2. Pasos de un Ataque a la Ciberseguridad

1. Footprinting y Reconnaisance

1.1 Introducción

Aunque el número de fases en un ataque a la seguridad informática y su nombre depende de la documentación consultada y está estandarizado, podemos definir a groso modo las siguientes etapas:

- 1. Reconocimiento
 - a. Activo/pasivo
- 2. Escaneo
- 3. Obtener acceso
 - a. Nivel de S.O./ Nivel de aplicación
 - b. Nivel de red
 - c. DoS
- 4. Mantener acceso
 - a. Subir/alterar/descargar programas o datos
- 5. Cubrir pistas

En este punto del curso vamos a centrarnos en la fase inicial de Reconocimiento.



1.2 Reconocimiento

El reconocimiento se refiere a la fase preparatoria en la que el atacante recopila tanta información (Information Gathering) como sea posible sobre un objetivo de evaluación antes de lanzar un ataque.

Se denomina information gathering a la instancia previa al intento de ejecutar una intrusión informática a un sistema por parte de alguien no autorizado. También es empleada (generalmente en organizaciones) por los profesionales éticos en caso de llevar a cabo una comprobación de seguridad. Information gathering implica llevar a cabo la tarea previa y minuciosa de inteligencia (similar a un reconocimiento del terreno), más precisamente a la recolección de datos acerca del objetivo o de algún componen te relacionado a este o a parte de él. Esta fase se compone, fundamentalmente, de investigación y análisis de datos recabados.

El sistema de información cuenta con incontables piezas y, por lo tanto, el factor permeable (brecha o agujero de seguridad) inicial de éste podría encontrarse en cualquiera de los niveles, comprendidos entre un fallo humano, uno de infraestructura (técnica), lógica y hasta externa por los agentes involucrados (por ejemplo, un proveedor de Internet o hosting inseguro o una sucursal con su red desprotegida) o distintos ambientes interconectados.

Los datos que buscan los intrusos antes de atacar pueden estar relacionados con algún empleado, ya sea ejecutivo u operario, con algún sistema o tramo de él o con algún procedimiento u operación que nos permita intervenir en él. También puede ser una dirección IP, un sitio, una red, una aplicación, un servicio (puerto abierto de autentificación o no), un protocolo, un determinado descuido de programación o de administración, un directorio, un documento, una plataforma o bien cualquier dato de ubicación física o denominación de algún sector de la misma organización.

Por supuesto, si puede directamente conseguir logins, lo intentará. No interesa si el dato es muy importante o casi insignificante. Todo es útil a la hora de la escalada en el sistema y la previa planificación de este embate (chequeo o simulación de ataque). Algunas de las preguntas útiles antes de proceder serían:

- ¿Qué sabemos de nuestro objetivo?
- ¿Dónde están sus redes, sitios o por dónde fluye su información?
- ¿Qué partes lo conforman?
- ¿Qué sistemas poseen y cómo están formados?
- ¿Cómo se llaman los integrantes de la organización?
- ¿Quiénes son sus empleados y qué hacen? ¿Cómo y dónde?
- ¿Qué información sobre ellos se puede consultar o conseguir en Internet?

Aquí tenemos una lista de alguna información crítica que debería obtenerse durante esta fase de reconocimiento:



Direcciones IPMáscara de subred
Topología de Red
Nombres de dominio
Nombres de dominio Nombres de usuario
 Nombres de grupos
 Tipo de arquitectura (x86, SPARC,)
 Familia y versión de sistema
operativo
 Servicios TCP y UDP que se están
ejecutando .
 Requisitos de complejidad de
contraseñas
 Frecuencia de cambio de contraseñas
• Retención de cuentas
caducadas/deshabilitadas
Seguridad física
-
Cortafuegos Sisteman de Datassión de latamente
Sistemas de Detección de Intrusos
 Dirección postal
 Número de teléfono
 Conocimientos en informática
 Secretos "oscuros"

Comúnmente, se denomina footprinting a esta recolección de información previa. De todos modos, cada atacante o consultor tiene sus propios métodos y recursos durante esta búsqueda. Cuanto más minuciosa e ingeniosa sea, más posibilidades tendrá de dar con un descuido, un objetivo o por lo menos, una pista para comenzar con otra etapa del ataque. Por ejemplo, un atacante real que posee en su haber algunas o la mayoría de las bases de datos de ISP (proveedores de Internet) del país, o tiene acceso a ellas, cuenta con una clara ventaja sobre el resto ya que, en éstas, posiblemente habrá mucha información útil relacionada con personas de la organización que pueden tener algún dato importante, como passwords, o permiten con seguirlo.

