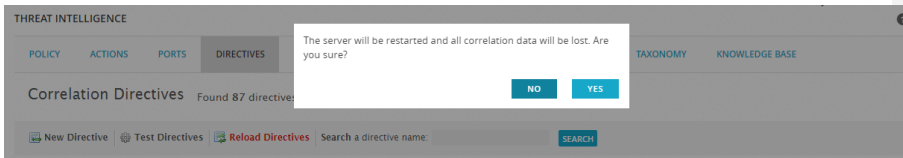


Ilustración 88 Creación de reglas de correlación. Recargar directivas.

Si se quiere una mayor información sobre las reglas de correlación implementadas como parte del desarrollo de PGR1, consulte el artículo de investigación asociado.

2.1.5 Definición de respuestas(10)

Las políticas cumplen un papel fundamental en la respuesta a un incidente. Permiten configurar cómo el sistema procesa los eventos una vez llegan a la plataforma.

Las políticas tienen condiciones y consecuencias:

- **Condiciones:** Determina qué eventos serán procesados por la política.
- **Consecuencias:** Define qué pasará cuando los eventos cumplan ciertas condiciones.

Para acceder a la vista de políticas nos dirigimos a **Configuration > Threat Intelligence > Policy**

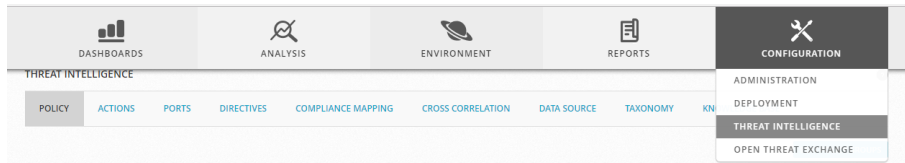


Ilustración 89 Vista de configuración

Las políticas pueden ser administradas y creadas usando grupos, se dividen en 3 secciones:

- **Default policy Group:** No contiene políticas por defecto. Se usa como contenedor de políticas creadas para procesar eventos externos. Los eventos externos son los recolectados dentro de la red a través de los sensores.
- **AV Default Policies:** Filtra los eventos de AVAPI un servicio de OSSIM que ejecuta diferentes tareas.
- **Policies for events generated in the server:** Es un contenedor que se usa para procesar los eventos de sistema, también son llamados eventos de directiva.

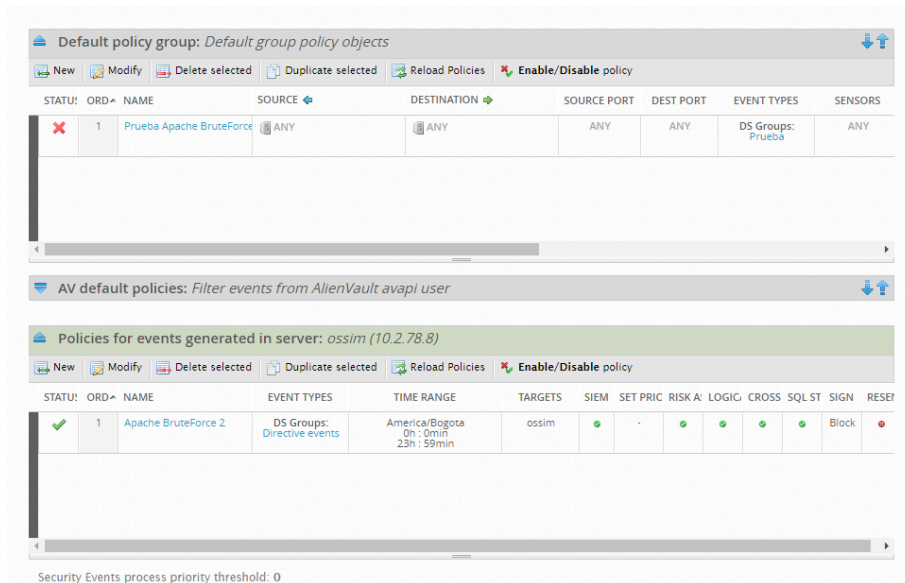


Ilustración 90 Vista de políticas de OSSIM

Las políticas tienen dos partes una condición y una consecuencia. Las condiciones determinan qué eventos van a ser procesados que pueden ser de dos tipos:

Tabla 7 Tipos de eventos que son procesados en OSSIM

Condición	Explicación
Data Sources	Es una aplicación o dispositivo que genera información que es recolectada y analizada. OSSIM incluye varios data sources que monitorean tráfico y activos para detectar eventos.
Taxonomy	Es una clasificación, usando categorías generales de los eventos. Se componen de categorías y subcategorías. Se utiliza el tipo de producto, categoría y subcategoría para crear una condición taxonómica

En OSSIM las políticas las podemos ver así:

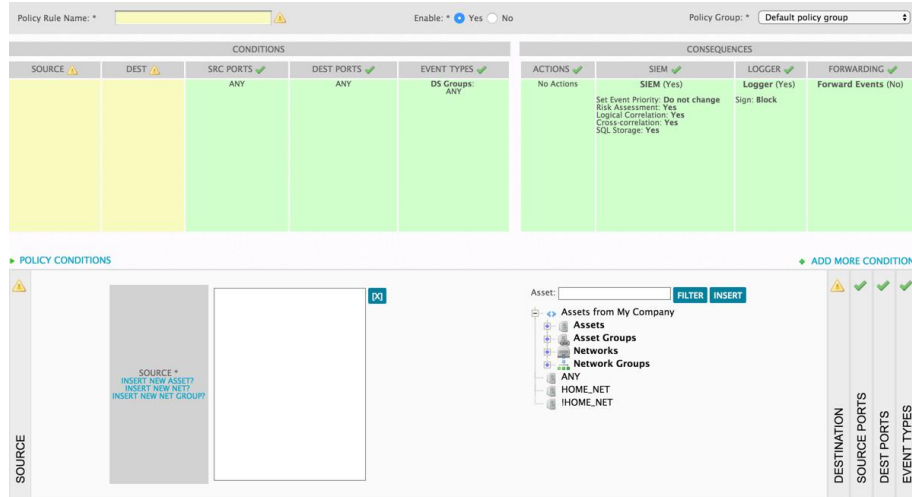


Ilustración 91 Visualización de una política

Los atributos necesarios para generar una condición se explican en la siguiente tabla:

Tabla 8 Atributos de una política en OSSIM- Condiciones

Atributo	Explicación
Source	Define la fuente del evento puede ser un activo, grupo de activo, una red o un conjunto de redes. Cuando se elige una fuente se dice que sólo los eventos de esa fuente serán procesados por la política.
Destination	Define el destino de los eventos.
Source Ports	Define los puertos TCP/UDP que son fuente de los eventos.

Destination	Definen los puertos TCP/UDP de destino de los eventos.
Ports	
Tipos de eventos	Define cuáles eventos van a ser procesados por la política, estos eventos se dividen en dos categorías.
Sensors	Define el sensor que recolecta y normaliza los eventos.
Reputation	Ayuda a filtrar eventos que tienen asociado cierto nivel de prioridad. Ayuda a definir el uso de OTX en la evaluación de direcciones IP en un evento.
Event Priority	Ayuda a filtrar qué eventos serán procesados por la política basados en la prioridad y confiabilidad del evento.
Time Range	Determina un rango de tiempo para que los eventos sean procesados.

Las consecuencias determinan las acciones que se toman con los eventos que cumplen las condiciones especificadas.

Tabla 9 Atributos de una política en OSSIM- Consecuencias

Atributo	Explicación
Actions	Es una consecuencia de una política, puede ser, enviar un correo, ejecutar un script que se encuentre en el servidor OSSIM , abrir un ticket.
SIEM	Permite realizar un procesamiento tipo SIEM.

Las acciones son la consecuencia de una política, cuando un evento cumple con las condiciones lanza una acción que puede ser:

- Enviar un correo
- Ejecutar un script o programa que esté dentro del servidor OSSIM.
- Abrir un nuevo ticket usando el sistema de tickets.

Action keywords: Son información de los eventos que pueden ser usados como parámetros de acción. Pueden ser utilizados como parámetros para email o como parámetros dentro de un programa.

Las acciones en OSSIM se pueden ver así:

You can use the following keywords within any field which will be substituted by its matching value upon action execution:

- DATE
- PLUGIN_ID
- PLUGIN_SID
- RISK
- PRIORITY
- RELIABILITY
- SRC_IP_HOSTNAME
- DST_IP_HOSTNAME
- SRC_IP
- DST_IP
- SRC_PORT
- DST_PORT
- PROTOCOL
- SENSOR
- BACKLOG_ID
- EVENT_ID
- PLUGIN_NAME
- SID_NAME
- USERNAME
- PASSWORD
- FILENAME
- USERDATA1
- USERDATA2
- USERDATA3
- USERDATA4
- USERDATA5
- USERDATA6
- USERDATA7
- USERDATA8
- USERDATA9

Action keywords

NAME * Notify CEO of Attack on Critical Servers

CONTEXT * My Company

DESCRIPTION *

TYPE * Send an email message

CONDITION Any Only if it is an alarm Define logical condition

FROM: *

TO: * bmehta@alienvault.com

SUBJECT: * Severity 8 event on mission-critical server

MESSAGE: * Attack discovered affecting DST_PORT DST_IP SRC_IP PRIORITY RISK

Ilustración 92 Vista de acciones

Para crear una política en OSSIM debemos seguir los siguientes pasos:

1. Dirigirnos a Configuration>Threat Intelligence>Policy>New (Recuerde primero configurar la acción). Se llenan los campos necesarios.

Default policy group: Default group policy objects

New Modify Delete selected Duplicate selected Reload Policies Enable/Disable policy

STATUS	ORD	NAME	SOURCE	DESTINATION	SOURCE PORT	DEST PORT	EVENT TYPES	SENSORS	TIME RANGE	TARGETS	SIEN
✓	1	Prueba Apache BruteForce	ANY	OracleDB	ANY	ANY	DS Groups: Apache BruteForce	ANY	America/Bogota 0h: 0min - 23h: 59min	ossim	
✓	2	DoS Policy Test	ANY	coral laboratoriois	ANY	ANY	DS Groups: DoS DoS Failures	ANY	America/Bogota 0h: 0min - 23h: 59min	ossim	
✗	3	IoT	ANY	ANY	ANY	ANY	DS Groups: IoT	ANY	America/Bogota 0h: 0min - 23h: 59min	ossim	

Ilustración 93 Vista de políticas señalando la acción 'New'

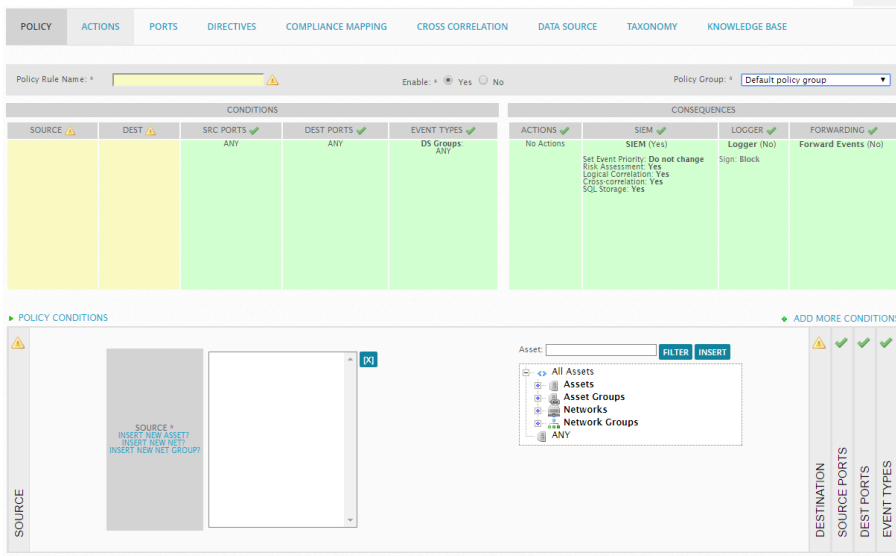


Ilustración 94 Vista de una nueva política

- Para establecer las consecuencias se debe dar click en el enlace consecuencias, al finalizar haga click en el botón "Update Policy".

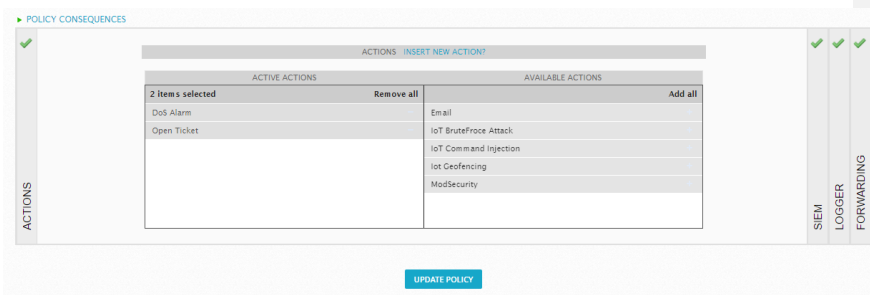


Ilustración 95 Vista de consecuencias

3. Haga click en el botón reload policies, para que sean cargadas en el servidor:

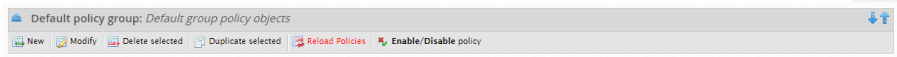


Ilustración 96 Estado de recarga de políticas OSSIM

Ilustración 97 Vista de políticas señalando la acción 'Reload Policies'

4. Compruebe que la política tenga un chulito verde a un lado, es decir que se encuentre activa en el sistema:



Ilustración 98 Comprobación de la creación de la política

2.1.6 Escenarios aplicados

En esta parte se hablará de los casos prácticos que implementamos para el Laboratorio de Informática, utilizando la lógica de OSSIM para detectar y prevenir escenarios concretos.

2.1.6.1 Escenario 1: Ataque distribuido de fuerza bruta con SSH y denegación de servicio.

En este escenario se buscaba prevenir un ataque distribuido de fuerza bruta con SSH y denegación de servicio que se estaba presentando sobre los servidores del Laboratorio de Informática. Un ataque distribuido es un ataque en el que varios computadores (normalmente sin que el dueño se entere) atacan a un mismo objetivo. Visto desde la parte del objetivo, parecería que son muchas peticiones de diferentes partes del mundo para acceder a su información, pero el motivo real es comprometer el mismo sin que el ataque sea fácil de detectar.

