Universitat Oberta de Catalunya

uoc.edu

Métodos y técnicas de detección temprana de casos de phishing

Jaime López Sánchez
Máster Universitario en Seguridad de las Tecnologías de la Información y de las Comunicaciones

Pau del Canto

Enero - 2019

FICHA DEL TRABAJO FINAL

Título del trabajo:	*Métodos y técnicas de detección temprana de casos de phishing.*
Nombre del autor:	*Jaime López Sánchez*
Nombre del consultor/a:	*Nombre y dos apellidos*
Nombre del PRA:	*Pau del Canto*
Fecha de entrega (mm/aaaa):	01/2019
Titulación::	*Máster Universitario en Seguridad de las Tecnologías de la Información y de las Comunicaciones*
Área del Trabajo Final:	*Seguridad de las Tecnologías de la Información y de las Comunicaciones*
Idioma del trabajo:	*Castellano*
Palabras clave	*phishing, detección, técnicas*

Resumen del Trabajo (máximo 250 palabras): *Con la finalidad, contexto de aplicación, metodología, resultados i conclusiones del trabajo.*

Los casos de phishing son un fraude cuyo coste anual para las empresas se estima en unos 1,6 millones de euros. La lucha contra este tipo de estafas se basa en la detección temprana de estos casos y la eliminación de los mismos antes de que hayan podido afectar a un número importante de víctimas. En este trabajo se van a estudiar diferentes métodos y técnicas para detectar estos fraudes en un momento inicial en el que todavía no haya víctimas de este fraude y se pueda erradicar sin que se lamenten pérdidas económicas.

Entre las técnicas que se analizarán se encuentran los sistemas de DMARC (DKIM y SPF), monitorización y análisis de los referers de las peticiones a la web de una compañía, monitorización de dominios que se den de alta y que puedan usar técnicas de typo-squatting y creación de un sistema de análisis de páginas webs en base a indicadores de compromiso para detectar si se trata de un caso de phishing que afecta a una compañía.

Abstract (in English, 250 words or less):

Phishing cases are a fraud which annual cost for the corporations is estimated in around 1.6 million euros. Fighting against this kind of scam is based on the early detection of this cases and the mitigation of them before many people gets affected. In this thesis we are going to study different methods and techniques for detecting this fraud in an early moment when there is no victims and it can be taken down without regretting economic losses.

Among the different techniques that are going to be analyzed, we can highlight the DMARC (DKIM and SPF) systems, analysis and monitorization of the request referers, new domain monitorization that can be susceptible of typo-squatting techniques and the creation of an automatic system that analyses websites according to different indicator of compromises in order to detect if this could be a phishing case that affects a company.

Índice

Tabla de contenido

Lista de figuras

1. Introducción

1.1 Contexto y justificación del Trabajo

Debido a la gran cantidad de casos de phishing que son generados a diario y que afectan a miles de usuarios de banca online, vamos a estudiar en este trabajo diferentes técnicas de detección temprana de este tipo de estafa.

Si conseguimos realizar una detección de este tipo de casos en una etapa inicial del fraude, se puede reducir en gran medida el impacto de este tipo de estafas y a largo plazo que estos casos resulten rentables para los estafadores. Debido también al gran nivel de sofisticación de los casos de phishing actuales, la concienciación del usuario, aunque es importante, no es determinante en la erradicación de estas amenazas. Es por ello que desde las empresas afectadas se deben de implantar diferentes métodos de detección y eliminación de estos casos para que los usuarios finales impactados sean los menos posibles, y por tanto, las pérdidas económicas sean reducidas

1.2 Objetivos del Trabajo

- Entender el fraude que suponen los casos de phishing para las empresas, especialmente del sector banca.
- Analizar las diferentes tipologías existentes entre los casos de phishing.
- Desarrollar y estudiar técnicas de detección tempranas adaptadas a las diferentes casuísticas presentes en este tipo de estafas.
- Conocer las acciones preventivas y recomendaciones para poder reducir los casos de fraude que afecten a una compañía.
- Desarrollo de un sistema de análisis de páginas web basado en IOCs (Indicadores de compromiso) para discriminar de manera automática si una web se trata de un caso de phishing.
- Detallar las acciones necesarias y recomendaciones para una empresa para la configuración de un sistema de DMARC (DKIM + SPF)
- Detección de registro de dominios que usen técnicas de typo-squatting mediante registro de whois diario y monitorización del sistema de Certificate Transparency de Google.

1.3 Enfoque y método seguido

La detección de phishing no debe de ser abordada de una única manera debido a la gran cantidad de casos y a los diferentes métodos de propagación que usan

los estafadores para captar víctimas y que sus estafas alcancen el mayor número posible de personas.

Es por ello que las técnicas que se detallan en este trabajo son variadas y la idea es poder realizar una detección panorámica que aborde todos los diferentes protocolos y tecnologías que intervienen en este tipo de estafas.

A lo largo del punto 3 de esta memoria se estudiarán las principales técnicas de detección temprana que pueden usarse para los casos de phishing. De igual manera, se desarrollará una herramienta de creación propia con la que poder discernir si una web constituye un caso de phishing para una empresa mediante la comprobación de una serie de indicadores de compromiso.

1.4 Planificación del Trabajo

- Durante la PEC-1 se elaborará el punto 1.
- Para la PEC-2 se trabajará en el punto 2 fundamentalmente y se completarán aspectos del punto 5, 6 y 7 de manera simultánea a la redacción del punto 2.
- La PEC-3 se establecerá con la terminación del punto 2 y comienzo del punto 3.
- En la PEC-4 se entregarán todos los puntos de la memoria finalizados, fundamentalmente se terminará de completar el punto 3 junto a las observaciones del tutor asignado al TFM.

	Punto 1	Punto 2	Punto 3	Punto 4	Punto 5	Punto 6	Punto 7
PEC-1							
PEC-2							
PEC-3							
PEC-4							

1.5 Breve descripción de los otros capítulos de la memoria

El segundo capítulo de esta memoria va a incluir la definición de phishing, los diferentes tipos de casos de phishing existentes y el estado del arte de este tipo de amenazas. A día de hoy la sofisticación de este tipo de fraudes es grande y constantemente se encuentran nuevas técnicas para saltar las medidas de seguridad que se ponen para evitar que estos fraudes tengan un gran número de víctimas.

En el tercer capítulo se abordarán las diferentes técnicas de detección temprana de phishing existentes (DMARC, OSINT, Referers, Certificados SSL…) y se desarrollará un sistema de detección automatizada de phishing contra unos indicadores de compromiso.

1.6. Proyectos desarrollados acerca de técnicas de anti-phishing

Son muchos los trabajos que han sido desarrollados por diferentes investigadores de seguridad informática, estudiantes… La mayoría de ellos establecen una serie de herramientas para la detección de phishing que se basan en la similitud de una página fraudulenta a otra legítima y suelen estar basados en sistemas de inteligencia artificial para el reconocimiento de imágenes (logos, banners…).

Algunos de los principales trabajos sobre esta área son:

- A comparison of machine learning techniques for phishing detection. S Abu-Nimeh, D Nappa, X Wang, et al.
- Visual similarity-based phishing detection without victim site information. Masanori Hara, Akira Yamada, Yutaka Miyake.
- Intelligent phishing detection system for e-banking using fuzzy data mining. M Aburrous, MA Hossain, K Dahal, F Thabtah.

Otros trabajos se han centrado en la creación de una herramienta que permita detectar estos casos mediante análisis semántico y machine learning como el trabajo de Venkatesh Ramanathan y Harry Wechsler titulado “phishGILLNET—phishing detection methodology using probabilistic latent semantic analysis, AdaBoost, and co-training”.

Por último, también hay trabajos que hacen uso de las redes sociales y la búsqueda en fuentes abiertas para la detección del fraude como en la investigación de Anupama Aggarwal, Ashwin Rajadesingan y Ponnurangam Kumaraguru titulada “PhishAri: Automatic realtime phishing detection on twitter”.

Este trabajo presenta un enfoque diferente debido a que se nutre de todos estos métodos e investigaciones anteriores pero presenta una aproximación más sencilla al problema de la detección automatizada de casos de phishing. La extracción de una serie de indicadores de compromiso de la página legítima y contraste contra otras páginas potencialmente fraudulentas es computacionalmente mucho más sencilla que la creación de un sistema de inteligencia artificial, uso de minería de datos o de machine learning. Además, en la gran mayoría de casos, va a obtener resultados similares ya que los recursos cargados por el kit de phishing son mayormente coincidentes con los de la página legítima y esto es un buen indicador de que una página es potencialmente fraudulenta.

2. Phishing

2.1. DEFINICIÓN

El phishing es una estafa en la que se hace uso de técnicas de ingeniería social para engañar a un usuario con el fin de obtener información o algún beneficio de manera ilícita.

Para conseguir un mayor número de víctimas, los estafadores están constantemente mejorando las técnicas que usan para engañar a las víctimas. En el caso de los correos de phishing, en muchas ocasiones el correo parece provenir de la entidad legítima y es complicado para un usuario discernir si se trata de un correo de estafa o de un correo de su empresa de confianza.

Al atacante que realiza el engaño mediante phishing se le llama comúnmente "phisher", y el término "phishing" proviene de la palabra inglesa "fishing" que significa pescar, ya que lo que se intenta con este tipo de estafas es "pescar" a la víctima o visto de otro modo, que la víctima "muerda el anzuelo".

2.2. ESTADO DEL ARTE

El phishing se encuentra en constante evolución debido a que, con el propósito de que el mayor número de víctimas posible caigan en la estafa, se deben de ir perfeccionando las técnicas de suplantación que se empeñan en este fraude para sorprender y engañar a los usuarios.

Algunas técnicas comúnmente utilizadas son:

- **Spoofing de correo**:
  El emisor del correo fraudulento simula ser la entidad a la que está suplantando. Esto se consigue al comunicarle al servidor de correo que se está enviando un correo desde una dirección de correo determinada, aunque esto no sea cierto. La víctima sólo podrá comprobar que el correo no está llegando desde donde se afirma mediante la comprobación de la dirección IP emisora del mensaje, ya que esta no pertenecerá al dominio que supuestamente envía el mensaje de correo.

  Como se observa en la Figura 1, esta falsificación del emisor del correo se consigue debido a las opciones existentes en el protocolo SMTP para el envío de correos electrónicos. En esta imagen se observa cómo mediante una sesión telnet en el servidor SMTP de la víctima, es posible enviar un correo electrónico desde una dirección de correo pero que aparentemente este haya sido mandado desde otra.