La recolección de datos previos al ataque generalmente comienza en algún tipo de base de datos y otros recursos que se dispongan. Después de cotejar las coincidencias de personas (existencia tanto en la base de datos como en la organización), tomará los datos personales y tratará de emplear como password sus fechas de nacimiento, sus números de documento de identidad, sus oficios o las mismas contraseñas allí utilizadas, pero esta vez en las cuentas de correo de la organización u otro servicio que requiera autenticación (ssh, ftp, telnet, ...)

Veamos algunos ejemplos. Si en su cuenta personal la víctima tiene una pregunta secreta relacionada con un libro (supongamos El principito) o encuentra en un foro que a esa persona le interesa ese libro, el intruso probará claves como las siguientes: elprincipito, zorro, invisiblealosojos, víbora, antoine, exupery, baobab, asteroide3251, b612, rosa, etcétera.

Incluso los datos personales (como nombre y apellido) servirán para deducir los usuarios de login y no solo la clave. El típico ejemplo es el UserID (el usuario que antecede al signo @) de las actuales cuentas de e-mail corporativas e institucionales, formadas por la primera letra del



primer nombre, seguido del apellido. Por ejemplo, aperez@dominiovictima.com. Información como ésta le servirá al intruso para sacar más información aún, quizá desde otros lugares. Esta persona, además, buscará en bases ilegales, como la de tarjetas de crédito, padrones de todo tipo (disponibles hasta en redes P2P), entidades privadas o bajadas de servidores de organizaciones que fueron previamente atacadas y comprometidas.

Para el intruso, una fuente de passwords o datos sensibles de ese estilo es de relevancia atemporal. Esto significa que no importa si la base de datos que tiene en su poder es antigua, ya que podrá usarse en un futuro como dato o como pista para conseguir el dato necesario actual.

Es común además que éste, más allá del análisis íntegro del sitio institucional del objetivo, busque también información de algún componente de la organización en portales o sitios relacionados con ofertas laborales (miles de currículum vítae), foros, blogs y comunidades online de todo tipo, eventos, registros de dominios, portales de educación, guías empresariales, telefónicas, profesionales e industriales, búsqueda de colegas, redes sociales tipo Linkedin, Facebook, Twitter, Myspace, avisos clasificados de todo tipo, policía y agencias de seguridad online.

Los buscadores también son una increíble fuente de clasificación, análisis, búsqueda y caché de información, confidencial o no, sobre un objetivo. Altavista fue el buscador preferido en los años 90, Yahoo lo fue más cerca del año 2000 y hoy lo es Google. Seguramente habremos escuchado hablar de Google hacking (lo trataremos en un tema posterior), es decir, utilizar el famoso buscador para encontrar datos relevantes del objetivo.

```
Vídeos Maps Noticias Shopping Gmail Más -

filetype:inc intext:mysql_connect password -please -could -port

Aproximadamente 2.210 resultados (0,44 segundos)
```

```
<?php /* $server = "localhost"; $database = "iie"; $user ...
www.iiethai.org/web/inc/config.inc - Traducir esta página
<?php /* $server = "localhost"; $database = "iie"; $user = ""; $password = "" ... "
iiethai"; $user = "root"; $password = "iie@bundith"; mysql_connect($server, $user, ...
<?php $dbname = 'grizle pn'; $hostname = 'humbug'; $username ...
www.grizle.co.uk/veryoldsite/common.inc - Traducir esta página
<?php $dbname = 'grizle_pn'; $hostname = 'humbug'; $username = 'grizle'; $password
= "13f4cff19bcb62e7"; $id_link = @mysql_connect($hostname, ...
$servidor="localhost"; $usuario="root"; $password="contracara ...
tarwi.lamolina.edu.pe/exceltosql/conexion.inc
$servidor="localhost"; $usuario="root"; $password="contracara"; $base="foro"; $SQLid
= mysql_connect($servidor,$usuario,$password); ...</pre>
```

Las búsquedas que hacen uso de Google para encontrar información sensible se suelen llamar Google Dorks. Podemos encontrar una gran cantidad de dorks en:

http://www.exploit-db.com/google-dorks/





Por su parte, el proyecto Google Hack Honeypot (http://ghh.sourceforge.net) merece especial atención.



