Spam: La sofisticación de un viejo conocido

LOS CIBERDELINCUENTES UTILIZAN TECNOLOGÍAS CADA VEZ MÁS INGENIOSAS PARA ELUDIR LAS DEFENSAS

Blas Simarro Lorite

SE MANAGER SPAIN AND PORTUGAL

McAfee Inc.

Durante un cierto tiempo, los proveedores de productos de seguridad lograron ganarle terreno al spam filtrando el correo basura mediante el análisis del texto. Pero en 2006 hicieron su aparición nuevas y sofisticadas bandas de 'spammers' con un nuevo modo de contrarrestar los filtros y desataron un torrente de campañas de spam con imágenes. En algunos casos, las imágenes contenidas en los mensajes están vinculadas a sitios web aparentemente inofensivos, pero normalmente emplean técnicas inteligentes de ingeniería social que en ocasiones llegan incluso a manipular el mercado bursátil.

El uso de imágenes en lugar de texto para transmitir un mensaje de spam se remonta a 2004, pero desde entonces la complejidad y el volumen de este tipo de spam ha experimentado un rápido aumento. El spam con imágenes alcanzó su popularidad en los círculos de ciberdelincuentes en 2006, año en que acaparó hasta el 30% de todo el spam frente a menos del 5% en 2005. A finales de 2006 ya constituía el 50% de todo el spam, lo que supone un incremento del 500% en un período de año y medio.

El spam actual tiene un aspecto totalmente nuevo. Es más colorido, a menudo granuloso, contiene menos texto y a veces está ligeramente sesgado.

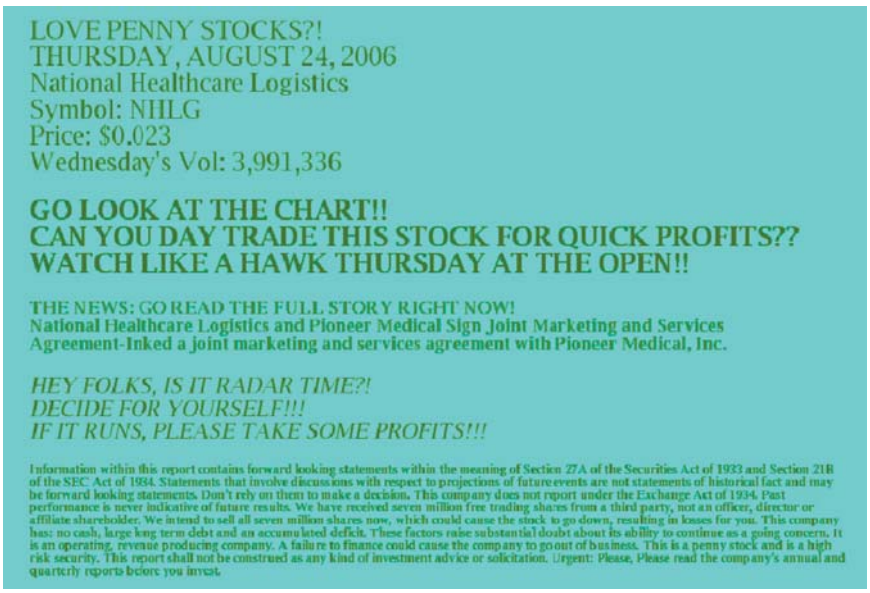
> El spam con imágenes es muy difícil de detectar porque cambia constantemente

Los remitentes de spam han agudizado su ingenio para eludir los filtros antispam. Para aquellos que buscan beneficios inmediatos, una imagen vale más que mil palabras. Para el usuario inexperto, el spam con imágenes tiene el mismo aspecto que cualquier otro mensaje de texto. La diferencia estriba en que el spam con imágenes es, como indica su nombre, un archivo gráfico .jpg o .gif formado por palabras integradas en la imagen. Muchas tecnologías de filtrado que analizan texto, nunca ven el contenido porque se trata de una imagen que les es invisible.

Las redes de ciberdelincuentes han recurrido a esta técnica para esquivar el radar mientras incitan a los destinatarios a ver pornografía, probar fármacos milagrosos y comprar acciones a precios muy bajos ("penny stocks").

Uno de los engaños de nueva generación, denominado "pump and dump", es una actividad delictiva en auge que no utiliza URL. Se trata de un ejemplo notable de ingeniería social y de manipulación del mercado. Los remitentes de spam adquieren acciones baratas y envían un mensaje de spam que incluye un símbolo bursátil real con la esperanza de que los más ingenuos compren y hagan subir el precio por acción. Inevitablemente, un determinado porcentaje de destinatarios cae en la trampa y compra acciones con la esperanza de obtener un beneficio. Cuando la cotización se dispara ("pump"), los remitentes de spam venden sus acciones ("dump"), cogen el dinero y corren. Los desafortunados que muerden el anzuelo pierden por lo general cerca del 7 % del dinero invertido.

