

Mecanismos de descubrimiento en las Redes de Contenido.

*Pablo Rodríguez-Bocca**

Héctor Cancela Bosi

*Instituto de Computación - Facultad de Ingeniería, Universidad de la República
Montevideo, Uruguay.*

Abstract

Una Red de Contenido es una red virtual que se monta sobre la infraestructura IP, donde se soportan búsquedas y enrutamiento en base exclusivamente al contenido. En los últimos años se han desarrollado gran cantidad de estas redes virtuales debido principalmente a su flexibilidad de diseño, incluyendo por ejemplo a las redes peer-to-peer, el sistema de nombre de dominios, etc.

En este trabajo se presenta un modelo detallado para la evaluación del diseño y arquitectura de las redes de contenido. Para mostrar la generalidad, se aplica el modelo a dos redes de contenido reales: Napster y Gnutella.

Keywords: *redes de contenido, peer-to-peer, diseño, performance, simulación.*

1. INTRODUCCIÓN

Una Red de Contenido es una red virtual que se monta sobre la infraestructura IP, donde se soportan búsquedas y enrutamiento en base exclusivamente al contenido. Estas redes tienen el atractivo y la flexibilidad de adecuarse a los requerimientos particulares de cada aplicación en cuanto a confiabilidad, performance, anonimato, seguridad, etc. En los últimos años se han desarrollado en distintos contextos gran cantidad de redes de contenido[12], incluyendo redes peer-to-peer, cooperative Web caching, content delivery network, subscribe-publish networks, content-based sensor networks, backup networks, distributed computing, collaboratives networks, instant messagings, multiplayer games y search engines.

Dada la complejidad y escala de la mayoría de estas redes es muy difícil modelar los distintos aspectos de sus diseños y arquitecturas. En este trabajo se presenta un modelo capaz de representar la dinámica del descubrimiento del contenido en la red, donde existe un compromiso entre la publicación y la búsqueda de información.

En las redes de contenido existen las fuentes de contenido (nodos que alojan contenido específico) y los solicitantes (nodos que requieren algún contenido). Toda red de contenido es en realidad una red de conocimiento, donde el conocimiento es la información de alojamiento de cada contenido específico (que relaciona el contenido con el nodo fuente que lo posee).

Generalmente esta meta-información se encuentra distribuida entre los nodos de la red, por lo que no existe una vista global. Las redes de contenido intentan maximizar la disponibilidad del contenido presente en la red, utilizando mecanismos de descubrimiento de contenido. Las limitaciones de ancho de banda de los nodos llevan a que se deba decidir si invertir en publicar los nuevos contenidos alojados en las fuentes o esperar a las búsquedas de los solicitantes. El diseño de la red depende entonces de la dinámica de las solicitudes y alojamientos del contenido en la red y del costo de transmisión de una publicación o una búsqueda.

En la sección 2 se define formalmente el problema del descubrimiento de contenido en redes distribuidas. Luego se instancia el problema general a ejemplos de redes de contenido reales (sección 3). Finalmente, en la sección 4, se discuten algunas conclusiones.

* Dirección: J. Herrera y Reissig 565, 11300 Montevideo, Uruguay. E-mails: prbocca AT fing.edu.uy.

2. EL PROBLEMA DEL DESCUBRIMIENTO DEL CONTENIDO

En esta sección se formaliza el problema de descubrimiento de contenido en redes distribuidas, en adelante denominado por las siglas CDP (Content Discovery Problem).

Se usarán las notaciones siguientes: \sum sumatoria real; $A \cdot B$ multiplicación de reales; Σ OR binario; $A+B$ OR binario; AB AND binario; $\neg A$ NOT binario; $A-B$ es igual a $A(\neg B)$; y $A \oplus B$ es el XOR binario.

2.1 El Tiempo

El modelo debe expresar el carácter dinámico de la red (nodos y contenido) y a su vez especificar la comunicación necesaria para descubrir la información. Por tanto la representación del tiempo juega un papel preponderante en el modelo. Se utiliza una discretización del tiempo en períodos llamados tiempo de ronda y denotada por t^r .

