Efectividad de campañas anti-*phishing*

Diego Pascaner, Patricia Prandini[1]

Posgrado en Seguridad Informática de la Universidad de Buenos Aires[2]

Resumen. El *phishing* es un tipo de ataque informático cuyo objetivo más habitual es robar información personal del usuario. El presente trabajo tiene como objetivo conocer el nivel de preparación de los usuarios para reconocer páginas web que formen parte de un ataque de *phishing* y medir la eficacia de una variedad de instrumentos utilizados en campañas de concientización.

Con este objetivo se desarrolló un aplicativo que invitaba a los participantes a identificar sitios maliciosos, proponía distintas instancias de capacitación y volvía a evaluar si se habían registrado mejoras.

El estudio se realizó sobre 250 participantes, con un promedio de edad de 40 años y un nivel de educación universitario o superior. Un 29% se dedicaba a la seguridad de la información, mientras que un 14% nunca había escuchado el término *phishing*.

El estudio demostró que un grupo importante de participantes desconocía cómo protegerse frente a este tipo de ataque y que su capacidad de detección de sitios falsos mejoraba considerablemente luego de una breve capacitación. No se encontraron diferencias significativas entre los tipos de instrumentos utilizados. En conclusión, el estudio mostró que todo esfuerzo, por mínimo que sea, para mejorar el nivel de preparación de los usuarios, resultará en mejoras en sus habilidades para evitar ser víctimas del *phishing*.

1 Introducción

Este Trabajo aborda el problema del *phishing* en la realidad actual del uso de Internet. En primer término, describe sus principales características, sus distintas técnicas y versiones y las medidas para prevenirlo, desde los diferentes roles involucrados. Luego plantea una serie de escenarios destinados a medir la eficacia de distintos instrumentos de capacitación de los usuarios.

Para el análisis planteado, el estudio incluyó el desarrollo de una aplicación web que permite evaluar el nivel de preparación de un usuario para detectar sitios fraudulentos y el que experimenta después de realizar una capacitación en particular, permitiendo cuantificar las mejoras promedio de los usuarios que recorrieron cada uno de los instrumentos configurados.

[1] El presente artículo refleja el Trabajo Final de Especialización de Diego Pascaner en el Posgrado de Seguridad Informática de la UBA, trabajo que fue realizado bajo la supervisión de la Mg. María Patricia Prandini.

[2] Programa Tripartito entre las Facultades de Ciencias Económicas, Ciencias Exactas y Naturales e Ingeniería de la Universidad de Buenos Aires.

adfa, p. 1, 2011.

Además del análisis de los resultados, el trabajo también desarrolla una sección de recomendaciones para la prevención de ataques de *phishing,* focalizada principalmente en la capacidad de los usuarios para reconocer páginas web maliciosas.

1.1 ¿Qué es el *phishing*?

El *phishing* es un tipo de ataque informático cuyo objetivo principal es robar información personal de un usuario. El vector de ataque suele partir de un mensaje a través del cual el atacante pretende engañar a la víctima haciéndose pasar por una institución válida (o cualquier posible emisor legítimo del mensaje), en el que se le solicitan datos personales. Dependiendo de la información que se persigue, se podría esperar que los usuarios respondan al mensaje entregando sus datos, o bien utilizar la primera comunicación como un punto de partida para el mismo fin. En el último caso, usualmente se incluye un acceso directo a un sitio web, que continúa simulando ser legítimo, en el que el usuario, si no se percata del engaño, ingresará la información. Para lograr su objetivo, los atacantes suelen ofrecer grandes beneficios o generar sensación de urgencia.[2]

Un ataque de *phishing* usualmente incluye los siguientes cinco pasos. Primero, identificar y seleccionar a las víctimas. Puede ser una campaña masiva, o bien un ataque dirigido en el que se selecciona un grupo específico de destinatarios. Segundo, configurar el emisor del mensaje. Entre las tareas que se desarrollan en esta etapa se encuentran: elegir una marca, crear un sitio de igual apariencia, armar un correo electrónico que parezca legítimo y, finalmente, hacer parecer que el emisor del mensaje es legítimo (por ejemplo, simulando provenir de un contacto de la víctima o de una dirección de correo de una marca). Tercero, distribuir el ataque. Se envían los correos con enlaces a los sitios *phishing* y se generan publicidades y participaciones en foros que también lleven al sitio apócrifo. Esto se realiza acorde al público seleccionado en el primer paso, en términos de idioma, tipo de invitación, intereses, etc. Cuarto, las víctimas caen en el engaño: ingresan a enlaces fraudulentos y proveen sus datos o responden los correos entregando la información. Quinto, monetizar el ataque. Se busca ganancia financiera vendiendo las credenciales obtenidas y accediendo a cuentas de pago de las víctimas. La información obtenida puede ser también el punto de partida para nuevos ataques.[3]

1.2 Evolución del *phishing*

Desde los primeros ataques de *phishing* hasta hoy, la realidad del uso de Internet cambió de forma considerable. Según el indicador "Individuos usando Internet (% de población)" medido por The World Bank[3], en 1995 menos del 1% de la población del mundo utilizaba Internet. Desde el año 2014, este valor superó el 40% y continúa en aumento (ver Fig. 1).[5]

