Anonimato y privacidad en las redes P2P

REDES LIBRES

Las "redes libres" son el resultado tecnológico de la necesidad de un modelo de comunicación en Internet que no sea sensible a la censura, y que pueda mantener el anonimato y la privacidad de cada uno de sus componentes.

POR DAVID GASCÓN

Referirse a redes de Pares "P2P" es sinónimo de hablar de una de las principales razones de acercamiento de la gente a Internet. Al igual que la mayoría de las conexiones en la red, se basa en interacciones directas entre las diferentes máquinas que participan en este fluir de cantidades ingentes de información.

Pero, al igual que cualquier otra interacción con la red, al operar en las redes de pares, vamos dejando *huellas*, pequeñas cantidades de datos que permiten a entidades privadas o gubernamentales seguir la pista a toda transacción que se realice. A diario se llevan a cabo acciones de investigación para transacciones de usuarios perfectamente legítimas en la red que en el mundo "real" exigirían una orden judicial o directamente violarían el derecho de la intimidad de los ciudadanos.

Y hay que desterrar el mito de que en la red nadie sabe quien eres. Todo lo que se hace queda registrado en algún lugar. Sólo hay que tener permiso para recabar la información.

Falacias sobre el Anonimato en la Red

Todo el mundo que posee conexión a Internet carece de anonimato en su navegación por la red, para ello desmentiré algunas ideas erróneas para la gente que piensa que posee anonimato en la red.

- *Yo no tengo contrato con un proveedor, me conecto mediante modem a un proveedor gratuito, por lo que nadie sabe quien soy.*

Falso: al hacerlo usas una conexión telefónica que identifica unívocamente al dueño de la línea.

- *Yo no tengo IP fija, mi proveedor me da una distinta cada vez que conecto, por lo que nadie puede saber quién soy*

Falso: tu proveedor guarda la información de cuál es la IP que usas en cada momento e información relativa a su uso.

- *El uso de proxies anónimos como Anonymizer me permite navegar anónimamente*

Falso: este tipo de servicios lo que consigue es concentrar un número alto de usuarios que desean anonimato en un único punto; centralizar a estos usuarios es una de las mayores trampas, pues facilita el trabajo de las instituciones que desean saber quién y para qué requiere anonimato.

En el modelo actual de comunicación todo se basa en "conexiones directas"; veamos en qué consisten y cuáles son

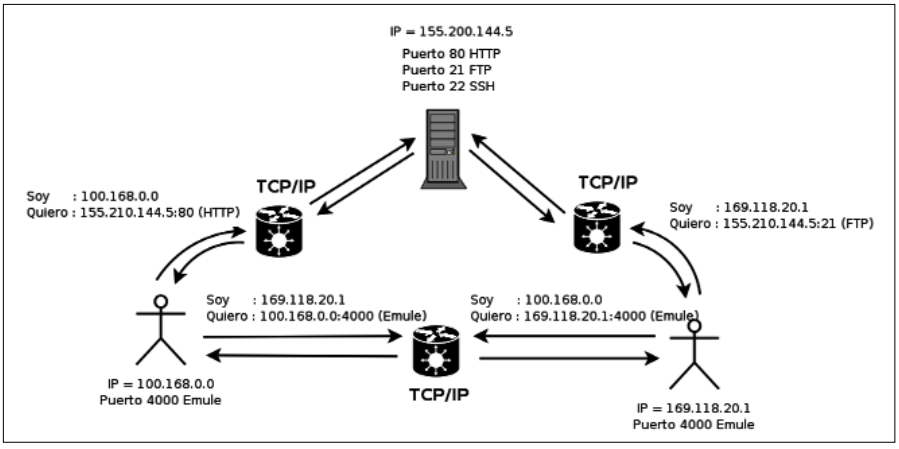

Figura 1: Conexiones directas entre cliente y servidor.