El gráfico siguiente muestra el ejemplo de una campaña de spam bursátil.


LOVE PENNY STOCKS?!
THURSDAY, AUGUST 24, 2006
National Healthcare Logistics
Symbol: NHLG
Price: $0.023
Wednesday's Vol: 3,991,336

GO LOOK AT THE CHART!!
CAN YOU DAY TRADE THIS STOCK FOR QUICK PROFITS??
WATCH LIKE A HAWK THURSDAY AT THE OPEN!!

THE NEWS: GO READ THE FULL STORY RIGHT NOW!
National Healthcare Logistics and Pioneer Medical Sign Joint Marketing and Services Agreement-Inked a joint marketing and services agreement with Pioneer Medical, Inc.

HEY FOLKS, IS IT RADAR TIME?!
DECIDE FOR YOURSELF!!!
IF IT RUNS, PLEASE TAKE SOME PROFITS!!!

Information within this report contains forward looking statements within the meaning of Section 27A of the Securities Act of 1933 and Section 21B of the SEC Act of 1934. Statements that involve discussions with respect to projections of future events are not statements of historical fact and may be forward looking statements. Don't rely on them to make a decision. This company does not report under the Exchange Act of 1934. Past performance is never indicative of future results. We have received seven million free trading shares from a third party, not an officer, director or affiliate shareholder. We intend to sell all seven million shares now, which could cause the stock to go down, resulting in losses for you. This company has: no cash, large long term debt and an accumulated deficit. These factors raise substantial doubt about its ability to continue as a going concern. It is an operating, revenue producing company. A failure to finance could cause the company to go out of business. This is a penny stock and is a high risk security. This report shall not be construed as any kind of investment advice or solicitation. Urgent: Please, Please read the company's annual and quarterly reports before you invest.

Efectos del spam con imágenes en la disponibilidad de la red

Normalmente, estos mensajes tienen tres o cuatro veces el tamaño de los mensajes de spam con texto, por lo que ocupan más espacio en el servidor y reducen el ancho de banda. Incluso si se detectan y se envían a una base de datos en cuarentena, cuyo tamaño suele ser fijo, existe el riesgo de que saturen el servidor hasta que se eliminen. La mayoría de las empresas prefieren utilizar un producto antispam capaz de reconocer y eliminar este tipo de spam en la puerta de enlace de correo, antes de que lleguen a los servidores.

Problemas de detección del spam con imágenes

El spam con imágenes es muy difícil de detectar porque cambia constantemente. Aunque en la mayoría de los casos las imágenes de una campaña de spam parecen idénticas, en realidad cada imagen es única. Los remitentes de spam protegen su trabajo con "ruido" de fondo aleatorio, es decir, manchas y

La complejidad y el volumen del spam ha experimentado un rápido aumento

puntos en el archivo de imagen, nombres de archivo y líneas de asunto aleatorios o mensajes destinados a burlar los filtros de hash (*"hash busters"*) a base de largos párrafos de texto sin sentido o copiado de obras literarias y páginas web. Estas técnicas disfrazan el spam y crean un número casi infinito de variantes del mismo mensaje. Algunos mensajes utilizan archivos .gif animados y de imagen multicapa para ocultar el spam en la imagen.

Otros incluyen un vínculo a un sitio web que puede detectarse con software antispam. Pero la mayoría recurre a otros mecanismos de respuesta del usuario para evitar la detección, como incorporar información en la imagen, por ejemplo, la URL que el destinatario debe introducir en el navegador, números de teléfono y símbolos bursátiles. Detectar el spam con imágenes es complicado principalmente por dos razones: en general, los mensajes de spam con imágenes tienen la misma estructura que los mensajes legítimos. Sus remitentes modifican aleatoriamente las imágenes para eludir los filtros antispam.