[3] http://www.worldbank.org/

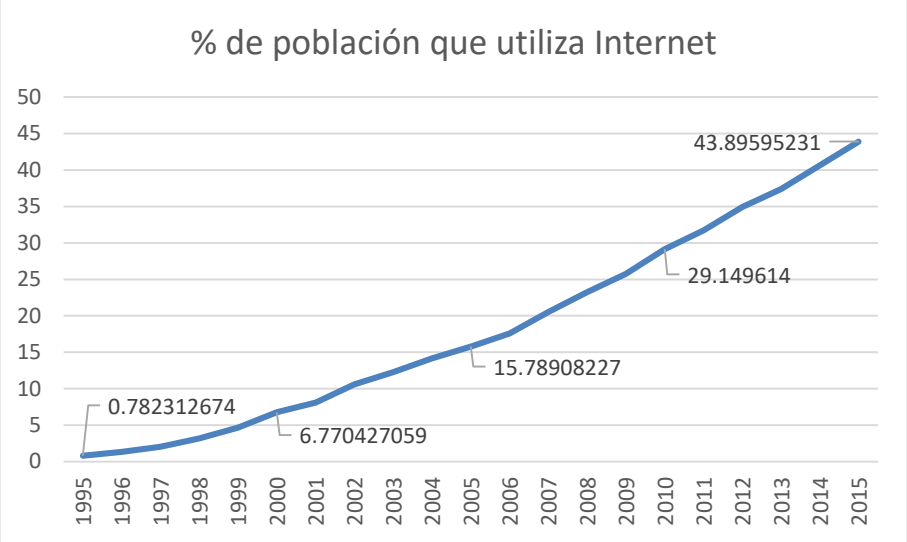

Fig. 1. - Evolución del porcentaje de población que utiliza Internet entre los años 1995 y 2015.

De la mano de este cambio de contexto de Internet, se advierte un crecimiento en la motivación para realizar ataques de *phishing* y, por lo tanto, resulta esperable observar un incremento de este tipo de actividad. La Fig. 2 refleja la variación en la cantidad de sitios únicos de *phishing* detectados según los reportes de tendencia de *Anti-phishing Working Group*[4].[9]

[4] https://www.antiphishing.org/

Fig. 2. - Evolución de los sitios únicos de *phishing* detectados entre los años 2005 y 2016.

En conjunto con el incremento en la cantidad, existe un proceso de perfeccionamiento en los aspectos técnicos del ataque, manteniéndose siempre estable el objetivo: lograr que uno o más usuarios entreguen información en forma inadvertida a una fuente no autorizada. Sin embargo, para mejorar la tasa de éxito y/o expandir los vectores de ataque, se desarrollaron nuevas técnicas, como las que se describen a continuación.

Phishing dirigido (en inglés *spear phishing*): en lugar de enviar grandes cantidades de mensajes con una pequeña probabilidad de respuesta, el atacante elige un objetivo (persona u organización), lo investiga y personaliza la comunicación. Esto dificulta que la víctima note que el mensaje es ilegítimo.

Inyección de contenido: es un ataque a los sitios web en el cual el atacante obtiene la capacidad de modificar parte del contenido de un sitio legítimo. Hasta aquí no se relaciona con el *phishing*, aunque se ha utilizado esta técnica para dirigir a los usuarios de un sitio legítimo a uno que no lo es.[10]

Punycode: es un mecanismo para ampliar los caracteres permitidos en una dirección de un sitio web. En general, un buen mecanismo de defensa ante el *phishing* es revisar la dirección del sitio que uno está navegando; si no es la del sitio que se desea, probablemente se esté ante un ataque. Aprovechando la tecnología del *punycode* se puede registrar un dominio cuya representación visual sea otra. Por ejemplo, alguien podría utilizar el dominio "xn--pple-43d.com" y los visitantes en su navegador visualizarían "apple.com"[5].[11]

1.3 Estado actual del *phishing*

Las organizaciones pueden optar por diversas alternativas técnicas para defenderse del *phishing*. Algunas son buenas prácticas genéricas de la seguridad de la información, como mantener los sistemas actualizados y con sus respectivos parches de seguridad, instalar soluciones *anti-malware* y monitorearlas y desarrollar políticas de seguridad (especialmente relacionadas con autenticación y contraseñas), entre otras. Otras defensas aplican de forma más directa al *phishing* (aunque también hagan su aporte en otros aspectos de la seguridad), como el uso de filtros de salida para la navegación de los usuarios o el análisis de los datos de entrada, especialmente en los correos.[12]

Adicionalmente se deben considerar variantes que pueden implementar tanto proveedores de servicios, como usuarios finales. Para el primer grupo, algunas opciones son el monitoreo de la marca en línea para detectar sitios "clonados", la creación de perfiles de comportamiento de los usuarios en pos de detectar anomalías o incluso, mejoras en los mecanismos de autenticación.[6] Los usuarios finales pueden utilizar herra-

[5] Poco tiempo después de ser descubierta esta posibilidad para ocultar un ataque de *phishing*, los principales navegadores hicieron mejoras para prevenir a los usuarios.

[6] Un ejemplo de un proveedor de servicio en línea que implementa soluciones *anti-phishing* es Google. En su sistema de correo electrónico Gmail, se incluye un algoritmo para la detección de *spam*, así como advertencias a usuarios cuando sospechan que está siendo víctima de un ataque de *phishing*. Se puede conocer más al respecto con la nota de Forbes disponible en:

mientas de análisis de correos, que se comportan como filtro, listas negras de direcciones de sitios o incluso de remitentes, aplicaciones que analizan el flujo de información o hasta la detección de sitios clonados por su parecido visual.[13]

A pesar del amplio abanico de soluciones técnicas *anti-phishing*, es importante destacar que, aún adoptadas en conjunto, nunca lograrán una efectividad del 100%. Esto puede deberse a cuestiones relativas a las configuraciones (bajar la sensibilidad de filtros para evitar falsos positivos) o a propiedades inherentes a la solución como el hecho de ser reactivas. Al respecto, Peter Wenham, director de seguridad de la información en la consultora *Trusted Management*, indica:

> "Abordar los engaños basados en correo electrónico y el *spam* comienza con reducir el volumen de *spam* a través de filtros y es completado educando a los usuarios para reconocer *spam* y engaños y borrarlos, desde la cúpula de una organización hasta los niveles más bajos. ¿Por qué? Porque no es posible detectar el 100% de este tipo de mensajes usando aplicativos especializados, complementos en los servidores de correo ni servicios externos. Finalmente, una persona deberá leer los mensajes que pasan los filtros electrónicos. Una aplicación, complemento en el servidor de correo o servicio externo debería reducir el nivel de *spam* cerca de un 95% (esto variará entre proveedor y proveedor y a lo largo del tiempo)."[7][14]

Tomando como base esta información, existen dos caminos no excluyentes entre sí a seguir para combatir el *phishing*: avanzar en las herramientas que reducen significativamente la cantidad de ataques que los usuarios reciben a diario y mejorar la concientización de los usuarios para que logren diferenciar mensajes legítimos de ataques con mejor tasa de éxito. Este trabajo se centra en el último de los dos caminos.