What is GHH?

Google Hack Honeypot is the reaction to a new type of malicious web traffic: search engine hackers. GHH is a "Google Hack" honeypot. I attackers that use search engines as a hacking tool against your resources. GHH implements honeypot theory to provide additional secu

Google has developed a powerful tool. The search engine that Google has implemented allows for searching on an immense amount of i billion pages [February 2005] and continues to grow daily. Mirroring the growth of the Google index, the spread of web-based applicatio administrative tools has resulted in an increase in the number of misconfigured and vulnerable web apps available on the Internet.

These insecure tools, when combined with the power of a search engine and index which Google provides, results in a convenient attack

GHH is powered by the Google search engine index and the Google Hacking Database (GHDB) maintained by the johnny.ihackstuff.com c

Honeynet Research with GHH

You can view research done with GHH in the Honeynet Project's "Know Your Enemy" paper on web application honeypots.

GHDB Honeypots Available:

```
GHDB Signature #365 Emulated (intitle: "PHP Shell *" "Enable stderr" filetype:php) GHDB Signature #833 (filetype:php HAXPLORER "Server Files Browser") GHDB Signature #733 ("Enter ip" inurl: "php-ping.php") GHDB Signature #365 (intitle: "BHP Shell *" "Fnable stderr" filetype:php)
```

Se denomina Honeypot al sistema (a modo de cebo) cuya intención es atraer a intrusos simulando ser sistemas vulnerables o débiles a los ataques. Es una herramienta de seguridad informática usada para recoger información sobre los atacantes y sus técnicas. Los Honeypots pueden distraer a los atacantes y advertir al administrador del sistema de un ataque, además de permitir un examen en profundidad del atacante, durante y después del ataque.

La gente de Google no es tonta y lo más probable es que cuando un atacante, buscando objetivos al azar, coloque algunos de los métodos descritos en la GHDB (Google Hack Database) sea redireccionado hacia alguna honeypot o a una mayoría de links con resultados 404 y a más honeypots. Esto permite controlar los ataques indiscriminados de script-kiddies, pero no así el ataque a organizaciones de modo focalizado.

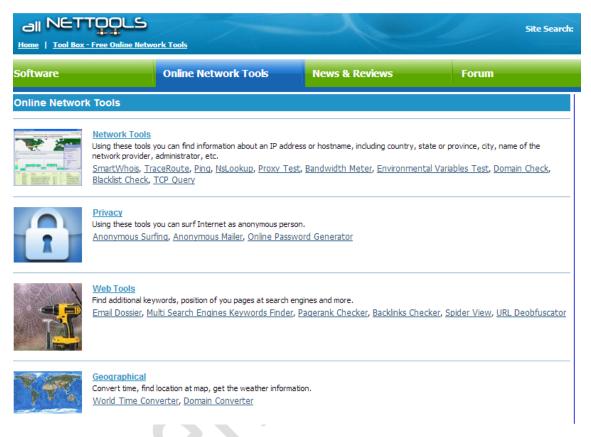
Para evitar problemas, en nuestros sitios web conviene utilizar un filtrador de robots de indexación, si no deseamos que Google u otro buscador cacheen el sitio o parte de ellos.

Algunas herramientas automatizan la búsqueda a través de Google, como QGoogle, GoogleScan, Google Enum o Site Digger. Incontables scripts en Perl y otras herramientas podrán ser encontradas en Securityfocus o Packetstorm para buscar en Google usuarios de sistema e información relaciona da al footprinting (Kali Linux tiene en su colección varias de éstas). Pero la búsqueda y el análisis manual es lo más recomendable para hacerlo de modo profesional dirigido a un objetivo en concreto.



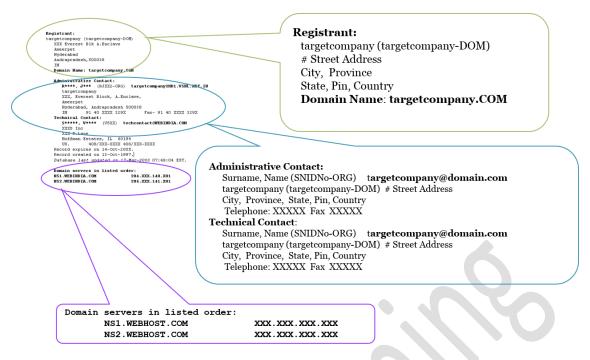
1.3 Otros Recursos Online

Hay otras bases de datos públicas y herramientas que brindarán datos en tiempo real en Internet. Entre estos últimos, los sitios más conocidos en el pasado fueron www.samspade.org y www.netcraft.com, que permitían saber el sistema operativo de los servidores, sus rangos de direcciones, qué sistema tenía históricamente, su uptime, IP, los nombres de administradores, teléfonos y direcciones físicas, entre otras cosas. Como ejemplo, veamos qué datos podemos obtener en el sitio www.all-nettools.com/toolbox, que tiene muchas herramientas.