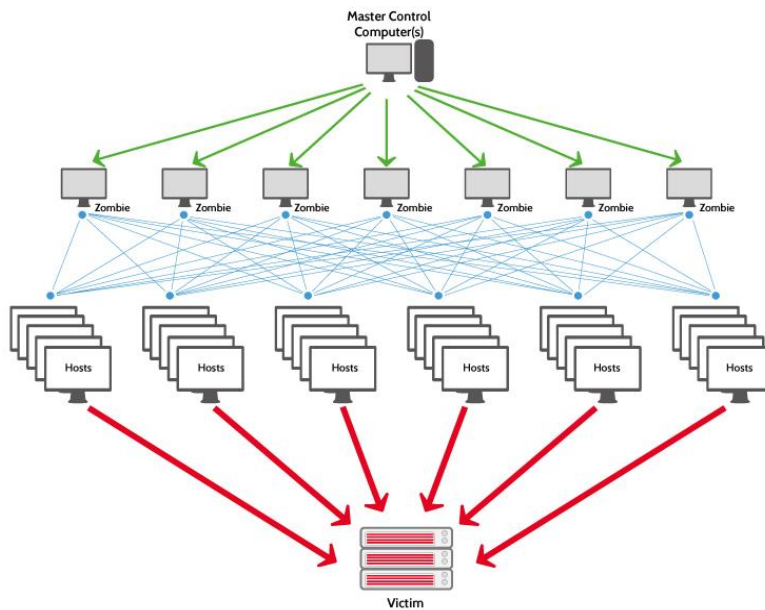


Ilustración 99 Esquema básico de un ataque distribuido

Se realizó un esquema de reglas de correlación anidada para detectar las diferentes fases de este ataque, como se puede observar en la siguiente ilustración:

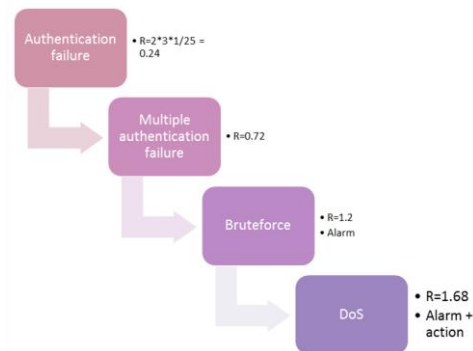
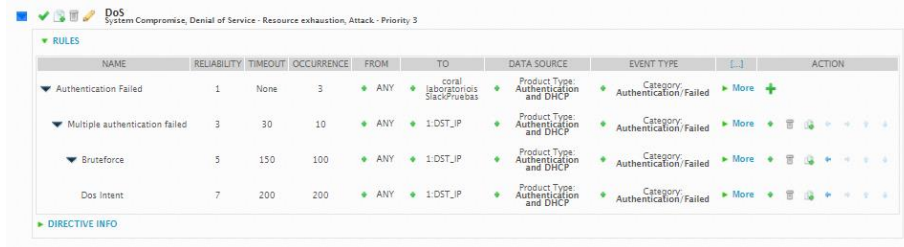


Ilustración 100 Esquema de reglas de correlación anidadas para el escenario 1

Al momento de ser visualizadas en OSSIM se ven de la siguiente manera:



The screenshot shows the OSSIM Rules configuration interface. At the top, it indicates 'DoS System Compromise, Denial of Service - Resource exhaustion, Attack - Priority 3'. Below this, there is a 'RULES' section with a tree view. The tree is expanded to show four rules:

NAME	RELIABILITY	TIMEDOUT	OCCURRENCE	FROM	TO	DATA SOURCE	EVENT TYPE	ACTION
Authentication Failed	1	None	3	ANY	coral laboratorios Slack/Firebas	Product Type: Authentication and DHCP	Category: Authentication/Failed	More +
Multiple authentication failed	3	30	10	ANY	1.DST_IP	Product Type: Authentication and DHCP	Category: Authentication/Failed	More + [Icons]
Bruteforce	5	150	100	ANY	1.DST_IP	Product Type: Authentication and DHCP	Category: Authentication/Failed	More + [Icons]
Dos intent	7	200	200	ANY	1.DST_IP	Product Type: Authentication and DHCP	Category: Authentication/Failed	More + [Icons]

Below the rules list, there is a 'DIRECTIVE INFO' section.

Ilustración 101 Esquema de reglas de correlación anidadas vistas en OSSIM

Luego, para probar que efectivamente las alertas se estaban creando, se procedió a realizar un ataque a través de la herramienta *hydra*, como se puede observar a continuación:

```
root@kali:~# ping 10.2.78.103
PING 10.2.78.103 (10.2.78.103) 56(84) bytes of data.
64 bytes from 10.2.78.103: icmp_seq=1 ttl=64 time=0.285 ms
64 bytes from 10.2.78.103: icmp_seq=2 ttl=64 time=0.292 ms
^C
--- 10.2.78.103 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1026ms
rtt min/avg/max/mdev = 0.285/0.288/0.292/0.017 ms
root@kali:~# hydra 10.2.78.103 ssh -s 22 -L user_wordlist -P pass_wordlist -f -vv
```

Ilustración 102 Ataque de autenticación SSH por fuerza bruta

SHOW 20 ENTRIES

DATE	STATUS	INTENT & STRATEGY	METHOD	RISK	OTX	SOURCE	DESTINATION
2017-05-11 03:57:00	open	Denial of Service - Resource exhaustion	Attack	LOW (11)		61.177.172.52:1872	coral
2017-05-11 03:49:58	open	Denial of Service - Resource exhaustion	Attack	LOW (11)		58.218.198.151:34732	coral
2017-05-11 03:43:21	open	Denial of Service - Resource exhaustion	Attack	LOW (11)		61.177.172.52:21550	laboratorois
2017-05-11 03:35:55	open	Denial of Service - Resource exhaustion	Attack	LOW (11)		58.218.198.151:42471	laboratorois
2017-05-11 03:28:41	open	Denial of Service - Resource exhaustion	Attack	LOW (11)		58.218.198.151:18291	laboratorois
2017-05-11 03:21:22	open	Denial of Service - Resource exhaustion	Attack	LOW (11)		58.218.198.151:58220	coral
2017-05-11 03:16:02	open	Denial of Service - Resource exhaustion	Attack	LOW (11)		61.177.172.52:24346	coral

Ilustración 103 Alertas creadas por OSSIM al detectar el ataque

Denial of Service - Resource exhaustion — Attack

Status	Risk	Attack Pattern	Created	Duration	# Events	Alarm ID	OTX Indicators
Open	LOW (11)	internal many-to-one	5 days ago	42 secs	311	298BF098366211E7A5F000CA8687B08	0

Source (2) 10.2.67.204

Host-10-2-67-204 (10.2.67.204) Location: Unknown

Asset Groups: Unknown

Networks: Local_10_2_0_0_16

OTX IP Reputation: No

VULNERABILITIES OPEN PORTS PROPERTIES NOTES

5 VULNERABILITIES

SCAN TIME	VULNERABILITIES	VULN ID	SERVICE	SEVERITY
No vulnerabilities found in the system				

SHOWING 0 TO 0 OF 0 VULNERABILITIES

Other Details: SIEM Events, Raw Logs, Honey-Pot, Whois, Reverse-DNS

Destination (1) 10.2.78.103

SlackPruebas (10.2.78.103) Location: Unknown

Asset Groups: Pruebas

Networks: Local_10_2_0_0_16

OTX IP Reputation: No

VULNERABILITIES OPEN PORTS PROPERTIES NOTES

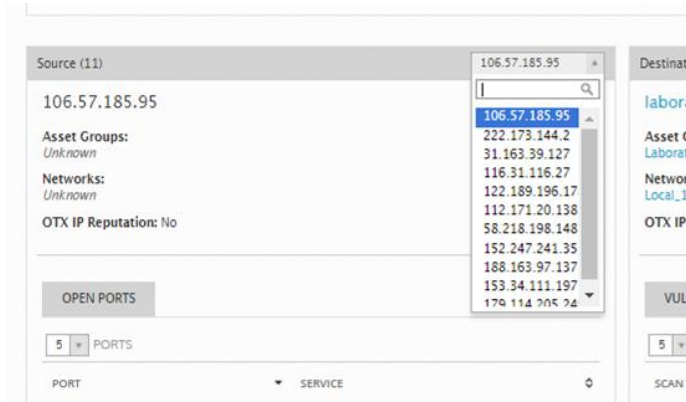
5 VULNERABILITIES

SCAN TIME	VULNERABILITIES	VULN ID	SERVICE	SEVERITY
2017-03-29 19:54:18	SSH Weak Encryption Algorithms Supported	105611	ssn (22/tcp)	Medium
2017-03-29 19:54:18	TCP timestamps	80091	general (80/tcp)	Medium
2017-03-29 19:54:18	SSH Weak MAC Algorithms Supported	105610	ssn (22/tcp)	Medium
2017-03-29 19:54:18	Services	10330	ssn (22/tcp)	Info
2017-03-29 19:54:18	SSH Protocol Algorithms Supported	105565	ssn (22/tcp)	Info

SHOWING 1 TO 5 OF 14 VULNERABILITIES

Ilustración 104 Detalle de una alerta proveniente de las reglas anteriormente creadas

Al dar clic en el apartado de **Source**, podemos apreciar que el ataque ha sido realizado desde varias IPs, comprobando que se ha detectado un ataque distribuido.



Finalmente, se estaba comprobando que en efecto había un ataque, pero aún no se había configurado una respuesta, así que se decidió optar por lanzar un script (que residía en el servidor objetivo) que bloqueará las IPs implicadas al momento de detectar el ataque. Para ello se definieron las políticas de la siguiente manera:

STATUS	ORD	NAME	SOURCE	DESTINATION	SOURCE PORT	DEST PORT	EVENT TYPES	SENSORS	TIME RANGE	TARGETS	SEN
✓	1	Apache BruteForce	ANY	laboratoriois	ANY	ANY	DS Groups: Apache BruteForce	ANY	America/Bogota (0h : 0min - 23h : 59min)	ossim	
✓	2	DoS Policy	ANY	laboratoriois	ANY	SSH	DS Groups: DoS Port Failures	ANY	America/Bogota (0h : 0min - 23h : 59min)	ossim	

Ilustración 105 Definición de políticas para el escenario 1

Política 1: Política de DoS

Esta política está asociada a los eventos de SSH, ya sean intentos de inicio de sesión fallidos o de usuario no existente, del activo laboratoriois por el puerto 22. La consecuencia asociada es una acción por medio de línea de comandos que bloquea la IP asociada al posible ataque.

Y para la respuesta se configuró dicho apartado de la siguiente manera:

You can use the following keywords within any field which will be substituted by its matching value upon action execution:

- DATE
- PLUGIN_ID
- PLUGIN_SID
- RISK
- PRIORITY
- RELIABILITY
- SRC_IP_HOSTNAME
- DST_IP_HOSTNAME
- SRC_IP
- DST_IP
- SRC_PORT
- DST_PORT
- PROTOCOL
- SENSOR
- BACKLOG_ID
- EVENT_ID
- PLUGIN_NAME
- SID_NAME
- USERNAME
- PASSWORD
- FILENAME
- USERDATA1
- USERDATA2
- USERDATA3
- USERDATA4
- USERDATA5
- USERDATA6
- USERDATA7
- USERDATA8
- USERDATA9

NAME *	Bloquear IP
DESCRIPTION *	Bloquea una ip dada
TYPE *	Execute an external program
CONDITION	<p><input type="radio"/> Any <input type="radio"/> Only if it is an alarm <input checked="" type="radio"/> Define logical condition</p> <p>Python boolean expression: <input type="text" value="RISK > 2"/> (*) <small>Up to 255 characters</small></p> <p>Only on risk increase: <input type="checkbox"/></p>
COMMAND: *	/opt/scripts/blockIP.sh DST_IP SRC_IP
TO: *	email;email;email

Ilustración 106 Configuración de la respuesta para el escenario 1

2.1.6.2 Escenario 2: Ataque de denegación de servicio distribuido, mediante HTTP Flood (saturación HTTP).

Para este caso, se decidió proteger un servidor web (Apache, software base), y el ataque a evitar es el llamado DDoS (Distributed Denial of Service), en el que, tal y como se explicó en el escenario 1, muchas IPs buscan conectarse al servidor web para que el servidor colapse, en lugar de que una sola se conecte muchas veces ya que es detectada y bloqueada con facilidad.

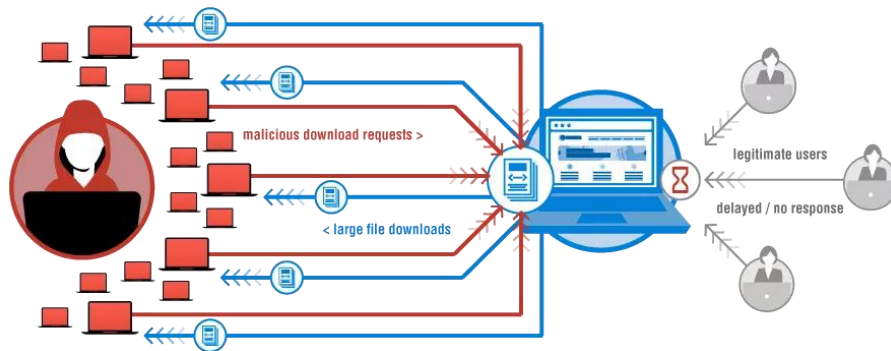


Ilustración 107 Ataque de denegación de servicio distribuido

Para la definición de reglas de correlación fue también similar al escenario 1, en la que se tienen diferentes fases. En OSSIM se pueden visualizar de la siguiente manera:

Bruteforce 404
Delivery & Attack, WebServer Attack, Multiple 404 headers - Priority 3

NAME	RELIABILITY	TIMOUT	OCCURRENCE	FROM	TO	DATA SOURCE	EVENT TYPE	ACTION
Bruteforce 404	2	None	1	ANY	laboratoriois OracleDB	apache (1501)	SIDs: 404	More +
Bruteforce 404	4	5	20	ANY	1.DST_IP	apache (1501)	SIDs: 404	More +
Bruteforce 404	6	15	280	ANY	1.DST_IP	apache (1501)	SIDs: 1-PLUGIN_SID	More +

DIRECTIVE INFO

Ilustración 108 Ilustración de las reglas de correlación anidada para el escenario 2

Al momento de ser puesta en marcha, un ataque en progreso se detecta de la siguiente manera:

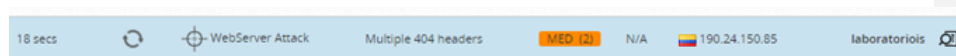


Ilustración 109 Alerta visualizada en OSSIM

Al momento de detallar la alerta, para ver la información se contiene, se ve de la siguiente manera:

Status	Risk	Attack Pattern	Created	Duration	# Events	Alarm ID	OTX Indicators
Open	MED (2)	external to internal one-to-one	16 mins ago	18 mins	86	298BF0983A6911E7A943000CCE112A0A	0

Source (1)	190.24.150.85	Destination (1)	190.24.150.88
190.24.150.85	Location: Colombia	laboratoriois (190.24.150.88)	Location: Colombia
Asset Groups: Unknown		Asset Groups: Laboratorio de informatica	
Networks: Unknown		Networks: Local_10_2_0_0_16	
OTX IP Reputation: No		OTX IP Reputation: No	

OPEN PORTS	
5	PORTS
PORT	SERVICE
0	ANY
SHOWING 1 TO 1 OF 1 PORTS < PREVIOUS 1 NEXT >	

VULNERABILITIES				
5	VULNERABILITIES			
SCAN TIME	VULNERABILITIES	VULN ID	SERVICE	SEVERITY
2017-05-14 21:06:55	phpinfo() output accessible	11229	http (80/tcp)	High
2017-05-14 21:06:55	Apache Tomcat servlet/JSP container default files	12085	http-proxy (8080/tcp)	High
2017-05-14 21:06:55	http TRACE XSS attack	11213	http (80/tcp)	High
2017-05-14 21:06:55	WordPress Multiple Vulnerabilities - July09	800657	http (80/tcp)	High
2017-05-14 21:06:55	WordPress MU Multiple Vulnerabilities - July09	800662	http (80/tcp)	High
SHOWING 1 TO 5 OF 135 VULNERABILITIES < PREVIOUS 1 2 3 4 5 NEXT >				

Ilustración 110 Detalle de la alerta del escenario 2

Y finalmente, la respuesta era subir el nivel de paranoia de ModSecurity. ModSecurity es un Web Application Firewall open source que permite proteger un Sistema de información web (para conocer más información de ModSecurity, puede acceder a <https://www.modsecurity.org/>), y el nivel de paranoia permite configurar qué tan “paranóico” es ModSecurity con respecto a las peticiones que ingresan. Si el nivel de paranoia es muy alto, ModSecurity puede ser capaz de proteger mucho más el sistema, pero, por ejemplo, también crear mucha información y generar indisponibilidad en el servicio mientras él la procesa. De esta manera, si parece haber un ataque de denegación de servicio distribuido, la idea sería subir el nivel de paranoia de ModSecurity para que él se encargue de proteger el sistema de información.

