# Provisión de Anonimato en Redes P2P

Juan Pedro Muñoz Gea, Pilar Manzanares López, Juan Carlos Sánchez Aarnoutse
Escuela Técnica Superior de Ingeniería de Telecomunicación. Universidad Politécnica de Cartagena
Campus Muralla de Mar. Edificio Antiguo Cuartel de Antigones
30202 Cartagena, Teléfono: 968 33 8871 Fax: 968 32 5973
E-mail: {juanp.gea, pilar.manzanares, juanc.sanchez}@upct.es

***Resumen***. *En este trabajo se presenta un nuevo mecanismo para proporcionar anonimato en redes peer-to-peer (P2P) de compartición de ficheros. El anonimato se consigue conectando los nodos origen y destino a través de un conjunto de nodos intermedios, creando un camino multi-salto. La principal contribución es un algoritmo distribuido capaz de garantizar el anonimato incluso cuando un nodo en el camino falla. El algoritmo tiene en cuenta los costes asociados con las comunicaciones multi-salto e intenta alcanzar un compromiso entre el grado de anonimato y los costes asociados.*

## 1 Introducción

Las redes P2P son las arquitecturas más populares para la compartición de ficheros. En algunos de estos escenarios, los usuarios están interesados en conseguir anonimato mutuo, es decir, ningún nodo en la red debería poder conocer quien es el origen o el destino exacto de un mensaje. Tradicionalmente, el anonimato se obtiene mediante la conexión de los nodos origen y destino a través de un conjunto de nodos intermedios, creando un camino multi-salto. Un mecanismo de anonimato con este funcionamiento es Crowds [1], y en él, el nodo origen reenvía su mensaje a un nodo elegido aleatoriamente con probabilidad $p$ (siendo $p$ un parámetro del sistema), o directamente al destino con probabilidad $1 - p$. Este proceso se repite de forma recursiva hasta que el mensaje llegada al destino.

Por lo tanto, los costes que conlleva utilizar caminos multi-salto para proporcionar anonimato son: consumo extra de ancho de banda y sobrecarga adicional de los nodos. Por otra parte, los nodos son propensos a desconexión imprevistas. En una red anónima esto es un grave problema porque la conexión entre dos nodos se interrumpirá con frecuencia, aunque ambos nodos estén activos.

En este trabajo presentamos un mecanismo de anonimato para una red P2P híbrida presentada en un trabajo previo [2], basado en Crowds para crear caminos multi-salto. Sin embargo, nuestro mecanismo introduce una longitud límite máxima en el proceso de creación del camino, con el objeto de limitar la sobrecarga de los nodos. La principal contribución del trabajo es un algoritmo distribuido para restaurar un camino cuando un nodo falla.

## 2 Arquitectura de Red

La figura 1 muestra la red híbrida propuesta en un trabajo previo. Todos los nodos están inmersos en una red estructurada P2P y se agrupan en varios subgrupos. Cada subgrupo está gestionado por un superpeer, y estos superpeers pueden comunicarse entre ellos utilizando un servicio multicast proporcionado por la red estructurada. Todos los nodos notifican sus recursos de CPU y ancho de banda a sus superpeers, y estos forman una lista ordenada de futuros candidatos a superpeer. Esta lista es transmitida a todos los miembros del subgrupo.

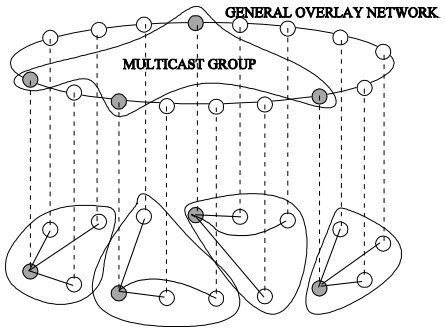


Fig. 1. Arquitectura de red.

## 3 Provisión de Anonimato

Nuestra solución se divide en 3 fases: publicación, búsqueda y descarga.

### 3.1 Fase de Publicación

Cuando un nodo desea publicar un contenido, en primer lugar elige un identificador de conexión aleatorio y al mensaje inicial le agrega un parámetro *TTL* (Time to Live). A continuación se utiliza el mecanismo de Crowds para crear un camino, pero en este caso cada nodo intermedio decrementa el *TTL* en 1, y cuando este parámetro es igual a 1, el mensaje se envía directamente al superpeer. Para mantener actualizados los caminos aleatorios de acuerdo a las desconexiones de los nodos cada nodo mantiene un temporizador, y cuando expira, el nodo comprueba la información almacenada para cada camino. Si un nodo detecta un fallo de conexión genera una notificación que es reenviada nodo a nodo, para que borren la información almacenada y el poseedor del contenido vuelva a publicarlo.

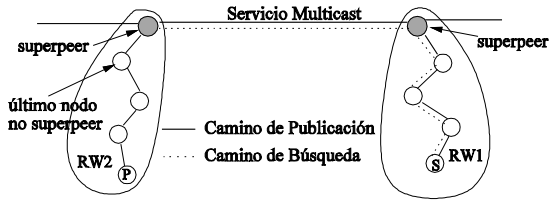


Fig. 2. Fase de Búsqueda.