2 Instrumento de medición

En la sección anterior se estableció que para combatir el *phishing* resulta fundamental la educación a los usuarios. Para esto se propone un sistema que ubique a los usuarios en situaciones de la vida diaria, en las que deban determinar si se encuentran ante un sitio legítimo o uno apócrifo. Luego se los capacitará utilizando diversos instrumentos y se los invitará a volver a realizar la evaluación, para verificar si mejora su capacidad para identificar posibles ataques.

---

https://www.forbes.com/sites/kevinmurnane/2017/06/05/google-enhances-gmail-with-new-anti-phishing-tools/#5287fb9c6991.

[7] Traducción propia. Original en fuente: "*Tackling e-mail-based scams and spam starts with reducing the volume of spam by filtering and is completed by the educating the users from the top of an organisation right down to the most junior levels to recognise spam and scams and to delete. Why? Because it is not possible to screen out 100% of these message types using specialist appliances, add-in applications to e-mail servers or external services. In the end a human will have to read the messages that get through the electronic screening. An appliance, e-mail server add-in or external service should reduce spam levels by about 95% (this will vary supplier to supplier and from time to time).*"

La actividad consistirá en la presentación de una secuencia de imágenes de un navegador de Internet mostrando distintos sitios web. El participante deberá determinar si cada uno de ellos es legítimo o no. Tendrá un tiempo máximo de ejecución, buscando imponer presión que aumente la similitud con una situación real en la que la atención de la persona no está puesta únicamente en el sitio que está navegando ni tampoco se está preguntando de forma constante si se encuentra frente a un ataque. A continuación, se describe el proceso en el que un usuario realiza la actividad.

2.1 Realización de la actividad

Cuando el usuario ingresa al enlace de la actividad, el sistema muestra una pantalla introductoria, que incluye la denominación del estudio, una explicación de sus objetivos y su contexto y las instrucciones para continuar. Toda esta información es acompañada por imágenes y se muestra organizada en pasos para no abrumar a los participantes con un volumen excesivo de texto. En la Fig. 3 se puede visualizar una imagen de esta pantalla.

Fig. 3. - Página de introducción del estudio.

Una vez que el usuario completa la introducción, se le presenta una lista de sitios de Internet, sobre los cuales tendrá que identificar posteriormente imágenes legítimas o apócrifas de los mismos. Se le solicita al usuario que marque los sitios de la lista con los que se siente familiarizado. Este paso no tiene ningún impacto en cómo continua la actividad y se realiza únicamente para determinar si la familiaridad previa con los sitios ayuda a los participantes a detectar eventuales ataques.

El siguiente paso consiste en una primera secuencia de ocho imágenes de sitios web. En esta sección, se muestra una imagen de un navegador en un sitio de Internet, botones para determinar si el sitio es válido o no y un cronometro del tiempo restante que comienza en 2 minutos[8]. Una vez que el usuario responde, se persiste su respuesta y se le informa si eligió correctamente. En caso responder en forma errónea, se descuentan 5 segundos. Para mejorar la experiencia de los participantes, se incluyeron botones para ambas respuestas, tanto sobre la imagen del sitio como por debajo. Además, al pasar el cursor por encima de la imagen se amplía esa sección para mejor la accesibilidad visual. La Fig. 4 muestra un ejemplo de esta pantalla, en la que se encuentra ampliada la parte superior de la imagen, que contiene la barra de direcciones.

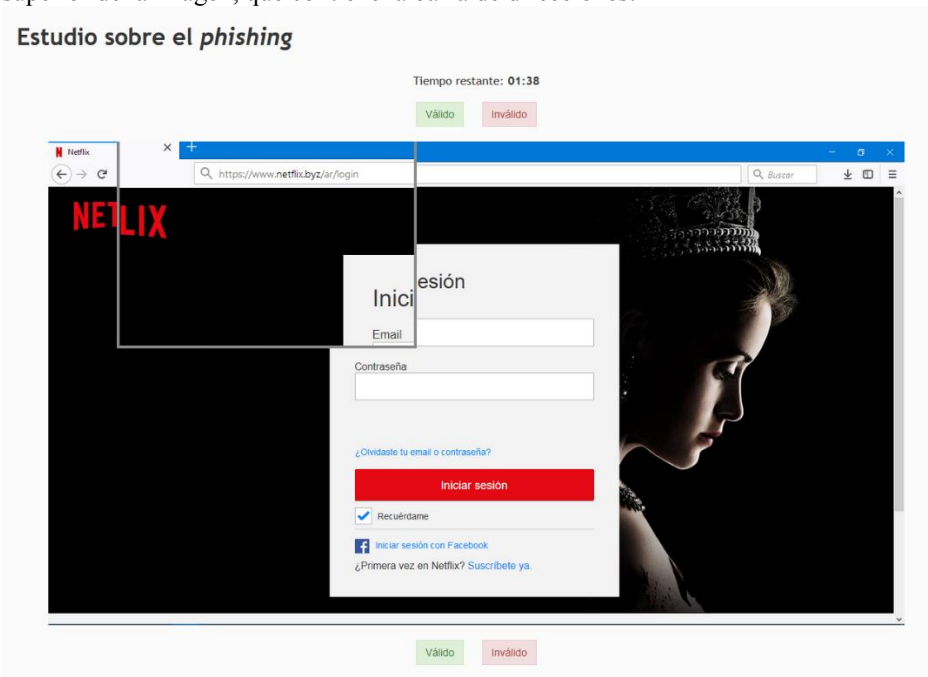

Fig. 4. - Página del estudio para que los participantes determinen la legitimidad de un sitio.