De esta manera, modelaremos una red sincrónica, en la cual el estado de la red se encuentra perfectamente definido en los tiempos múltiplos de t^r . Por tanto, todos los eventos transcurren exclusivamente durante los períodos de ronda. Como se verá más adelante, un evento puede ser la conexión/desconexión de un nodo, el alojamiento/desalojamiento de un contenido por parte de un nodo, la comunicación a través de mensajes, etc.

2.2 Los Nodos y el Contenido

Los nodos de la red pertenecen al conjunto V . Debido a la autonomía de los nodos, en general estos tienen un comportamiento de conexión y desconexión. El estado de un nodo se expresa con la variable aleatoria: $X_i^t \in \{0,1\} \forall t, \forall i \in V$, donde $X_i^t = 1$ sii el nodo i se encuentra conectado en el tiempo $t.t^r$. Por tanto existe un evento de conexión en el período $[t.t^r, (t+1)t^r)$ sii $(X_i^{t+1} - X_i^t) = 1$ y de desconexión sii $(X_i^t - X_i^{t+1}) = 1$ ¹.

El contenido de la red se representa con el conjunto C . Un subconjunto de nodos $K \subseteq V$, llamados terminales, tiene la posibilidad de solicitar contenidos a la red. Un nodo solicita un contenido en la ronda t según la variable aleatoria: $S_{i,k}^t \in \{0,1\} \forall t, \forall i \in V, \forall k \in C$, donde $S_{i,k}^t = 1$ sii el nodo i solicita el contenido k en el período $[t.t^r, (t+1)t^r)$. Los contenidos son alojados físicamente en los nodos. Uno de los puntos más importantes para preservar la autonomía es que los nodos puedan elegir cuando comenzar a alojar y terminar de alojar un contenido dado. El estado de un contenido en un nodo se expresa con la variable aleatoria: $A_{i,k}^t \in \{0,1\} \forall t, \forall i \in V, \forall k \in C$, donde $A_{i,k}^t = 1$ sii el nodo se encuentra alojando el contenido k en el tiempo $t.t^r$. Por tanto existe un evento de alojamiento en el período $[t.t^r, (t+1)t^r)$ sii $(A_{i,k}^{t+1} - A_{i,k}^t) = 1$ y de desalojamiento sii $(A_{i,k}^t - A_{i,k}^{t+1}) = 1$ ².

2.3 Estado de la Red

La meta-información que cada nodo conoce es una tabla de índices, donde un índice es una relación nodo-contenido representando el conocimiento que dicho nodo tiene de dicho contenido. El estado de la red en un momento dado es el conjunto de todas las tablas de índices de cada nodo: $I_{i,j,k}^t \in \{0,1\} \forall t, \forall i, j \in V, \forall k \in C$, donde $I_{i,j,k}^t = 1$ si en el tiempo $t.t^r$, el nodo i aloja el índice $j-k$, y $I_{i,j,k}^t = 0$ en caso contrario. Un nodo conoce perfectamente el contenido que él mismo aloja, por tanto, su tabla de índices representa en el caso particular de los índices propios al contenido alojado, esto es: $I_{i,i,k}^t = A_{i,k}^t \forall t, \forall i \in V, \forall k \in C$. No existen restricciones sobre la validez de los índices alojados para el caso general, por tanto es posible

¹ Se considera improbable, e innecesario de modelar, que un nodo se conecte y desconecte en una ronda.

² Se considera improbable, e innecesario de modelar, que un nodo aloje y desaloje un contenido en una ronda.

