2022

MACHINE LEARNING

Detección de ataques de phishing

Caso de uso.

Alumno. José David Ortiz Cruz

BOOTCAM
CYBER III

Descripción del caso de uso

¿Cuál es el problema?

Los Equipos y Unidades pertenecientes a las Fuerzas y Cuerpos de Seguridad del Estado que tienen como cometido la investigación de aquellos delitos cometidos a través de Internet, llevan a cabo labores de respuesta ante incidentes y análisis forense tendentes a la recolección de pruebas e indicios del ciberataque objeto de investigación, que puedan contribuir a la atribución del ataque a los autores del hecho investigado.

Como parte de estas tareas, se considera de máxima prioridad a fin de detectar e identificar modus operandi similares, el vector de entrada de los ciberataques que tengan como víctimas a las pequeñas y medianas empresas.

Según estudios internos por parte del Ministerio del Interior, el 93 % de estos ciberataques tienen como vector de entrada el correo electrónico.

Por ello, siempre que se da un ciberataque a una empresa, agentes especializados en investigación tecnológica recaban las cabeceras de los correos electrónicos existentes en las bandejas de entrada de los buzones de la totalidad de direcciones de correo tanto corporativas como personales que puedan estar relacionadas con la entidad objeto del ciberataque.

Esto se debe a que con posterioridad se analizarán dichas cabeceras en busca de elementos típicamente observados en campañas de phishing tales como enlaces de descarga, ficheros documentales adjuntos con código malicioso inyectado en los mismo, campos de origen del mensaje deliberadamente ocultos, etc…

¿Cómo se está afrontando ahora?

Dichas cabeceras a día de hoy requieren de un análisis manual y que el personal que lo realice tenga una formación y capacitación técnica altamente desarrollada.

La actual demanda de profesionales en ciberseguridad dificulta encontrar y formar perfiles adecuados para este cometido, más aún siendo el principal requisito previo, ser miembro en activo de las fuerzas y cuerpos de seguridad del estado previo pase del correspondiente concurso-oposición.

Una organización media que dispone de 14 cuentas de correo electrónico, con un volumen de recepción aproximado de 90 correos electrónicos al mes cada uno de ellos, en un análisis del trimestre anterior a la ocurrencia del ataque, conlleva el análisis de 3.780 cabeceras de correo electrónico.

Dada la falta de personal especializado así como la necesidad actual de llevar a cabo el referido análisis de forma manual, conlleva que un único analista dedicará a razón de diez cabeceras analizadas por hora, 378 horas o lo que es lo mismo, algo más de tres meses de exclusiva dedicación a esta tarea, lo que resulta obviamente inviable, optando por limitar el análisis a los días y buzones de correo que más probabilidad tengan de haber sido los vectores de entrada concretos.

Esta operativa de investigación resta una importante capacidad a los analistas de observar movimientos laterales de atacantes, intentos anteriores que no hubiesen fructificado o

poder determinar si la intrusión fue motivada por una campaña de phishing aleatoria o fue un ataque dirigido a la organización víctima.

¿Acción que buscamos poder hacer para solucionar el problema?

Resulta de máximo interés el desarrollo de una herramienta que permitiese analizar de manera automatizada las cabeceras de los correos electrónicos. Si bien, se haría necesario un posterior análisis manual sobre los casos detectados por la herramienta, estos serian una cuantía infinitamente inferior al bruto de los datos obtenidos, permitiendo así al analista ganar velocidad y efectividad en su labor de investigación.

KPIs – Indicadores de negocio

Partiendo del ejemplo anterior, dónde un analista tardaría 378 horas en analizar 3.780 cabeceras de correo electrónico, si estas se viesen reducidas tras el análisis automatizado de la herramienta a los correos que mostrasen indicios habituales de ser ataques de phishing, los cuales suelen concretarse en un análisis manual en torno al 5 % de la muestra de entre los cuales posteriormente se aíslan los 3 o 4 correos electrónicos que realmente han sido enviados con ánimo de ejecutar el ciberataque, el analista tan solo tendría que analizar 189 correos electrónicos.

Esto redundaría en que pasaríamos de las 378 horas a algo más de 19 para poder tener un análisis completo de la información extraída en la entidad atacada. El tiempo de dedicación del analista se vería reducido de más de tres meses a dos días y medio de jornada laboral, lo que permitiría obtener resultados en casos de emergencia en menos de 24 horas, aumentando en un 95 % la capacidad de respuesta de la Unidad de investigación tecnológica que dispusiese de la herramienta.

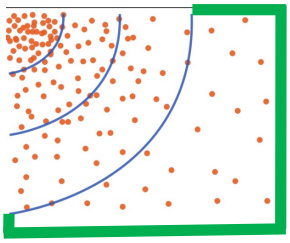
¿Cuáles son los mínimos que se esperan de este caso de uso?

