Fundamentos de Redes (2017-2018)
Doble Grado en Ingeniería Informática y Matemáticas
Universidad de Granada

Invisible Internet Project (I2P)

Antonio Coín Castro
Alberto Jesús Durán López

27 de noviembre de 2017

Índice

0. Motivación

El incremento en el uso cotidiano de las nuevas tecnologías conlleva una serie de riesgos ocultos a la mayoría de los usuarios, que normalmente no son conscientes de los mismos, sino que se centran en el uso de las mismas sin preocuparse de otros factores.

Es por esto que vemos de vital importancia poner de manifiesto los puntos negativos que afectan a los usuarios de Internet, sobre todo en cuanto a privacidad se refiere. Este problema en particular es preocupante, pues cada vez que nos conectamos a la red cedemos una gran cantidad de datos que en manos equivocadas pueden suponer un grave peligro para nosotros.

Afortunadamente, existen diversos métodos que, en mayor o menor medida, nos ayudan a prevenir esta cesión indiscriminada de datos, sin privarnos de la facilidad y comodidad del uso corriente de Internet. En este trabajo nos centraremos en un software en particular, llamado **Invisible Internet Project (I2P)**, y también hablaremos en general sobre redes de anonimato. Sin embargo, antes explicaremos con un poco más de detalle los riesgos concretos en un acceso corriente a Internet, así como otros métodos menos sofisticados para protegernos.

1. Anonimato en el acceso corriente a Internet

¿Alguna vez te has planteado qué ocurre cuando abres el navegador y te conectas a una web de Internet? En teoría es un proceso sencillo de cara al usuario, algo que hacemos de forma casi automática. Sin embargo, lo que en realidad ocurre en las capas interiores es un proceso complejo, y tratando de entenderlo podemos condicionar nuestro comportamiento en la red para intentar proteger nuestra privacidad.

Para ilustrar el proceso de conexión de forma muy simplificada, podemos pensar en una plaza o un parque de una ciudad, y personas entrando y saliendo del mismo. No es extraño pensar que al conectarnos a Internet realizamos un proceso análogo a entrar en la plaza, pues podemos ver que hay otros usuarios conectados, y relacionarnos con ellos si así lo queremos. Pero no en realidad no es exactamente lo mismo, pues en el primer caso somos conscientes de los riesgos que conlleva entrar al parque, y solo compartimos la información que nosotros queremos en cada momento. En Internet, automáticamente estamos cediendo una gran cantidad de información, como si fueramos al parque con una hoja con nuestros datos personales en alto, y cualquier persona pudiera, desde la distancia, saber a qué nos dedicamos, por qué hemos venido al parque, nuestra dirección, etc.

Uno de los mayores problemas es la huella digital que dejamos inconscientemente al realizar cualquier conexión, compuesta por diferentes elemntos que pueden ser potencialmente usados para identificarnos. Un elemento destacable es nuestra *dirección IP*, la cual es fácilmente detectable y nos identifica de forma unívoca en Internet. Además, es posible **geolocalizar** el dispositivo desde el cual se está accediendo, simplemente con saber la dirección IP (por ejemplo, en `whatismyipaddress`).

Es cierto que mucha gente se conecta a Internet con una IP temporal asignada para una sola sesión. Sin embargo, dichas direcciones son registradas por los proveedores de

servicio (*ISP*), por lo que sería posible descubrir quién usó una cierta dirección IP en un intervalo tiempo concreto.

Existen otros problemas derivados, como el envío del historial de búsqueda, cookies de rastreo, y otra información almacenada en nuestro navegador que nos puede identificar en la red. En general, compartir esta información no tiene por qué ser dañino para nosotros, pero en un entorno tan grande como Internet, toda precaución es buena.

2. Métodos para mantener el anonimato

A la hora de ocultar nuestro rastro en Internet existen diferentes alternativas, con diferentes niveles de efectividad.

2.1. Proxy

Se trata de un servidor intermediario entre las conexiones de un cliente y un servidor, centrado en la navegación. De esta manera, al visitar una página web se puede enviar y recibir información a través de un proxy, y así ocultar nuestra dirección IP a la página de destino, que verá la IP del proxy. Puede utilizarse para acceder a contenido bloqueado en algunos países, como medio de evitar la censura.

2.2. VPN

Es un acrónimo para *Virtual Private Network*, red privada virtual. Se trata de un medio de extender una red privada a través de una red pública, de forma que simula que un usuario conectado a través de VPN forme parte físicamente de la red privada. Esto proporciona diversas ventajas:

- Se puede acceder remotamente a los recursos de la red privada como si se formase parte de ella.
- Se camufla la dirección IP, pues la dirección que se ve es aquella que proporciona la red virtual.
- Se garantiza la confidencialidad, de forma que los paquetes intercambiados en la red están encriptados.
- Existe un sistema de autentificación para conectarse a la red e impedir accesos no deseados.
- También se proporcionan mecanismos para mantener la integridad de los mensajes y detectar mensajes fraudulentos.