Cuando el usuario finaliza con la primera secuencia de sitios o se termina el tiempo disponible para evaluar las ocho imágenes, se muestran los resultados, que reflejan el desempeño del participante: la cantidad de sitios legítimos, cuántos fueron contestados correctamente, cuántos fueron erróneos y cuántos no se llegó a contestar, si fuera el caso. Se muestra la misma información en el caso de los sitios ilegítimos. Finalmente, se indica la cantidad de respuestas correctas. Si el usuario cometió errores marcando como legítimos sitios ilegítimos, se le muestran los motivos por los que los sitios seleccionados no eran válidos, es decir los indicios que no detectó para identificar la imagen del sitio atacado.

[8] La aplicación permite también configurar actividades que no limiten el tiempo disponible; en esos casos no se mostrará el cronometro.

El usuario continúa luego con la sección de la encuesta. El sistema solicita responder todas las preguntas y muestra de a una a la vez. Todas son de opción múltiple (lista de respuestas predefinidas) pero en algunas se puede seleccionar solo una única respuesta y en otras, varias.

Una vez completada la encuesta, el sistema persiste las respuestas. En este punto existen dos posibles instancias para la continuación del estudio. En caso de que en la primera secuencia de imágenes de sitios el usuario haya llegado a tiempo a responder todos los casos y lo haya hecho correctamente, se le muestra una pantalla explicando que es el fin de la actividad por haber distinguido con éxito las imágenes de los sitios web atacados de los sitios legítimos. Luego se sugiere mantener la atención y se brindan algunas recomendaciones al respecto.

En caso de que en la primera secuencia de imágenes el usuario no hubiera respondido todos los casos con éxito, la actividad continúa ofreciéndole una instancia de capacitación. Ésta será, entre las configuradas para la evaluación, la que haya sido determinada por el sistema en el ingreso del usuario a la primera pantalla. El criterio utilizado para esto es la capacitación que haya aparecido menos veces entre todas las ejecuciones de la actividad. Para esta aplicación se seleccionó un artículo de lectura, una infografía y un video[9]. Se sugiere al usuario tomarse su tiempo con la capacitación e incorporar los conocimientos ofrecidos. Al completarla, se prosigue con una segunda secuencia de ocho imágenes de sitios, similar a la etapa inicial. Una vez completada esta segunda sección, y luego de ver sus resultados, se indica el fin de la actividad y agradece al usuario por su participación.

2.2 Ejecución de la actividad y análisis de datos

Entre las devoluciones de algunos participantes respecto al tenor del estudio, un número importante manifestó que la actividad le había servido para aprender, darse cuenta de lo poco que conocían el tema e identificar los errores que cometían.

Sin embargo, se evidenció un alto porcentaje de abandono de la actividad antes de finalizarla, como se puede observar en la Fig. 5. Al respecto, algunos participantes evidenciaron como un aspecto negativo la imposibilidad de acceder a través de sus teléfonos celulares, mientras que otros indicaron que los textos iniciales eran demasiado extensos para leer. En total, el 51% de los participantes que ingresaron a la actividad, la abandonaron sin haber completado las instrucciones iniciales.

El gráfico de embudo (Fig. 5) muestra en cantidades y porcentajes, fraccionando el proceso de realización de estudio en cuatro etapas: ingreso a la dirección web que se les envió; lectura de las instrucciones; realización del primer paso de la actividad, indicando cuáles de los sitios presentados le resultan conocidos; y finalmente, recorrido completo del estudio.

[9] Los instrumentos de capacitación son parte de la configuración de la actividad. Se podrían incluir, por ejemplo, varias capacitaciones de lectura y ningún video ni imagen o la combinación que se desee.

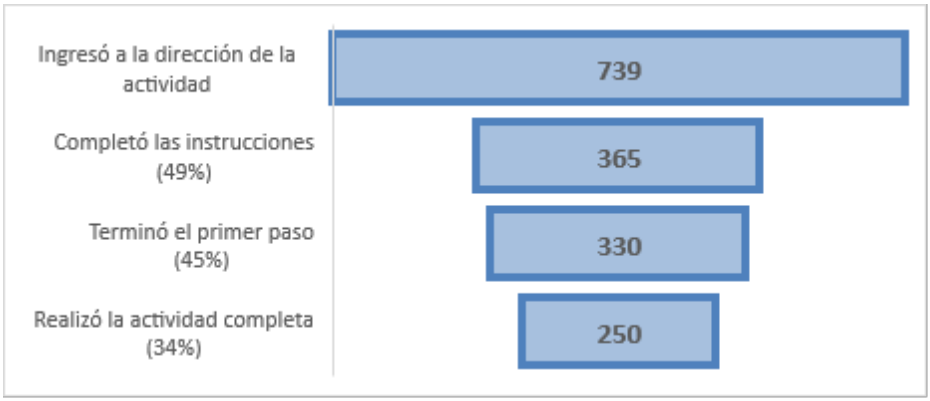

Fig. 5. - Gráfico de conversión del estudio.

Existen otros factores que podrían explicar la diferencia entre las primeras dos etapas. Por ejemplo, que a los participantes les haya resultado demasiado larga la introducción o que entraran a ver con qué se encontraban y dejaran su participación, para otro momento, sin volver a retomarla.

Para el análisis de los resultados se tienen en cuenta únicamente las participaciones completas en la actividad; descartándose todas las inconclusas. Utilizando este criterio, se considera un total de 250 participaciones sobre los 739 ingresos a la página inicial.

Al comenzar a estudiar la información recolectada, se analizó la composición del grupo de participantes. El promedio de edad fue aproximadamente 40 años con una concentración significativamente mayor entre los 25 y 50 años. La población que realizó la actividad presentó un nivel de educación elevado: un 63% de los participantes contaba con título universitario o superior. Esto posiblemente se deba a que el estudio fue realizado mayormente entre los alumnos y egresados del Posgrado en Seguridad Informática de la UBA. Particularmente, 72 participantes (29%) indicaron que la seguridad de la información es su actividad principal. Finalmente, resulta importante considerar los niveles de conocimiento iniciales del *phishing*. Un 80% de los participantes sabía de qué se trataba, mientras que un 14% era la primera vez que escuchaba el término. El alto nivel educativo de los participantes posiblemente tenga una influencia significativa en estos últimos resultados.