Para cumplir esto se instaló ModSecurity y se creó una política y un script para aumentar el nivel de paranoia, de esta manera OSSIM podrá lanzar dicho script

para cuando se genera el ataque. La respuesta se configuró entonces de la siguiente manera:

Política 2: Apache Bruteforce

Esta política está asociada a los eventos de Apache con código 404 (Not Found), 403(Forbidden) o 302(Redirection), del activo laboratoriois por el puerto 80. (Ilustración 105).

You can use the following keywords within any field which will be substituted by its matching value upon action execution:

- DATE
- PLUGIN_ID
- PLUGIN_SID
- RISK
- PRIORITY
- RELIABILITY
- SRC_IP_HOSTNAME
- DST_IP_HOSTNAME
- SRC_IP
- DST_IP
- SRC_PORT
- DST_PORT
- PROTOCOL
- SENSOR
- BACKLOG_ID
- EVENT_ID
- PLUGIN_NAME
- SID_NAME
- USERNAME
- PASSWORD
- FILENAME
- USERDATA1
- USERDATA2
- USERDATA3
- USERDATA4
- USERDATA5
- USERDATA6
- USERDATA7
- USERDATA8
- USERDATA9

NAME *	ModSecurity
DESCRIPTION *	Activar
TYPE *	Execute an external program
CONDITION	<input checked="" type="radio"/> Any <input type="radio"/> Only if it is an alarm <input type="radio"/> Define logical condition
COMMAND: *	/opt/scripts/apache-modsecurity.sh
TO: *	email;email;email

Ilustración 111 Detalle de la respuesta del escenario 2

2.1.6.3 Escenario 3: Detección de inicios de sesión sospechosos por geolocalización

La idea de este escenario era detectar sobre un aplicativo web si se han realizado inicios de sesión sospechosos por geolocalización. Por ejemplo, si una persona

habitualmente se conecta desde Colombia, sería muy sospechoso que iniciara sesión desde Angola.

Inicio de sesión del mismo usuario

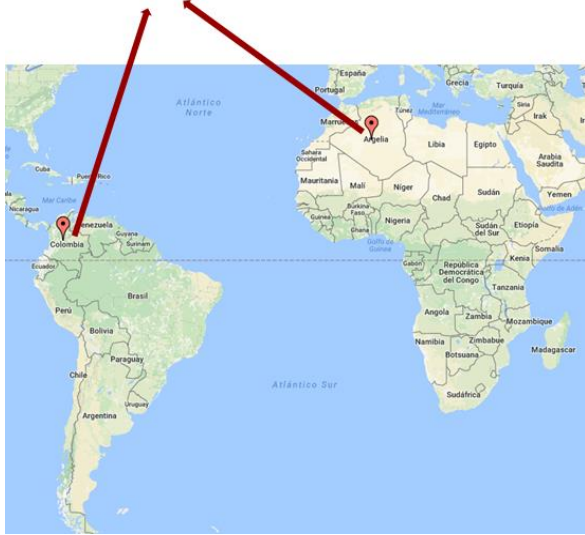


Ilustración 112 Detección de inicio de sesión sospechoso

Para esto, lo primero que se hizo fue intervenir el código del aplicativo para poder enviar un evento al OSSIM relatando la detección del inicio de sesión de un usuario puntual, indicando en él también el país desde donde se realizó dicho inicio. En el código se vería de la siguiente manera:

```

TokenTransfer token = login(username, password, user);

LOG.info(String.format("%s inicia sesi\u00F3n", username));
logSession.sendLog(username, IP, "AE13", DetectionPoint.Category.AUTHENTICATION);
response = Response
    .status(Response.Status.OK)
    .entity(new GenericEntity<TokenTransfer>(token) {
    })
    .header("Access-Control-Allow-Headers",
        "X-extra-header").build();

return response;

```

Ilustración 113 Intervención del código de la aplicación para enviar el evento de inicio de sesión

El evento detallado que llega a OSSIM se ve de la siguiente manera:

AppSensor: Event Triggered ACTIONS ▾

DATE	2017-05-11 11:39:14 GMT-5:00	CATEGORY	N/A
ALIENVAULT SENSOR	ossim [10.2.78.8]	SUB-CATEGORY	N/A
DEVICE IP	10.2.67.164 [eth0]	DATA SOURCE NAME	AppSensor
EVENT TYPE ID	1	DATA SOURCE ID	1000001
UNIQUE EVENT ID#	366811e7-aeb0-000c-29bb-f09b5e81a0fc	PRODUCT TYPE	Application
PROTOCOL	TCP	ADDITIONAL INFO	N/A

PRIORITY	RELIABILITY	RISK	OTX INDICATORS
2	2	LOW	0

SOURCE Host:10.2.67.164 [10.2.67.164]

Hostname: Host-10-2-67-164	Location: N/A
MAC Address: AB:20:66:09:EF:34	Context: N/A
Port: 0	Asset Groups: N/A
Latest update: N/A	Networks: Local_10_2_0_0_16
Username & Domain: N/A	Logged Users: N/A
Asset Value: 2	OTX IP Reputation: No

SERVICE ▾ PORT ◯ PROTOCOL ◯

No services available

SHOWING 0 TO 0 OF 0 SERVICES FIRST PREVIOUS NEXT LAST

DESTINATION Host:10.2.67.164 [10.2.67.164]

Hostname: Host-10-2-67-164	Location: N/A
MAC Address: AB:20:66:09:EF:34	Context: N/A
Port: 0	Asset Groups: N/A
Latest update: N/A	Networks: Local_10_2_0_0_16
Username & Domain: N/A	Logged Users: N/A
Asset Value: 2	OTX IP Reputation: No

SERVICE ▾ PORT ◯ PROTOCOL ◯

No services available

SHOWING 0 TO 0 OF 0 SERVICES FIRST PREVIOUS NEXT LAST

USERNAME	USERDATA1	USERDATA2	USERDATA3	USERDATA4	USERDATA6
edwin.ceron@mail.escuelaing.edu.c	AE13	Authentication	3	event_detection	SISTEMAS_TEST

RAW LOG

```

May 11 16:39:13 i1mac04_CEF: 010NASP|appsensor|1.0|AE13|Authentication|3|cat=event_detection deviceExternalId=SISTEMAS_TEST src=127.0.0.1 dst=10.2.67.164 asset=edwin.ceron@mail.escuelaing.edu.c

```

Ilustración 114 Evento detallado del inicio de sesión del escenario 3

Debido a las limitaciones técnicas no se pudo implementar la regla de correlación asociada a este caso. Ya que no nos fue posible simular el ambiente en que dos usuarios se conectaran de ubicaciones geográficas, por ello sólo se realizó una definición en OSSIM.

2.2 Data science applied in Cybersecurity

Al realizar la implementación previa de OSSIM en el Laboratorio de Informática pudimos identificar algunos retos asociados a la administración de eventos de seguridad en las plataformas SIEM, como los grandes volúmenes de eventos de seguridad que obligaban a la organización a tener un experto vigilando las diferentes alarmas que genera el SIEM, además cada evento de seguridad cuenta con diferentes propiedades y atributos lo que dificulta la tarea anterior y aumenta la diversidad de eventos.

Algunos eventos pueden tener relación entre ellos y otros no, por ello las reglas de correlación se hacen insuficientes para detectar anomalías, dejando pasar por alto eventos que podrían tener un mayor riesgo y, por último, algunos eventos de seguridad son útiles y otros no, haciendo perder tiempo al operador en eventos que no son cruciales para el cumplimiento de su labor de protección de las infraestructuras críticas de la organización.

Estos eventos de seguridad cumplen con las 5 v's del Big Data, velocidad, refiriéndose a la rapidez con la que los eventos son generados; volumen, por la cantidad de datos que son generados cada segundo; valor, que tiene en cuenta la importancia de los datos que son extraídos; variedad, por la cantidad de tipos de datos que se pueden extraer; y veracidad, por la calidad y confianza que generan los datos. Por esto y por los retos mencionados anteriormente, se decidió aplicar Data Science sobre los datos extraídos de OSSIM para resolver preguntas asociadas a los datos recolectados aplicando algoritmos de Machine Learning.

La ciencia de los datos es una disciplina que se forma con la fusión de la Ciencia en Computación, las Matemáticas, la Estadística y el conocimiento de campo, que busca principalmente explotar los datos para la toma de decisiones. Responde diferentes preguntas como lo son:

¿Es A o B?

Esta pregunta predice un nombre o categoría. Utiliza procesos de clasificación para la toma de decisiones o identificar situaciones que involucren dos o más casos.

¿Es extraño?

Se emplean técnicas de Detección de Anomalías para identificar eventos inesperados o comportamientos poco habituales.

¿Cuánto? o ¿Cuántos?

Analiza datos históricos reales o datos actuales para realizar predicciones numéricas para el futuro mediante el uso de modelos predictivos.

¿Cómo está organizado?

Identificación de diferentes perfiles a través de la segmentación de datos, que permiten conocer su estructura y predecir mejor eventos y comportamientos.

¿Qué se debe hacer a continuación?

Aplicación de técnicas de aprendizaje reforzado, cuya toma de decisiones se basa en el aprendizaje continuo.

2.2.1 Conceptos clave

Hablando del ciclo de vida de la ciencia de los datos, cabe realizar la aclaración de que estas fases no se realizan en un orden específico, ya que al ser la metodología

TDS ágil, permite la realización del ciclo de data science en cualquier sentido y pudiendo regresar a una fase “anterior” en cualquier momento.

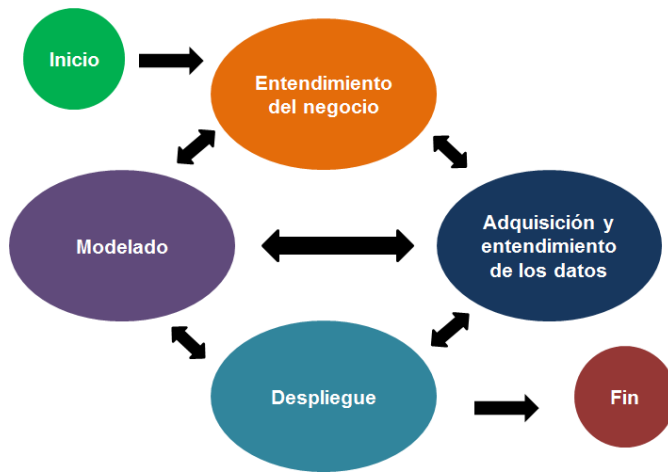


Ilustración 115 Ciclo de vida de la ciencia de los datos

2.2.1.1 Business Understanding

El entendimiento del negocio es la parte clave de todo el proceso de data science, ya que aquí es donde entenderemos y definiremos qué es lo que requiere el negocio y cómo mediante la aplicación de data science se puede lograr una mejora en los diferentes procesos de la compañía.

Para lograr el entendimiento de negocio se deben realizar visitas de observación, leer manuales, realizar entrevistas y entender sus restricciones y beneficios, ya que cada compañía es un mundo diferente.

Una vez realizado lo anterior se deben identificar variables clave, las cuales serán el objetivo del modelo y de las cuales también se obtendrán las mediciones para determinar el éxito del proyecto. Además de esto, en esta fase debemos identificar

claramente cuáles son las fuentes de datos que vamos a utilizar para alimentar el modelo de data science.

2.2.1.2 Data acquisition and understanding

En esta fase del ciclo de data science, definiremos cómo se obtendrán los datos, el medio y la forma de entrega, un ejemplo del medio son los archivos comprimidos y cifrados entregados en una memoria USB, en cuanto a la forma un ejemplo podría ser el formato CSV.

Una vez acordada la adquisición de los datos, lo siguiente es entender los datos recibidos, para esto se debe entender cada atributo, sus posibles valores, determinar su importancia para el negocio y para el modelo. En esta fase también se realiza la depuración de los diferentes atributos en función a su utilidad en el modelo de Data Science.

2.2.1.3 Modeling

En esta fase se busca alimentar los diferentes modelos matemáticos existentes con los datos que han sido previamente adquiridos y filtrados. Una vez los modelos matemáticos entreguen los resultados del entrenamiento, la idea es pulir los modelos a través de las diferentes variables que pueden ser modificadas del mismo y volver a probar con la información de la fase anterior para obtener mejores resultados. Este último proceso se realiza hasta que se esté satisfecho con los resultados.

Weka es una herramienta desarrollada por el grupo de Machine Learning de la Universidad de Waikato, contiene diferentes herramientas para preprocesamiento de datos, clasificación, regresión, clusterización, reglas de asociación y visualización. También se encuentra en la capacidad de desarrollo de esquemas

de Machine Learning, se puede utilizar por medio de Java o su herramienta disponible para Linux, Windows y Mac.(11)

En nuestro caso, se utilizó la herramienta Weka para evaluar, modificar y configurar los diferentes modelos matemáticos. Para cargar un modelo en Weka se realizan los siguientes pasos:

1. Estando en Weka Explorer (luego de ser instalado y ejecutado) se procede a dar clic en “Open file”

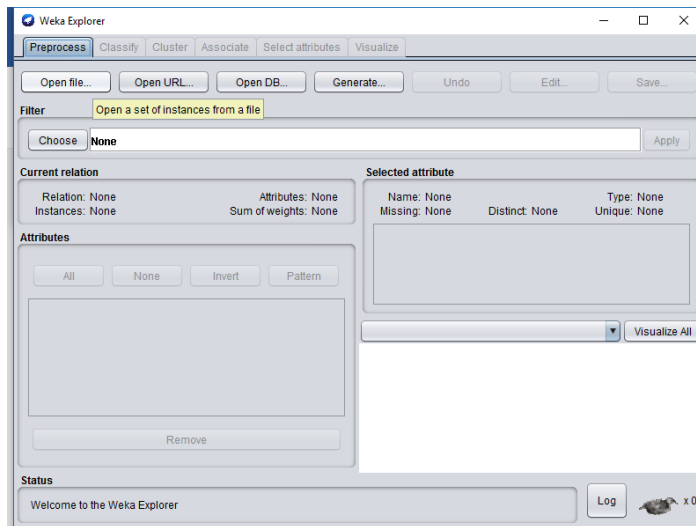


Ilustración 116 Vista de Exploración de Weka

2. Se elige el modelo a ser evaluado (con extensión.arff) y se pasa a la pestaña de “Classify” si se quiere usar un modelo para clasificar, o a “Cluster” si se busca un modelo para clusterizar la información.

Es necesario explicar la forma de interpretar los resultados para así poder entender los que se mostrarán más adelante, por tanto, se mostrarán los campos más relevantes:

Error relativo

El error relativo nos permite ver la magnitud del error independientemente de la escala, ya que no es lo mismo ver o entender el error obtenido si los valores con los que se está probando el modelo son muy pequeños o muy grandes y es en estos casos que el error relativo nos permite tener un mejor dimensionamiento del error.