(aunque no deseable) que $I_{i,j,k}^t = 1$ a pesar de que $I_{j,j,k}^t = 0$. Esto determina que sea necesario por parte de cada nodo tomar la decisión de que un índice de la tabla expire. El criterio de expiración de índices depende de la política de la red, es decir, de cada implementación en particular, esto se representa con la variable: $E_{i,j,k}^t \in \{0,1\} \forall t, \forall i, j \in V, \forall k \in C$, donde $E_{i,j,k}^t = 0$ representa que en el período $[t.t^r, (t+1).t^r)$ el nodo i borrará el índice $j-k$ de su tabla (si $E_{i,j,k}^t = 1$ entonces no se realiza acción sobre el índice).

2.4 Descubrimiento del Contenido: Mensajes y Restricción de Ancho de Banda

El descubrimiento de índices entre nodos se realiza mediante dos primitivas de comunicación: la publicación y la búsqueda. La publicación es el envío de uno o varios índices: $P_{m,i,j,k}^t \in \{0,1\} \forall t, \forall m, i, j \in V, \forall k \in C$, donde $P_{m,i,j,k}^t = 1$ si el nodo m le envía al nodo i el índice $j-k$, en el período $[t.t^r, (t+1).t^r)$ ($P_{m,i,j,k}^t = 0$ en caso contrario). Se considera que los nodos publican solo aquellos índices que se encuentran en su propia tabla, lo que se expresa por $P_{m,i,j,k}^t \leq I_{m,j,k}^t \forall t, \forall m, i, j \in V, \forall k \in C$. Esto implica que no existe otra forma de comunicación entre nodos que la aquí modelada, y que los nodos sólo dan información veraz. La búsqueda es la solicitud de uno o varios contenidos: $B_{i,j,k}^t \in \{0,1\} \forall t, \forall i, j \in V, \forall k \in C$, donde $B_{i,j,k}^t = 1$ si el nodo i le solicita al nodo j el contenido k , en el período $[t.t^r, (t+1).t^r)$ ($B_{i,j,k}^t = 0$ en caso contrario). La respuesta a una búsqueda es una publicación.

La única restricción explícita que presenta la red es la limitante sobre la cantidad de mensajes que puede manejar cada nodo de la red en una ronda de tiempo t^r , que debe reflejar el ancho de banda de los nodos (en general asimétrico). Las primitivas de comunicación emplean un tamaño dado por las constantes α (bytes necesarios para publicar un índice) y β (bytes necesarios para buscar un índice). Esta restricción se formula con las ecuaciones:

$$BW_{IN,i} \cdot t^r \geq \alpha \cdot \sum_{m,j,k}^+ P_{m,i,j,k}^t + \beta \cdot \sum_{m,k}^+ B_{m,i,k}^t \quad \forall t, \forall m, i, j \in V, \forall k \in C$$

$$BW_{OUT,i} \cdot t^r \geq \alpha \cdot \sum_{j,m,k}^+ P_{i,j,m,k}^t + \beta \cdot \sum_{j,k}^+ B_{i,j,k}^t \quad \forall t, \forall m, i, j \in V, \forall k \in C$$

donde en una ronda, todos los nodos deben recibir menos bytes que su ancho de banda entrante y deben enviar menos bytes que su ancho de banda saliente.

2.5 Objetivo de la Red

El objetivo de la red es maximizar el descubrimiento de contenido, es decir encontrar la mayor cantidad de índices correctos respecto a los contenidos solicitados (bajo la hipótesis de estado estacionario de la red). Según el tipo de contenido y las características particulares de la red, tiene sentido esperar por la información de alojamiento una cantidad máxima de rondas r_{\max} . En forma matemática el objetivo del problema es:

$$\max E \left\{ \sum_{i,k}^+ \left[S_{i,k}^t \sum_j \left(\sum_{r=1}^{r_{\max}} I_{i,j,k}^{t+r} I_{j,j,k}^{t+r} \right) \right] \right\}$$

Para todos los nodos i , que solicitan algún contenido k , maximizar el valor esperado de la obtención de índices $j-k$ válidos en el período de r_{\max} rondas.