El primer resultado interesante es que 72 participantes respondieron correctamente todos los sitios de la primera sección y, por lo tanto, no realizaron la capacitación ni la segunda secuencia de sitios. Esta cantidad representa un 29% del total de participantes. Sin embargo, la distribución de este grupo no muestra diferencias significativas en cuanto a conocimientos previos del *phishing* o al porcentaje que se dedica a la seguridad de la información.

La primera comparación que se realizó es el total de imágenes de sitios que los usuarios no hicieron a tiempo de responder antes y después de la capacitación. Esto se hizo con el objetivo de evaluar si la capacitación logró mejorar el tiempo promedio que utiliza un usuario para identificar la legitimidad de un sitio. Para esto sólo se consideró a los 178 participantes que pasaron por la segunda secuencia de sitios. En este caso, se

puede observar un resultado que, a priori, puede considerarse opuesto al esperado: antes de la capacitación hubo 95 (7%) sitios que, tomando el conjunto de los participantes, no se llegaron a evaluar, mientras que después de la capacitación hubo 128 (9%) no evaluados. Una posible explicación es que al principio solamente miraban la apariencia del sitio y luego de ser capacitados, prestaban mayor atención a los aspectos técnicos específicos como la dirección web y el certificado de HTTPS; lo cual lleva más tiempo. Sin embargo, tomando la cantidad de participantes y no de imágenes de sitios, de la primera secuencia de sitios hubo 37 participantes (21%) a los que no les alcanzó el tiempo para realizar las 8 respuestas; mientras que después de la capacitación, este número se redujo a 28 (16%). Esta situación podría indicarnos que algunos de los participantes ganaron velocidad en la evaluación de cada escenario, mientras que otro grupo reducido de usuarios vio esa velocidad reducida, posiblemente a causa de aplicar en cada caso los conocimientos incorporados.

Otro aspecto interesante que diferencia la primera tanda de la segunda, son los sitios respondidos correctamente. Sobre los 178 participantes que realizaron la segunda secuencia de sitios, en la primera tanda de imágenes se respondieron correctamente 897 sitios (63% del total o 67% si se descartan los que quedaron sin respuesta por falta de tiempo). En la segunda secuencia en cambio, se observó un 15% de incremento habiendo evaluado correctamente 1037 sitios (73% del total de sitios y 80% sobre los sitios respondidos). Resulta destacable que si se consideran únicamente aquellas personas que no conocían lo que es el *phishing* el incremento asciende a un 33% con 174 respuestas correctas en la primera secuencia (44%) y 231 (58%) sitios evaluados correctamente en la segunda tanda.

El promedio de mejora de todos los participantes, entre la primera y la segunda etapa, fue de un 31%. Adicionalmente, se pudo observar que el 83% de los participantes mostró un cambio positivo o se mantuvo igual luego de la capacitación, siendo que el 63% mostró mejoras y el 20% restante se mantuvo con la misma cantidad de aciertos. Dentro de ese 83%, el promedio de mejora global asciende al 45%. Es posible que esta diferencia se deba a que algunos participantes no hayan completado la capacitación, o bien que el tipo de capacitación que recibieron no fuera el adecuado para ellos. No se encontró en los datos recolectados ninguna característica que segmente de forma clara al grupo que mejoró su cantidad de aciertos respecto al que empeoró.

El hecho de estar familiarizado con un sitio no se presentó como una ventaja a la hora de detectar su legitimidad. Tanto para los sitios señalados como utilizados frecuentemente como para los desconocidos, ocurrió un 76% de acierto al responder. Concretamente de 1235 respuestas a sitios que los participantes conocían, 944 fueron correctas. Para los sitios que a los participantes no les resultaban familiares, hubo un total de 1738 respuestas y de ellas, 1319 correctas.

Asimismo, resulta interesante analizar las mejoras de los participantes para detectar los distintos tipos de engaño que aparecieron durante el estudio. La Tabla 1. - Mejoras según el tipo de error presentado, muestra las mejoras porcentuales para cada tipo de error tanto de forma global, así como desagregado según la capacitación recibida.

Tipo de error	Mejora promedio para detectar el engaño luego de la capacitación			
	Global	Infografía	Video	Artículo de lectura
Solicitar datos personales sin cifrar el tráfico (sin HTTPS)	54%	58%	46%	57%
Error de tipeo en la dirección web	45%	46%	53%	37%
Uso del nombre del sitio en la dirección web pero no como dominio	7%	-1%	3%	17%

Tabla 1. - Mejoras según el tipo de error presentado. Se muestran las mejoras globales y clasificadas por tipo de capacitación recibida.

3 Recomendaciones para combatir el *phishing*

En base al trabajo realizado, en esta sección se elabora una serie de recomendaciones para que los usuarios de Internet puedan estar más prevenidos frente a eventuales ataques de *phishing*.

En este sentido, el aspecto fundamental que deben comprender es la necesidad de mantener una actitud de desconfianza ante cualquier mensaje o imagen conteniendo un enlace o al pretender navegar en cualquier sitio web. En otras palabras, hasta no estar seguro de su legitimidad, debe asumir que se trata de un mensaje o sitio web fraudulento. A continuación, y a manera de ejemplo, se explican situaciones de distinto nivel de sofisticación, que podrían ocurrir de manera habitual.

El escenario más usual y quizás el más simple de detectar, es la recepción de un mensaje (correo electrónico, SMS, mensajes instantáneos o cualquier medio similar) de una dirección desconocida, de apariencia legítima que incluya, por ejemplo, nombres de un contacto o de entidades conocidas. Algunos signos específicos de alerta son: el ofrecimiento de situaciones poco probables y muy ventajosas, la amenaza de consecuencias drásticas si no se siguen las indicaciones brindadas, la inclusión de enlaces a sitios y la solicitud de datos personales o de seguridad.