Error absoluto medio

El error absoluto medio es la sumatoria de la diferencia entre el valor predicho y el valor verdadero dividido entre la cantidad de datos (promedio)

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |\hat{\theta}_i - \theta_i|$$

Ilustración 117 Fórmula del error absoluto medio

Raíz del error cuadrático medio

Es una medida muy similar al error absoluto medio, pero en este caso se amplifica y penaliza con mayor fuerza aquellos errores de mayor magnitud.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta_i)^2}$$

Ilustración 118 Fórmula de la raíz del error cuadrático medio

Error absoluto relativo

el error absoluto relativo toma el error absoluto total y lo normaliza dividiendo por el error total absoluto del predictor simple.

$$\text{RAE} = \frac{\sum_{i=1}^N |\hat{\theta}_i - \theta_i|}{\sum_{i=1}^N |\bar{\theta} - \theta_i|}$$

Ilustración 119 Fórmula del error absoluto relativo

Raíz del error relativo al cuadrado

el error cuadrado relativo toma el error cuadrado total y lo normaliza dividiendo por el error cuadrado total del predictor simple.

$$\text{RRSE} = \sqrt{\frac{\sum_{i=1}^N (\hat{\theta}_i - \theta_i)^2}{\sum_{i=1}^N (\bar{\theta} - \theta_i)^2}}$$

Ilustración 120 Fórmula del error relativo al cuadrado

Coficiente de correlación

El coeficiente de correlación habla de qué tan relacionados están los valores predichos con respecto a los reales. Esto está representado por números entre -1

y 1, donde 1 indica una relación lineal muy fuerte, 0 indica que no existe relación alguna y -1 indica que existe una relación inversa.

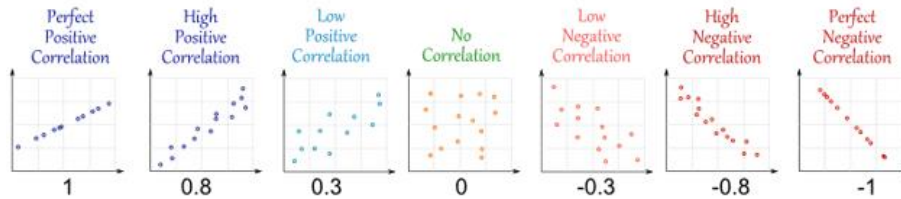


Ilustración 121 Explicación gráfica del coeficiente de correlación

2.2.1.4 Implementation

En esta fase el modelo es exportado a una aplicación funcional para que la organización empiece a usarlo, entrenarlo y darle uso a los resultados que la misma entrega, ya que en la fase anterior el modelo aún no ha sido exportado y solamente se ejecuta en un ambiente controlado para verificar los resultados que se obtienen.

2.2.2 Caso 1: Sistema de Información del Laboratorio de Informática

2.2.2.1 Business Understanding

Para el Laboratorio de Informática de la Escuela Colombiana de Ingeniería “Julio Garavito” se tenía una gran ventaja ya que todos los estudiantes que realizaron este proyecto estuvieron involucrados en el negocio directamente, ya que trabajaron como monitores en el Laboratorio y por ende tuvieron acceso al Sistema de Información objetivo, luego el entendimiento fue más sencillo.

Para contextualizar, el Laboratorio de Informática tiene como misión "Apropiar y divulgar las tecnologías de hardware y software para ponerlas a disposición de toda la comunidad, fomentando el uso de las alternativas de software de libre

distribución. Administrar y mantener la infraestructura computacional y de comunicaciones con las últimas tecnologías de hardware y software para suministrar a la comunidad universitaria un conjunto de servicios de excelente calidad.”. El Sistema de Información del Laboratorio de Informática contiene, entre otros, los siguientes activos identificados:

1. Información digital: Información personal de trabajadores, recursos físicos disponibles, horas trabajadas, manuales, documentos contractuales, etc.
2. Software: Web server, base de datos.
3. Hardware: Servidores físicos y virtualizados, Storage externo, etc.

El escenario que suponíamos para un ingreso normal al Sistema de Información era el siguiente:

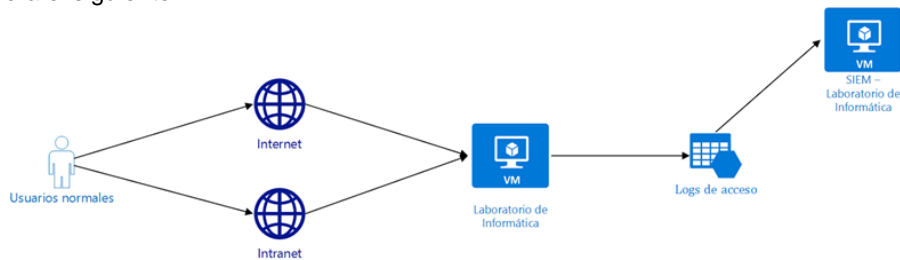


Ilustración 122 Escenario idóneo de acceso al Sistema.

Sin embargo, el sistema en realidad se comportaba de la siguiente manera:

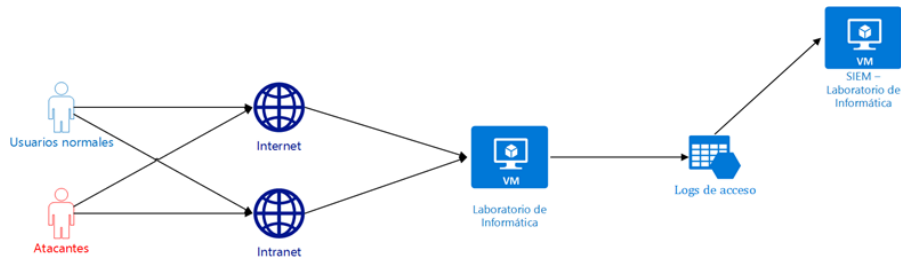


Ilustración 123 Escenario real de acceso al Sistema

Esto se concluyó a partir de la identificación de varios ataques que se llevaban sobre el activo y se detectaron a través del SIEM OSSIM instalado.

A partir de lo dicho anteriormente definimos tres preguntas de data science a resolver:

1. ¿Cuántos eventos se reportarán en la siguiente hora?
2. ¿Cómo predecir si un usuario presenta un comportamiento anómalo en su interacción con el Sistema?
3. ¿Qué tipo de intentos de autenticación SSH se realizan sobre un activo de alta criticidad?

2.2.2.2 Data acquisition and understanding

La adquisición de los datos se realizó directamente sobre OSSIM. Los pasos fueron los siguientes:

1. Se accede a OSSIM y al menú Analysis> Security Events(SIEM), desplegamos el menú de exportación de eventos haciendo clic al botón que aparece a la izquierda, escogemos el formato en el que queremos exportar los eventos, en este caso será en CSV.

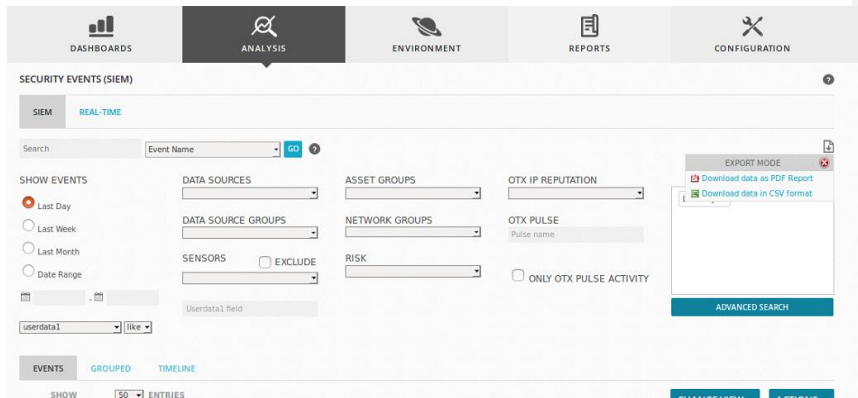


Ilustración 124 Vista de eventos de seguridad del SIEM

2. Aparecerá el siguiente menú donde seleccionamos la cantidad de eventos que queremos exportar, en este caso serán 500, hacemos clic en el botón CSV para exportar:



Ilustración 126 Menú de exportación de eventos

3. Al dar clic se deben esperar algunos segundos o minutos dependiendo de la cantidad de eventos que se quisieron exportar:

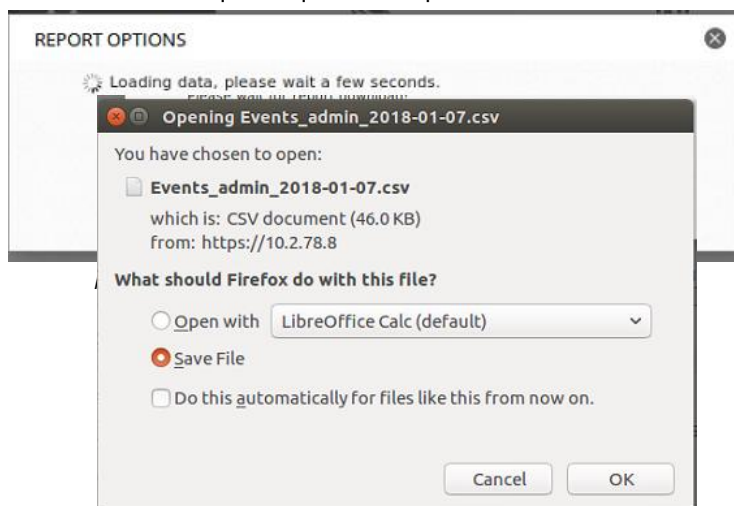


Ilustración 127 Ventana de archivo del navegador.

4. Una vez hecho esto se obtienen los logs de acceso al Sistema de Información en forma de CSV.

Una vez los datos habían sido obtenidos había que entenderlos. Para ello se realizó un análisis estadístico empleando Excel, de donde se obtuvieron diferentes gráficas asociadas a las tendencias de los eventos, gráficos de calor para ver la relación que tenían los diferentes atributos y tablas dinámicas, para entender la relación entre atributos y ocurrencias. Este análisis se presenta a continuación:

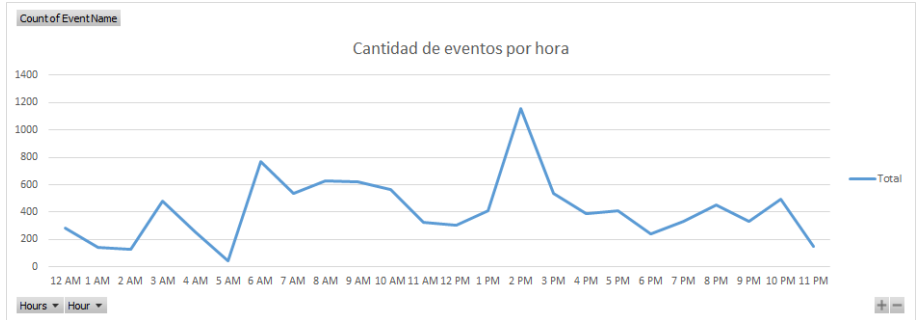


Ilustración 128 Cantidad de eventos por hora laboratorio de infomática

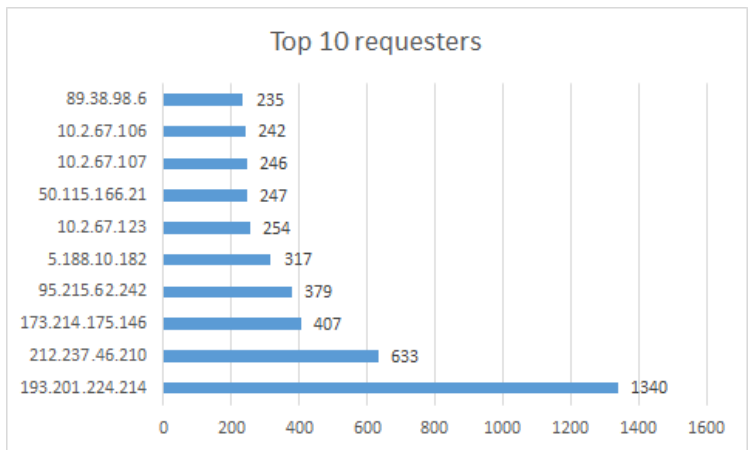


Ilustración 129 Top 10 direcciones IP con más peticiones al servidor del laboratorio de infomática

OSSIM por defecto que no representaban variación entre los registros. Finalmente se obtuvo una tabla como la que se observa a continuación:

Dst IP	Src IP	Event Type ID	Date GMT-5:00	Userdata1	Userdata2	Userdata3
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /favicon.ico HTTP/1.1		209
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon.png HTTP/1.1		218
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon-precompose		230
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /favicon.ico HTTP/1.1		209
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon.png HTTP/1.1		218
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon-precompose		230
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon-120x120.png		226
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon-120x120-pre		238
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /favicon.ico HTTP/1.1		209
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon-120x120.png		226
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon-120x120-pre		238
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /favicon.ico HTTP/1.1		209
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon.png HTTP/1.1		218
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon-precompose		230
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon-120x120.png		226
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon.png HTTP/1.1		218
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon-precompose		230
10.250.0.88	152.173.89.39	404	19/09/2017 23:39	GET /apple-touch-icon-120x120.png		226
10.250.0.88	152.173.89.39	404	19/09/2017 23:38	GET /apple-touch-icon-120x120-pre		238
10.250.0.88	152.173.89.39	404	19/09/2017 23:38	GET /apple-touch-icon-120x120-pre		238
10.250.0.88	181.165.83.93	404	19/09/2017 23:21	GET /favicon.ico HTTP/1.1		209 http://laboratorio.is.escuelair
10.250.0.88	181.165.83.93	404	19/09/2017 23:21	GET /favicon.ico HTTP/1.1		209 http://laboratorio.is.escuelair
10.250.0.88	186.119.107.162	404	19/09/2017 23:21	GET /favicon.ico HTTP/1.1		209 http://laboratorio.is.escuelair
10.250.0.88	186.119.107.162	404	19/09/2017 23:21	GET /favicon.ico HTTP/1.1		209 http://laboratorio.is.escuelair
10.250.0.88	186.119.107.162	206	19/09/2017 23:21	GET /wp-content/uploads/2016/09/	464102	http://laboratorio.is.escuelair
10.250.0.88	186.119.107.162	206	19/09/2017 23:21	GET /wp-content/uploads/2016/09/	464102	http://laboratorio.is.escuelair

Ilustración 132 Registros depurados – peticiones legítimas

A partir de aquí se realizaron varias copias de los datos para intentar resolver distintas preguntas de data science previamente identificadas:

¿Cuántos eventos se reportarán en la siguiente hora?

Para esta pregunta se necesitaba saber cuántos eventos se habían recibido por hora, luego había que realizar una modificación adicional a la tabla, y quedaría de la siguiente forma:

Horas	Fecha	Hora	Cant Eventos			fecha	Dias	cant Eventos
1	10-Sep	24	12	24	12	10-Sep	7	486
2		1	4	1	4	11-Sep	1	218
3		3	20	3	20	14-Sep	4	894
4		4	6	4	6	16-Sep	6	118
5		5	20	5	20	17-Sep	7	332
6		6	22	6	22	18-Sep	1	1408
7		8	4	8	4	19-Sep	2	1164
8		9	42	9	42	20-Sep	3	514
9		10	16	10	16	22-Sep	5	548
10		11	32	11	32	23-Sep	6	880
11		12	44	12	44	24-Sep	7	500
12		13	54	13	54	25-Sep	1	796
13		14	12	14	12	26-Sep	2	954
14		15	4	15	4	27-Sep	3	698
15		16	32	16	32			
16		17	12	17	12			
17		18	56	18	56			
18		19	26	19	26	snl		
19		20	38	20	38	10-Sep	1	
20		21	20	21	20	11-Sep	2	
21		22	2	22	2	14-Sep	3	
22								

Ilustración 133 Registros depurados – cantidad de eventos por hora

Lo siguiente es realizar la traducción de los datos para poderlos poner en el formato que los modelos matemáticos entienden. Para ello se desarrolló un traductor que fue utilizado en todas las preguntas de Data Science y es parte del aplicativo desarrollado. Una vez los datos fueron pasados por el traductor quedaron de la siguiente manera:

```

1 @RELATION apache
2
3 @ATTRIBUTE hour NUMERIC
4 @ATTRIBUTE Quantity NUMERIC
5
6 @DATA
7 24,12
8 1,4
9 3,20
10 4,6
11 5,20
12 6,22
13 8,4
14 9,42
15 10,16
16 11,32
17 12,44
18 13,54
19 14,12
20 15,4
21 16,32
22 17,12
23 18,56
24 19,26
25 20,38
26 21,20
27 22,2
28 23,8
29 24,26
30 2,8
31 3,6
32 4,22
33 5,8
34 6,26
35 7,122
36 9,6
37 10,26
38 11,32
39 12,32
40 13,116
41 14,120
42 15,40
43 16,108
44 17,298

```

Ilustración 134 – Traducción eventos

¿Cómo predecir si un usuario presenta un comportamiento anómalo en su interacción con el Sistema?