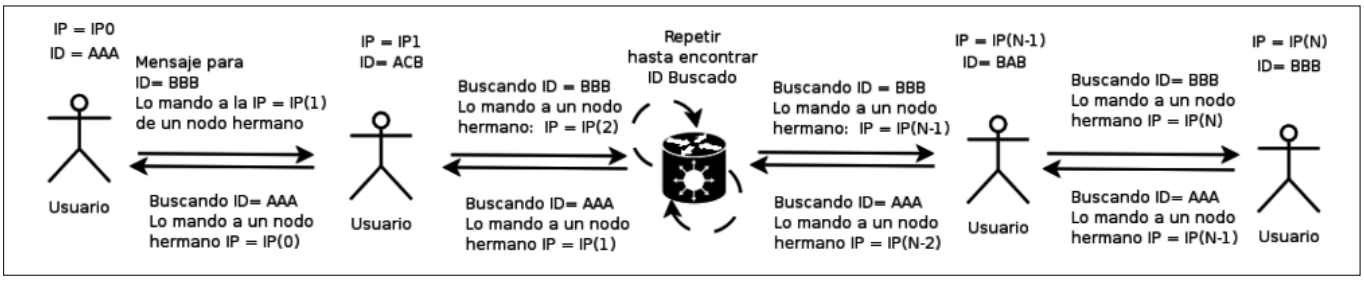

Figura 2: Conexiones indirectas entre nodos de la red libre.

los problemas que conllevan a la hora de pensar en términos de anonimato y privacidad.

Todo el mundo sabe quien eres

Comencemos con un ejemplo: Cuando el nodo A quiere adquirir un servicio del nodo B, genera una petición que lleva la dirección donde tiene que dirigirse, en este caso, la IP y puerto del nodo B, y la dirección y puerto que especifican dónde se ha de devolver la información al nodo A.

En el camino intermedio entre el nodo A y el B, son los protocolos como TCP/IP y UDP/IP los que se encargan del correcto enrutado de la información hasta el destino, pero siempre usando las direcciones IP de ambos clientes.

Únicamente con la información que guarda esa petición podemos identificar quién está haciendo uso del servicio, y cuando digo "quién" me refiero a la persona que tiene contratada el servicio de acceso a Internet, pues los Servidores de Internet (ISPs) guardan una referencia completa (por ley) de los accesos que sus usuarios hacen cada vez que realizan una petición a la Red.

A su vez, el servidor que ofrece el servicio almacena la petición que hizo el nodo A, con la posibilidad, dependiendo del servicio, de que se guarde información del uso del mismo como las páginas que visitó, los ficheros que descargó o en qué sala de chat estuvo hablando.

La Figura 1 refleja el modelo que acabamos de ver sobre las *conexiones directas* entre máquinas. En ella vemos reflejada cuál es la información necesaria para poder acceder a los servicios de una máquina en la red y qué otras puedan mantener conexiones con nuestro nodo.

Como se puede ver, en el paso intermedio, relegamos en protocolos como TCP/IP para llegar al nodo final a través de las distintas máquinas especializadas en enrutado de datos a través de la red.

En este modelo únicamente son los *routers* los que se encargan de mover la información y hacer que llegue al destino.

Vemos cómo los nodos clientes hacen peticiones a un servidor, mientras entre ellos realizan intercambio de ficheros mediante una plataforma P2P. Podemos observar cómo las peticiones están hechas usando las direcciones IP reales de origen y destino. En dicha figura además se ve cómo el servidor se encarga de centralizar las peticiones de los clientes y servirlas.

Este es un modelo *sensible a la censura*, ya que valdría con eliminar la máquina servidora para privar a todos los usuarios de la información y servicios que está ofreciendo, y con controlar las máquinas enrutadoras para controlar el tráfico de entrada o salida a una determinada red.

La idea que ha rodeado a Internet desde sus inicios es la eficiencia de las comunicaciones, sin embargo esta misma topología y forma de funcionar ha hecho que países como China sean capaces de censurar cualquier tipo de información externa, ya que todas las peticiones pasan por sus máquinas de enrutado antes de llegar al destino, y es aquí donde se pueden rechazar las comunicaciones de los nodos con el exterior de China.

Internet hoy en día es una red que no nos asegura el anonimato ni la privacidad. Además la información y los servicios están centralizados en máquinas concretas con lo que, para cortar esa fuente de comunicación es suficiente con eliminar la comunicación de esas máquinas con el resto de la red.

Por todas estas razones nacieron las redes anónimas, las cuales me gusta denominar *"Redes Libres"* ya que devuelven a sus usuarios la posibilidad de moverse libremente por ellas y acceder a la información de forma que nadie puede ejercer control sobre ellos o sobre la información que se está transmitiendo. Veamos en qué consisten.

Una nueva topología de la Red

Antes de empezar con los entresijos de esta nueva *topología de red*, me gustaría aclarar algunos conceptos que harán que el lector entienda mejor algunas de las decisiones tomadas para su implementación.