2.6 Transición de Estado

Se han definido los conceptos generales para el modelo del problema de descubrimiento. Es importante determinar la transición de estados de forma general:

$$I_{i,j,k}^{t+1} = \begin{cases} A_{i,k}^{t+1} X_i^{t+1} & \text{si } i = j \\ \left[I_{i,j,k}^t + \sum_m P_{m,i,j,k}^t \right] E_{i,j,k}^t X_i^{t+1} & \text{si } i \neq j \end{cases}$$

Si $i = j$ entonces, la ecuación refiere al contenido alojado en el nodo, en ese caso el estado es igual al anterior a menos que el nodo i

decida quitar el contenido $A_{i,k}^{t+1} = 0$ o desconectarse de la red $X_i^{t+1} = 0$. Para el caso general $i \neq j$, se mantiene un índice o se descubre un índice (si algún otro nodo lo envía $\sum_m P_{m,i,j,k}^t$) siempre que el mismo no haya expirado $E_{i,j,k}^t = 0$ o el nodo no se desconecte de la red $X_i^{t+1} = 0$.

El modelo general resultante se presenta en la Tabla 1.

$$\begin{aligned} & \max E \left\{ \sum_{i,k} \left[S_{i,k}^t \sum_j \left(\sum_{r=1}^{t_{\max}} I_{i,j,k}^{t+r} I_{j,j,k}^{t+r} \right) \right] \right\} \\ & \text{sujeto a:} \\ & I_{i,j,k}^{t+1} = \begin{cases} A_{i,k}^{t+1} X_i^{t+1} & \text{si } i = j \\ \left[I_{i,j,k}^t + \sum_m P_{m,i,j,k}^t \right] E_{i,j,k}^t X_i^{t+1} & \text{si } i \neq j \end{cases} \\ & BW_{IN,i} \cdot t^t \geq \alpha \cdot \sum_{m,j,k} P_{m,i,j,k}^t + \beta \cdot \sum_{m,k} B_{m,i,k}^t \quad \forall t, \forall m, i, j \in V, \forall k \in C \\ & BW_{OUT,i} \cdot t^t \geq \alpha \cdot \sum_{j,m,k} P_{i,j,m,k}^t + \beta \cdot \sum_{j,k} B_{i,j,k}^t \quad \forall t, \forall m, i, j \in V, \forall k \in C \\ & P_{m,i,j,k}^t \leq I_{m,j,k}^t \quad \forall t, \forall m, i, j \in V, \forall k \in C \\ & X_i^t, S_{i,k}^t, A_{i,k}^t \in \{0,1\} \text{ v.a. conocidas} \quad \forall t, \forall i \in V, \forall k \in C \\ & I_{i,j,k}^t \in \{0,1\} \text{ estado} \quad \forall t, \forall i, j \in V, \forall k \in C \\ & E_{i,j,k}^t, P_{m,i,j,k}^t, B_{i,j,k}^t \in \{0,1\} \text{ variables de decisión} \quad \forall t, \forall m, i, j \in V, \forall k \in C \end{aligned}$$

Tabla 1. Definición de CDP (Content Discovery Problem).

3. APLICACIÓN DEL MODELO A REDES REALES

Para poder evaluar la expresividad del modelo en esta sección se especifican los ejemplos más citados en la literatura de redes de contenido.

3.1 Napster

Napster[5][6] presenta una topología muy sencilla donde todos los nodos solicitan contenido, excepto el servidor central. Sea el servidor central de Napster como el nodo 0, por tanto $K = V \setminus \{0\}$. Los clientes de Napster se identifican al ingreso a la red en el servidor central, informándole los archivos que comparte: $P_{i,0,i,k}^t = I_{i,i,k}^t (X_i^t - X_i^{t-1})$. Cuando una fuente cambia el contenido ofrecido, esto se le informa al servidor central: $P_{i,0,i,k}^t = (A_{i,k}^t \oplus A_{i,k}^{t-1}) X_i^t$ y $E_{0,i,k}^t = -(A_{i,k}^{t-1} - A_{i,k}^t) X_i^t$. Se supone que el costo de informar que ya no se posee un contenido es igual al costo de informar que se aloja uno nuevo, es decir una publicación. Cuando un solicitante busca un contenido, lo hace en el servidor central, el cual conoce completamente los archivos compartidos reduciendo la búsqueda a una búsqueda local eficiente: $B_{i,0,k}^t = S_{i,k}^t$ y $P_{0,i,j,k}^{t+1} = B_{i,0,k}^t I_{0,j,k}^t$.