• **SmartWhois**: encuentra información acerca de una dirección IP, hostname, incluyendo país, provincia, ciudad, nombre del proveedor de Internet, su administrador.





- **TraceRoute**: devuelve la máquina y la IP de cada salto que da un paquete desde la máquina original hasta la de destino por Internet. Además, también informa el tiempo en milisegundos que tarda éste.
- Ping: envía un echo request a una máquina específica en la red. Esto puede ser utilizado para comprobar la comunicación entre dos máquinas o para ver si el host específico está corriendo o existe.
- **NsLookup**: resuelve un hostname a dirección IP o viceversa.
- **ProxyTest**: comprueba si un Proxy es realmente anónimo. Este trata de reconocer la verdadera dirección IP incluso si ésta se encuentra detrás de un Proxy httpd.
- Environmental Variables Test: muestra varias configuraciones remotas del navegador y de nuestra máquina.

Muchos de los datos obtenidos a través de estas páginas, también pueden conseguirse directamente mediante la utilización de comandos de consola en un sistema operativo como Windows, Linux o Unix.



```
_ | | | | | | | | |
 cav cmd
Microsoft Windows XP [Versión 5.1.2600]
(C) Copyright 1985—2001 Microsoft Corp.
 C:\WINDOWS>tracert
Uso: tracert [-d] [-h saltos_máximos] [-j lista_de_hosts] [-w tiempo_de_espera]
nombre_destino
Opciones:
                                                        No convierte direcciones en nombres de hosts.
Máxima cantidad de saltos en la búsqueda del
          −d
−h saltos_máximos
                                                              objetivo.
Enrutamiento relajado de origen a lo largo de la
          -j lista-de-host
          lista de hosts.
-w tiempo_espera Cantidad de milisegundos entre intentos.
C:\WINDOWS>ping
Uso: ping [-t] [-a] [-n cuenta] [-l tamaño] [-f] [-i TTL] [-v TOS]
[-r cuenta] [-s cuenta] [[-j lista-host] ; [-k lista-host]]
[-w tiempo de espera] nombre-destino
                                        Ping el host especificado hasta que se pare.

Para ver estadísticas y continuar - presionar Control-Inter;

Parar - presionar Control-C.

Resolver direcciones en nombres de host.

Número de peticiones eco para enviar.

Enviar tamaño del búfer.

Establecer No fragmentar el indicador en paquetes.

Tiempo de vida.

Tipo de servicio.

Ruta del registro para la cuenta de saltos.

Sello de hora para la cuenta de saltos.

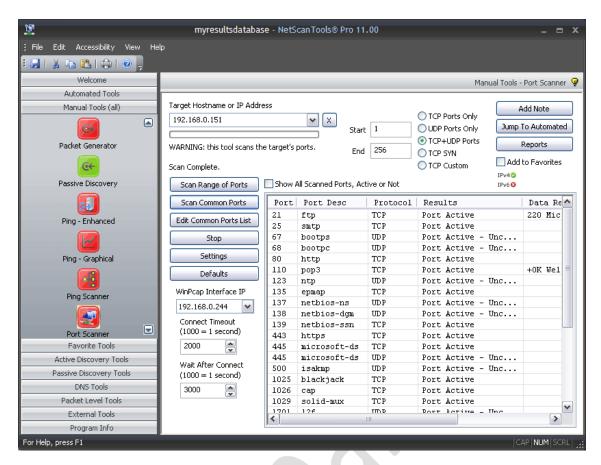
Afloja la ruta de origen a lo largo de la lista- host.

Restringir la ruta de origen a lo largo de la lista- host.

spera Tiempo de espera en milisegundos para esperar cada respuesta.
Opciones:
            cuenta
tamaño
           TTL
TOS
cuenta
           count
lista-host
lista-host
            tiempo de espera Ti
respuesta.
 C:\WINDOWS>_
```