Lo primero es reafirmar que el modelo actual es sensible a la censura, ya que las *conexiones directas* son la mejor manera de conseguir redes eficientes. El modelo de hoy en día es un modelo donde los algoritmos de enrutado buscan un solo propósito: mover información de la manera más rápida minimizando al máximo las latencias entre petición y servicio.

Como contrapartida, las redes que presento, no son tan eficientes como las actuales, ya que tienen que arreglárselas para dotar a cada uno de los nodos de anonimato y privacidad partiendo del sistema de conexiones actuales entre nodos. Ello se conseguirá gracias a un nuevo factor que cambia por completo los modelos actuales de enrutado de datos; desde ahora podemos decir que *cada servidor es un Nodo y cada Nodo es un Servidor*.

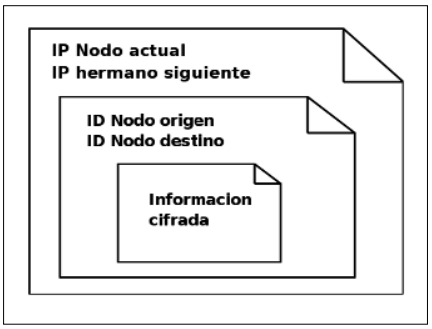

Figura 3: Estructura general de los mensajes.

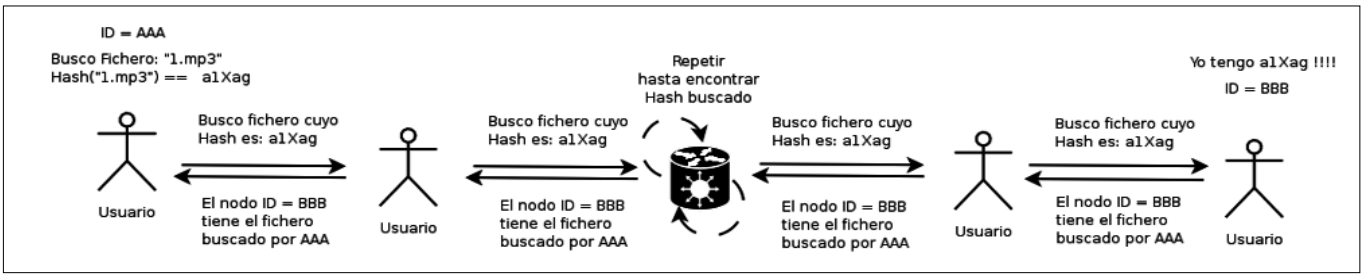

Figura 4: Sistema de búsqueda de ficheros manteniendo la privacidad de la información buscada.

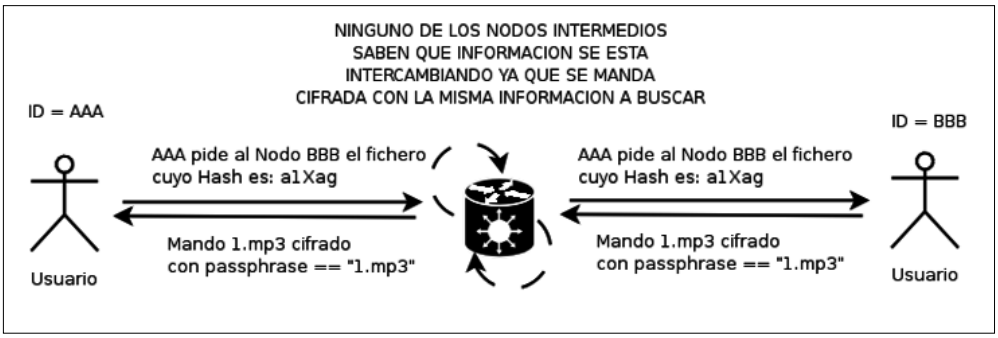

Figura 5: Cifrado de la información enviada.

En las Redes Libres todas las máquinas tienen un papel dual:

- **Como Cliente:** Hacer peticiones a la Red, y gestionar las respuestas.
- **Como Servidor:** Enrutar los datos que nos mandan otros nodos y que no son para nosotros.