Finalmente cuando un nodo se desconecta de la red, este le avisa al servidor central que ya no se encuentra más en la red. El contenido $k=0$, se considera un contenido especial, y es utilizado para anunciar por ejemplo el cambio de estado en un nodo: $P_{i,0,i,0}^t = (X_i^{t-1} - X_i^t)$ y $E_{0,i,k}^t = -(X_i^{t-1} - X_i^t)$.

3.2 Gnutella

Hacemos referencia al protocolo de la red Gnutella[2][3][4] en su versión original (protocolo v0.4), actualmente la red incorpora mejoras como el concepto de nodos superpeers

(conocidos en Gnutella como ultrapeers). En esta situación $K = V$, todos los nodos son terminales, con la capacidad de solicitar, enrutar y ofrecer contenido.

En una red basada en *broadcasts* ningún contenido es publicado de forma proactiva por parte de las fuentes, sino que los nodos solicitantes realizan búsquedas específicas de contenido a todos sus pares. En este modelo los pares de un nodo, también llamados vecinos, son aquellos nodos de los cuales se conoce el contenido $k=0$: $N^t(i) = \{j/I_{i,j,0}^t = 1\}$.

Los pares que reciben las consultas en realidad actúan como nodos enrutadores volviendo a propagar la consulta y como nodos fuentes contestando al solicitante que les hizo la consulta si poseen el contenido buscado. Para evitar bucles y esperas muy largas, las búsquedas se propagan con un mecanismo de tiempo de vida limitado, TTL (generalmente TTL=7).

Para entender la notación considérese que TTL=2, entonces para todo camino en la red de largo 2: $\forall \langle j, j_2 \rangle / j \in N^t(i), j_2 \in N^{t+1}(j)$. Si el nodo i realiza una solicitud en el tiempo t del

contenido k , entonces inmediatamente buscaría en sus vecinos por el contenido $B_{i,j,k}^t = S_{i,k}^t$.

Estos vecinos directos responderían si tienen el contenido en el tiempo siguiente

$P_{j,i,j,k}^{t+1} = B_{i,j,k}^t I_{j,j,k}^t$ o reenviarían la búsqueda al resto de los vecinos $B_{j_2,j_2,k}^{t+1} = B_{i,j,k}^t$, los cuales

podrían responder por el camino en que se realizó la búsqueda $P_{j_2,j_2,k}^{t+2} = B_{j_2,j_2,k}^{t+1} I_{j_2,j_2,k}^{t+1}$ y

$P_{j,i,j_2,k}^{t+3} = P_{j_2,j_2,k}^{t+2}$. En el caso general el mecanismo es el mismo, para todo camino de largo

menor o igual a TTL $\forall \langle j, j_2, \dots, j_{TTL-1}, j_{TTL} \rangle / j \in N^t(i), j_2 \in N^{t+1}(j), \dots, j_{TTL-1} \in N^{t+TTL-2}(j_{TTL-2}), j_{TTL} \in N^{t+TTL-1}$.

En este caso la comunicación se forma por los siguientes paquetes:

Largo 1: $\forall j / j \in N^t(i)$

$$B_{i,j,k}^t = S_{i,k}^t$$

$$P_{j,i,j,k}^{t+1} = B_{i,j,k}^t I_{j,j,k}^t$$

....