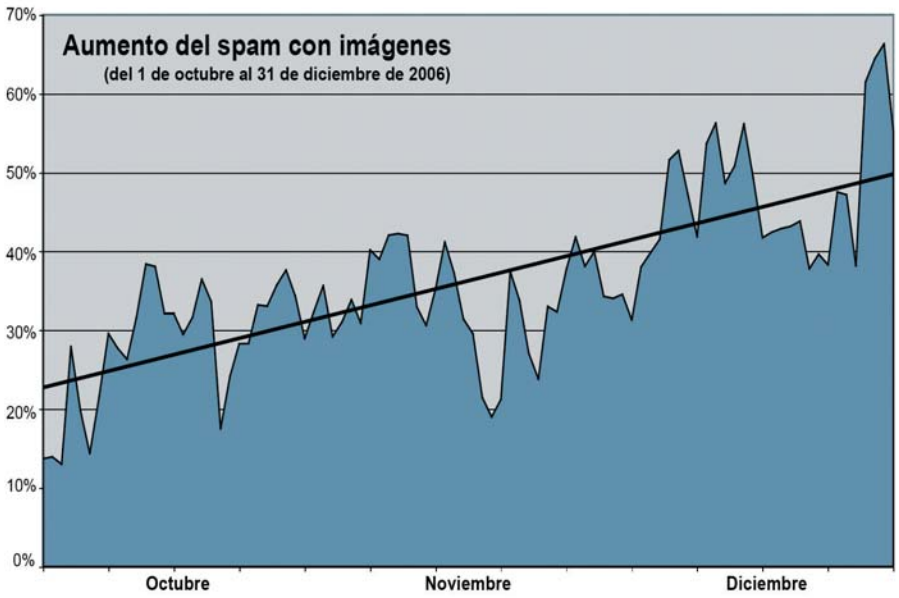
Imitación del correo legítimo

En algunos casos, los remitentes de spam intentan sortear los filtros antispam estructurando el mensaje de spam para que parezca un mensaje legítimo procedente de programas de correo habituales, como Microsoft® Outlook®, Outlook Express o Thunderbird. Cuando alguien redacta un mensaje, normalmente escribe el texto y a veces adjunta una imagen. Los remitentes de spam hacen exactamente lo mismo.

Incluyen encabezados, líneas de asunto y texto complejo, A veces, el spam con imágenes contiene elementos falsos de conversación para que parezca que el spam responde a un mensaje anterior.

Cambios aleatorios

Los 'spammers' recurren a una variada gama de técnicas de modificación de imágenes para sortear los filtros antispam, el reconocimiento óptico de caracteres (OCR) y los métodos de análisis de imágenes. Además de agregar ruido aleatorio a las imágenes, los remitentes a menudo camuflan el texto con diversas técnicas o dividen la imagen en un "mosaico" de imágenes más pequeñas para que el conjunto parezca un rompecabezas. A simple vista, una nueva andanada de spam



con imágenes tiene exactamente el mismo aspecto que el ataque precedente, pero las propiedades de la nueva imagen son completamente distintas.

Problemas de la detección OCR

La ocultación deliberada hace más difícil distinguir qué píxeles son texto y cuáles son color o ruido de fondo. Para encubrir el texto, los remitentes de spam con frecuencia emplean letras en relieve, tamaños de letra diferentes y colores aleatorios que dificultan la lectura hasta para el ojo humano.

Más allá del análisis de imágenes

Como ya hemos visto, el spam con imágenes es un blanco móvil, ya que los remitentes de spam cambian con frecuencia sus herramientas y el contenido y formato de los mensajes, por lo que no basta con una sola técnica para detectarlos todos. Los métodos actuales de detección de spam con imágenes pueden quedarse obsoletos el año que viene o incluso el próximo mes -como he comentado en referencia a la tecnología OCR-. La detección precisa de este tipo de spam exige una combinación de tecnologías avanzadas y el continuo proceso de investigación del 'estado del arte' de las técnicas de ocultación empleadas por los spammers.

Los "spammers" recurren a una variada gama de técnicas de modificación de imágenes para sortear los filtros antispam

Veamos a continuación algunas de las técnicas de detección más eficaces:

Reputación del remitente basada en IP

Con este método de detección, se puede rechazar el spam de forma inmediata sin necesidad de aceptar y analizar cada mensaje. Cuando se detecta un ataque, se bloquea automáticamente el spam que se origina en la dirección IP del remitente malintencionado antes de que entre en el servidor de correo. Las direcciones IP se agregan y eliminan dinámicamente en función del comportamiento del remitente.