Esto hace cambiar la perspectiva que hemos tenido hasta el momento, pues ahora todos nosotros como **nodos** de la red tenemos la obligación de participar en el enrutado de la información de nuestros nodos hermanos de forma que se puedan crear las capas de abstracción necesarias para conseguir el anonimato, ya que como consecuencia inicial de este método la única forma de poder "seguir la información" sería pinchando todos los nodos de la red, lo cual es una tarea imposible que hace que el modelo actual consiga su propósito.

A primera vista, si se está familiarizado con los protocolos de transporte TCP/IP y UDP/IP, podríamos pensar que los nodos "enrutadores" podrían conocer igualmente la dirección origen y destino. Sin embargo esto no es así, ya que hemos introducido una capa de abstracción entre la petición y el contenido de la capa de transporte de las tramas.

Desde el nodo inicial que origina la petición se sitúan como dirección inicial y final *identificadores* que sólo tienen sentido dentro de la red libre. Este identificador es único para cada nodo y es diferente en cada sesión que iniciamos en la red.

El nodo inicial lanza una petición a sus nodos hermanos. Para ello encapsula la información incluyendo los IDs de inicio y final dentro del cuerpo del mensaje (al igual que si empaquetara información para ser enviada). Después, encima de esa información, añade el resto de cabeceras necesarias para que el transporte hacia el nodo hermano sea exitoso.

El comportamiento general de un nodo ante la llegada de un mensaje es:

- Primero mira si el ID destino coincide con el suyo
- Si es así, responde al servicio y lanza la respuesta a sus hermanos con la dirección de destino el ID del nodo origen.
- Si no es esa ID vuelve a lanzar la petición a todos sus hermanos de forma que continúe su camino en busca del nodo cuya ID coincide con la buscada.

En la Figura 2 se puede ver un ejemplo de cual es la información que van a usar a la hora de llevar a la práctica el algoritmo de enrutado de las redes libres. Como se puede observar nunca se revela la IP del origen o del destino a los nodos intermedios de la red, sino que se usa un identificador (ID), lo que nos va a permitir que nadie pueda conocer quién hace la petición o quién la responde.

Asuntos Privados

Hasta ahora hemos hablado de "anonimato" en la red; sin embargo estas plataformas también nos proporcionan privacidad o, lo que es lo mismo, además de que no sepan quién manda algo, que ignoren qué es lo que se está mandando por la red a no ser que sean ellos mismos los destinatarios. Esto se consigue mediante el uso de los sistemas de *cifrado*.

La forma más segura consiste en aplicar varias capas de cifrado a la información que viaja por la red, de forma que sólo el