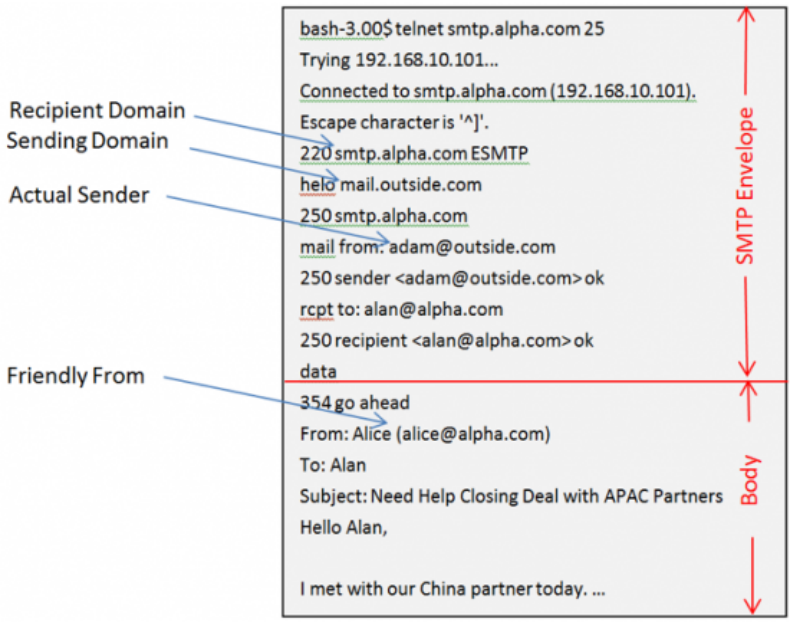


Figura 1 – Ejemplo de email spoofing (2)

- **Typosquatting**:
  Los *phishers* a menudo registran dominios en internet similares a los de la empresa a la que quieren suplantar. A simple vista la víctima que acceda a la página web con nombre similar al esperado, puede confundirla con la de la empresa legítima e introducir los datos de inicio de sesión sin más reparo. Habitualmente los dominios que usan técnicas de "*typosquatting*" difieren del original en una letra, bien suprimiendo una letra con respecto al original o duplicando una de ellas. También son habituales los casos en los que, aunque el dominio registrado no se parezca en nada al que se quiere suplantar, se generan subdominios de gran tamaño para simular el nombre original de la página (véase **Figura 2**).

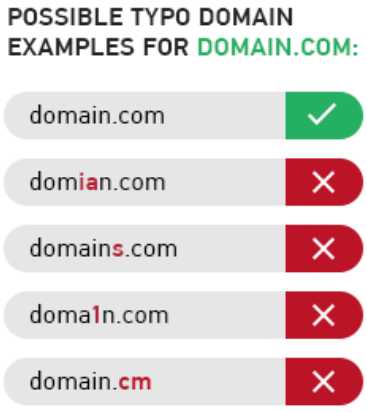


Figura 2 – Ejemplos de dominios de typosquatting **(1)**

- **IDN homograph:**
  En abril de 2017, el investigador de seguridad informática Xudong Zheng descubrió una funcionalidad presente en todos los navegadores de internet modernos que podía ser usada con fines ilícitos y para poder enmascarar una página fraudulenta y que pareciera la web legítima.

  A este tipo de engaño se le llamó ataque homográfico (IDN homograph) y consiste en registrar un dominio web que utiliza caracteres en alfabetos distintos al latino. Estos caracteres son traducidos por el navegador a su equivalente en el alfabeto latino y son mostrados al usuario de manera que este no es capaz de diferenciar entre un dominio con este tipo de caracteres y otro con caracteres latinos.

  Un ejemplo de este tipo de dominios es “xn--pple-43d.com” el cual está compuesto por el carácter cirílico “a” (U+0430) seguido “pple”. Este dominio en un navegador web se verá como “apple.com” como se observa en la Figura 3, aunque como hemos visto no es la web legítima de la marca de productos informáticos.

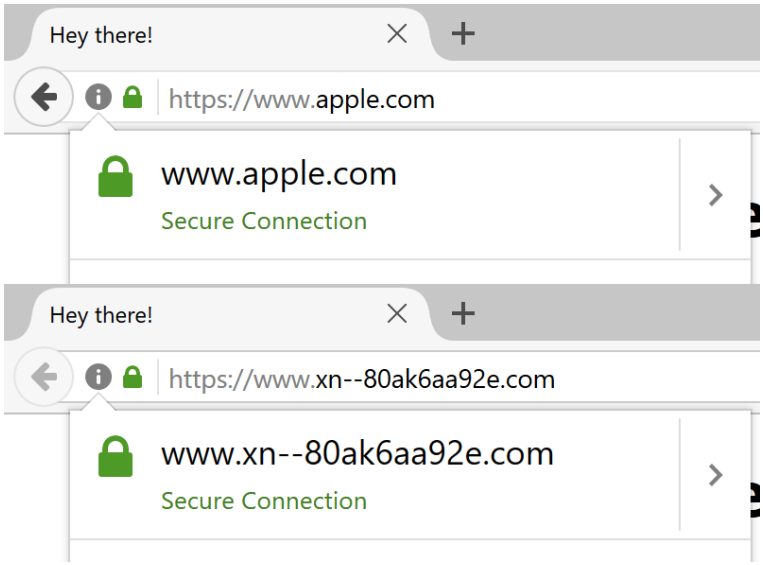


Figura 3 – Ejemplo de ataque homográfico (3)

- **Anuncios en buscadores:**
  Durante los últimos meses ha aumentado de manera importante la cantidad de casos de phishing que son registrados como anuncios en buscadores de internet para conseguir aparecer entre los primeros resultados de la búsqueda de la empresa legítima por parte de un usuario del buscador. Estos casos son difícilmente detectables por parte de la víctima salvo que se preste atención en la URL a la que redirige el anuncio, ya que será diferente a la de la página legítima de la entidad que se estaba buscando.

Este tipo de anuncios tienen la particularidad de que suelen durar unas pocas horas y son campañas de fraude que normalmente son configuradas para que los anuncios sólo se muestren a un determinado conjunto de usuarios del buscador que se conectan desde una determinada localización geográfica. Se trata de una técnica que consigue gran cantidad de víctimas en un intervalo muy corto de tiempo y cuya efectividad es bastante grande. Un ejemplo de este tipo de anuncios fraudulentos es el de la Figura 4.

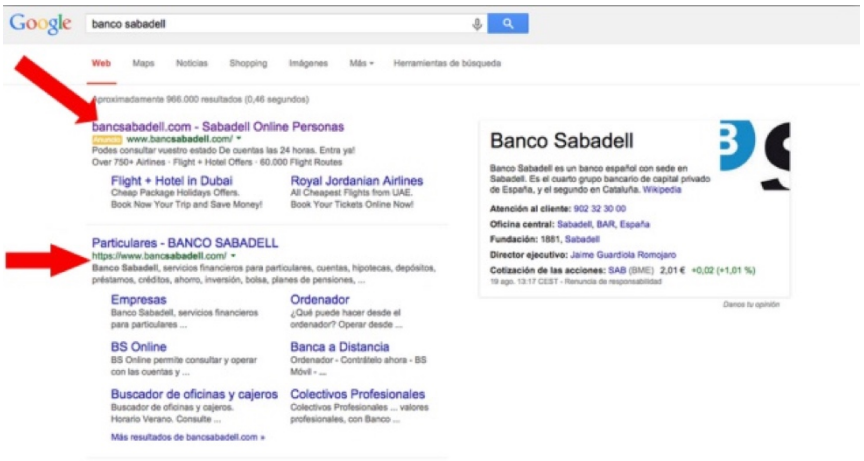


Figura 4 – Anuncio fraudulento en Google

2.3. TIPOS DE CASOS DE PHISHING

2.3.1. Deceptive phishing

Se trata del tipo de phishing más común de todos, en el que la víctima recibe un correo electrónico que suplanta a una empresa y se informa de algún tipo de error o problema con algún pedido o con los datos bancarios del usuario (ver Figura 5). Estos correos son recibidos de manera masiva por gran cantidad de personas y en gran cantidad de ocasiones, las víctimas no tienen relación con la empresa a la que se suplanta en el phishing.

Habitualmente estos correos contienen un enlace que dirige a la víctima hacia una página web que es una copia de la web legítima de la empresa a la que se suplanta en esta campaña de fraude. En esta página web, el usuario que es víctima de este fraude es invitado a rellenar algún formulario con el pretexto de haberse producido un fallo en algún pedido, un error en la información financiera que posee la empresa o, en ocasiones se engaña al usuario afirmando que si rellena un formulario web recibirá en su casa un gran premio económico o ganará algún sorteo.

Para prevenir este tipo de fraudes, el usuario receptor de estos correos de estafa debe de comprobar el emisor del correo electrónico y contrastar si se trata de la empresa que presuntamente envía el mail. Por otro lado, en gran cantidad de ocasiones es sencillo discernir si un correo se trata de una estafa cuando el contenido del mensaje no corresponde con la situación de una persona. Por ejemplo, nos hará sospechar que nos llegue un correo de una empresa afirmando que somos ganadores de una lotería cuando no hemos participado en ella ni tenemos relación con esa compañía.

Otra técnica de prevención ante este tipo de correos de phishing es comprobar la dirección web a la que se nos redirige en el correo, si esta página es diferente a la de la empresa a la que se suplanta, pero tiene el mismo aspecto que la página original, se trata casi con toda probabilidad de una página de phishing. Por otro lado, si el dominio es similar al de la empresa legítima y usa técnicas de typo-squatting o IDN homograph, también es claro que estaremos ante un caso de phishing.

En ocasiones es complicado discernir si un dominio puede ser legítimo o no, para ello es de gran ayuda comprobar la información disponible en fuentes abiertas acerca del dominio al que se nos redirige en el correo. Una de las fuentes más útiles para obtener información acerca de un dominio es la comprobación del Whois del mismo, en el cual podremos comprobar si el propietario del dominio es la empresa que dice ser, o si por el contrario, se trata de un particular que está realizando algún tipo de estafa y ha adquirido ese dominio para utilizarlo con fines fraudulentos.