Para esta pregunta ya se tenían las columnas necesarias para poder saber si una petición había sido legítima o no, y el paso siguiente era etiquetar manualmente cada uno de los registros con dicha determinación, para que el sistema aprenda cuál es el patrón de una petición ilegítima /legítima y en el futuro sepa diferenciar las nuevas peticiones que ingresen al sistema. De esta manera, se realizó el etiquetado y la tabla quedó de la siguiente forma (note la última columna que se agregó):


```

@DATA
0,0,404,0,0,0,0,F
0,0,404,0,1,1,0,F
0,0,404,0,2,2,0,F
0,0,404,0,0,0,0,F
0,0,404,0,1,1,0,F
0,0,404,0,2,2,0,F
0,0,404,0,3,3,0,F
0,0,404,0,4,4,0,F
0,0,404,0,0,0,0,F
0,0,404,0,3,3,0,F
0,0,404,0,4,4,0,F
0,0,404,0,0,0,0,F
0,0,404,0,1,1,0,F
0,0,404,0,2,2,0,F
0,0,404,0,3,3,0,F
0,0,404,0,1,1,0,F

```

Ilustración 137 Traducción archivo comportamiento legitimo

2.2.2.3 Modeling

A partir de este punto se entiende que ya se tienen los datos depurados para poder ser ingresados a los diferentes modelos, y luego determinar el mejor para cada caso particular. Al tener varias preguntas a resolver los resultados son distintos para cada una, así que esta fase se dividirá por cada pregunta nuevamente.

Para el proceso de modelado en Weka se siguen los siguientes pasos:

- Debe abrirse el programa y seleccionar la pestaña “Explorer”



Ilustración 138 Interfaz inicial WEKA

- Ahora debemos seleccionar el archivo previamente traducido para esto damos clic en la pestaña “open file...”

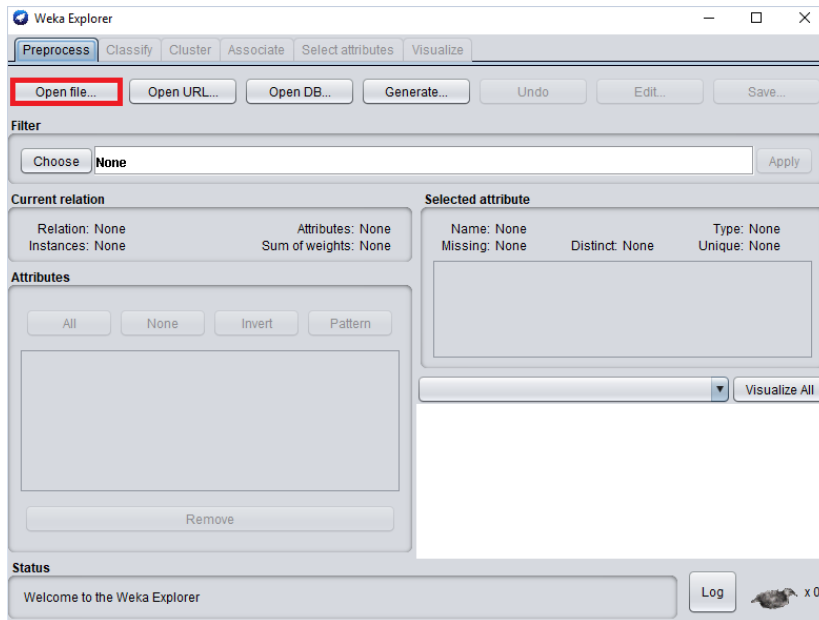


Ilustración 140 Selección de archivo WEKA

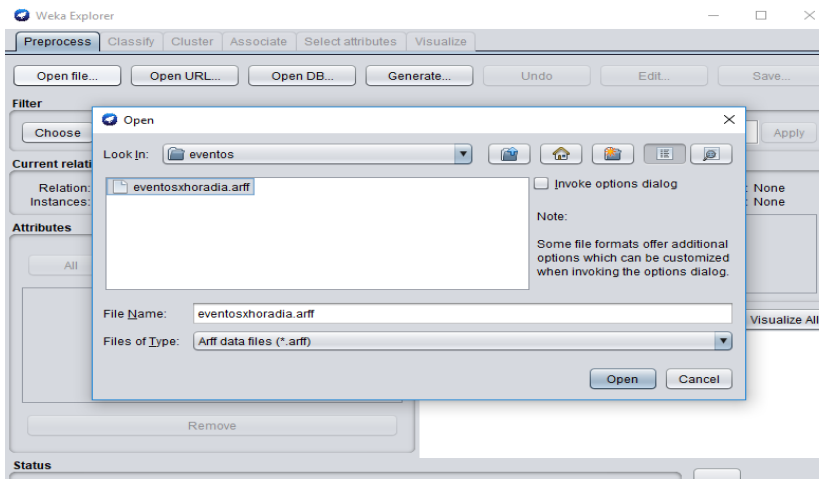


Ilustración 139 Interfaz de selección de archivo WEKA

- Una vez seleccionado el archivo seleccionamos la pestaña “Classify” o “Cluster” dependiendo de si deseamos realizar una clasificación o una agrupación de los datos, en este caso mostraremos el proceso de clasificación.
- En la pestaña de “Classify”, lo primero que demos hacer es escoger el modelo que deseamos utilizar para esto le damos clic en choose y se mostrará el siguiente menú, en el cual están clasificados los diferentes modelos por su metodología

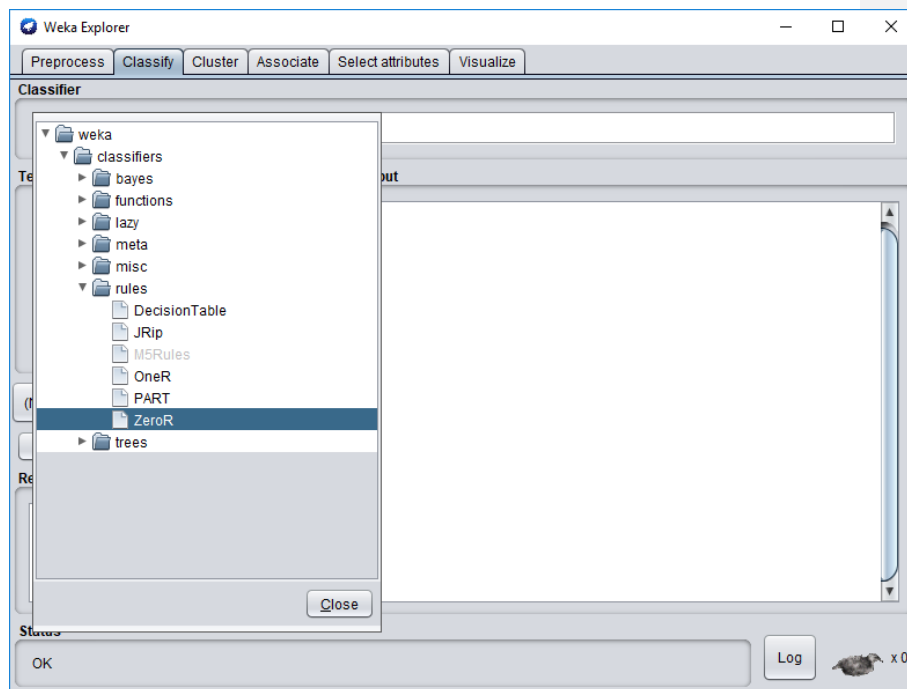


Ilustración 141 Selección de modelo matemático WEKA

- Una vez elegido el modelo si se desea realizar una configuración de este se debe dar clic sobre el nombre y se desplegará el siguiente menú, el cual es diferente para cada modelo

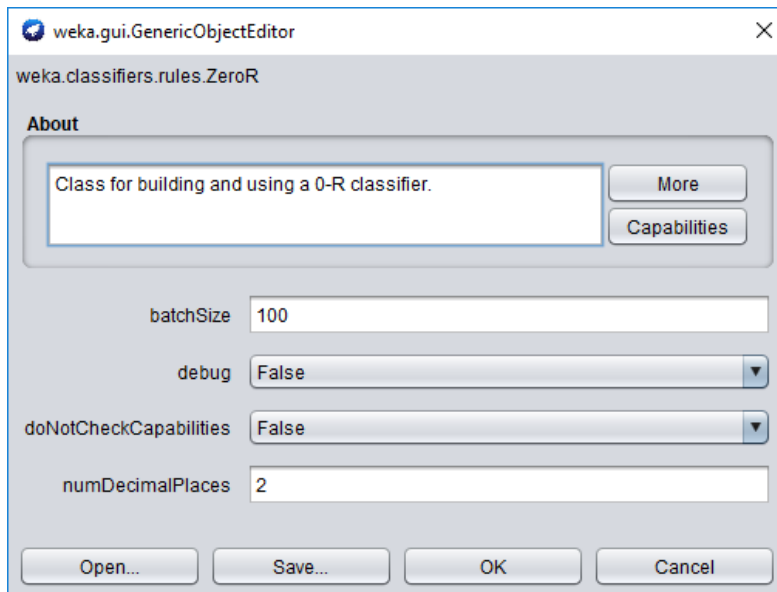


Ilustración 142 Interfaz de configuración de modelo matemático WEKA

- Después de esto debemos elegir la forma en que se entrenará el modelo hay cuatro posibles formas:
 - Use training set: esta opción entrena al modelo con todo el archivo.arff y lo evalúa con el mismo.
 - Supplied test set: esta opción entrena al modelo con todo el archivo.arff y lo prueba con un archivo externo.
 - Cross-validation: esta opción
 - Percentage Split: esta opción permite entrenar al modelo con un porcentaje de los casos y evaluarlo con el restante.

- Después de esto debemos realizar un cambio para poder visualizar los resultados esto se hace en la pestaña “More options...”

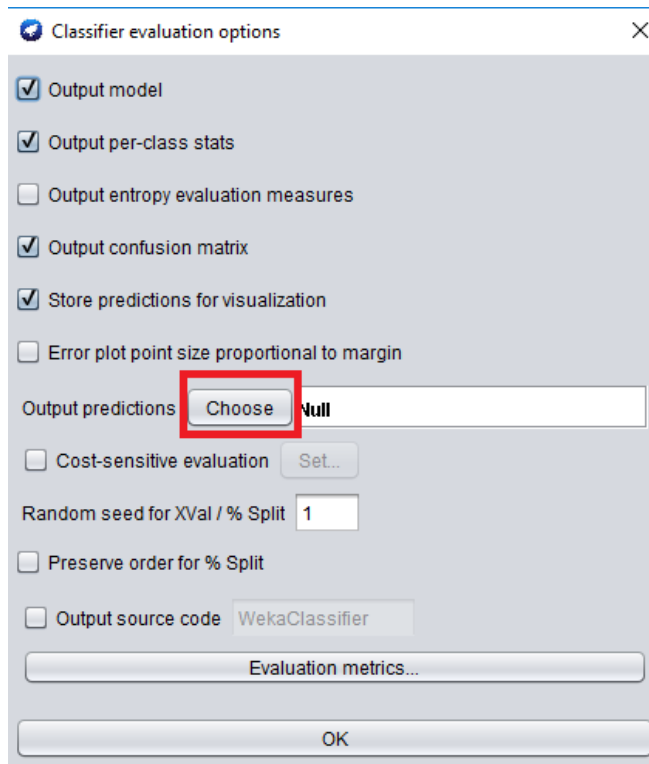


Ilustración 143 Opciones del clasificador WEKA

- Se desplegará una ventana y podremos elegir la forma en la que se imprimirán los resultados hay 4 posibles formas CSV, HTML, XML, PlainText. En este caso escogeremos PlainText.

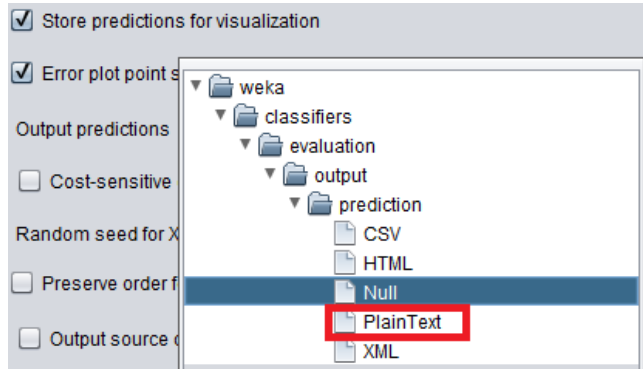
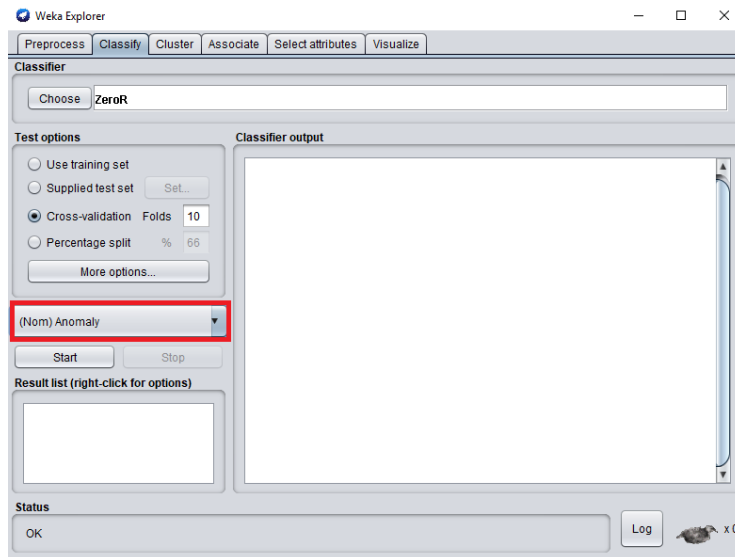


Ilustración 144 Interfaz selección formato de salida WEKA

- Una vez realizado esto guardamos y procedemos a elegir la variable que se va a predecir, para esto damos clic en la pestaña donde aparecen las variables que estaban en el archivo.arff



- Una vez seleccionada la variable que queremos predecir damos clic al botón start y empezara el proceso de entrenamiento y evaluación y obtendremos una pantalla como la siguiente

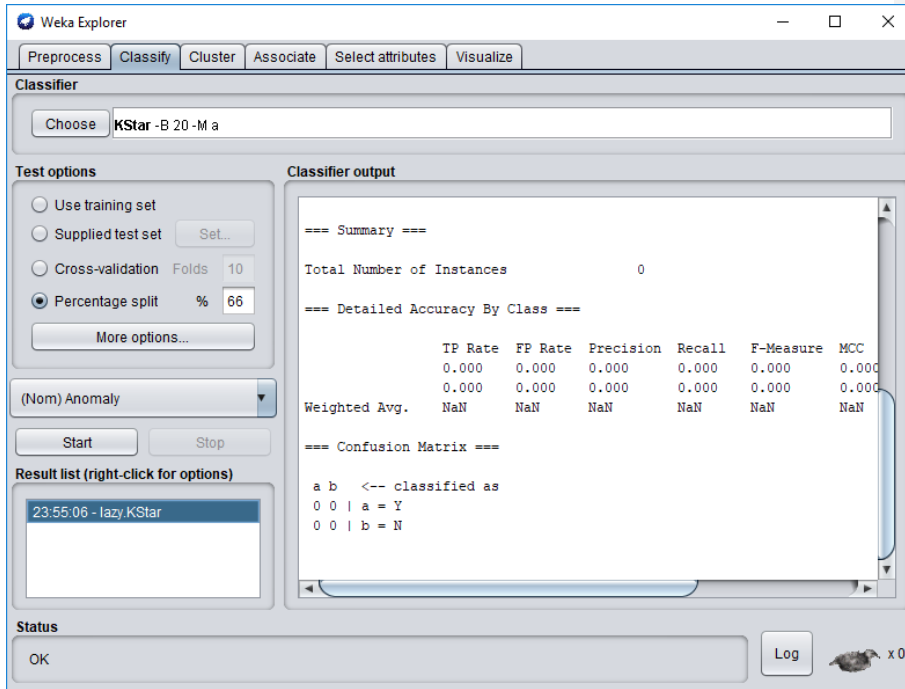


Ilustración 145 Pantalla de resultados WEKA

Modelado

Para esta pregunta se probaron varios modelos de Data Science a continuación, se muestran los resultados más relevantes.