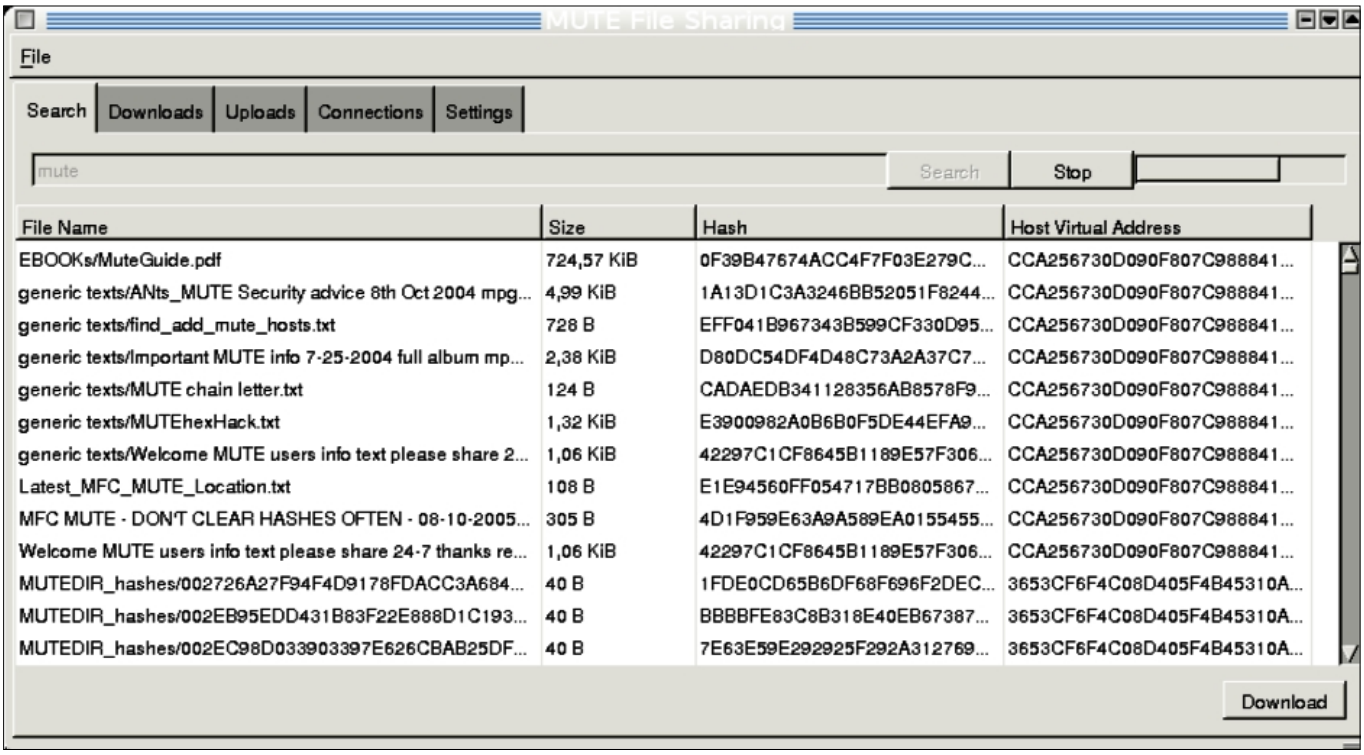

Figura 6: Ejemplo real de plataforma de redes libres: Mute.

nodo destinatario del mensaje sea capaz de descifrarlo.

Mediante la Figura 3 podemos hacernos una idea de los distintos niveles de "metainformación" que rodean a los datos que se transmiten en una red libre. Lógicamente, al añadir capas extra se incrementa el tiempo de "análisis" de cada una de las diferentes tramas, por lo que se ve aumentada la latencia de la transmisión de las tramas por la red, así como el uso de recursos en cada uno de los nodos que participan de forma activa. Este es el precio que hay que pagar para obtener la protección de la privacidad en red.

Hay muchísimas formas de conseguir cifrado a través de la red. Algunas de ellas son el uso de claves públicas, privadas entre nodos hermanos, o el uso de algoritmos de hashing como MD5 sobre las palabras clave de forma que obtengamos "passphrases" para el cifrado.

Veamos un ejemplo de esta última idea (ver Figuras 4 y 5). Supongamos que queremos buscar un archivo llamado "1.mp3". Lo primero que vamos a hacer es calcular el hashing del nombre del fichero mediante el algoritmo elegido para ello (MD5, RSA, ElGamal, etc), por ejemplo, *Hash("1.mp3") == "a1Xag"*, este identificador del nombre del fichero hará que cuando hagamos la petición, sólo los nodos que posean ese fichero sepan lo que está buscando el nodo que inició el mensaje (el cual recordemos, nadie, excepto él mismo sabe quién es).

Una vez se establece la conexión entre los nodos lo que se hace es usar la palabra clave original "1.mp3" como "passphrase" de cifrado, de forma que sólo pueden descifrar la información el nodo inicial y el nodo final. Para este paso usamos un algoritmo de cifrado simétrico, de forma que el nodo origen pueda transformar los datos cifrados que le van llegando a su forma original. Algunos de los posibles algoritmos que se pueden usar son: DES, AES, Blowfish …

La Figura 6 muestra un ejemplo real de una de las plataformas de redes libres. En ella podemos observar ejemplos reales de *hashing de nombres ficheros* y del uso de *Identificadores*, dos de las piezas claves en el funcionamiento de las redes libres como hemos visto hasta el momento.

Proyecto APEIRON

Dada la importancia de salvaguardar el derecho que tenemos a la libertad de expresión y para luchar contra la creciente censura a la que se ve sometida la Red, hemos creado el proyecto APEIRON [1], como punto de encuentro para la gente que queremos luchar activamente por ello.

En la sección de "Documentación" quien esté interesado puede encontrar información detallada sobre muchas plataformas de redes libres como Mute, Freenet, Gnunet, Konspire, etc. Cada una de las cuales difiere del resto por centrar su desarrollo en alguna característica concreta de los modelos de transporte y distribución de la información.

Mediante la lista de correo se puede participar de forma activa, además de estar informado de los diferentes eventos, charlas y talleres que vamos organizando por las diferentes universidades y ferias tecnológicas. ■

RECURSOS

[1] Proyecto APEIRON: *http://apeiron.laotracara.com*