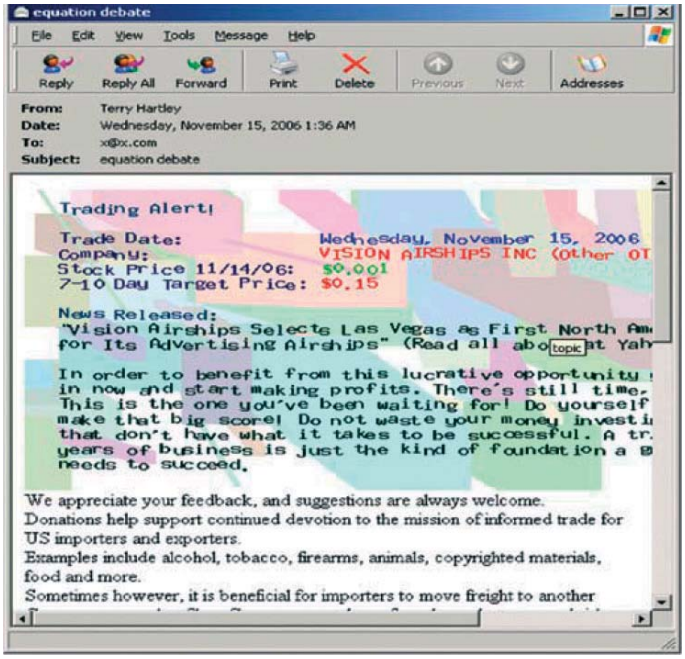
Propiedades de la imagen

El examen de las propiedades de la imagen es eficaz y preciso. Se aplican reglas que extraen las propiedades internas de las imágenes del spam de manera más eficiente que el OCR. El spam con imágenes posee numerosas características que le son propias y que pueden utilizarse como firmas para su detección. Por ejemplo, para averiguar si un mensaje contiene spam, se buscan las técnicas de ocultación y otras firmas. Se mide la cantidad de información que contiene una imagen (número de colores, tamaño y altura de la imagen en píxeles, etc.). Con esta información, se puede determinar con exactitud si una imagen es spam. Combinando esta técnica con otras reglas, es posible llegar a cotas de detección cercanas al 99%. Las imágenes insertadas en correo

electrónico legítimo tienden a ser más complejas y llevan asociadas más dimensiones que las imágenes de spam. Pensemos en una fotografía normal integrada en un mensaje de correo legítimo, por ejemplo, un automóvil o un lugar de vacaciones. Si medimos el número de colores que contiene una fotografía, el resultado normalmente alcanza cientos o miles de colores. Por su parte, en la clásica imagen de spam bursátil el número de colores ronda probablemente diez, porque las imágenes son muy simples y el texto está en blanco y negro. Si seguimos analizando otras dimensiones de la imagen y las alineamos en los ejes de abscisas y ordenadas ("X" e "Y") de un gráfico, observaremos que las características asociadas a las imágenes normales se distribuyen en una zona muy amplia, mientras que las asociadas a las imágenes de spam se aglutinan en sectores concretos. Si las características de una imagen concuerdan con las del agrupamiento típico del spam con imágenes, no cabe duda de que pertenece a una campaña de spam.

Reputación del nombre de dominio

En el caso del spam con imágenes que contiene enlaces a sitios web, es necesario detectarlo con fiabilidad comparando la URL del mensaje con sitios de spam conocidos. Los remitentes de spam intentan atravesar los filtros cambiando las URL con frecuencia, pero es mucho más difícil cambiar los servidores subyacentes. Aquí es donde entra en juego la reputación del nombre de dominio para bloquear las campañas de spam. Para que el chequeo sea posible, es necesario haber rastreado por anticipado a los remitentes de spam y los dominios que registran para nuevas campañas e incluirlos en una lista de vigilancia. En el momento en que se lanza un nuevo ataque de spam desde estos dominios, una rápida consulta a la lista identificará de manera certera el mensaje de spam.



> Es necesario analizar el formato, la disposición y la estructura de los mensajes de correo electrónico

Estructura del mensaje

Es necesario analizar el formato, la disposición y la estructura de los mensajes de correo electrónico.

A menudo detectamos los mensajes de spam por su estructura singular. Estas reglas han tenido mucho éxito a la hora de localizar spam bursátil porque sus remitentes suelen utilizar una plantilla para enviar el mensaje y sólo cambian la imagen que contiene. Al examinar numerosas características de cada mensaje, este método de detección representa un modo rápido y fiable de identificar y bloquear spam con imágenes.

Reglas de encabezados

Al igual que otros tipos de spam, el spam con imágenes incluye encabezados de texto normal que contienen gran cantidad de información sobre el contenido y el remitente del mensaje. Estas reglas identifican incoherencias en los encabezados de los mensajes y localizan otros rasgos específicos del spam.

***Hash* de imágenes**

Se trata de la creación de una firma digital de la imagen real. Esta técnica es muy veloz y eficaz en términos de cálculo, pero sólo es útil cuando los remitentes de spam bombardean la misma imagen durante varios días, hábito común entre algunos reputados remitentes de spam de Europa Oriental.

Las tecnologías aquí presentadas ilustran los métodos exigibles a una solución de filtrado de spam de nivel empresarial. Por cierto, el último 'sabor' de spam, con ficheros PDF...