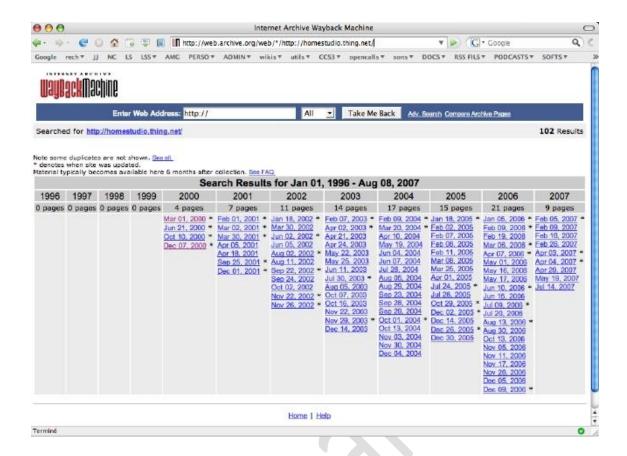
Otro modo de obtenerlos es con herramientas de interfaz gráfica en dichos sistemas. La diferencia entre utilizar estas últimas y hacerlo de forma online (a través de sitios) es que el intruso hábil difícilmente lo haga desde su máquina (a menos que sea a través de un servidor proxy o intermediario) para no dejar rastros en el objetivo de los comandos ejecutados y su búsqueda. El intruso inteligente jamás dejará su dirección IP real en algún log del objetivo, a diferencia del profesional ético que no ve ningún problema en hacerlo.





Para obtener información del tipo contenido histórico de sitios, podemos visitar, por ejemplo, www.archive.org. Éste nos muestra, en una línea temporal, las diferentes páginas web que tuvo una actual. Lo importante de esto es que muestra variaciones de contenido en el tiempo (imaginemos nombres de contactos de empleados, cuentas de e-mail, archivos o directorios sensibles, etcétera).







1.4 Encabezados de Correo Electrónico

Con el correo electrónico nos encontramos con varias situaciones que debemos de tener muy en cuenta. Empezando por la privacidad del mensaje. Habrá situaciones en las que no queramos que el contenido de un mensaje sea leído por nadie más que por la persona a la que va dirigida. Antes escribíamos cartas que no se podían leer sin romper el sobre o postales que podía leer cualquiera que cayese en sus manos. Normalmente también firmábamos las cartas ordinarias para que se supiera que éramos realmente nosotros quien las escribía y no otro. En el correo electrónico pasa con mayor razón ya que no es difícil aparentar otra identidad o que intenten falsificar la nuestra. Ya no vale que pongamos nuestro nombre al pie del mensaje o el remite con nuestros datos correctos. Tenemos que utilizar para ello otra forma que dé veracidad a la personalidad del que remite el mensaje. Necesitamos pues una "firma digital", Por ultimo también hay situaciones en las que no queremos que se conozca quién ha enviado un mensaje, o más comúnmente, que no se puedan obtener nuestros datos de forma automática cuando enviamos los mensajes a listas, foros o grupos de noticias.

Cuando enviamos un correo electrónico, lo que estamos haciendo, siguiendo con el símil del correo ordinario, es enviar una postal, no una carta. Esa carta puede pasar por múltiples servidores y redes que la pueden leer y por supuesto, está suficientemente identificada para saber quién la ha enviado (en esto, como en todo, también existen las falsificaciones de identidad).

Un mensaje de correo electrónico se compone de una cabecera con los datos identificativos del mensaje y el texto del mensaje. Así que vamos a centrarnos en cada una de las partes para hacer de nuestro correo un medio más seguro.

Cabecera del mensaje

Es evidente que se necesitan varios datos identificativos para que un mensaje llegue a su destinatario. Además, aquí no existe un solo servicio de correo, sino miles de ellos por los que puede pasar nuestro mensaje. Cada servidor por el que vaya pasando el mensaje irá incluyendo sus propios datos en la cabecera.

Esa cabecera es perfectamente legible por todos los ordenadores por los que vaya pasando (y por supuesto por nosotros mismos). Vamos a analizar qué es lo que hay en una cabecera con un ejemplo. Nota: La cabecera hay que leerla de abajo a arriba.