M5P

=== Summary ===

Correlation coefficient	0
Mean absolute error	27.5688
Root mean squared error	49.6611
Relative absolute error	100 %
Root relative squared error	100 %
Total Number of Instances	268

Ilustración 146 Resultado del modelo MP5 aplicado a los datos de eventos por hora

Bagging

=== Summary ===

Correlation coefficient	0.3848
Mean absolute error	25.0195
Root mean squared error	46.066
Relative absolute error	90.7529 %
Root relative squared error	92.7606 %
Total Number of Instances	268

Ilustración 147 Resultado del modelo Bagging aplicado a los datos de eventos por hora

REPTree

=== Summary ===

Correlation coefficient	0.3044
Mean absolute error	25.7053
Root mean squared error	47.3037
Relative absolute error	93.2405 %
Root relative squared error	95.253 %
Total Number of Instances	268

Ilustración 148 Resultado del modelo REPTree aplicado a los datos de eventos por hora

Como podemos observar, en este caso los resultados obtenidos para el Laboratorio de Informática no fueron los mejores, se obtuvo un gran error absoluto y relativo y por tanto se concluyó que las posibles causas de estos resultados podrían ser la falta de datos históricos, ya que al ser un modelo de predicción la cantidad de posibles escenarios requería una gran cantidad de datos históricos que no se lograron obtener; otra posible causa podría ser que los datos no seguían un patrón por lo que no se ajustan a ningún modelo de Data Science, lo cual pasa en muchos casos.

¿Cómo predecir si un usuario presenta un comportamiento anómalo en su interacción con el Sistema?

Para esta pregunta, se probaron diferentes modelos, los cuales se muestran a continuación:

NaiveBayes

=== Summary ===

Correctly Classified Instances	311	91.4706 %
Incorrectly Classified Instances	29	8.5294 %
Kappa statistic	0.8108	
Mean absolute error	0.0773	
Root mean squared error	0.2568	
Relative absolute error	17.753 %	
Root relative squared error	55.6199 %	
Total Number of Instances	340	

Ilustración 149 Resultado del modelo NaiveBayes aplicado a los datos de comportamiento anómalo

DecisionTable

=== Summary ===

Correctly Classified Instances	324	95.2941 %
Incorrectly Classified Instances	16	4.7059 %
Kappa statistic	0.8886	
Mean absolute error	0.1285	
Root mean squared error	0.2287	
Relative absolute error	29.4858 %	
Root relative squared error	49.5402 %	
Total Number of Instances	340	

Ilustración 150 Resultado del modelo DecisionTable aplicado a los datos de comportamiento anómalo

Bagging

=== Summary ===

Correctly Classified Instances	320	94.1176 %
Incorrectly Classified Instances	20	5.8824 %
Kappa statistic	0.8561	
Mean absolute error	0.0915	
Root mean squared error	0.2023	
Relative absolute error	21.0009 %	
Root relative squared error	43.8237 %	
Total Number of Instances	340	

Ilustración 151 Resultado del modelo Bagging aplicado a los datos de comportamiento anómalo

KStar

=== Summary ===

Correctly Classified Instances	333	97.9412 %
Incorrectly Classified Instances	7	2.0588 %
Kappa statistic	0.9519	
Mean absolute error	0.0289	
Root mean squared error	0.1189	
Relative absolute error	6.6415 %	
Root relative squared error	25.7572 %	
Total Number of Instances	340	

Ilustración 152 Resultado del modelo KStar aplicado a los datos de comportamiento anómalo

Se puede observar que KStar fue el modelo que mejor resultados obtuvo, por lo que fue el modelo elegido para realizar una posible implementación en el Laboratorio de Informática.

¿Qué tipo de intentos de autenticación SSH se realizan sobre un activo de alta criticidad?

En este caso se utilizaron técnicas de clusterización en WEKA para ver el agrupamiento de los diferentes eventos.

Realizar clusterización en WEKA:

1. Seguimos el proceso de abrir archivo expuesto anteriormente y seleccionamos la ventana de clustering:

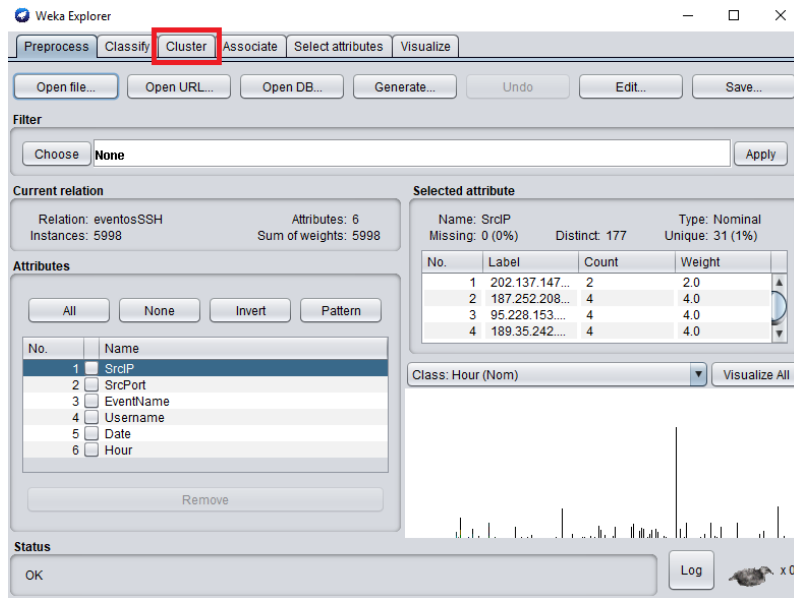


Ilustración 153 Vista de preprocesamiento de datos

- Al dar clic en el botón aparecerá la pantalla de la imagen a continuación, allí en el botón “choose” podremos ver los diferentes métodos de clusterización disponibles. Dentro de ellos los que nosotros utilizamos SimpleKMeans, EM, HierarchicalClusterer, entre otros.

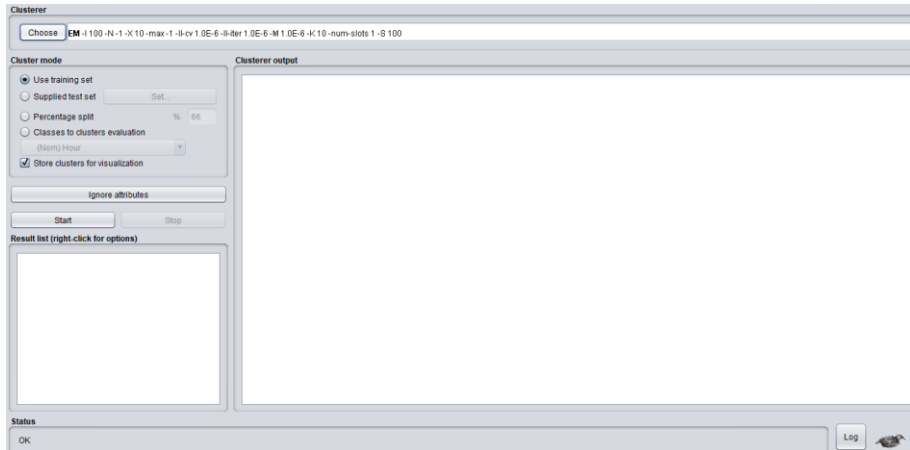


Ilustración 154 Vista de clusterización de WEKA

3. Al dar clic en el espacio de texto ubicado a la derecha del botón “choose” podremos ver los atributos que podemos modificarle, esto depende del algoritmo que vamos a utilizar.

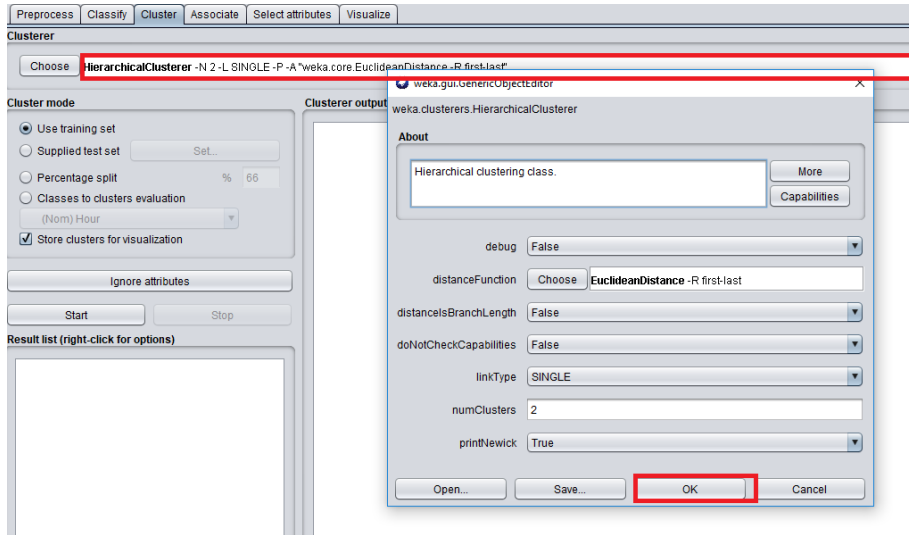


Ilustración 155 Ventana de modificación de atributos

4. Una vez configurado el algoritmo, hacemos clic en el botón “Start” y esperamos a que nos arroje los resultados. Una vez terminado dependiendo del algoritmo utilizado nos arrojará una información, para interpretar los resultados se debe revisar la variable “Sum of squared errors”, la correcta clasificación se toma como la que menor suma de errores cuadrados tenga.

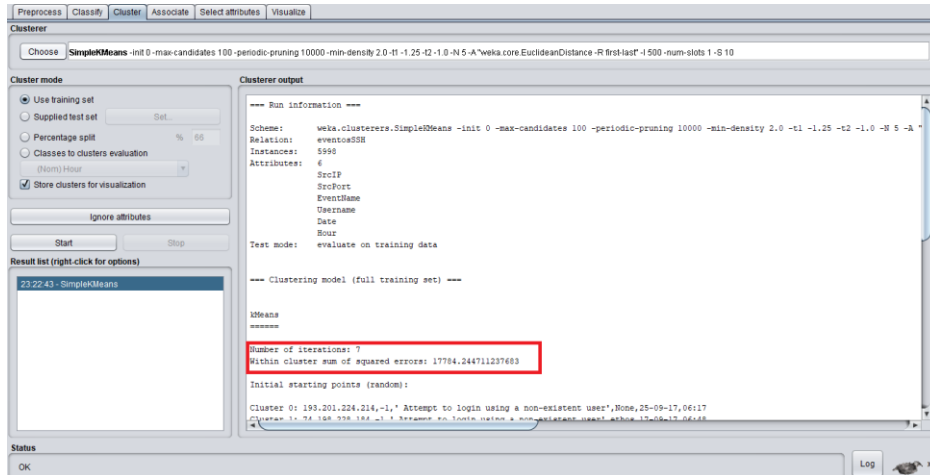


Ilustración 156 Resultado de clusterización de eventos SSH.

5. Para tener una mejor visualización de los clústeres generados hacemos click sobre el modelo en el panel izquierdo, y seleccionamos la opción “visualize cluster assignments” y aparecerán una serie de gráficas donde podremos ver la agrupación por atributos.

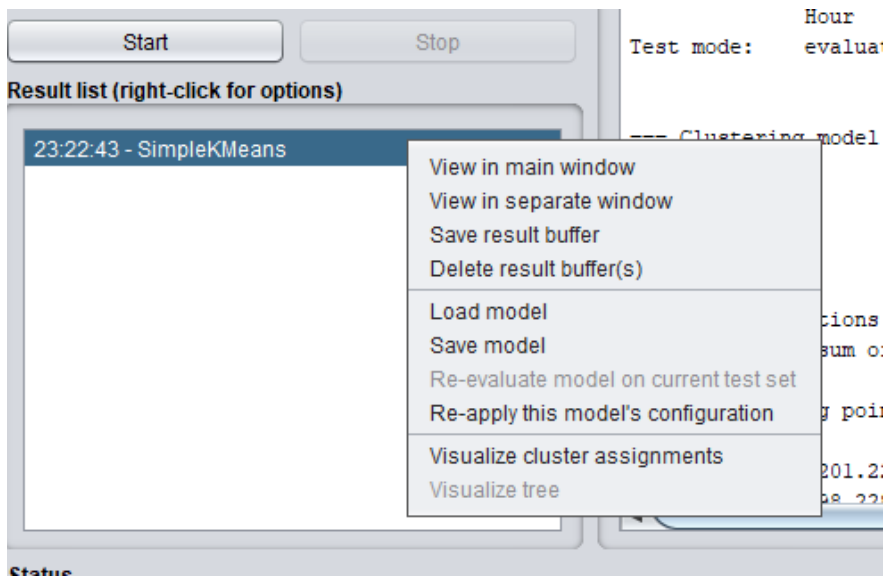


Ilustración 157 Menú de opciones sobre los modelos generados

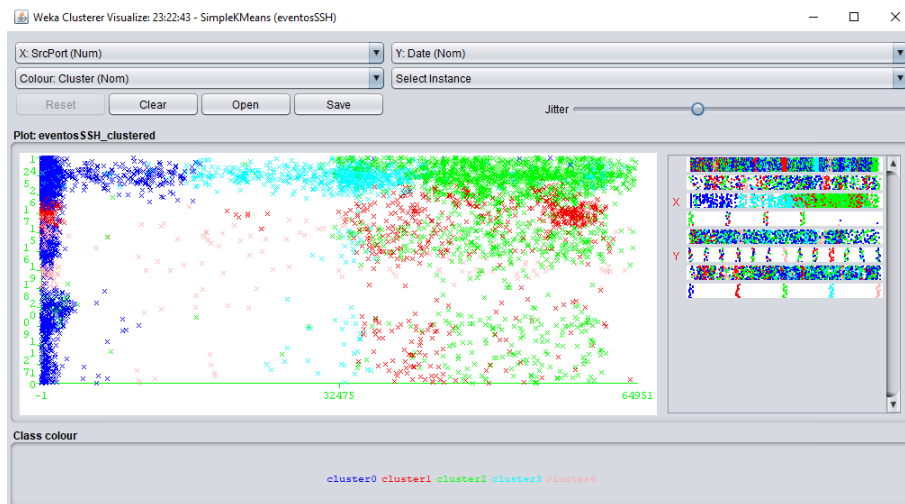


Ilustración 158 Gráfica de distribución de clústeres según puerto y fecha.