Un caso similar, aunque más complejo de detectar, es la recepción de un mensaje desde una dirección electrónica conocida. Si bien los usuarios viven esto en forma habitual, podrían encontrarse ante una situación de un ataque, por ejemplo, si el emisor tuviera algún dispositivo infectado con *malware* o si alguna de sus cuentas de mensajería hubiera sido robada. Un primer filtro rápido para clasificar la mayoría de los mensajes es si se encuentra mal escrito o si es distinto de lo que se esperaría del emisor (uso de palabras extrañas, español neutro, otro idioma, etc.), además de las alertas señaladas en el caso anterior.

Otro escenario más complejo de detectar es cualquier enlace encontrado en Internet, incluso en aquellos sitios considerados confiables. Los atacantes tienen distintas formas de hacer aparecer un enlace a un sitio apócrifo en una página legítima y confiable. Siempre que se pueda, se debe ingresar de forma directa al sitio que buscamos y no a través de enlaces, ni siquiera si se trata de sitios confiables y de uso habitual, como por ejemplo Google.[10] Muchos usuarios tienen la costumbre de escribir en su navegador el nombre del sitio al que desean ingresar en lugar de su dirección web, generando una búsqueda en el motor que esté configurado por defecto y desde ahí, acceden al sitio que aparece normalmente, como primer resultado.

A partir de la recomendación de desconfiar de los mensajes recibidos y de los enlaces en sitios de Internet, a los usuarios sólo cabe inspeccionar cada sitio y determinar su veracidad, aspecto que fue cubierto por el estudio realizado.

Para detectar si un enlace es legítimo es necesario comprender dos aspectos: sus componentes y cómo se conforma una dirección web. Para ver la verdadera dirección web a la que apunta un enlace se debe poner el cursor encima de su contenido y el navegador mostrará en su parte inferior la dirección. A continuación se muestra una imagen con la estructura de una dirección web, en la que se detallan sus componentes.

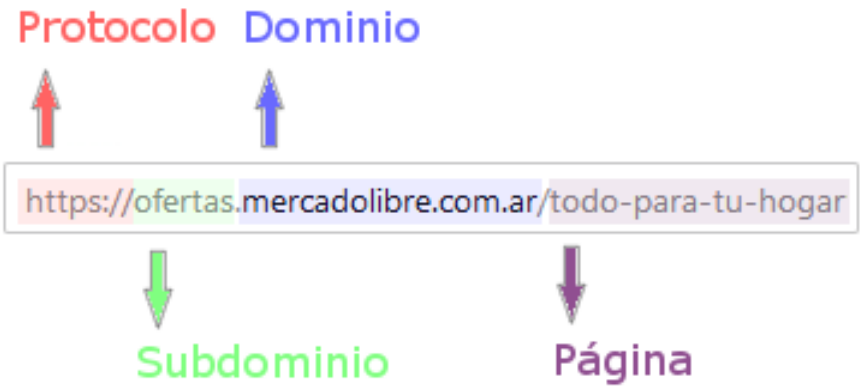

Fig. 6. - Componentes de una dirección web

Además de distinguir direcciones de sitios legítimos, los usuarios deben estar en condiciones de detectar direcciones de correo ilegítimas. Este caso es más simple ya que todo lo posterior al carácter "arroba" (@) de una dirección, es el dominio en cuestión.

[10] En el 2016, en Argentina, ocurrió una estafa millonaria posicionando un sitio de *phishing* previo al sitio legítimo en el buscador Google. Se puede ver más información en el artículo "Confirman que fue *phishing* lo que permitió el robo de $3,5 millones en Argentina" de welivesecurity https://www.welivesecurity.com/la-es/2017/08/07/confirman-phishing-robo-argentina/.

Finalmente, se cierran las recomendaciones de esta sección con un conjunto adicional de buenas prácticas ante un posible mensaje o sitio web apócrifo. Para lograr distinguir frente a qué situación se encuentran, la mejor opción es contactar al supuesto emisor por un medio alternativo y en caso de haberse hecho efectivo el ataque por no haberlo detectado a tiempo, se recomienda reportar el caso frente a la organización o persona cuya identidad fue suplantada y ante las autoridades competentes. Luego, para evitar un mayor compromiso de la información robada, se deben adoptar las medidas pertinentes, como modificar las credenciales de acceso o anular los datos de una tarjeta de crédito.

Por su parte, las organizaciones deben estar atentas frente a posibles ataques de *phishing* porque además de la posibilidad de perder información, suelen ser el punto de partida hacia ataques más sofisticados. Una organización madura en este aspecto debe en primer lugar, realizar mediciones respecto a la capacidad de sus empleados y clientes o usuarios, para detectar eventuales ataques de *phishing* y en base a este diagnóstico, se debe reducir cualquier falencia. Para esto, en cuanto al personal, se pueden realizar jornadas presenciales o virtuales de concientización, distribuir boletines de forma periódica y realizar ejercicios de simulación en los que cuando los empleados sean víctimas de un ataque ficticio preparado de exprofeso, reciban una notificación con las explicaciones pertinentes al caso. En Internet hay disponible abundante material respecto a las campañas que pueden llevar a cabo las organizaciones[11] e incluso, herramientas que permiten gestionarlas de forma simple y centralizada.

Como resumen de esta sección, se muestran a continuación dos gráficos que sintetizan las recomendaciones detalladas.

[11] Por ejemplo el Instituto Nacional de Ciberseguridad de España (INCIBE) comparte de forma gratuita un kit de concienciación disponible en: https://www.incibe.es/protege-tu-empresa/kit-concienciacion. *Stop. Think. Connect.* también tiene abundante material disponible incluyendo posters, hojas de consejos, videos, entre otras, que se puede acceder en: https://stopthinkconnect.cc/resources/landing/.