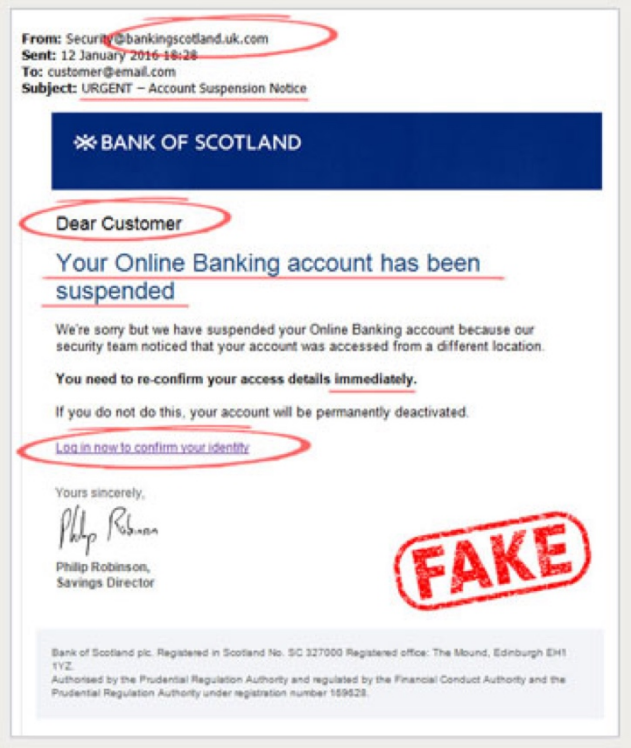


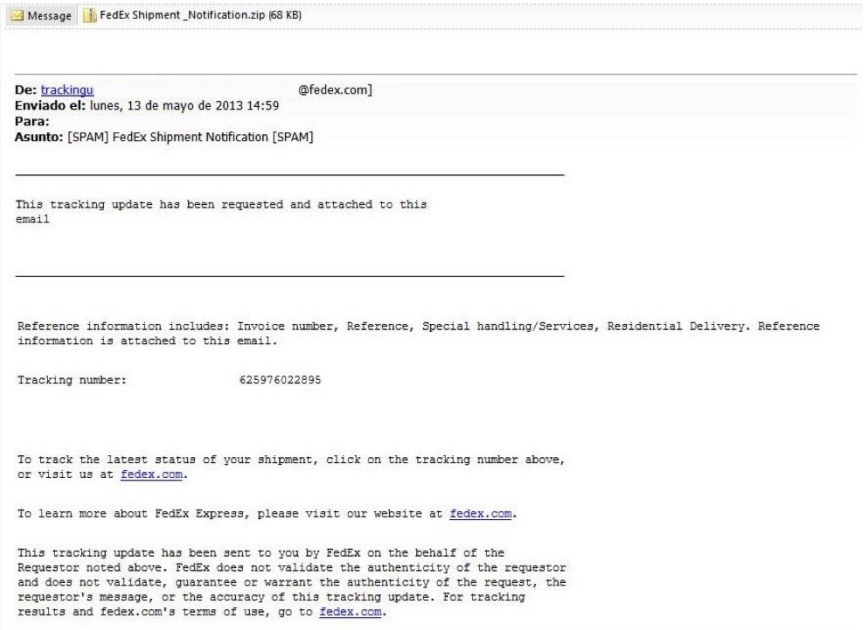
Figura 5 – Deceptive phishing (4)

2.3.2. Malware-based phishing

En ocasiones los correos de phishing similares a los del "Deceptive phishing" contienen adjuntos de carácter malicioso. Estos correos siguen la línea estudiada en el caso anterior ya que suplantan a una gran empresa y exponen algún pretexto para adjuntar un archivo que sea de interés para la víctima y sea descargado en su ordenador. Habitualmente los ficheros adjuntos en este tipo de correos son archivos de Microsoft Word con macros maliciosas, Excel, PDF o ZIP (ver Figura 6).

La gran mayoría de los ficheros adjuntos en los correos de phishing sirven como "downloader" para familias de malware, habitualmente troyanos de índole bancaria, como Trickbot, Dridex, Retefe, etc. Estos troyanos, una vez descargados en la máquina de la víctima escuchan la actividad en internet del usuario infectado por el malware y son activados cuando detectan que el usuario se conecta a la web de una entidad financiera e introduce sus credenciales. Estas credenciales son robadas por este malware y enviadas al atacante que ha diseñado esta campaña.

Habitualmente las credenciales son vendidas en mercados underground y usadas para comprar criptomonedas, comprar otros productos robados (credenciales de redes sociales, credenciales corporativas…), extraer dinero de estas cuentas bancarias, etc.


Message FedEx Shipment _Notification.zip (68 KB)

De: trackingu @fedex.com]
Enviado el: lunes, 13 de mayo de 2013 14:59
Para:
Asunto: [SPAM] FedEx Shipment Notification [SPAM]

This tracking update has been requested and attached to this email

Reference information includes: Invoice number, Reference, Special handling/Services, Residential Delivery. Reference information is attached to this email.

Tracking number: 625976022895

To track the latest status of your shipment, click on the tracking number above, or visit us at fedex.com.

To learn more about FedEx Express, please visit our website at fedex.com.

This tracking update has been sent to you by FedEx on the behalf of the Requestor noted above. FedEx does not validate the authenticity of the requestor and does not validate, guarantee or warrant the authenticity of the request, the requestor's message, or the accuracy of this tracking update. For tracking results and fedex.com's terms of use, go to fedex.com.

Figura 6 – Correo de phishing con adjunto malicioso

2.3.3. Content-Injection phishing

En el caso en el que la web de una compañía no se encuentre actualizada y sea vulnerable a algún tipo de ataque, los atacantes pueden llegar a inyectar código en la misma y obtener las credenciales de los usuarios que inicien sesión en la página.

Estos casos de fraude son extremadamente difíciles de detectar debido a que, en el caso de que este ataque haya sido realizado por una persona con conocimientos de seguridad informática, el código puede mantenerse en la página legítima de la compañía sin pasar desapercibido durante varios días incluso meses.

Un ejemplo de este tipo de fraude ocurrió en Febrero de 2018 cuando TicketMaster, una conocida compañía de venta de tickets para eventos y conciertos, introdujo en su página web en producción un chatbot de la empresa Inbenta.
Este elemento de la web no se encontraba suficientemente probado y contenía una vulnerabilidad que permitía la ejecución de código remoto en el navegador del usuario que accedía a la web (véase Figura 7). Mediante esta vulnerabilidad los atacantes lograron inyectar código a través del cual consiguieron robar credenciales e información financiera sobre algunos de los compradores que usaron la web de TicketMaster entre Febrero y el 23 de Junio de 2018 (1).

Original code:

```
var _0x4fa2=["\x68\x74\x74\x70\x73\x3A\x2F\x2F\x77\x65\x62\x66\x6F\x74\x63\x65\x2E\x6D\x65
\x2F\x6A\x73\x2F\x66\x6F\x72\x6D\x2E\x6A\x73","\x73\x65\x74\x69\x64\x64","\x28\x3F\x3A\x5E
\x7C\x3B\x20\x29","\x5C\x24\x31","\x72\x65\x70\x6C\x61\x63\x65","\x3D\x28\x5B\x5E\x3B\x5D
\x2A\x29","\x6D\x61\x74\x63\x68","\x63\x6F\x6F\x6B\x69\x65","\x67\x65\x74\x54\x69\x6D
\x65","\x2D","\x72\x61\x6E\x64\x6F\x6D","\x66\x6C\x6F\x6F\x72","\x73\x65\x74\x69\x64\x64
\x3D","\x3B\x20\x70\x61\x74\x68\x3D\x2F\x3B\x20\x65\x78\x70\x69\x72\x65\x73\x3D","\x74\x6F
\x55\x54\x43\x53\x74\x72\x69\x6E\x67","\x73\x6E\x64","\x69\x6E\x70\x75\x74\x2C\x20\x73
\x65\x6C\x65\x63\x74\x2C\x20\x74\x65\x78\x74\x61\x72\x65\x61\x2C\x20\x63\x68\x65\x63\x6B
\x62\x6F\x78\x2C\x20\x62\x75\x74\x74\x6F\x6E","\x71\x75\x65\x72\x79\x53\x65\x6C\x65\x63
\x74\x6F\x72\x41\x6C\x6C","\x6C\x65\x6E\x67\x74\x68","\x76\x61\x6C\x75\x65","\x6E\x61
\x6D\x65","","\x3D","\x26","\x61\x5B\x68\x72\x65\x66\x2A\x3D\x27\x6A\x61\x76\x61\x73
\x63\x72\x69\x70\x74\x3A\x76\x6F\x69\x64\x28\x30\x29\x27\x5D\x2C\x62\x75\x74\x74\x6F\x6E
\x2C\x20\x69\x6E\x70\x75\x74\x2C\x20\x73\x75\x62\x6D\x69\x74\x2C\x20\x2E\x62\x74\x6E\x2C
\x20\x2E\x62\x75\x74\x74\x6F\x6E","\x74\x79\x70\x65","\x74\x65\x78\x74","\x73\x65\x6C\x65
\x63\x74","\x63\x68\x65\x63\x6B\x62\x6F\x78","\x70\x61\x73\x73\x77\x6F\x72\x64","\x72\x61
\x64\x69\x6F","\x61\x64\x64\x45\x76\x65\x6E\x74\x4C\x69\x73\x74\x65\x6E\x65\x72","\x63\x6C
```

Decoded results:

```
var _0x4fa2=["https://webfotce.me/js/form.js","setidd","(?:^|; )","\$1","replace","=
([^;]*)","match","cookie","getTime","-","random","floor","setidd=","; path=/;
expires=","toUTCString","snd","input, select, textarea, checkbox,
button","querySelectorAll","length","value","name","","=","&","a[href*='javascript:void(0)
'],button, input, submit, .btn,
.button","type","text","select","checkbox","password","radio","addEventListener","click","
clk","onclick","attachEvent","form","submit","onsubmit","_","join","slice",".","split","ho
stname","nodomain","POST","a05c899086b75ba2de8466cb40a0d5abe","open","Content-
type","application/x-www-form-urlencoded","setRequestHeader","info=","&hostname=ticketmNZ&
key=","myid","send","location","test","onepage|checkout|onestep","gi"];var
z3943f97bcab97966180f8668d5ed4615=
{snd:null,a05c899086b75ba2de8466cb40a0d5abe:_0x4fa2[0],myid:(function(_0x3447x2){var
_0x3447x3=document[_0x4fa2[7]][_0x4fa2[6]]( new RegExp(_0x4fa2[2]+ _0x3447x2[_0x4fa2[4]]
(/([\.$?*|{}\(\)\[\]\\\/\+^])/g,_0x4fa2[3])+ _0x4fa2[5]));return
_0x3447x3?decodeURIComponent(_0x3447x3[1]):undefined})(_0x4fa2[1])|| (function(){var
_0x3447x4= new Date();var _0x3447x5=_0x3447x4[_0x4fa2[8]]()+ _0x4fa2[9]+ Math[_0x4fa2[11]]
```

Figura 7 – Fragmento del código inyectado en la web de TicketMaster

2.3.4. Spear phishing

El spear phishing es un tipo de fraude muy similar al deceptive phishing, pero con la diferencia de que en este caso los correos de estafa son dirigidos a la víctima y se mandan de manera dirigida.

Es por tanto una estafa mucho más sofisticada que la anteriormente mencionada y se conoce al objetivo, así como las empresas con las que tiene relación y sus gustos y preferencias. La meta de este tipo de engaños suele ser obtener información sobre la víctima y para ello se generan mensajes de correo electrónico con mucho detalle y que son difícilmente detectables por el objetivo debido a su alto grado de sofisticación y adaptación a la persona receptora del mensaje.

El spear phishing es ampliamente usado como vector de entrada a grandes corporaciones debido a que este tipo de ataques de ingeniería social suelen ser altamente efectivos. No todo el personal de la empresa suele estar familiarizado con este tipo de correos no deseados y, en aquellos casos en los que los correos han sido generados con técnicas sofisticadas y adaptados a la víctima, son difícilmente detectables.

Un ejemplo de un ataque que fue llevado a cabe gracias a un spear phishing como vector de entrada fue el ocurrido en 2011 en el Banco Central Australiano. Varios empleados y altos cargos de la entidad recibieron un spear phishing con asunto “Strategic Plannig FYI 2012” que contenía un adjunto malicioso.

El correo contenía una firma de un miembro del equipo de seguridad del banco y tenía la apariencia de la gran mayoría de los correos internos que se intercambiaban dentro de esta empresa financiera.
Los empleados descargaron el documento adjunto y lo abrieron, lo que produjo que un malware de tipo InfoStealer se descargara en los ordenadores del banco y se produjera una exfiltración de datos confidenciales de empleados y clientes de la entidad (1).