La herramienta desarrollada debería ser capaz de detectar la máxima cantidad posible de cabeceras de correo electrónico que compartan parámetros comunes con ataques de phishing habituales, no siendo necesario que el nivel de aciertos sobre estas detecciones fuese amplio, dado que las muestras detectadas por la herramienta serán analizadas seguidamente por el analista.

Validación: ¿Qué criterio se va a usar para decidir si la solución es aceptable?

Como se ha expuesto en el punto anterior, la prioridad debería ser la detección, muy por encima de la capacidad de acierto de la herramienta.

La zona señalada en verde de la siguiente imagen resultaría muy representativa del objetivo perseguido, dónde la herramienta consigue detectar el 90 % de las cabeceras susceptibles de ser un ataque de phishing, consiguiendo acertar en 1 de cada 9.



Experimentación: ¿Cómo vamos a corroborar el funcionamiento?

Se debería llevar a cabo el desarrollo de un modelo predictivo experimental de machine learning mediante el entrenamiento de aprendizaje supervisado sobre el mismo, utilizando para ello un algoritmo de regresión logística. Todo ello, permitirá trabajar a la herramienta sobre una base donde se le proporciona al algoritmo de aprendizaje automático un conjunto de datos conocidos que incluye las entradas y salidas deseadas, y el algoritmo debe encontrar un método para determinar cómo llegar a esas entradas y salidas.

El algoritmo realizará predicciones será corregido por el analista, y este proceso seguirá hasta que el algoritmo alcance un alto nivel de precisión y rendimiento.

El tiempo necesario de este entrenamiento será definido por los resultados que se vayan obteniendo.

Se utilizará como dataset de entrenamiento, una selección de ficheros .eml recogidos durante investigaciones reales anteriores sobre los cuales ya se ha realizado un estudio manual determinándose cuales de ellos eran considerados phishing.

Productivización: ¿Qué salida debe tener la solución que se desarrolle?

Lo ideal sería que la herramienta fuese desarrollada para que su uso pudiese llevarse a cabo por personal no especializado, siendo necesario unos mínimos conocimientos de ofimática para ejecutar la herramienta.

Esta debería estar montada sobre un modelo cliente-servidor con una interfaz gráfica de acceso desde el navegador que permitiese la carga de ficheros .eml.

Una vez cargados dichos ficheros y ejecutada la acción predictiva de la herramienta, esta determinaría y aislaría las cabeceras de correo electrónico contenido en los ficheros .eml susceptibles de contener un ataque de phishing, almacenando copia de las mismas en un directorio específico.

Que la acción se desarrolle de la forma anteriormente descrita, permitirá al personal no especializado hacer entrega exclusivamente de los ficheros que deberán ser analizados manualmente al analista técnicamente capacitado, reduciendo sustancialmente la carga de trabajo de éste y posibilitando un completo análisis de la línea de investigación seguida.

Equipo de trabajo

Identificación de personas colaboradoras

Para el desarrollo del proyecto se hará necesaria la coordinación entre un científico de datos, un desarrollador web que domine tecnológicas y frameworks tanto de backend como de frontend, así como un especialista en ciberseguridad con capacitación en análisis forense y conocimientos en malware y reversing que proporcione la base de muestras reales “datasets” sobre las que se desarrolle el aprendizaje del algoritmo.

Igualmente, resultaría de interés la participación en el proyecto de alguno de los miembros de las Unidades de Investigación Tecnológica de las Fuerzas y Cuerpos de Seguridad del Estado que con posterioridad podrían ejercer como formadores de sus compañeros en el uso de la herramienta e interpretación de sus resultados.

Detalle del caso de uso

Detalle funcional

La herramienta sería utilizada en el entorno de producción del Ministerio de Interior por lo que en todo momento deberá verse sometida a los principios regidos por el Reglamento General de Protección de Datos así como a lo establecido en la Ley de Enjuiciamiento Criminal sobre los criterios y operativas en la investigación de delitos en el ámbito tecnológico dónde se hace necesaria la permanente tutela judicial de las investigaciones.

Por otra parte, al tratarse de un organismo perteneciente a la Administración General del Estado, el uso de la herramienta, una vez desarrollada, deberá ser autorizada por el Centro Criptológico Nacional.

Identificación de orígenes de datos

Los datos que se utilizaran para dar funcionalidad a la herramienta serán los correos electrónicos almacenados en ficheros de extensión .eml que previamente serán extraídos por los especialistas forenses de los equipos informáticos de las organizaciones afectadas.

Desarrollo del caso de uso

Puntos intermedios o seguimiento

Sería adecuado poder verificar que correos analizados no son compatibles con ser elementos parte de un supuesto ataque de phishing al objeto de poder descartarlos y reducir así la carga de trabajo manual del analista.

Aporte esperado

El problema a solventar por el equipo de desarrollo consistirá en dotar a los investigadores de una herramienta simple y funcional que acelere notablemente la labor de análisis de estos y reduzca su necesidad de capacitación y especialización técnica.

El alumno

José David Ortiz Cruz