La interpretación de los clústeres se debe realizar de forma manual, es decir, revisar qué datos fueron clasificados en qué clúster y con ello ver cuáles son los grupos generados.

En nuestro caso utilizando el algoritmo SimpleKMeans se lograron identificar 6 clústeres con las siguientes características:

```

Number of iterations: 16
Within cluster sum of squared errors: 9004.852897746763
Missing values globally replaced with mean/mode

Cluster centroids:

```

Attribute	Full Data (14750)	Cluster# 0 (4217)	1 (1674)	2 (1873)	3 (2125)	4 (2050)	5 (2811)
Country	Ukraine	Ukraine	France	Germany	Japan	Argentina	Brazil
SrcIP	2614672925.3056	3224804486.1119	2438333849.8459	482545708.7747	224115588.0414	3053383137.7176	3187492029.9669
SrcPort	21679.2594	24206.4174	6561.586	19733.0587	41795.696	11816.7278	20173.0181
Date	293.0103	295.9656	288.4295	289.8361	289.5539	298.0288	292.3728
Hour	45641.5892	47757.9654	43213.4767	35936.7325	34944.3388	46148.4	58096.158

Ilustración 159 Resultado clusterización con SimpleKMeans

Categorización de adversarios

1. Clúster 0: Adversarios ubicados principalmente en Ucrania con las siguientes ventanas de ataque: 6:00-7:54, 17:58-19:09. 14:21-15:24 en las fechas 16, 25, 27, de septiembre.
2. Clúster 1: Adversarios ubicados principalmente en Francia, Holanda e Italia con las siguientes ventanas de ataque: 13:10-15:50 atacando principalmente los fines de semana.
3. Clúster 2: Adversarios ubicados en Alemania y Croacia con ataques distribuidos a lo largo del día en la fecha 27 de septiembre.
4. Clúster 3: Adversarios ubicados en Japón con ataques distribuidos a lo largo del día durante la semana exceptuando el miércoles.
5. Clúster 4: Adversarios ubicados en Argentina con las siguientes ventanas de ataque: 5:55- 23:00 atacando el día 25 de septiembre.

6. Clúster 5: Adversarios ubicados principalmente en Brasil con ataques distribuidos a lo largo del día en las fechas 10 y 25 de septiembre.

2.2.2.4 Implementation

Este ejercicio se realizó con fines académicos, donde se buscaba aplicar el TDSP en un entorno real, siguiendo el proceso y entendiendo cómo se obtenían datos de calidad para ser sometidos a algoritmos de Machine Learning. Por dicha razón, no se realizó la implementación de los modelos en el Laboratorio de Informática y se pasó directamente al siguiente caso.

2.2.3 Caso 2: Comando Conjunto Cibernético de las Fuerzas Militares

2.2.3.1 Business Understanding

El Comando Conjunto Cibernético de las Fuerzas militares (CCOC), tiene como objetivo la protección de la infraestructura crítica nacional, además de ser el ente encargado de la defensa cibernética del Estado Colombiano, respondiendo a ataques informáticos que lo afecten y defiende las redes informáticas militares.

Para esta primera fase con el CCOC, se plantearon varias reuniones en las cuales, se entendió, cómo y por qué surgió, la misión de la institución, sus clientes y cómo se trabaja con cada uno de ellos, además de conocer sus sistemas SIEM, los cuales usaban para realizar la recolección de los datos. Adicionalmente se realizó una visita con el fin de entender como era el proceso de identificación de incidentes que realizaba cada operador en el CCOC.

De las anteriores visitas se concluyeron cuáles eran las preguntas que se resolverían, con cuáles entidades y qué fuentes de información utilizaríamos para resolverlas.

A continuación, se muestran las preguntas de data science a resolver:

- ¿Cómo predecir si se está presentando un comportamiento sospechoso sobre el activo?
- ¿Cómo saber si se presentan anomalías en el uso habitual del sistema?

2.2.3.2 Data acquisition and understanding

La adquisición de los datos se realizó mediante la recolección de los eventos de una de las plataformas SIEM del CCOC, estos datos eran recolectados por operadores del CCOC y enviados a nosotros cifrados y en formato CSV.

Una vez los primeros datos fueron entregados se procedió a entender cada uno de los atributos y sus posibles valores y el porque de ellos, estos atributos los podemos ver a continuación:

Tabla 10 Atributos de eventos recibidos

Atributo	Descripción	Posibles datos
Manager Receipt Time	Fecha en la que el SIEM recibe el evento.	1 al 3 de noviembre 13, 17, 18, 27, 30 y 31 de octubre
Name	Nombre del evento	Accept, Allow, Drop
Type	Tipo de evento	Base, Aggregated
End Time	Hora en la que el evento termina y es registrado en el IPS	1 al 3 de noviembre 13, 17, 18, 27, 30 y 31 de octubre
Transport Protocol	Protocolo de transporte	TCP, vacíos
Aggregated Event Count	Cantidad de eventos que fueron agrupados por ser los mismos	1 a 10 eventos agregados
Relevance	Determina que tan aplicable es el ataque para el host.	Valor por defecto: 10 Altamente relevante.

Priority	Depende del puerto y las vulnerabilidades asociadas. Prioridad del evento	Número de 0 a 10 calculado.
Agent Severity	Qué tan severo consideró el creador del evento que fue de forma normalizada	Low, High, Unknown. (Low, medium, high, very-high-unknown)
Device Event Category		SecurityLog
Device Action		Accept, Allow, Drop
Device Address	Ip del dispositivo	
Device Vendor	Identifica el dispositivo	Check Point
Attacker Address	Ip del atacante	Números diversos
Attacker Port	Puerto del atacante	Números diversos
Attacker Geo Country Name	País desde el que proviene el atacante.	No hay datos
Target Address	Dirección Ip del target	172.18.42.64, 190.248.87.2
Target Port	Puerto del target	443, diversos.
Target Service Name	Servicio que está siendo atacado	https, números diversos
Target Geo Location Info	Información adicional sobre la localización	Vacío, Medellín
Target Geo Country Name	País del que proviene el target	Vacíos, Colombia
Request Url	Url desde donde se realiza la petición	Vacíos, 113 & Acceso_SGDA, 114 & Acceso_SGDA, 115 & Acceso_SGDA, 36 & Navegación para el CGFM, 37 & Navegación para el CGFM, 38 &

Device Custom String1	Datos adicionales, en este caso el nombre de la regla en el Firewall.	Navegación para el CGFM, Autorización de servicios J2, &. No hay datos
Target Service Name	Servicio que está siendo atacado	https, números diversos
Target Geo Location Info	Información adicional sobre la localización	Vacío, Medellín
Target Geo Country Name	País del que proviene el target	Vacíos, Colombia
Request Url	URL desde donde se realiza la petición	Vacíos
Device Custom String1	Datos adicionales, en este caso el nombre de la regla en el Firewall.	Vacíos, 113 & Acceso_SGDA, 114 & Acceso_SGDA, 115 & Acceso_SGDA, 36 & Navegación para el CGFM, 37 & Navegación para el CGFM, 38 & Navegación para el CGFM, Autorización de servicios J2, &.
Device Custom String3	Datos adicionales	{5C25CE02-3C7C-4B10-B290-5C434B0D80B9}, {84AD1002-64B8-4E25-BCD8-7B9FA723AE8C}, vacíos
Device Custom String4	Datos adicionales	Standart, Vacíos

Los atributos relevance y priority requieren de una mayor aclaración por lo que a continuación se muestra cómo se calculan.

Relevance (R)

How applicable is the attack against the target host?

Effect	Relevance provides full or partial support for incoming agentSeverity					
Requirement	Heavily dependent on port and vulnerability scanning data					
Factors	Default Value		Port Scan?	Vuln Scan?	Port Open?	Is Vuln?
	10	+	-5	-5	+	+5
Possible values	10	5	0			
	Highly Relevant	Partially Relevant	Irrelevant			

Ilustración 160 Formula matemática para calcular la variable Relevance


Priority formula

$$\text{agentSeverity} * \left(\frac{R}{(R + MC) - \left(\frac{R * MC}{10}\right)} \right) * \left(1 + \left(\frac{S * 3}{100} \right) \right) * \left(1 + \left(\frac{C - 8}{10} \right) * 20\% \right) = \text{Priority}$$

↓
↓
↓

Vulnerability **Threat** **Impact**

% 1.00 - 1.30 .84 - 1.04



11 © Copyright 2014 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice.

Ilustración 161 Formula matemática para calcular la variable Priority

La Ilustración 161 muestra la formula matemática utilizada por el SIEM para establecer el nivel de prioridad de un evento, donde la severidad del agente se

multiplica por el MC (Modelo de confianza) y R (Relevancia), la severidad y la criticidad.

Una vez entendido cada uno de los atributos se procedió a realizar un análisis estadístico, el cual se puede ver en el archivo anexo: Análisis estadístico CCOC, el cuál fue entregado como producto al CCOC y se realizó mediante el uso de tablas dinámicas de la herramienta Excel.

También se realizaron mejoras al traductor con el fin de que fuera estándar para cualquier archivo que se le introdujera.

2.2.3.3 Modeling

En la fase de modelado se identificaron los atributos que se utilizarían para resolver cada una de las preguntas, también se adicionaron campos como Ilegítimo, el cual fue llenado con ayuda de un operador del CCOC. Después de esto se pasaron los archivos ya depurados por el traductor y estos archivos resultantes fueron usados para probar los modelos.

Los modelos utilizados y sus resultados se muestran a continuación:

¿Cómo predecir si se está presentando un comportamiento sospechoso sobre el activo?

DecisionStump		
=== Summary ===		
Correctly Classified Instances	2760	99.711 %
Incorrectly Classified Instances	8	0.289 %
Kappa statistic	0.9942	
Mean absolute error	0.0065	
Root mean squared error	0.0537	
Relative absolute error	1.3079 %	
Root relative squared error	10.7358 %	
Total Number of Instances	2768	

Ilustración 162 Resultado del modelo DecisionStump aplicado a los datos de comportamiento sospechoso

J48

=== Summary ===

Correctly Classified Instances	2760	99.711 %
Incorrectly Classified Instances	8	0.289 %
Kappa statistic	0.9942	
Mean absolute error	0.0065	
Root mean squared error	0.0537	
Relative absolute error	1.3092 %	
Root relative squared error	10.7359 %	
Total Number of Instances	2768	

*Ilustración 163 Resultado del modelo J48 aplicado a los datos de comportamiento sospechoso***NaiveBayes**

=== Summary ===

Correctly Classified Instances	2760	99.711 %
Incorrectly Classified Instances	8	0.289 %
Kappa statistic	0.9942	
Mean absolute error	0.003	
Root mean squared error	0.0541	
Relative absolute error	0.5999 %	
Root relative squared error	10.8144 %	
Total Number of Instances	2768	

*Ilustración 164 Resultado del modelo NaiveBayes aplicado a los datos de comportamiento sospechoso***SMO**

=== Summary ===

Correctly Classified Instances	2766	99.9277 %
Incorrectly Classified Instances	2	0.0723 %
Kappa statistic	0.9986	
Mean absolute error	0.0007	
Root mean squared error	0.0269	
Relative absolute error	0.1445 %	
Root relative squared error	5.3769 %	
Total Number of Instances	2768	

Ilustración 165 Resultado del modelo SMO aplicado a los datos de comportamiento sospechoso

KStar		
=== Summary ===		
Correctly Classified Instances	2767	99.9639 %
Incorrectly Classified Instances	1	0.0361 %
Kappa statistic	0.9993	
Mean absolute error	0.0013	
Root mean squared error	0.022	
Relative absolute error	0.2632 %	
Root relative squared error	4.4064 %	
Total Number of Instances	2768	

Ilustración 166 Resultado del modelo Kstar aplicado a los datos de comportamiento sospechoso

2.2.3.4 Implementation

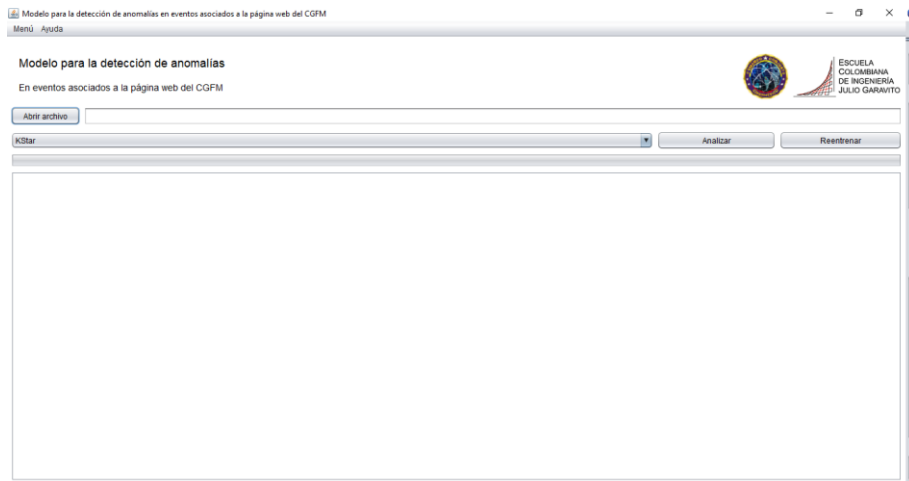


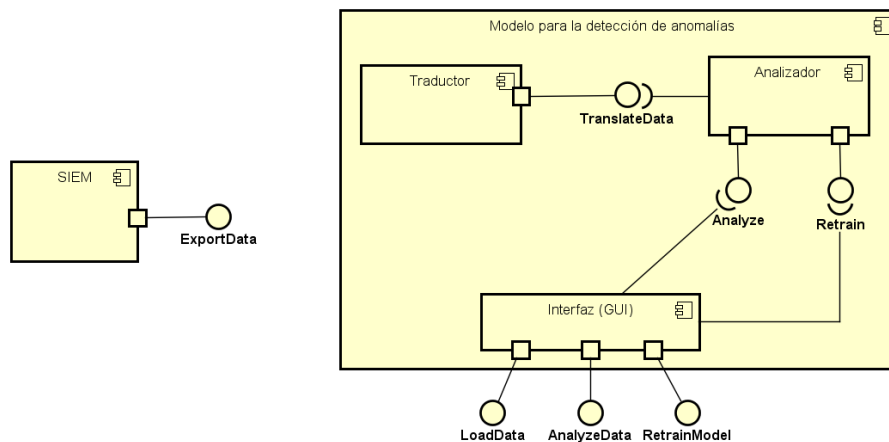
Ilustración 167 Pantalla principal de la aplicación

Para la fase de implementación se revisó cómo facilitar al CCOC el análisis permanente de sus datos. A este respecto se determinó que ellos requerían apoyo en la normalización de los datos y el entrenamiento del modelo por medio de una aplicación donde pudieran cargar sus datos y realizar el análisis predictivo o, reentrenar los modelos generados, apoyando el proceso de análisis del operador.

Entendido esto, la implementación se hizo a través de un aplicativo en Java que permite:

1. Cargar los datos que se extraigan del SIEM del CCOC para ser analizados. En el proceso la aplicación filtra automáticamente las columnas innecesarias, realiza la conversión para que la información sea leída por Weka, se analiza el archivo y las instancias que puedan ser predichas se predicen y se muestra el resultado. Las instancias que no puedan ser clasificadas se exportan a un archivo CSV para que el operador del CCOC las clasifique y las importe al aplicativo para que la aplicación aprenda aún más.
2. Cargar los datos que se extraigan del SIEM del CCOC o que se exportaron en la funcionalidad anterior para que el modelo aprenda. En el proceso la aplicación filtra automáticamente las columnas innecesarias, realiza la conversión para que la información sea leída por Weka y se construye un nuevo modelo a partir de la información nueva y la histórica.