2.3. Redes de anonimato

Una red de anonimato permite a los usuarios acceder a la web de forma que se intenta bloquear cualquier seguimiento de su identidad en Internet. Para ello, se canaliza el tráfico a través de una red global de servidores voluntarios.

Son el objetivo central de este trabajo, y las discutiremos con detalle posteriormente.

2.4. Otros

Existen otros mecanismos, menos sofisticados pero muchas veces muy efectivos, como por ejemplo utilizar buscadores que respetan la privacidad (DuckDuckGo), o emplear sistemas operativos orientados específicamente a la seguridad y la privacidad (Tails).

3. Invisible Internet Project. Estructura, funcionamiento y seguridad

I2P es una herramienta de **software libre** que ofrece una capa de red de abstracción **distribuida** para comunicaciones entre ordenadores, la cual permite a las aplicaciones que la utilizan transmitir mensajes de forma anónima y segura. Se trata de una red superpuesta (del inglés *overlay network*), que funciona con nodos virtuales en una red construida sobre Internet, cuyas conexiones en la red subyacente pueden ser estar ubicadas en zonas geográficamente muy alejadas. Su misión es conseguir que los datos lleguen desde el origen al destino aunque no exista una conexión directa entre los nodos.

En este caso, la selección de ruta o *routing* que ofrece I2P se basa en una serie de nodos conectados a la misma red, que crean caminos temporales y unidireccionales (túneles), siguiendo la técnica de enrutamiento *garlic*, que veremos a más adelante

Además, todas las comunicaciones están encriptadas *punto a punto*, lo que añade una capa extra de seguridad a nuestras aplicaciones.

3.1. Nodos y túneles

En la terminología de *I2P*, el software que implementa la capa de red se denomina **router**, y un ordenador ejecutándolo se denomina **nodo**. Para garantizar que los mensajes enviados sean anónimos, se construye para cada cliente una serie de *túneles* de llegada y de salida. Estos túneles son una secuencia de nodos que pasan los mensajes de unos a otros, de manera unidireccional. La red es, por tanto, basada en mensajes, de forma similar al protocolo **IP**.

El mecanismo es el siguiente: cuando un cliente quiere enviar un mensaje a otro cliente, lo pasa a uno de sus nodos de salida que apunta a uno de los nodos de llegada del otro cliente, de forma que después de viajar por el túnel, el mensaje llega a su destino. Cada participante de la red elige la longitud de los túneles, de forma que se alcance un equilibrio entre el anonimato, la latencia y el *throughput* que lo satisfaga. El resultado es que el número de conexiones es el mínimo necesario para satisfacer las necesidades del emisor y del receptor.

A la hora de implementar la conexión, la primera vez que un cliente quiere contactar a otro cliente, realizan una consulta a la "base de datos de redes", que es una tabla *hash* distribuida basada en el algoritmo Kademlia, que explicaremos a continuación.

Cuando un router I2P quiere contactar con otro router, necesita saber una información mínima, como la *identidad* del router (una clave de encriptación de 2048 bits) o la dirección de contacto donde se encuentra.

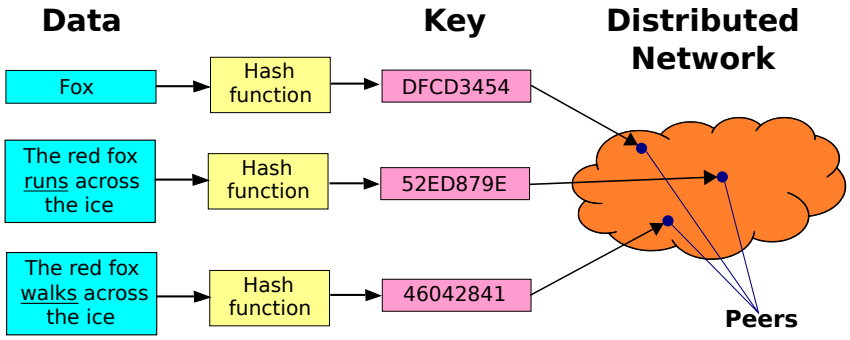

Figura 1: Tabla hash distribuida (DHT)

La búsqueda en la tabla se realiza para encontrar de forma eficiente los túneles de entrada del nodo destino, pero los siguientes mensajes entre el emisor y el receptor normalmente incluyen esta información, por lo que no son necesarias nuevas búsquedas.