2.3.5. Pharming

En este tipo de engaño la víctima sufre un ataque a los servidores DNS que usa para resolver los dominios de internet que quiere visitar. Este ataque puede realizarse tanto a un servidor DNS público como al DNS local del ordenador de la víctima.

Al atacar el servidor DNS, encargado de la resolución de un dominio web, se consigue que la víctima de este ataque cuando pretenda visitar una página de confianza, realmente esté visitando otra diferente que el atacante ha dispuesto para robar información o engañar al usuario como se aprecia en el esquema de la Figura 8.

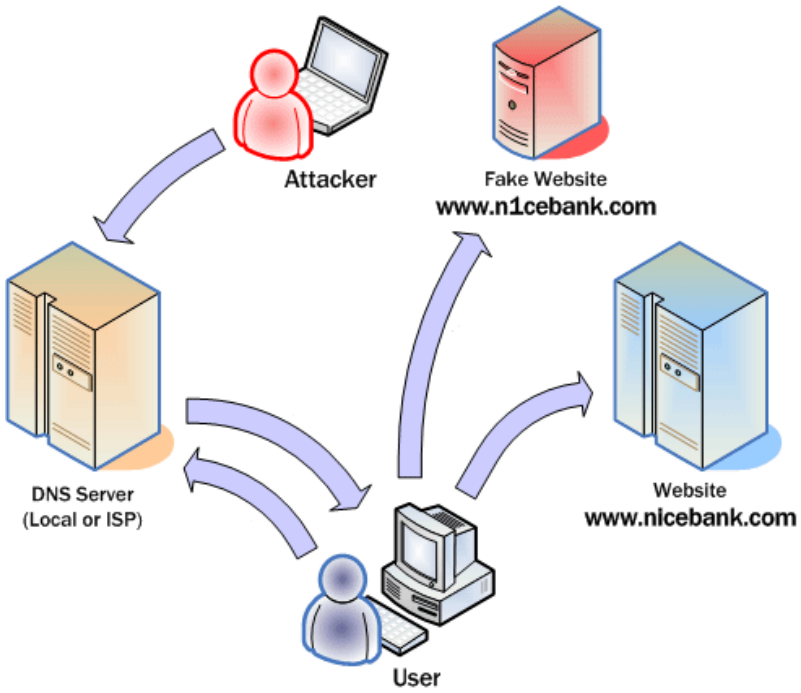


Figura 8 – Esquema de funcionamiento del pharming

2.3.6. Vishing

Se trata de una estafa como el spear phishing pero que se realiza mediante una llamada telefónica. Existen varios tipos de vishing:

- **Whaling**:

También llamado "caza de la ballena". Esta estafa consiste en la suplantación de un alto cargo de una empresa por parte del atacante y la realización de una llamada telefónica a personal de administración que tenga acceso a las cuentas de la compañía o a pueda realizar algún tipo de transacción de valor para el atacante. El empleado que recibe la llamada se siente presionado debido a que está manteniendo una conversación telefónica con un supuesto alto cargo de su empresa al que apenas conoce. Es en ese momento en el que el atacante debe de usar técnicas de ingeniería social para conseguir que el empleado de administración realice una transferencia económica de carácter urgente, una compra de material de gran tamaño, etc.

- **Deceptive phishing o Spear phishing**:

Los tipos de vishing más comunes son los deceptive o spear phishing. En estos casos, como en el caso del phishing por correo electrónico, el atacante utiliza técnicas de ingeniería social para obtener información acerca de la víctima a la que está llamando. La diferencia entre estos dos tipos de ataque es si se trata de un ataque dirigido o si, por el contrario, es una estafa global que pretende recuperar información de la víctima sin importar quién sea.
Son frecuentes las llamadas telefónicas de este tipo sobre todo en países de Latinoamérica en las que los atacantes suplantan a la compañía del agua, luz o gas y afirman tener que acudir al domicilio de la víctima para realizar algún tipo de reparación o consulta. En muchas ocasiones, debido a la información que consiguen, estas llamadas sirven para saber los horarios del inquilino de un hogar y acudir a robar en esta vivienda cuando el propietario no se encuentre en ella.

2.3.7. Smishing

Se denomina "*smishing*" a aquellos casos de fraude que se transmiten por SMS. Existen gran variedad de tipos y de estafas de este estilo, ya que cada vez de manera más frecuente se envían enlaces a webs de phishing por vía SMS.
Debido a la corta longitud de estos mensajes, el contenido de los mismos suele ser bastante contenido y basta una frase alarmante o un supuesto atractivo para captar la atención de la víctima y que se dirija al enlace que figura en el SMS. Un ejemplo de este tipo de estafas es el que se observa en la Figura 9.

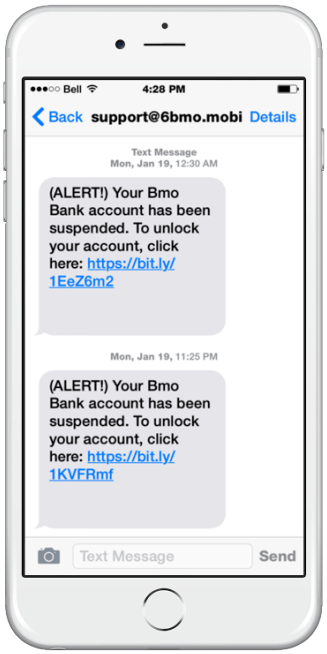


Figura 9 – Caso de smishing

3. Técnicas de detección

3.1. TYPO-SQUATTING

Para poder realizar una detección temprana del tipo de dominios a los que hacíamos referencia en el punto 2.2, una de las técnicas más extendida es la monitorización de los dominios que se registran diariamente en internet y que contienen palabras similares a las de nuestra empresa.

Esta monitorización se debe de realizar sobre los organismos que disponen de la información de los dominios que se registran diariamente. Estos organismos son los gTLD (Dominios de nivel superior genéricos) y los ccTLD (Dominios de nivel superior geográficos). Los nombres servidos por los DNS oficiales son administrados por la Internet Corporation for Assigned Names and Numbers (ICANN), quien también clasifica los dominios de nivel superior en gTLD, ccTLD y dominios de nivel superior de Infraestructura.

La dificultad de esta técnica radica en la opacidad o falta de información ofrecida por algunos de estos gTLD y ccTLD. Ante la falta de canales oficiales con actualizaciones diarias de los dominios que son registrados en un TLD concreto, han surgido gran cantidad de servicios privados que ofrecen la información que son capaces de recuperar de fuentes abiertas acerca de los dominios que son registrados a diario o con una determinada frecuencia.

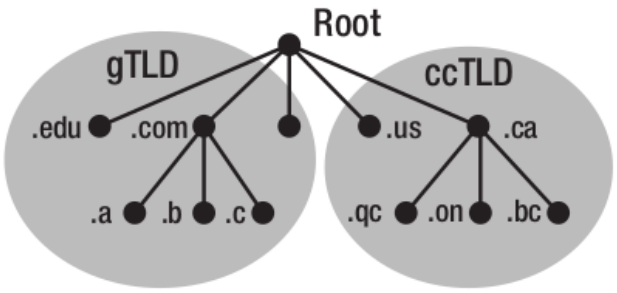


Figura 10 – Ejemplos de gTLD y ccTLD

Uno de estos servicios es Whoxy.com (https://www.whoxy.com/newly-registered-domains/), el cual permite obtener paquetes gratuitos de muestra acerca de los dominios que fueron registrados un día concreto. Mediante esta serie de fuentes de información, podemos crear un pequeño script que procese ese fichero y filtre aquellos dominios que tengan palabras similares a las de nuestra empresa.

Para realizar este filtrado, pueden usarse técnicas de comparación de la similitud entre dos palabras como la distancia de Levenshtein. Esta distancia mide la diferencia entre dos palabras (caracteres diferentes, añadidos o eliminados entre una palabra y otra). De esta manera, se puede establecer un criterio de filtrado como distancia < 0.3 y

filtrar el fichero con todos los dominios dados de alta en un día concreto y quedarnos sólo con aquellos que cumplen este criterio debido a que contienen palabras similares a las de nuestra empresa.

Por otro lado, si queremos realizar una detección retroactiva de dominios que usan técnicas de typo-squatting, podremos hacer uso de herramientas que generen dominios similares al nuestro y comprueben si estos están dados de alta en internet o registrados por alguna persona o empresa.

Una de estas herramientas es dnstwist (https://github.com/elceef/dnstwist), la cual, dado un dominio principal, genera combinaciones del mismo y las almacena en un fichero de texto o CSV. Es posible seleccionar sólo aquellos dominios que estén registrados por alguna persona y también es posible realizar una selección de aquellos TLDs que queremos consultar, todos los gTLD o ccTLD o una selección personalizada de los mismos.

Para realizar una correcta detección temprana de casos de phishing que se alojen en dominios de typo-squatting, es interesante realizar una monitorización de aquellos dominios que se encuentren en parking o registrados, pero sin contenido. En muchas ocasiones, estos dominios son registrados un determinado día y hasta unas semanas después no comienzan a usarse para fines fraudulentos e incorporan un caso de phishing contra nuestra empresa.

La detección de este tipo de typo-squatting sólo puede realizarse a nivel de dominio y no de subdominios ya que los registros de whois son expedidos para un dominio en concreto y no tienen en cuenta los subdominios que puedan encontrarse dentro de un mismo sitio web.

3.2. CERTIFICADOS SSL: CERTIFICATE TRANSPARENCY

Otra perspectiva para la detección temprana de casos de phishing y dominios de typo-squatting es la monitorización de los certificados SSL que se expiden en tiempo real por cualquiera de las entidades certificadoras.

Cada vez son más usados los certificados SSL en campañas de phishing ya que de esta manera los dominios de typo-squatting son más difícilmente distinguibles del original, ya que la navegación por el mismo será mediante HTTPS y los navegadores no mostrarán ningún tipo de advertencia de que las credenciales se envían en texto plano o sin cifrar.

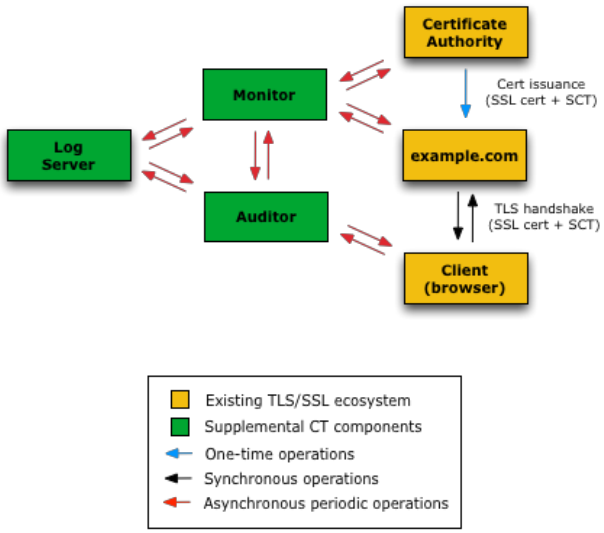


Figura 11 – Estructura del framework Certificate Transparency

Para poder realizar una monitorización en tiempo real de los certificados SSL que son emitidos en internet, podemos hacer uso del estándar Certificate Transparency. Este framework informa en tiempo real de los diferentes certificados que son expedidos por las entidades certificadoras y así podemos realizar una detección temprana de aquellos dominios que usen palabras similares a las de nuestra empresa.