En el siguiente diagrama se muestra la arquitectura de la aplicación:



Comentado [CPSC4]: Explicar los componentes de la arquitectura.

Ilustración 168 Diagrama de Componentes de la aplicación desarrollada

En la Ilustración 168 se muestra el diagrama de arquitectura de la aplicación, por un lado, se encuentra el SIEM que le permite al operador exportar los eventos de seguridad en un archivo csv, y en el caso de la aplicación, permite al operador cargar los datos y realizar análisis predictivo o reentrenar el modelo. La aplicación tiene 3 componentes fundamentales, la interfaz que es la cara al usuario, el analizador que es el que realiza el análisis predictivo por medio del uso del API de Weka, y el traductor, que se encarga de traducir los eventos a líneas que los modelos matemáticos entiendan.

Realizar una predicción:

1. Abra el programa y haga clic en el botón de abrir archivo, aparecerá un diálogo donde se debe cargar un archivo .csv. Procure que los eventos que aparezcan en dicho archivo tengan al menos los siguientes atributos: Type, Transport Protocol, Aggregated Event Count, Priority, Agent Severity, Device Event Category, Device Action, Device Vendor, Attacker Address, Attacker Port, Attacker Geo Country Name, Target Address, Target Translated

Address, Target Port, Target Service Name, Target Geo Location Info, Target Geo Country Name, Request Url, Device Custom String1, Device Custom String3, Device Custom String4.

Modelo para la detección de anomalías

En eventos asociados a la página web del CGFM

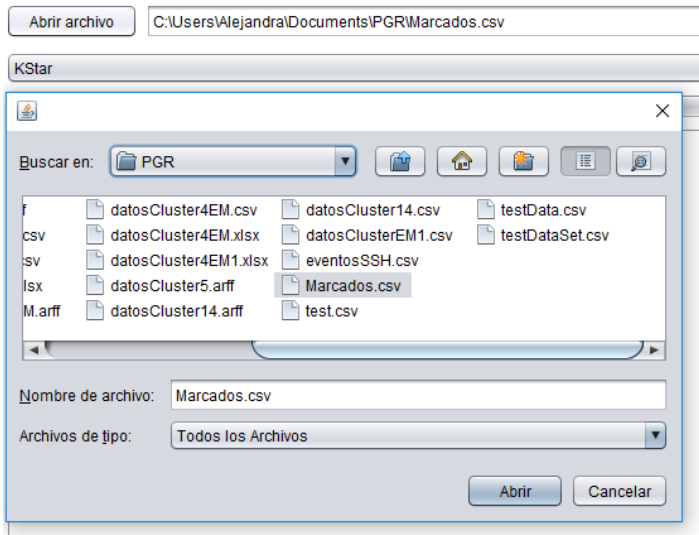


Ilustración 169 Diálogo para escoger archivo para realizar el análisis.

2. Una vez seleccionado el archivo el programa pregunta qué delimitador tiene el archivo csv, en este caso será un “;”.

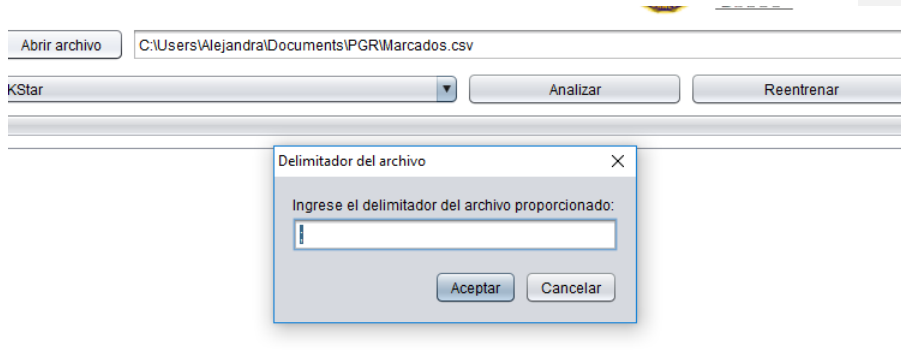


Ilustración 170 Diálogo de delimitador del programa.

3. Hacemos clic en el botón analizar, nos dirá las instancias que conoce y que desconoce. Las instancias desconocidas se guardarán en el archivo “unknownInstances.csv” ubicado en la carpeta de instalación del programa.

Modelo para la detección de anomalías

En eventos asociados a la página web del CGFM

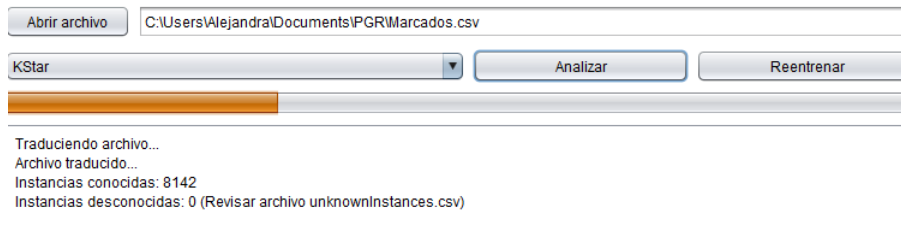


Ilustración 171 Ventana de análisis

4. Al finalizar nos mostrará una tabla con el identificador del evento, el resultado de la predicción y su probabilidad:



Ilustración 172 Finalización del análisis

Realizar el reentrenamiento del modelo:

1. Seguimos el mismo proceso para abrir un archivo csv, este archivo debe tener los campos que se listan a continuación: Type, Transport Protocol, Aggregated Event Count, Priority, Agent Severity, Device Event Category, Device Action, Device Vendor, Attacker Address, Attacker Port, Attacker Geo Country Name, Target Address, Target Translated Address, Target Port, Target Service Name, Target Geo Location Info, Target Geo Country Name, Request Url, Device Custom String1, Device Custom String3, Device Custom String4, Illegitimate Request.
Este último, campo debe ser diligenciado por el operador, siendo “Y” si es sospechoso y “N” si no es sospechoso. Ya que el modelo se va a reentrenar con el criterio del operador.
2. Hacemos clic en el botón reentrenar y nos mostrará un mensaje de espera:

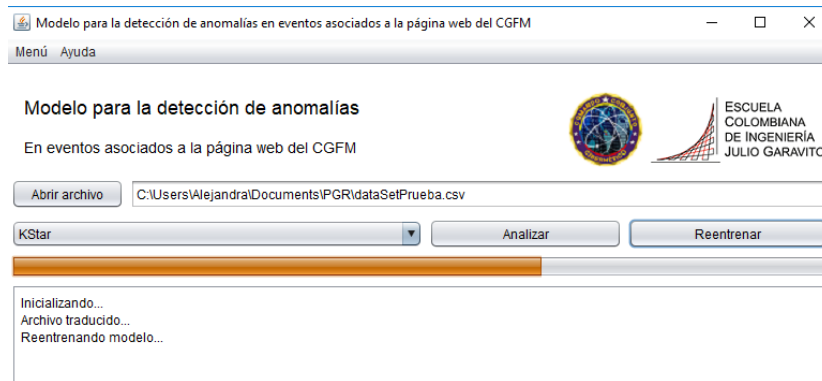
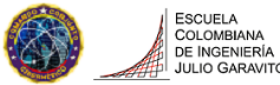


Ilustración 173 Vista de reentrenamiento del modelo

3. Al finalizar nos dará un resumen de los resultados del reentrenamiento del modelo con el número de instancias correctamente clasificadas, se debe tener en cuenta que este modelo es histórico.

Modelo para la detección de anomalías

En eventos asociados a la página web del CGFM



Abrir archivo

KStar

```

Inicializando...
Archivo traducido...
Reentrenando modelo...
Evaluando el 66.0% de los datos suministrados
Resultado del entrenamiento para: KStar
-----
Resultados
Correctly Classified Instances  2770    99.7479 %
Incorrectly Classified Instances  7    0.2521 %
Kappa statistic                0.9932
K&B Relative Info Score        351133.2981 %
K&B Information Score          3366.1394 bits  1.2121 bits/instance
Class complexity | order 0     3392.1298 bits  1.2215 bits/instance
Class complexity | scheme      59.5055 bits   0.0214 bits/instance
Complexity improvement (Sf)    3332.6243 bits  1.2001 bits/instance
Mean absolute error            0.0051
Root mean squared error        0.0532
Relative absolute error        0.9105 %
Root relative squared error     9.3381 %
Total Number of Instances     2777
Fmeasure: 0.998331346841478
Precision: 0.9990458015267175
Recall: 0.9976179132920439=== Confusion Matrix ===

```

Ilustración 174 Salida de reentrenamiento del modelo.

Este aplicativo se entregó al CCOC, para que ellos puedan utilizarlo y afinarlo según lo requieran.

Por último, queremos que los lectores se animen a trabajar en este campo ya que la implementación de Data Science, Machine Learning y en general la inteligencia artificial, hoy en día se está aplicando a todos los campos de la computación tal como lo es la seguridad informática, sin embargo, aún queda mucho trabajo por realizar para lograr que estas implementaciones sean eficientes y efectivas.

Conclusiones

De este trabajo podemos concluir varias cosas:

- En primera instancia los sistemas SIEM idealmente deberían poder detectar ataques y generar respuestas efectivas, pero en la práctica una respuesta activa puede llegar a ser muy intrusiva dentro del sistema y generar incidentes en la operación de la organización, es por ello que estos sistemas suelen utilizarse como sistemas de auditoría limitando sus capacidades.
- En cuanto al envío de eventos de seguridad, la instalación de agentes como método para el envío de eventos dentro de los sistemas operativos nos permite aumentar la visión al poder no sólo enviar logs sino también realizar un escaneo del estado del sistema, pero en la práctica puede llegar a consumir muchos recursos dentro de la máquina; uno de los retos es generar un mecanismo que permita el envío de eventos de seguridad, tanto logs como estado de la máquina, sin ser intrusivos y consumiendo menos recursos. El uso de puntos de detección para el aseguramiento de aplicaciones de usuario es una buena práctica y debe ser considerada en el proceso de desarrollo de software, enfocándose en la auditoría de lo esencial y sin llegar al punto de sobrecargar el sistema con el envío de demasiados eventos. Por otra parte, la creación de plugins (adaptadores) para conectar las diferentes aplicaciones con el SIEM es un reto para las organizaciones que poseen aplicaciones desarrolladas a la medida.
- Las reglas de correlación dentro de un SIEM pueden llegar a ser muy limitadas, es necesario generar métodos de correlación que permitan detectar ataques mucho más elaborados como lo son los APT's (Advanced Persistent Threat), utilizando algoritmos de aprendizaje automático y generando reglas según ese aprendizaje. Por otra parte, la identificación del tipo de evento entrante al sistema se ve muy limitada ya que la mayoría de

plataformas implementan expresiones regulares para realizarla; es bien sabido que estas son muy limitadas y no nos permiten realizar todas las acciones que se desearían, es por esto que el desarrollo de lenguajes formales para la identificación de eventos sería una buena práctica ya que además de brindar más libertad, contribuiría a la eliminación de la necesidad de desarrollar adaptadores para cada aplicación.

- Generar la mejor respuesta a un incidente es uno de los mayores retos, por ello se debe profundizar mucho más en el tema para generar respuestas efectivas, mitigando el riesgo asociado y minimizando el impacto que podría tener en la organización.
- La realización de modelos generales de data science, es un proceso muy complicado debido a que existe una delgada línea entre la generalidad y la particularidad en los procesos que implican manejo de datos como lo son la traducción y la configuración de los modelos. Sin embargo, etapas como el entendimiento del negocio y el entendimiento de los datos siempre son particulares para cada caso por lo que la realización de un modelo general no es posible aún.
- Data science aplicado a ciberseguridad ayuda a optimizar procesos, ahorrando tiempo a el operador y dinero a las empresas, también permite identificar perfiles de adversarios y con dicha información generar mejores estrategias de defensa.
- Para que el proceso de Data science sea aplicado correcta y eficazmente se debe asegurar que los datos ingresados a los diferentes modelos deben ser de calidad, no deben introducir ruido y se deben tener una cantidad considerable de datos dependiendo de cada caso para lograr un correcto entrenamiento de los modelos.

- La aplicación de los modelos de machine learning en entornos académicos afrontan retos como, la calidad, cantidad y diversidad de datos, ya que en nuestro caso estos factores fueron críticos en el proceso.

Bibliografía

1. Alienvault. About Reporting in AlienVault USM Appliance [Internet]. www.alienvault.com/documentation. [cited 2018 Jan 4]. p. 1. Available from: <https://www.alienvault.com/documentation/usm-appliance/reports/about-usm-reports.htm>
2. Alienvault. Open Threat Exchange (OTX) | AlienVault [Internet]. www.alienvault.com. [cited 2018 Jan 4]. p. 1. Available from: <https://www.alienvault.com/open-threat-exchange>
3. AlienVault Open Threat Exchange User Guide, rev. 4 AlienVault Open Threat Exchange (OTX)[™] User Guide. AlienVault Open Threat Exch User Guid [Internet]. 2016 [cited 2018 Jan 4]; Available from: <https://www.alienvault.com/doc-repo/OTX/user-guides/AlienVault-OTX-User-Guide.pdf>
4. Alienvault. SNMP Configuration in AlienVault USM Appliance [Internet]. www.alienvault.com/documentation. [cited 2018 Jan 4]. p. 1. Available from: <https://www.alienvault.com/documentation/usm-appliance/kb/2016/01/snmp-configuration-v52.htm>
5. OSSEC Project. Getting started with OSSEC — OSSEC [Internet]. ossec.github.io. [cited 2018 Jan 6]. p. 1. Available from: <https://ossec.github.io/docs/manual/non-technical-overview.html>
6. Imperva. Top Ten Database Security Threats [Internet]. Forecast. 2015 [cited 2017 Apr 7]. p. 10. Available from: https://www.imperva.com/docs/gated/WP_Top_5_Database_Security_Threats.pdf
7. Williams J, Wichers D. OWASP Top 10 - 2013 [Internet]. 2013 [cited 2017 Apr 10]. Available from: https://www.owasp.org/images/f/f8/OWASP_Top_10_-_2013.pdf
8. Alienvault. About Correlation Directives in AlienVault USM Appliance [Internet]. www.alienvault.com/documentation. [cited 2018 Jan 5]. p. 1. Available from: <https://www.alienvault.com/documentation/usm-appliance/correlation/about-correlation-directives.htm>
9. Alienvault. Customizing Correlation Directives and Cross Correlation Rules

Edition Date of Issue Description of Change(s).
www.alienvault.com/documentation [Internet]. [cited 2018 Jan 5];1. Available
from: <https://www.alienvault.com/doc-repo/USM-for-Government/all/Customizing-Correlation-Directives-or-Cross-Correlation-Rules.pdf>

10. Alienvault. About the Policy View in AlienVault USM Appliance [Internet].
www.alienvault.com/documentation. [cited 2018 Jan 11]. p. 1. Available from:
<https://www.alienvault.com/documentation/usm-appliance/policy-management/about-pol-view.htm>
11. Universidad de Waikato. Weka 3 - Data Mining with Open Source Machine
Learning Software in Java [Internet]. www.cs.waikato.ac.nz. 2018 [cited 2018
Jan 16]. Available from: <https://www.cs.waikato.ac.nz/ml/weka/>