Dirección de respuesta	Return-Path: <mipc@dominio.org></mipc@dominio.org>
Servidores por los que ha pasado	Received: from mipc ([213.96.68.187])
	by aseara.com (8.9.3/8.9.3) with ESMTP id
	NAA20948
	for <usuario@destino.com>; Sun, 7 Oct 2001</usuario@destino.com>
	13:46:50 +0200
	Received: from [127.0.0.1] by mipc
	(ArGoSoft Mail Server, Version 1.3 (1.3.0.1)); Sun, 7
	Oct 2001 14:53:06 +0200
Identificación	Message-ID:
	<000901c14f2f\$02494ce0\$0700a8c0@dominio.ws>
	From: "Origen" <origen@dominio.org></origen@dominio.org>
	To: <usuario@destino.com></usuario@destino.com>
	Subject: prueba
	Date: Sun, 7 Oct 2001 14:52:47 +0200



Características del mensaje MIME-Version: 1.0

Content-Type: text/plain; charset="iso-8859-1"

Content-Transfer-Encoding: 7bit

X-Priority: 3

X-MSMail-Priority: Normal

X-Mailer: Microsoft Outlook Express 5.50.4133.2400 X-MimeOLE: Produced By Microsoft MimeOLE

V5.50.4133.2400

Status: R X-Status: N

En el ejemplo podemos ver entre otros datos cómo se ha enviado el correo y qué características tiene, con qué programa se ha hecho (X-Mailer: Microsoft Outlook Express 5.50.4133.2400), en qué fecha se ha enviado (Date: Sun, 7 Oct 2001 14:52:47 +0200), el asunto del mensaje (Subject: prueba), a quién va dirigido (To: usuario@destino.com), quien lo manda (From: "Origen" origen@destino.org) y qué identificador tiene ese correo (Message-ID: <000901c14f2f\$02494ce0\$0700a8c0@dominio.ws>)

Los siguientes campos, que empiezan por Received, nos indican todos los servidores por los que ha pasado el mensaje. El primer received y el último (siempre empezando por abajo) apuntan al servidor del emisor y del destinatario del mensaje. En este caso, y como ejemplo, hemos puesto uno con solo dos servidores de correo por los que ha pasado el mensaje (en este caso el emisor lo ha enviado desde un servidor que se ha montado en su propio PC sin utilizar ningún proveedor.

El primer received nos dice que lo ha enviado el PC llamado MiPC utilizando un programa servidor de correo de Argosoft corriendo sobre la dirección de localhost (127.0.0.1) y fue enviado por éste servidor unos cuantos segundos después de recibirlo por parte del cliente de correo.

El segundo Received nos dice que el mensaje ha sido recogido por el servidor de aseara.com, que se lo ha enviado a él un ordenador que se llama mipo que tenía en esos momentos una dirección IP determinada, para entregarlo al buzón del destinatario del mensaje.

Cuerpo del mensaje

Ahora entramos ya en el mensaje en sí. Como hemos indicado antes, todo correo enviado de la forma "normal" se envía en claro. Es decir, estamos enviando una postal, no una carta que tiene el texto protegido por un sobre para que no se lea, por lo que no tenemos ninguna garantía de que nuestra privacidad esté a salvo ya que, con medios y técnicas adecuadas, lo pueden leer mientras esta "por el camino". La única solución posible es enviar los mensajes cifrados. ¿Cómo?, pues para eso están los programas de criptografía como PGP o GNUPG. Hablaremos de criptografía en otro tema del curso.

Es posible que queramos enviar el mensaje en claro. Pues también tenemos que pensar en los problemas que eso pueda acarrear. Tenemos dos formas de enviar (o recibir) el mensaje. En texto plano o en formato html. Leer el correo en formato html funciona prácticamente igual que cuando visitamos una web. Por poner un ejemplo, históricamente algunos anunciantes, de los



que nos llenan el buzón, han utilizado sistemas como el web bugs o gráficos descargados directamente desde su servidor para saber si hemos abierto el mensaje, que eficacia tiene, etc.. Como html que es, prácticamente se puede hacer las mismas cosas que con una página web. Pueden abrir en frames ocultas páginas web en las que va a quedar registrada nuestra IP, incluyen javascripts... en fin, es largo la cantidad de cosas que se puede hacer con el correo en formato html.

Si, por el contrario, configuramos nuestro cliente de correo para recibir el correo en formato texto, nada de eso puede suceder. El texto es eso, solo texto. El único inconveniente será que aquel que envíe el correo en html, veremos también los tag del mismo y a veces se hace muy difícil la lectura. Por lo tanto un consejo, es enviar el correo en formato texto.