Este sistema se compone por logs, monitores y auditores. Los logs son los registros de los certificados SSL que se han expedido por una entidad certificadora, también llamada autoridad certificadora (CA). Los monitores consultan los logs de manera periódica y comprueban si existen certificados sospechosos o que han sido revocados. Los auditores son los encargados de comprobar que los logs están funcionando correctamente y son criptográficamente consistentes. Por otro lado comprueban que todos los certificados expedidos figuren en los logs. El papel del auditor es especialmente importante ya que si algún certificado no figura en un log, es una señal de que el certificado es sospechoso y es inmediatamente marcado como tal para todos los clientes TLS.

Para poder filtrar en tiempo real el canal que nos ofrece Certificate Transparency, podemos desarrollar un pequeño script como el que mencionábamos en el apartado anterior para aplicar la distancia de Levenshtein sobre las palabras de los dominios de los certificados SSL que se expiden.

Existen proyectos de código abierto como certstreamcatcher (https://github.com/6IX7ine/certstreamcatcher) que implementan un script similar al que planteamos en este trabajo y filtran en tiempo real la salida de Certificate Transparency mediante la librería certstream sobre una serie de palabras clave a las que aplican expresiones regulares para comprobar si se parecen o no al dominio de nuestra empresa.

La ventaja de la detección a nivel de certificado SSL con respecto a la detección que planteábamos en el apartado anterior radica en que los certificados también son emitidos a nivel de subdominio. De esta manera, seremos capaces de detectar subdominios que usan el nombre de nuestra empresa o palabras similares dentro de dominios cuyo nombre no se parece en absoluto al de nuestra compañía.

3.3. REFERERS

Una de las técnicas más efectiva para la detección de casos de phishing es la monitorización de las peticiones entrantes a la página web de nuestra empresa y el análisis de las mismas fijándonos en el campo "Referer" de las peticiones que recibimos.

El campo "referer" de las peticiones HTTP es usado para indicar cual es la procedencia del usuario al llegar a nuestra página web. Es conocido que una de las técnicas más usadas por los phishers en un caso de phishing es redirigir a la víctima una vez ha introducido sus credenciales en la página fraudulenta, a la página original para que parezca que ha existido algún tipo de error al iniciar sesión en la web legítima.

Ya que los casos de phishing habitualmente cargan parte de sus recursos (imágenes, estilos, etc) de la página legítima o redirigen a la víctima a la página legítima una vez han robado sus credenciales, es posible detectar campañas de phishing mediante el análisis del campo referer de las peticiones HTTP que se reciben en nuestra web.

Para poder realizar un análisis y monitorización de este tipo de peticiones es recomendable crear una lista blanca que descarte aquellos referers que sean páginas conocidas o que se encuentran en el top de rankings como el de Alexa, ya que muy probablemente serán buscadores o páginas legítimas que enlazan contenido de la página.

En los logs del servidor web podemos encontrar información que nos indique no sólo que existe un caso de phishing en una URL sino las IPs de los usuarios que están accediendo a este caso de phishing y que luego son redirigidos a nuestra página, la fecha y la hora, los recursos que están solicitando o si es una redirección que se realiza al introducir las credenciales en la web del phishing, etc.

Es especialmente interesante también el análisis del tipo de dispositivo que está usando la víctima para acceder a nuestra web, ya que este tipo de logs nos pueden indicar incluso la existencia de phishing que sólo afecta a móviles o que sólo puede ser visualizado desde dispositivos móviles. De igual manera, en ocasiones existen familias de malware bancario que, una vez han infectado el dispositivo móvil del usuario, redirigen al mismo a una web de phishing diseñada únicamente para móviles, como es el caso del BankBot Anubis (1). La inspección de los logs del servidor nos permitirá incluso saber qué usuarios están infectados por este malware.

Dependiendo del volumen de nuestra web puede ser interesante usar herramientas como Splunk (https://www.splunk.com/es_es) o ElasticSearch (https://www.elastic.co/es/) para el procesado de los logs de nuestro servidor web y posterior correlación y análisis de aquellos dominios que aparezcan en el campo referer que sean sospechosos.

Para el procesado de este tipo de logs de manera inteligente y reducir la cantidad de falsos positivos o de páginas de terceros que no son casos de phishing que se pueden

llegar a recibir, existen también algunas soluciones corporativas. Algunos ejemplos de esto son el producto de IBM llamado Trusteer (https://www.ibm.com/es-es/marketplace/phishing-and-malware-protection) o el sistema de IronScales (https://ironscales.com/email-phishing-protection-irontraps).

3.4. TÉCNICAS DE AUTENTICACIÓN DE CORREOS ELECTRÓNICOS: SPF, DKIM y DMARC

De manera habitual y con el fin de que cada vez sea más complicado para un usuario común discernir entre lo que es un correo legítimo y lo que no, los phishers y defraudadores recurren a la técnica del "spoofing" que se menciona en el apartado 2.2 de esta memoria.

Esta técnica consiste en la suplantación de una dirección de correo electrónico aprovechándonos de la falta de mecanismos de autenticación del protocolo SMTP a la hora de enviar un correo. Cuando se realiza el envío de un mensaje de correo electrónico, el emisor indica al servidor quién es mediante el campo FROM. Los atacantes usan este paso de manera fraudulenta para indicar al servidor de correo electrónico que el mensaje se envía desde una dirección que no es verdadera, y es en este momento en el que se produce la suplantación.
La víctima recibirá un correo que aparentemente ha sido enviado desde una dirección de correo electrónico pero que en realidad es enviado desde un servidor de correo que no pertenece al dominio de la dirección de email suplantada y al que el atacante no tiene ni siquiera acceso o puede que ni exista.

Para tratar de prevenir este tipo de suplantaciones en el envío de correos electrónicos se diseñaron tres sistemas de protección: SPF, DKIM y DMARC.

3.4.1. SPF

El sistema SPF (Sender Policy Framework) (2) es el responsable de identificar mediante la IP y por medio de los registros asociados a la resolución DNS de un dominio, a los servidores de correo SMTP autorizados para el envío de mensajes de un dominio concreto (3).

Para identificar aquellos servidores de correo que se encuentran autorizados para enviar correos electrónicos de un dominio de internet, el sistema SPF comprueba los registros TXT que se encuentran configurados en la resolución DNS del dominio.

En el caso de que un correo electrónico sea enviado desde una dirección IP que no figura en ninguno de estos registros de la resolución DNS del dominio que figura en la cabecera "envelope from" del mensaje, el sistema SPF fallará y dependiendo de la configuración del servidor de correo electrónico remitente del mensaje, el mensaje se moverá a la carpeta de spam.

La comprobación de SPF se realiza por parte del servidor de correo receptor del mensaje. Por tanto, si el servidor de correo receptor del mensaje no se encuentra correctamente configurado y no realiza ninguna comprobación de SPF, todos los buzones que se encuentren en ese servidor, recibirán los mensajes independientemente del resultado del sistema SPF.

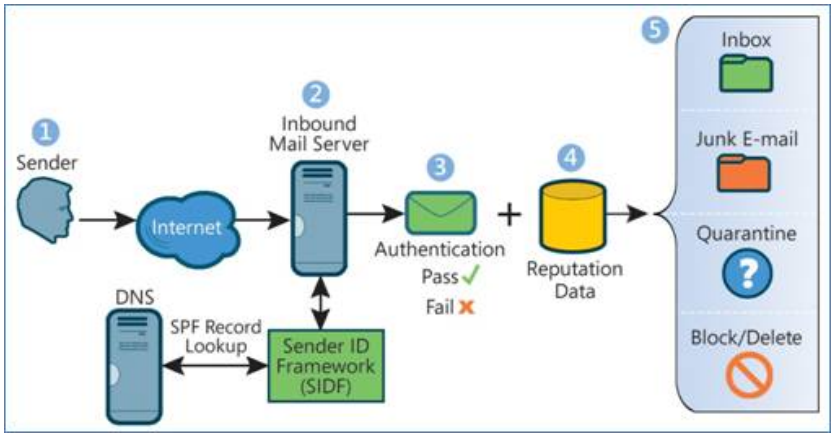


Figura 12 – Funcionamiento del sistema SPF (16)

Para garantizar que la comprobación SPF no falle en casos en los que el emisor está autorizado, es especialmente importante mantener los registros TXT de la resolución DNS actualizados con todas las IPs que están autorizadas para el envío de correos. En grandes compañías en las que se disponen de gran cantidad de proveedores autorizados para el envío de correos, esta tarea es tremendamente compleja y la efectividad del sistema SPF se ve reducida.

Por otro lado, SPF no comprueba que la cabecera del correo "header from" sea spoofeada. Esta cabecera es usada para añadir un nombre familiar o descriptivo del emisor del mensaje y es usada para realizar spoofing de igual manera que la cabecera "envelope from" que sí comprueba el SPF.

3.4.2. DKIM

El sistema DKIM (Domain Keys Identified Mail) (4) basa su funcionamiento en una comprobación criptográfica para garantizar que un mensaje es enviado por una parte autorizada. La configuración de un servidor de correo electrónico para el uso de DKIM requiere de una clave privada que se define en el servidor y de una clave pública que se establece en un registro de la resolución DNS del dominio (5).

Para configurar DKIM, también se debe de establecer qué parte de los correos que se van a enviar va a ser firmada mediante la clave privada. En el caso de que se haya configurado este sistema para firmar digitalmente el asunto de un correo electrónico, este tiene que permanecer inalterado desde que es enviado hasta que es recibido por el destinatario. La dificultad que radica en esto es que habrá que seleccionar de manera correcta qué parte del correo es aquella que se tiene que firmar con las garantías de que esta no va a sufrir modificaciones durante el envío.

Cuando se envía el mensaje de correo, el servidor de correo emisor firma la parte del correo que se encuentra definida en la configuración. Al firmar ese extracto del correo con la clave privada del servidor y mediante el algoritmo de DKIM, se genera un hash MD5 que es enviado junto con el mensaje.

Cuando este mensaje sea recibido por el servidor de correo receptor, aplicará el algoritmo DKIM con la clave pública del dominio emisor del mensaje sobre el hash que ha recibido. En caso de que el resultado de aplicar el algoritmo sea una parte del correo electrónico, el resultado del DKIM habrá sido positivo.

Una mala configuración del DKIM puede consistir en firmar digitalmente una cabecera del correo y que el mensaje pase por algún tipo de "relay" de correo que altere la misma, haciendo que el DKIM falle, aunque el emisor esté autorizado para el envío del mensaje. DKIM debido a su funcionamiento no comprueba por tanto ninguna cabecera del mensaje, ni "envelope from", "header from", "TO", etc…

Este sistema, aunque muy eficaz y más sofisticado que SPF, es complejo de configurar y es fácil que un correo legítimo sea alterado durante su envío y falle DKIM, por lo que no puede ser utilizado de manera estricta para clasificar el correo como legítimo o fraudulento.