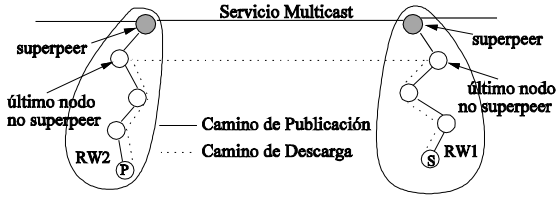


Fig. 3. Fase de Descarga.

## 3.2 Fase de Búsqueda

Cuando un nodo (*S*) desea hacer una búsqueda, en primer lugar selecciona un identificador de conexión y a continuación el mensaje de búsqueda es reenviado hacia el superpeer utilizando el camino asociado. El superpeer puede localizar los contenidos que coinciden dentro de su subgrupo, y además utiliza el servicio multicast para distribuir este mensaje a todos los superpeers. Los superpeers con una coincidencia en la búsqueda responden con un mensaje que contiene el identificador de conexión establecido por el poseedor del contenido (*P*) y la identidad del último nodo no-superpeer en el camino hacia el poseedor del contenido. Antes de enviar el mensaje de respuesta, el superpeer comprueba el camino completo hacia el poseedor del contenido mediante un ping y si falla envía un mensaje broadcast con el identificador de conexión asociado. El nodo asociado volverá a publicar todos sus contenidos y los nodos del camino borrarán la entrada correspondiente.

## 3.3 Fase de Descarga

En esta fase se conectan directamente los últimos nodos no-superpeers de los caminos de publicación del solicitante (*S*) y el poseedor (*P*) del contenido. De esta forma se libera a los superpeers de todo el tráfico de descarga de contenidos.

# 4 Simulaciones

Se ha desarrollado un simulador de eventos discretos en lenguaje C para evaluar nuestro sistema. En los resultados representados en la figura 4 el sistema no implementa el mecanismo de fiabilidad, y muestra el número de solicitudes que no pueden llevarse a cabo en una hora porque, al menos un nodo en el camino de descarga ha caído cuando se ejecuta la búsqueda. Este resultado se representa en función del tiempo (en horas) y observamos que fluctúa entorno a 11.000. Por lo tanto, se demuestra que en una red P2P anónima real es necesario proporcionar un mecanismo para reconstruir eficientemente los caminos.

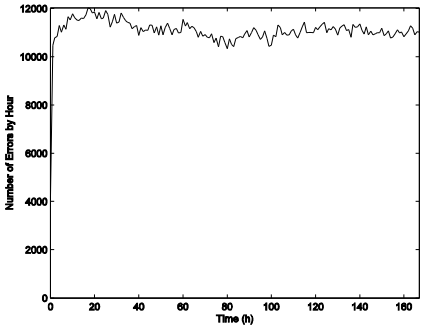


Fig. 4. Número de solicitudes que fallan en una hora.

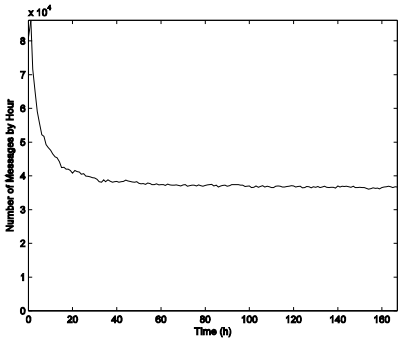


Fig. 5. Número medio de mensajes de control.

En nuestra red las descargas nunca fallan porque se implementa un mecanismo de fiabilidad. La figura 5 representa el número medio de mensajes de control en función del tiempo, fruto del mecanismo de fiabilidad. Inicialmente, debido a las restricciones de simulación el temporizador expira simultáneamente en muchos nodos, pero en estado estable el número medio de mensajes está entorno a 40.000. Si suponemos que cada mensaje de control tiene una longitud de 50 bytes, el tráfico de control sólo supone una tasa de tráfico de 4.4 kbps.

# 5 Conclusiones

En este trabajo se presenta un mecanismo de anonimato para una red P2P. Las simulaciones tratan de evaluar los costes de provisión de anonimato. Nuestra propuesta consigue anonimato mutuo con una tasa de tráfico de control de 4.4 kbps.

# Referencias

[1] M. K. Reiter, A. D. Rubin. “Crowds: Anonymity for web transactions”. Communications of the ACM 42(2), 32-48 (1999).

[2] J. P. Muñoz-Gea, J. Malgosa-Sanahuja, P. Manzanares-Lopez, J. C. Sanchez-Aarnoutse, A. M. Guirado-Puerta. “A hybrid topology architecture for p2p file-sharing systems”. Proceedings of the First International Conference on Software and Data Technologies (ICSOFT2006), Setúbal, Portugal (2006)