Por último, cabe mencionar que los mensajes enviados entre los routers están definidos por el **protocolo I2NP**, cuya especificación completa se puede consultar en [**?**].

Algoritmo Kademlia

La base de datos de red (*netDB*) se distribuye mediante una técnica simple llamada "*floodfill*", donde un subconjunto de los routers mantienen la base de datos (cada uno tiene una parte).

A la hora de realizar una búsqueda desde un nodo, se realiza una consulta al router "floodfill"más cercano. La medida de cercanía se computa siguiendo una modificación del algoritmo Kademlia. Este algoritmo emplea la función XOR entre dos claves de routers para determinar cómo de cerca están uno de otro.

La función XOR ($\oplus$) funciona como una *distancia* en el espacio de las claves de los nodos:

- La distancia entre un nodo y él mismo es 0.
- La distancia es simétrica: $A \oplus B = B \oplus A$.
- Se verifica la desigualdad triangular: $A \oplus B \leq (A \oplus C) + (C \oplus B)$

Una red con $2n$ nodos que emplee este algoritmo en su forma más básica necesitará solo n pasos (en el peor caso) para encontrar un nodo en concreto.

3.2. Enrutamiento tipo *garlic*

En la tecnología I2P, es necesario comprender qué significa *Garlic*, cuyo término, al no ser preciso, se le puede atribuir a 3 conceptos diferentes.

Cifrado por capas

En esta técnica se construyen los correspondientes túneles en los que la información, previamente cifrada por el emisor, viaja a través de una serie de nodos. Para explicar la comunicación entre 2 usuarios, nos ayudaremos de un pequeño esquema:

Figura 2: Comunicación a través de los nodos del túnel

Si Alice quiere enviarle un mensaje a Bob, envía el mensaje a uno de sus túneles de salida que está conectado a uno de los túneles de entrada de Bob.

El primer nodo del túnel se conoce como *gateway*, y su función es la de fragmentar y empaquetar mensajes I2P en mensajes de túnel de tamaño fijo, que más adelante son encriptados. Los mensajes del túnel están formados por:

- ID del tunel: identificadores del túnel de un tamaño de 4 bytes.
- Vector de inicializacion de 16 bytes: hace que cada mensaje sea único.
- Instrucciones de envío.
- Relleno (datos inútiles desechables criptográficamente)

Una vez que los mensajes han sido procesados, el *gateway* crea un valor IV aleatorio de 16 bytes, encriptándolo junto al mensaje y enviando al siguiente nodo la tupla {ID del túnel, IV, mensaje del túnel}.

Este proceso es repetido por cada nodo hasta el nodo final del túnel de salida, que recuperará los datos prepocesados iniciales. En cada paso, se desencripta solo la capa superior del mensaje, que proporciona información sobre el siguiente nodo al que debe ser transmitido.

Agregación de múltiples mensajes juntos

En *Onion Routing* los mensajes se encapsulan en capas de cifrado, a diferencia de *Garlic Routing* que encripta varios mensajes juntos para que el análisis de tráfico realizado por los atacantes resulte más difícil, así como para aumentar la velocidad de transferencia de datos.

Cifrado utilizando ElGamal/AES

Es una herramienta para cifrar la información que usa I2P cuyo funcionamiento explicaremos más adelante.

3.3. Encriptación de la información

Se trata del mecanismo empleado por cada uno de los routers en cada canal de comunicación para cifrar los paquetes que viajan entre cada uno de ellos. Todos los mensajes son cifrados utilizando el algoritmo de cifrado *ElGamal* y posteriormente se utiliza *AES+SessionTag*.

Una de las consecuencias más importantes de la encriptación es que, al estar oculto el contenido de los mensajes enviados (la información que se quiere transmitir, y también otros campos necesarios para establecer la comunicación), **se enmascara la dirección IP del nodo que envía el mensaje**.

- **ElGamal:** Algoritmo de criptografía asimétrica. Su funcionamiento se basa en cálculos sobre logaritmos discretos usando un número primo y dos enteros.
- **AES:** Es lo que se conoce como un cifrado simétrico por bloques, es decir, cifra y descifra los datos en bloques de 128 bits cada uno. Para ello utiliza una clave criptográfica específica que puede ser de 128, 192 o 256 bits de tamaño. Dicha clave se especifica al nombrar el cifrado usado (AES128, AES256...). Además, a cada bloque de texto se le aplica una operación XOR con el bloque previo ya cifrado, haciendo cada mensaje único con el uso de un vector de inicialización en el primer bloque. Esto es lo que se conoce como *Cipher Block Chaining - CBC*.