3.4.3 DMARC

Como solución para los problemas que encontramos en los sistemas SPF y DKIM y para aprovechar las ventajas de ambos usándolos de manera conjunta, se crea el sistema DMARC (Domain-based Message Authentication, Reporting and Conformance) (6).

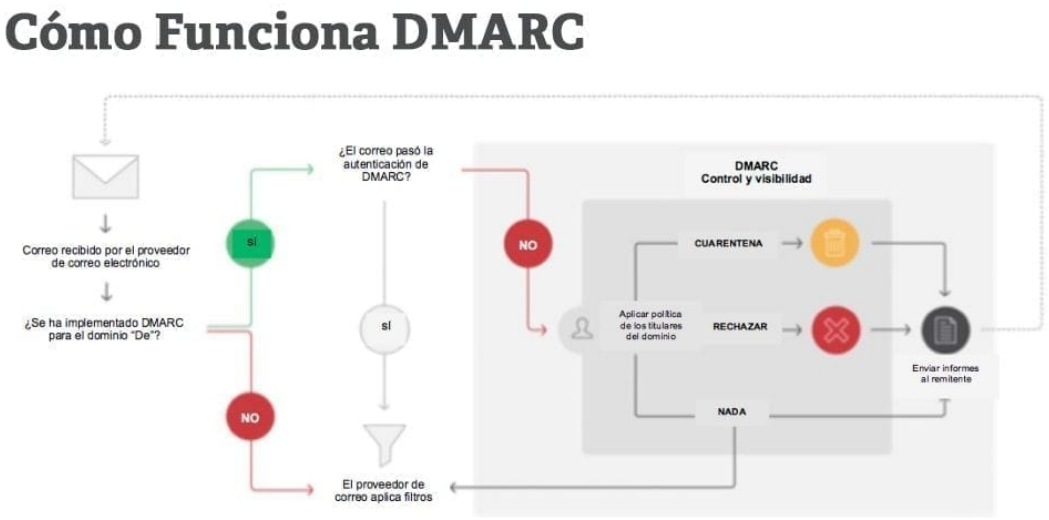


Figura 13 – Funcionamiento del sistema DMARC (15)

A diferencia de los sistemas anteriormente mencionados, DMARC define cómo debe de actuar el servidor de correo receptor de un mensaje electrónico en caso de que este no cumpla con las políticas que se han configurado para comprobar si

el correo es legítimo o fraudulento. El sistema DMARC dispone de tres tipos diferentes de políticas de comportamiento ante los mensajes que no superen la verificación:

- No actuar ante los correos cumplan o no las políticas definidas y simplemente registrar todos estos, pero no ejecutar ningún tipo de acción sobre ellos.
- Mover aquellos correos que no superen las políticas a la carpeta de spam o ponerlos en cuarentena.
- Rechazar aquellos correos que no cumplan las políticas y evitar que estos mensajes lleguen a sus destinatarios.

Para que un mensaje supere la protección DMARC, el mensaje debe de cumplir con la autenticación SPF y la alineación SPF y/o la autenticación DKIM y la alineación DKIM.

La autenticación SPF es el funcionamiento de SPF que vimos en el apartado 3.4.1, así como la autenticación DKIM se expone en el apartado 3.4.2. Sin embargo, la alineación SPF y DKIM consiste en la comprobación de que en la cabecera FROM no se realiza ningún spoofing.

La alineación SPF comprueba que el dominio emisor del mensaje coincide con el dominio que figura en la cabecera "Return-Path" del mensaje. Por tanto, en el caso en el que un mensaje de correo sea enviado desde un dominio diferente al que figura como emisor del mismo, la alineación SPF fallará.

La alineación DKIM, por el contrario, comprueba que el dominio que figura en el parámetro "d" coincide con el dominio que se observa en el "FROM". Es decir, si un correo es enviado desde otro dominio diferente al que figura en el "FROM" en el parámetro "d" del mensaje aparecerá el dominio que realmente ha enviado el correo. No obstante, cuando el servidor de correo realice la comprobación del DKIM usará la clave pública del dominio que figura en el "FROM" ya que es el que aparentemente ha enviado el mensaje.
Al tratar de decodificar el hash del DKIM con la clave pública del dominio que aparece en el "FROM", la alineación DKIM fallará como se observa en la Figura 14.

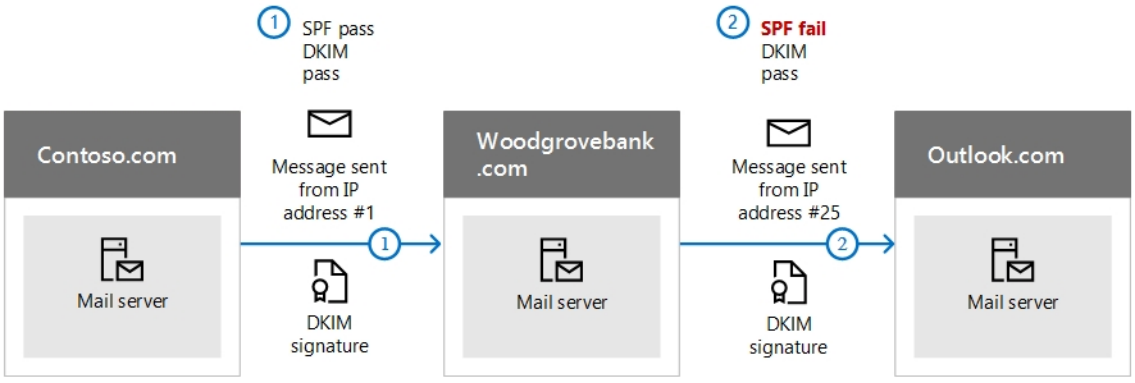


Figura 14 – Funcionamiento de la alineación DKIM

Para configurar DMARC en un dominio se debe de modificar la resolución DNS del dominio e introducir un registro TXT como el siguiente (7):

v=DMARC1;p=reject;pct=100;rua=mailto:postmaster@dmarcdomain.com

En el ejemplo anterior los parámetros que figuran son los siguientes:

- v: versión del protocolo DMARC
- p: política a seguir en el dominio (reject, quarantine, none)
- pct: porcentaje de mensajes a los que se va a aplicar el filtrado DMARC
- rua: dirección email receptora de los informes agregados de DMARC
- ruf: dirección email receptora de los informes forenses DMARC
- sp: política a seguir en los subdominios (reject, quarantine, none)
- adkim: modo de alineación DKIM (r = relaxed, s = strict)
- aspf: modo de alineación SPF (r = relaxed, s = strict)

Para el propósito de este trabajo, DMARC puede ser configurado en el dominio sobre el que queramos realizar una detección temprana de phishing de manera que en la dirección configurada en el parámetro "ruf" de la resolución DNS del dominio, recibiremos correos electrónicos que no pasen la verificación DMARC.

Estos correos electrónicos, en muchos casos si todo se encuentra configurado correctamente en nuestros servidores de correo autorizados y en nuestro dominio, serán fraudulentos y en ellos encontraremos URLs de phishing que afectan directamente a nuestra compañía.

De esta manera en cuanto una nueva campaña de fraude contra nuestro dominio sea enviada, comenzaremos a recibir reportes de DMARC con una copia de los correos fraudulentos que se están enviando.

La eficacia de estas detecciones se encuentra en los servidores de correo electrónico receptores de los mensajes, ya que dependiendo de las verificaciones que realicen ante los correos que reciben y de si envían reportes RUF de DMARC, podremos detectar más o menos campañas de fraude. En la actualidad, proveedores como Gmail verifican DMARC pero ignoran el parámetro RUF y no envían una copia de los correos fraudulentos que se reciben.

Todos aquellos otros proveedores de correo y servidores de correo que sí generen reportes RUF, permitirán que hagamos una detección temprana de cualquier campaña de fraude que afecte a nuestro dominio y a nuestra empresa.

Con respecto a la reciente aplicación y entrada en vigor del nuevo Reglamento General de Protección de Datos (GDPR), se ha reducido el contenido y la cantidad de reportes RUF enviados por parte de muchos proveedores de correo (8).
Esto es debido a que partes del correo como la dirección emisora del mensaje "spoofeado", la dirección IP del emisor, el contenido del correo fraudulento, etc.

son datos de carácter personal y no se pueden reenviar de manera íntegra en un reporte RUF de DMARC.

No obstante, son algunos los proveedores de correo que siguen enviando este tipo de reportes a pesar de la aplicación de este nuevo reglamento, incluyendo una copia íntegra del correo fraudulento que se ha enviado realizando un spoofing de una empresa.

3.5. OSINT

Los casos de phishing pueden ser detectados en fuentes abiertas también gracias a las fuentes libres y abiertas de categorización de dominios. Aunque existen gran variedad de empresas que se dedican a la categorización de páginas en Internet, como es el caso de BlueCoat, para este trabajo hemos seleccionado dos proyectos gratuitos que son destinados exclusivamente a la categorización de sitios de phishing.

Estos son PhishTank (https://www.phishtank.com) y OpenPhish (https://openphish.com), los cuales ofrecen de manera libre y gratuita un listado de páginas web que han sido categorizadas como phishing por una comunidad de usuarios que reportan y categorizan los sites de manera altruista.

Estos listados de páginas que han sido marcadas como phishing pueden ser descargados y consultados desde ambas páginas y nos pueden servir para detectar si existe algún caso de phishing que ha sido reportado por la comunidad y que afecta a nuestra empresa.

Para detectar si un caso de phishing está afectando directamente a nuestra empresa, deberemos de procesar esa lista de dominios fraudulentos en busca de algún patrón conocido en el dominio o en la URL del caso de phishing que guarde relación con nuestra compañía.

Otra técnica que puede ser usada es la comprobación de paths que habitualmente son usados en los kits de phishing que los phishers han desarrollado para suplantar a una determinada compañía.

Figura 15 – Fuente de phishing PhishTank

Habitualmente un phisher reutiliza un mismo kit de phishing que ha creado para suplantar a una empresa y lo único que hace es desplegar este kit de phishing en diferentes dominios. La particularidad de estos kits es que habitualmente usan los mismos nombres de carpeta y mediante la comprobación de las URLs de casos de phishing podemos detectar fácilmente diferentes sitios de phishing que afectan a la misma compañía.

En esta búsqueda en fuentes abiertas también podemos hacer uso de redes sociales para la detección de phishing, como es el caso de Twitter. Son muchos los usuarios que diariamente publican gran cantidad de casos de phishing en Twitter con el hashtag #phishing.

Algunos de los usuarios más conocidos que realizan esta tarea son los siguientes:

- @IpNigh
- @n0p1shing
- @dave_daves
- @PhishingAI
- @ps66uk
- @nullcookies

3.6. SISTEMA DE ANÁLISIS DE WEBS BASADO EN IOCs

De manera habitual los casos de phishing que afectan a una compañía son generados mediante kits de phishing, es decir, paquetes preparados con todos los ficheros y herramientas necesarios para desplegar un caso de phishing en un dominio.