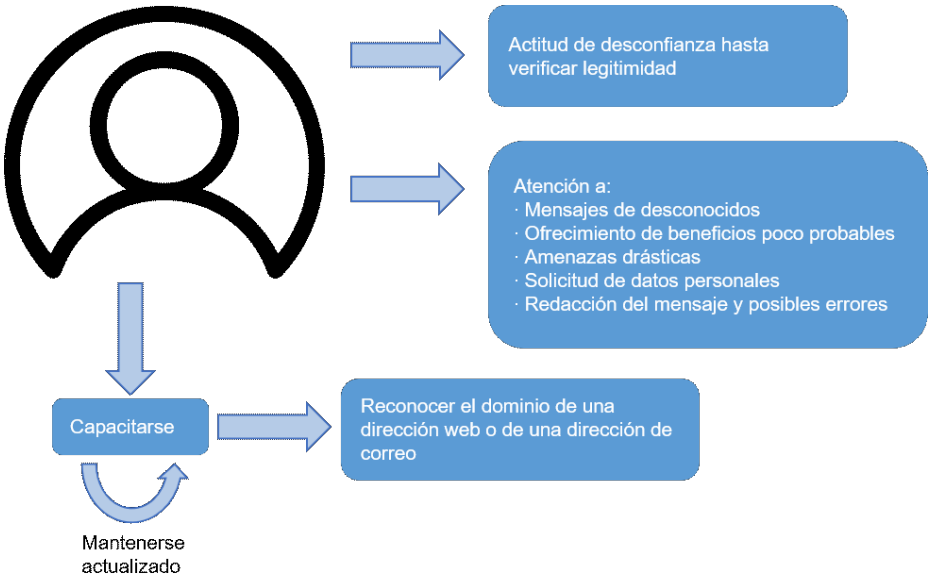

Fig. 7. - Recomendaciones para combatir el *phishing* (usuarios)

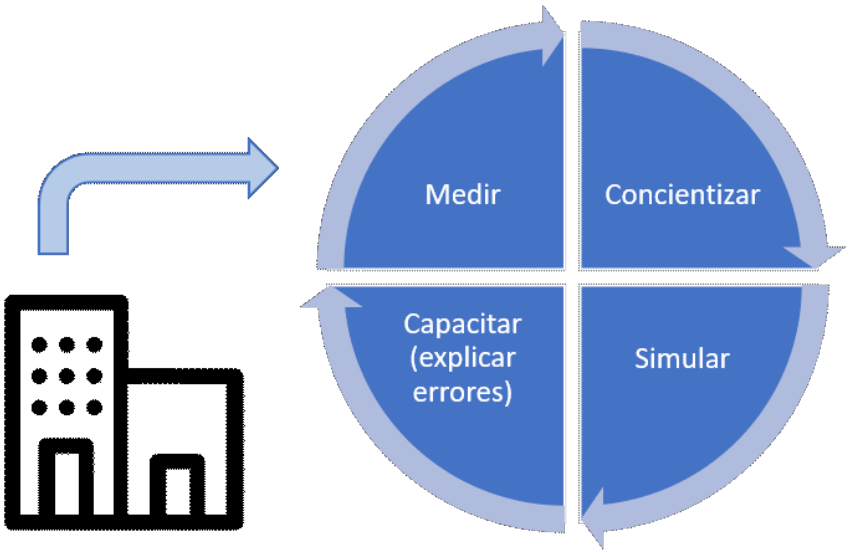

Fig. 8. - Recomendaciones para combatir el *phishing* (organizaciones)

4 Continuación del trabajo

En este momento se encuentra en progreso el desarrollo de una Tesis de Maestría que propone profundizar lo realizado en este trabajo. Ésta tendrá por objetivo ampliar el estudio considerando un mayor espectro de situaciones en los que el usuario puede estar expuesto a este tipo de ataques, por ejemplo a través de mensajes de correo electrónico falsos, generar un mayor sustento para las conclusiones a través de una muestra significativamente más grande y abarcativa e introducir mejoras en la herramienta desarrollada con el fin de facilitar su uso, por ejemplo permitiendo su realización a través de celulares y otros dispositivos.

5 Conclusiones

El estudio realizado permitió comprobar que la herramienta desarrollada cumplió con las expectativas iniciales. Los resultados del análisis de datos reafirman la importancia de concientizar y capacitar a los usuarios para evitar ser víctimas de un ataque de *phishing*. Adicionalmente, se pudo comparar distintos tipos de instrumentos de capacitación.

Si bien la muestra utilizada mostró algunas limitaciones en cuanto al tamaño y a la composición de la población bajo estudio, es posible formular las siguientes conclusiones.

En primer lugar, se observó la falta de conocimiento del tema; incluso estando en plena era de la información y habiendo seleccionado una población calificada, un 20% respondió que nunca había escuchado sobre el *phishing* o bien, que conocía el término pero no su significado.

A pesar de que en el marco de los participantes que se dedican a la seguridad de la información, el desconocimiento del tema fue inferior, este conocimiento pareciera ser meramente teórico ya que en la práctica consiguieron promedios de aciertos levemente por encima del resto de los participantes. Más aún, entre los que hicieron la primera secuencia de imágenes de sitios completa y sin ningún error, la proporción de participantes dedicados a la seguridad de la información es apenas 2 puntos porcentuales superior. Esto lleva a pensar que las concientizaciones deben tener en cuenta otros factores además de los aspectos técnicos, como invitar a permanecer alerta, a desconfiar y a alocar el tiempo suficiente, evitando actuar apresuradamente, entre otros.

Cabe acotar que entre el grupo que no necesitó realizar la capacitación, las características promedio en cuanto a conocimientos previos del *phishing*, edad, cantidad de tiempo que dedican a Internet, etc., son similares a las del total de la muestra. Esto podría indicar que quizás puedan existir otras características personales como la capacidad de atención o de observación, que se relacionen más directamente con la facilidad para detectar engaños de *phishing*.

Respecto a la velocidad de los participantes para determinar si un sitio es legítimo o apócrifo, no se obtuvo evidencia determinante en cuanto a la existencia de una mejora real. Después de la capacitación los participantes parecen haber tomado más tiempo para responder, lo que podría ser atribuido a una mayor atención y/o a una más precisa

búsqueda de indicios que antes pasaban por alto. Sin embargo, en la segunda secuencia de imágenes fueron menos los participantes que no pudieron completar los 8 sitios en los 2 minutos, indicando que en este sentido sí existió una mejora.