Procedimiento

La primera vez que un usuario quiere enviar un mensaje para otro usuario, ambos cifran las claves para una clave de sesión AES256 con *ElGamal* y se adjuntan los bloques que han sido cifrados con AES256/CBC, que se van cifrando de nuevo cada vez que va pasando por cada nodo del túnel con el siguiente algoritmo:

- Se cifra el IV (vector de inicialización) recibido con AES256/ECB usando su clave IV, obteniendo así el nuevo IV.
- Usa el IV obtenido con la clave de capa del participante para cifrar los datos.
- Cifra el IV actual con AES256/ECB usando la clave IV de nuevo.
- Envía la tupla {ID del túnel, IV, mensaje del túnel} al siguiente nodo.

Cuando se produce la comunicación de dos usuarios, se crean un número de *SessionTag* (etiquetas de sesión) que pueden usarse si el remitente desea cifrar un mensaje *garlic* para otro router. Dichas etiquetas de sesión pueden ser usadas únicamente una vez para evitar vulnerabilidades, y además tienen un ciclo de vida muy corto.

4. Otras redes de anonimato: la red Tor

Una de las redes anónimas más conocidad y utilizadas es la **red Tor**. Esta red cuenta con una base de usuarios notablemente más extensa que I2P, pero a la hora de implementar las comunicaciones hay algunas diferencias.

4.1. *Onion routing*

La red Tor utiliza un mecanismo conocimo como *onion routing* para transmitir mensajes, en contraposición al ya mencionado *garlic routing* que emplea I2P. En esencia, el enrutamiento tipo *garlic* es una modificación del *onion*, pues se basa en la misma idea.

Con la técnica de *onion routing*, se cifran los mensajes en varias capas, de forma que cada nodo desencripta únicamente una capa, que contiene las instrucciones para enviar el mensaje al siguiente nodo, sin tener acceso al mensaje en sí. El nombre proviene de la analogía de las capas de una cebolla, pues en cada paso se "quita" una capa.

La técnica de *garlic routing* es similar, con la diferencia de que permite enviar varios mensajes agrupados en un mismo paquete, que se denomina *garlic clove* (diente de ajo). Así, por ejemplo, un nodo puede enviar un mensaje a otro, agrupando también en dicho mensaje la dirección de alguno de sus túneles de entrada. De esta forma se permite que el nodo receptor envíe una respuesta al nodo emisor sin tener que realizar una búsqueda en la base de datos (*netDB*).

4.2. Túneles bidireccionales

La red Tor implementa túneles bidireccionales, mientras que I2P hace uso de túneles unidireccionales. Aunque no está demostrado que los segundos proporcionen más seguridad que los primeros, en principio con túneles unidireccionales se necesitaría el doble de esfuerzo para acceder a la misma cantidad de datos que se podría acceder con túneles bidireccionales (mensaje-respuesta).

4.3. Directorio centralizado

Tor no es una red P2P distribuida pura, pues emplea una serie de servidores de confianza, establecidos por la propia red, que se encargan de mantener la base de datos en red. La IP de estos servidores es pública, y los nodos de la red deben pasar por ellos para establecer conexiones. Esto puede ser explotado de diferentes maneras por atacantes potenciales, e incluso puede ser utilizado para imponer **censura** en la red: basta con bloquear de forma global el acceso a dichas direcciones IP.

4.4. *Outproxies*

Tor está diseñada y optimizada para permitir el acceso a Internet a través de la red interna. Los nodos con salida a Internet se denominan *exit nodes*, y suponen una vulnerabilidad añadida. Un atacante malicioso posicionado en un nodo de salida podría realizar lo que se conoce como *eavesdropping* (escuchar a escondidas), para capturar el tráfico entre dicho nodo y el servidor de destino.

I2P también proporciona mecanismos de salida a Internet. Sin embargo, la atención se centra en comunicaciones dentro de la red, por lo que el número de nodos con salida a Internet es bastante más pequeño.

5. Software I2P

Una vez tengamos configurado I2P en nuestro equipo, podremos conectarnos a la red y así aprovechar todo el software construido sobre ella que se encuentra disponible. Entre las diferentes aplicaciones, destacan:

- **Susimail:** Interfaz web simple para emails, centrada en la seguridad y el anonimato.
- **Syndie:** Principal aplicación de distribución de contenidos diseñada para blogs, noticias y foros para I2P.
- **I2P Messenger:** Cliente de mensajería instantánea con cifrado de extremo a extremo. Los mensajes no se almacenan en servidores para asegurar el anonimato.
- Navegación web mediante webs anónimas.
- Compartición de archivos mediante el uso de BitTorrent dentro de la red I2P.
- También en Android...
  Nightweb, aplicación que usa I2P y BitTorrent para compartir entradas de blogs, fotos y otros contenidos similares.

Referencias

[1] I2NP Specification