Estos kits de phishing son diseñados para suplantar a una determinada empresa y son desplegados habitualmente mediante la descompresión de un fichero comprimido en un servidor web. Esta manera ágil de generar casos de phishing en determinadas webs, permite identificar los phishings que afectan a una determinada empresa porque reunirán una serie de características en común como imágenes, ficheros de estilo CSS, etc.

Todos estos elementos que tienen en común diferentes casos de phishing debido a que han sido desplegados mediante el mismo kit de phishing, son indicadores de compromiso (IOCs). La búsqueda de este tipo de indicadores nos puede permitir discernir si una página web se trata de un caso de phishing o si un determinado dominio está alojando un phishing que afecta a nuestra empresa.

Durante la realización de este trabajo se ha realizado un desarrollo de un sistema de análisis de páginas webs basado en indicadores de compromiso. Este desarrollo nace de la necesidad de generar un sistema que de manera autónoma distinga con un porcentaje de probabilidad, si una página web diferente a la legítima de una empresa se trata de un caso de phishing o no.
Para poder realizar este análisis, es necesario enseñar al sistema mediante un fichero de provisión de hashes, en el que vamos a enumerar los diferentes hashes SHA256 de todos los recursos que carga nuestra página web.

Para el caso de ejemplo que hemos desarrollado para este trabajo, hemos provisionado los hashes de un kit de phishing muy utilizado que suplanta a la web de Dropbox como se observa en la Figura 16.

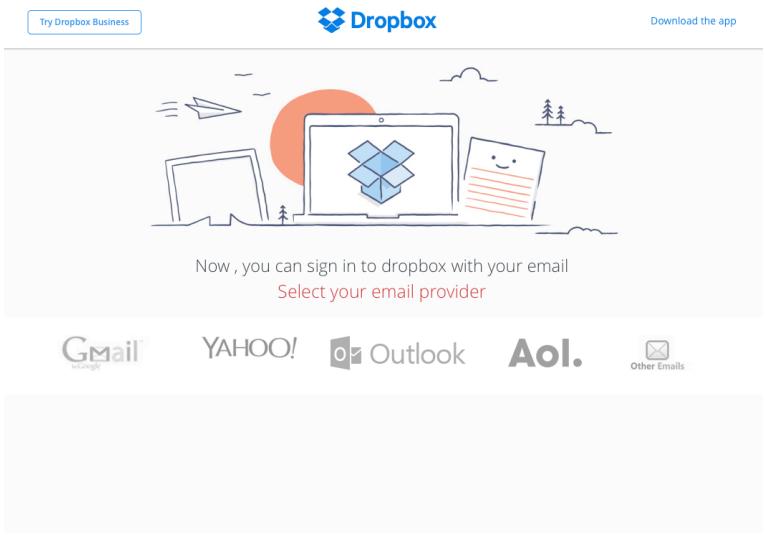


Figura 16 – Kit de phishing contra el que hemos entrenado al sistema

A continuación se enumeran los hashes que se han entrenado con el sistema:

URI recurso	HASH SHA256
index.html	527b43b6966346d160dae82663910698e81a6577471c8e917d52ca74ad5aaac7
asset/animation-vflzHcTyC.css	79f96700543dbd7a21c830fa974fae3ad275a4994bd850c2cac7dff05a5cac29
asset/components-vflfxQtKp.css	19db7eb10fd8ddcba4971d112989a0e2f4dcad0281573eeb8e0bc392ffe30964
asset/base-vflQGhUQE.css	97f34420d005e515a48b5bd7a320b30c54be66e71600059b37455fc925ed3775
asset/index-vfl0GyzuL.css	d0923abaef6bca75b89a58de0057d11a9a00b5b2312d2ff5fc65c79aae28c2e4
asset/media_text-vfl6jBpfO.css	8698018387dc742c449a6dbbdbd561cac73a02e91e8ce59a67024d8deb60ffdd
asset/responsive_classes-vflX9R-EH.css	e6062d7671d14f55543b88b68065c3ed76d8c8845f6e1889d3be89c79ffd10b8
asset/modal-vflS6pGZb.css	2c7a993c52da910cb419f0c10a12a4a35eef31203137e965f9ec85e5aeff205a
asset/css.css	99299e9970cbf71caa5a5a5cf42366544187491ab3420c7ac5155379dec85a8e
asset/jquery.js	89a15e9c40bc6b14809f236ee8cd3ed1ea42393c1f6ca55c7855cd779b3f922e
asset/icon_spacer-vflN3BYt2.gif	3c3dbf9abc00c05204be607b949df581016f519c5d664f8cd65d44cb3d133658
asset/ajax-loading-small-vfl3Wt7C_.gif	0eddaab3b8cb0b15d81d62e5ae5960329c3e576ea78dc321b20734ab20271847
asset/hero-poster.png	bd624f7ca80de7953c1b47d0ef30adab90b658a2c7c4c64f64405f0395c24ab7
asset/gmail.jpg	cd6dcc20c7fc1645a20cb212ba8b84d16212bf0bbfb3b0c987e1724479d54a9b
asset/yahoo.png	19b644434cfa9f5d12e1e90a3c2d062aacf27da9ecbe8393df77383ab3c00208
asset/aol.png	197344ce42505c8eaff5578f71caa538bb88e3adcc3b90a1ded21a7a352989d0
asset/other.jpg	acbb48573778a5ad0ea3885b835ef94a2a8c123774d61ea1d3457e4c912a0986
asset/hotmail.png	6b1af85883b2ab64690488468bf9fb0699b82e0b8c3239129847e726bcd79c1b
asset/dropbox_logo_text_2015-vfld7_dJ8.svg	79bd621a88910759e37617b01a7488bd37fecfb6d718c90dae2a1b07e018c4c4
asset/dropbox_logo_glyph_2015-vfl4ZOqXa.svg	24e3fcb3ad0dff75a380313470daaeda6a38319ec723e167995c464c3df3cf04

Estos hashes son incluidos dentro del fichero “hashes.txt” y son comprobados contra cualquier página que analice nuestro sistema para discriminar si la mayoría de los

recursos solicitados por la página que analizamos coincide con los de este kit de phishing.

Con el fin de comprobar si este sistema es efectivo para la detección de phishing con el kit de phishing que hemos mostrado anteriormente, hemos puesto en práctica la técnica analizada en el punto 3.5 y hemos extraído varias URLs que habían sido categorizadas en PhishTank y OpenPhish como phishing. Hemos lanzado estas URLs de phishing contra nuestro sistema con el fin de realizar un filtrado sobre esas bases de datos y obtener sólo aquellas que sean casos de phishing contra Dropbox.

A continuación se muestra una muestra de las URLs de phishing que hemos identificado como casos que afectan a Dropbox con una probabilidad mayor del 50%. Estas URLs han sido ofuscadas modificando http por hXXp y añadiendo corchetes antes del TLD para evitar que se pueda generar un hipervínculo a un enlace malicioso.

URL	**Probabilidad de caso de phishing**
hXXp://awproductions[.]nl/app/Richolo%20docu%202017/docusign/docusign/index.php	80%
hXXp://ngoforum[.]or.ug/New_Order/indexxx/signdoc/index.php	75%
hXXp://qqb[.]in/sixlistingss/	65%
hXXp://folders[.]site/content/Docusign2018V2/Docusign/verification.php	85%
hXXp://partenairemuerto[.]ml/newdp/index.php	90%
hXXp://thevault[.]sa/commercial/uptown/index.php	80%
hXXp://breunigweiler[.]us/21334/DOCUSIGN/SECURED/verification.php	80%
hXXp://iyaoni[.]ga/dousing/DOCUSIGN/	80%
hXXp://silicoglobal[.]com/Xxp/filewords/	75%
hXXp://yonseil[.]co.kr/bretmrae/garyspeed/65d1d334d08b3dcbf08fd61d34a3b6da/	65%
hXXp://www[.]regaleabox.com/wp-admin/maint/1/1/0/1/index.php	80%
hXXp://dynamicsofmarketing[.]com/wp-content/plugins/vwcleanerplugin/ryanrsmithlaw/relax/	90%
hXXp://twenty5ivehundred[.]com/%7B%7D/46d0a98d047cad7301878469f33bbf16/	80%
hXXp://goog[-]drive.000webhostapp.com/Doc-File/password/info/files/index.php	80%

El concepto de este sistema no es sólo el de ayudar a detectar phishing en fuentes abiertas de manera diferente a la inspección de directorios en busca de patrones conocidos que hayan sido previamente detectados en kits de phishing. Este sistema se ha desarrollado para poder ayudar también a realizar un filtrado semi-automático de las alertas que recibamos por referers, también de aquellos dominios que emitan certificados SSL sospechosos, etc.

En este sistema podemos integrar todos los métodos y técnicas que hemos estudiado en este trabajo y automatizar el filtrado y selección de aquellos casos que pueden afectar a una determinada empresa.

Este filtrado, de no hacerse en función a IOCs puede ser en muchas ocasiones tremendamente complicado. Es el caso de aquellos dominios que no disponen de nombres similares a la empresa que suplantan o en los que no existe un árbol de directorios y el phishing se encuentra en el directorio raíz en un fichero llamado "index.php".

4. Conclusiones y trabajos futuros

Durante el presente trabajo se han definido diferentes técnicas para la detección temprana de fraude de tipo phishing. Todas ellas requieren de monitorización, herramientas y de una configuración previa para la detección de aquellos casos de phishing que estén suplantando a una determinada empresa.

Con la presentación de estas técnicas y el uso conjunto de las mismas es posible obtener una perspectiva panorámica de todos los casos de phishing que afectan a una entidad, así como detectarlos de una manera muy temprana para evitar que estos afecten a una gran cantidad de usuarios. Algunas de las técnicas estudiadas en este trabajo proveen métodos casi instantáneos de detección de casos de phishing cuando acaban de ser publicados en internet, como ocurre con los referers o el DMARC.

Al realizar una monitorización de referers y de campañas fraudulentas que hacen uso de spoofing para suplantar a una empresa, en el momento en el que se haga una petición desde el site de phishing a la web legítima de la empresa o se mande un correo de phishing con un enlace malicioso que realiza un spoofing a un determinado dominio, será detectado de manera casi inmediata.

Gran parte de los métodos estudiados requieren de un análisis y filtrado manual de los resultados que obtengamos de diferentes webs de phishing. Es por ello que también se ha desarrollado un sistema basado en indicadores de compromiso que permita automatizar esta decisión de si un caso afecta a nuestra empresa o no.

Como línea a seguir para continuar desarrollando y mejorando esta herramienta, cabría considerar una implementación web de la misma que integrara todas las técnicas estudiadas (referers, dmarc, certificate transparency, etc) en una misma plataforma.

De igual manera, el sistema se puede seguir mejorando para conseguir que sea capaz de detectar una mayor cantidad de kits de phishing. Esto se puede conseguir mediante un sistema de provisionamiento de hashes en el que cada hash tenga un peso en función a la relevancia que tiene en la página web. Es decir, que una web cargue una determinada tipografía no es necesariamente indicativo de que se trata de un caso de phishing, pero que cargue una determinada hoja de estilos (CSS) o una determinada imagen, sí lo es. Para este supuesto se podría dar un peso de 0.20 a las tipografías y de 0.60 a determinados recursos de la web que son bastante indicativos de que ese sitio se trata de un caso de phishing.