El aspecto más destacable del estudio realizado es el que permite concluir que a partir de la exposición a uno de los tres instrumentos de capacitación utilizados, los participantes experimentaron una mejora sustancial en su habilidad para evaluar sitios web. Es dable destacar en este sentido, que la actividad completa estaba pensada para realizarse en no más de 15 minutos, por lo que la duración de la capacitación era bastante acotada e incluso así, fue suficiente para generar estas mejoras. El crecimiento es aún mayor si se toma a los participantes que no sabían qué era el *phishing* antes de comenzar. Esto muestra cuán sencillo es incorporar los aspectos mínimos y suficientes para defenderse de los ataques más básicos.

Resulta interesante observar que con independencia del instrumento utilizado para capacitar, siempre se observan mejoras en la cantidad de respuestas correctas, si bien los incrementos variaron según cada modalidad. Se considera que el tamaño de la muestra y el hecho de tener una única versión de cada instrumento de capacitación (una infografía, un video y un artículo de lectura) no permiten concluir fehacientemente cuál de los instrumentos fue más efectivo y menos aún, segmentar el público según cuán efectiva haya sido cada capacitación, tratando de encontrar características en común en cada uno de estos grupos. Sin embargo, la herramienta desarrollada permite obtener este tipo de información que, en un proyecto a gran escala sería más precisa en sus conclusiones.

Adicionalmente, se observó que estar familiarizado con un sitio web no necesariamente influye en la capacidad del usuario para distinguir si es legítimo o un ataque de *phishing*.

En base a las mejoras de los participantes para identificar cada tipo de engaño utilizado, es notable la diferencia entre los dos engaños que no requieren conocimiento técnico y el que sí. Verificar que existe un certificado HTTPS mostró una gran mejora mientras que revisar que no haya errores de tipeo en la dirección web, tuvo una incidencia menor. Sin embargo, que el nombre del sitio sea realmente parte del dominio y no de otro componente de la dirección web (subdominio, página o un parámetro) tuvo un porcentaje aún menor de mejora. Se destaca también que para este último caso, existió una diferencia notable a favor de la capacitación a través del artículo de lectura. Los resultados obtenidos mostraron que para los engaños más simples, la infografía y el video tienen un comportamiento similar o inclusive mejor que el artículo de lectura, pero respecto a los conocimientos más técnicos, la lectura muestra una mejor incorporación por parte de los participantes. Estos resultados deben ser confirmados a través de la realización de un nuevo estudio con una muestra más significativa y diversa.

Como conclusión final, puede afirmarse que el estudio realizado confirma que todo esfuerzo, por mínimo que sea, realizado a nivel organizacional o personal para mejorar las capacidades de detección de sitios web falsos, redundará en una mejora en las habilidades de los usuarios para evitar ser víctimas del *phishing*.

6 Bibliografía

[1] *Efectividad de campañas anti-phishing* TFE Diego Pascaner 2018.

[2] «Phishing.org - What is Phishing,» [En línea]. Available: http://www.phishing.org/what-is-phishing. [Último acceso: 07 2017].

[3] «PHISH GUTS - The Anatomy of a Phishing Attack,» [En línea]. Available: http://www.cyren.com/tl_files/downloads/Phishing_infographic_print_r2.pdf?utm_medium=digital_ad&utm_source=APWG_resources. [Último acceso: 04 2018].

[4] «Phishing.org - History of Phishing,» [En línea]. Available: http://www.phishing.org/history-of-phishing. [Último acceso: 07 2017].

[5] «The World Bank - Individuals using the Internet (% of population),» [En línea]. Available: http://data.worldbank.org/indicator/IT.NET.USER.ZS. [Último acceso: 07 2017].

[6] «Dashlane - Infographic online overload its worse than you thought,» [En línea]. Available: https://blog.dashlane.com/infographic-online-overload-its-worse-than-you-thought/. [Último acceso: 07 2017].

[7] «Cisco - The Zettabyte Era: Trends and Analysis,» [En línea]. Available: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html. [Último acceso: 07 2017].

[8] «KPCB - Internet Trends,» [En línea]. Available: http://www.kpcb.com/internet-trends. [Último acceso: 07 2017].

[9] «Anti-phishing Working Group - APWG Phishing Attack Trends Reports,» [En línea]. Available: https://www.antiphishing.org/resources/apwg-reports/. [Último acceso: 07 2017].

[10] «Phishing.org - Phishing Techniques,» [En línea]. Available: http://www.phishing.org/phishing-techniques. [Último acceso: 07 2017].

[11] «Xudong Zheng - Phishing with Unicode Domains,» [En línea]. Available: https://www.xudongz.com/blog/2017/idn-phishing/. [Último acceso: 07 2017].

[12] «Digital Guardian - Phishing Attack Prevention: How to Identify & Avoid Phishing Scams,» [En línea]. Available: https://digitalguardian.com/blog/phishing-attack-prevention-how-identify-avoid-phishing-scams. [Último acceso: 08 2017].

[13] A. P. E. Rosiello, «Blac Hat - Anti-Phishing Security Strategy Brief,» [En línea]. Available: http://www.blackhat.com/presentations/bh-europe-08/Rosiello/Presentation/bh-eu-08-rosiello.pdf. [Último acceso: 08 2017].

[14] «Computer Weekly - How can organisations guard against phishing scams?,» [En línea]. Available: http://www.computerweekly.com/opinion/How-can-organisations-guard-against-phishing-scams. [Último acceso: 08 2017].

[15] «Cyren - Phishing Infographic,» 2016. [En línea]. Available: http://www.cyren.com/tl_files/downloads/Phishing_infographic_print_r2.pdf. [Último acceso: 08 2017].

[16] «Wombat Security - State of the Phish 2017,» [En línea]. Available: https://info.wombatsecurity.com/state-of-the-phish. [Último acceso: 08 2017].