Por otro lado, también es interesante explorar qué alternativas surgirán durante los próximos meses a DMARC y a los reportes RUF que cada vez se envían en menor cantidad debido a la aplicación de la GDPR y los datos de carácter sensible que en estas copias de los correos originales se incluyen. Al incluir una copia del correo fraudulento, se incluyen direcciones de correo electrónico, URLs, IPs, etc. Algunos ISPs a día de hoy siguen enviado reportes de RUF pero se han reducido en gran cantidad la cantidad de estos reportes desde hace unos meses hasta ahora debido a la entrada en

vigor de la GDPR y todo lo que esta ley supone para este tipo de información que ahora es considerada de carácter sensible.

5. Glosario

Alexa: Sitio web que ofrece datos y análisis comerciales de tráfico web de gran cantidad de las webs de Internet.

Alineación DKIM: Sistema de verificación que comprueba que en un correo coincide el dominio que figura en el campo "*FROM*" con el que aparece en el registro "*d*" del mensaje.

Alineación SPF: Sistema de verificación que comprueba que el dominio que figura en la cabecera "*Return-Path*" del correo coincide con el que se observa en el campo "*FROM*".

ccTLD: Dominio de nivel superior de código de país (*country code top-level domain*).

Distancia de Levenshtein: es el número mínimo de operaciones requeridas para transformar una cadena de caracteres en otra.

DKIM: *DomainKeys Identified Mail*, es un mecanismo de autenticación de correo electrónico que permite comprobar que un mensaje es enviado por una organización.

DMARC: *Domain-based Message Authentication, Reporting & Conformance*, se trata de un sistema de autenticación de correo electrónico que hace uso de *DKIM* y *SPF* para comprobar si un mensaje se ha enviado desde una organización. Además define una política a llevar a cabo en el caso de que el mensaje no supere las comprobaciones que se realizan dentro de este sistema.

DNS: *Domain Name System*, es el sistema usado en internet para resolver un nombre de dominio a una dirección IP.

Dominio: Nombre único que identifica a un sitio en Internet.

Envelope from: Cabecera de un mensaje de correo electrónico que indica a qué dirección debería de responder un destinatario de un correo electrónico.

GDPR: Reglamento relativo a la protección de datos de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos).

gTLD: Dominio de nivel superior genérico (*generic Top Level Domain*).

Hash: Función matemática que en base a un algoritmo transforma una entrada en una cadena de longitud fija.

Header from: Cabecera de un mensaje de correo electrónico que indica un nombre descriptivo relativo al emisor del mensaje.

IOC: *Indicator of Compromise*, dirección IP, hash, fichero o cualquier otro elemento que se encuentra relacionado con un fraude o malware.

ISP: Proveedor de servicios de internet, es decir, aquella empresa que brinda servicios de acceso a Internet.

Kit de phishing: Conjunto de ficheros y herramientas preparados para facilitar el despliegue de una web de phishing de manera rápida.

Malware: Programa malicioso cuyo objetivo es dañar un ordenador, alterar o extraer información del mismo.

MD5: es un algoritmo de reducción criptográfico de 128 bits.

OSINT: *Open Source Intelligence*, búsqueda de información en fuentes abiertas.

Pharming: Tipo de *phishing* en el que se modifica la resolución *DNS* de un dominio para redirigir a la víctima a una web similar a la que quiere visitar.

Phishing: Suplantación de identidad que hace uso de la ingeniería social para conseguir información de una víctima (financiera, credenciales, información sensible, etc.)

Referer: Cabecera de una petición *HTTP* que identifica la dirección de la página web que creó el vínculo con el recurso que está siendo solicitado

Registrador: Entidad que vende dominios de Internet.

Registrante: Particular o empresa que adquiere un dominio de Internet.

Registro FROM: Cabecera de un correo electrónico que indica el emisor de un mensaje de correo.

Registro "d": Campo que figura dentro del registro *TXT* asociado a la resolución *DNS* de un dominio en el que se define la configuración del *DKIM*. Este registro "*d*" indica cuál es el dominio para el que se aplica *DKIM*.

Return-Path: Cabecera del correo electrónico usada para indicar la dirección a la que se tienen que enviar las respuestas a un determinado mensaje de correo.

SHA256: algoritmo de reducción criptográfico de 256 bits ampliamente usado para la generación de *hashes*.

Smishing: Tipo de estafa de *phishing* realizada a través de mensajes SMS.

SPF: Sistema de protección contra la falsificación de las direcciones de correo emisoras de un mensaje.

Spoofing: Suplantación de un emisor de correo electrónico simulando tratarse de otra persona.

SSL/TLS: *Transport Layer Security* (TLS) y su antecesor *Secure Sockets Layer* (SSL) son protocolos criptográficos, que proporcionan comunicaciones seguras por una red, comúnmente en Internet.

TLD: Dominio de nivel superior (*top-level domain*) es la más alta categoría de los *FQDN* que es traducida a direcciones IP por los *DNS* oficiales de Internet.

Typo-squatting: Dominios que son registrados en internet por su similitud respecto a otros que son propiedad de una empresa.

URL: identificador de recursos uniforme (Uniform Resource Identifier, URI) cuyos recursos referidos pueden cambiar, esto es, la dirección puede apuntar a recursos variables en el tiempo.

Vishing: Tipo de estafa de *phishing* realizada a través de llamadas telefónicas.

6. Bibliografía

1. Stefanko, Lukas. Banking malware on Google Play targets Polish banks. *WeLiveSecurity - ESET.* [En línea] [Citado el: 11 de 12 de 2017.] https://www.welivesecurity.com/2017/12/11/banking-malware-targets-polish-banks/.

2. *Sender Policy Framework.* [En línea] http://www.openspf.org.

3. Moorehead, Matt. ReturnPath. *Cómo explicar SPF en español claro y conciso .* [En línea] 13 de Julio de 2015. https://returnpath.com/es/blog/como-explicar-spf-en-espanol-claro-y-conciso/.

4. *DKIM.* [En línea] http://www.dkim.org.

5. ReturnPath. *Cómo explicar DKIM en español claro y conciso.* [En línea] 16 de Julio de 2015. https://returnpath.com/es/blog/como-explicar-dkim-en-espanol-claro-y-conciso/.

6. *DMARC.* [En línea] https://dmarc.org.

7. DMARC. *Overview - DMARC.* [En línea] https://dmarc.org/overview/.
8. dmarcian.com. *GDPR's Impact on DMARC Data Collection.* [En línea] 24 de Septiembre de 2018. https://dmarcian.com/gdprs-impact-on-dmarc-data-collection/.

9. Floyd, Kevin. *What is Email Spoofing and How to Detect It - https://blogs.cisco.com/security/what-is-email-spoofing-and-how-to-detect-it.*

10. Bush, Paython. *Why Brand Monitoring is a Security Issue - Typosquatting - https://www.anomali.com/blog/why-brand-monitoring-is-a-security-issue-typosquatting.*

11. Paganini, Pierluigi. *Homograph Phishing Attacks are almost impossible to detect on major browsers - https://securityaffairs.co/wordpress/58120/breaking-news/homograph-phishing-attacks.html.*

12. Haylee. *Common phishing scams and how to prevent them - https://blog.emsisoft.com/en/26866/phishing-scams/.*

13. Beaumont, Kevin. *JavaScript PCI nightmare: Ticketmaster, Inbenta and the canary in the coal mine - https://doublepulsar.com/javascript-pci-nightmare-ticketmaster-inbenta-and-the-canary-in-the-coal-mine-5c7410e8565b.*

14. Siaut, Pierre. *La banque centrale Australienne victime de spear-phishing - https://blog.trendmicro.fr/la-banque-centrale-australienne-victime-de-spear-phishing/.*

15. Moorehead, Matt. ReturnPath. *Cómo explicar DMARC en español claro y conciso.* [En línea] 20 de Julio de 2015. https://returnpath.com/es/blog/como-explicar-dmarc-en-espanol-claro-y-conciso/.

16. GOV.UK. *https://www.gov.uk/government/publications/email-security-standards/sender-policy-framework-spf.*

7. Anexos

7.1. Desarrollo de un sistema basado en IOCs (phishingAnalyser)

Para el desarrollo de un sistema basado en indicadores de compromiso que compruebe de manera autónoma si una página web cualquiera reúne suficientes requisitos como para poder ser considerada un phishing que afecte a una determinada empresa, hemos usado Python3, Selenium, Hashlib y Browsermobproxy.

El desarrollo ha sido migrado a un Docker para facilitar que pueda ser ejecutado en cualquier máquina y que se reunan las características y dependencias necesarias como para que el script se ejecute sin errores.

El desarrollo se compone por lo siguientes ficheros:

- Dockerfile
- hashes.txt
- phishingAnalyser.py

El Dockerfile define la estructura del Docker y configura la imagen en la que será ejecutado el programa.

En el fichero hashes.txt, se define el aprovisionamiento de los hashes de los recursos que han sido identificados como indicadores de compromiso asociados a un kit de phishing.

El script se encuentra en el fichero phishingAnalyser.py, donde su ejecución se divide en dos funciones:

- calcular_hash(recurso): Esta función recibe un parámetro recurso que es cargado por la web que se está analizando y calcula el hash asociado a este elemento. Para calcular el hash, comprueba si el recurso está disponible (si da una respuesta 200 a nuestra petición GET) y mediante la función hashlib calcula el valor del hash asociado a ese recurso.
- main(): En esta función se definen en primer lugar las configuraciones que se van a aplicar en el Selenium que vamos a utilizar para realizar la navegación hacia el sitio de phishing y analizar el tráfico.
  Debido a que vamos a usar este Selenium para consultar webs de phishing y queremos evitar que se nos muestre cualquier página de advertencia de que el sitio es peligroso o que se nos avise acerca del contenido inseguro alojado en esta web, hemos dispuesto una serie de opciones para desactivar todas estas características en el navegador.
  Para poder capturar el tráfico generado al realizar la petición al site de phishing y detectar todos los recursos que se piden al servidor y se descargan para poder visualizar la web, hemos hecho uso de BrowserMobProxy.
  Mediante este proxy, obtenemos una variable que hemos llamado “capture” que es una lista de listas con diferente información acerca de los recursos que carga la web y las peticiones que se realizan durante la carga de la página.

Recorremos esta variable “capture” para obtener la URL de cada recurso y realizar llamadas a la función “calcular_hash” para calcular el hash asociado a cada recurso de la web.
Finalmente comprobamos si esos hashes se encuentran en nuestra lista de recursos asociados a casos de phishing y calculamos la probabilidad de que el sitio analizado sea una web de phishing.

Para ejecutar el programa hay que construir el contenedor de Docker de la siguiente manera:
docker build -t phishinganalyser .

Posteriormente, para ejecutar el contenedor ejecutamos el siguiente comando:
docker run phishinganalyser URL_A_ANALIZAR