Largo TTL: $\forall \langle j, j_2, \dots, j_{TTL-1}, j_{TTL} \rangle / j \in N^t(i), j_2 \in N^{t+1}(j), \dots, j_{TTL-1} \in N^{t+TTL-2}(j_{TTL-2}), j_{TTL} \in N^{t+TTL-1}$

$$B_{j_{TTL-1}, j_{TTL}, k}^{t+TTL-1} = S_{i,k}^t$$

$$P_{j_{TTL}, j_{TTL-1}, j_{TTL}, k}^{t+TTL} = B_{j_{TTL-1}, j_{TTL}, k}^{t+TTL-1} I_{j_{TTL}, j_{TTL}, k}^{t+TTL-1}$$

$$P_{j_{TTL-1}, j_{TTL-2}, j_{TTL}, k}^{t+TTL+1} = P_{j_{TTL}, j_{TTL-1}, j_{TTL}, k}^{t+TTL}$$

Largo 2: $\forall \langle j, j_2 \rangle / j \in N^t(i), j_2 \in N^{t+1}(j)$

$$B_{j_2, j_2, k}^{t+1} = S_{i,k}^t$$

$$P_{j_2, j_2, k}^{t+2} = B_{j_2, j_2, k}^{t+1} I_{j_2, j_2, k}^{t+1}$$

$$P_{j,i, j_2, k}^{t+3} = P_{j_2, j_2, k}^{t+2}$$

$$P_{j_2, j_2, k}^{t+2TTL-2} = P_{j_2, j_2, k}^{t+2TTL-3}$$

$$P_{j,i, j_{TTL}, k}^{t+2TTL-1} = P_{j_2, j_2, k}^{t+2TTL-2}$$

Gnutella escala muy poco, para aumentar el grado de los nodos participantes y de esta forma permitir a la red que utilice un TTL menor se utiliza una mensajería de Ping y Pong, el Ping es una inundación de consultas de estados de conexión, mientras que el Pong es su respuesta. La lógica de comunicación es igual a la de las consultas (con TTL y con la respuesta recorriendo el camino inverso), solo que en este caso el contenido por el cual se pregunta es el $k=0$. En este caso la comunicación se forma por los siguientes paquetes:

Largo 1: $\forall j / j \in N^t(i)$

$$B_{i,j,0}^t = S_{i,0}^t$$

$$P_{j,i,j,0}^{t+1} = B_{i,j,0}^t X_j^t$$

....

Largo TTL: $\forall \langle j, j_2, \dots, j_{TTL-1}, j_{TTL} \rangle / j \in N^t(i), j_2 \in N^{t+1}(j), \dots, j_{TTL-1} \in N^{t+TTL-2}(j_{TTL-2}), j_{TTL} \in N^{t+TTL-1}$

Largo 2: $\forall \langle j, j_2 \rangle / j \in N^t(i), j_2 \in N^{t+1}(j)$

$$B_{j_2, j_2, 0}^{t+1} = S_{i,0}^t$$

$$P_{j_2, j_2, 0}^{t+2} = B_{j_2, j_2, 0}^{t+1} X_{j_2}^{t+1}$$

$$P_{j,i, j_2, 0}^{t+3} = P_{j_2, j_2, 0}^{t+2}$$

$$\begin{aligned}
B_{j_{TTL-1}, j_{TTL}, 0}^{t+TTL-1} &= S_{i,0}^t & \dots \\
P_{j_{TTL}, j_{TTL-1}, j_{TTL}, 0}^{t+TTL} &= B_{j_{TTL-1}, j_{TTL}, 0}^{t+TTL-1} X_{j_{TTL}}^{t+TTL-1} & P_{j_2, j, j_{TTL}, 0}^{t+2TTL-2} = P_{j_3, j_2, j_{TTL}, 0}^{t+2TTL-3} \\
P_{j_{TTL-1}, j_{TTL-2}, j_{TTL}, 0}^{t+TTL+1} &= P_{j_{TTL}, j_{TTL-1}, j_{TTL}, 0}^{t+TTL} & P_{j, i, j_{TTL}, 0}^{t+2TTL-1} = P_{j_2, j, j_{TTL}, 0}^{t+2TTL-2}
\end{aligned}$$

De esta forma todos los nodos que contestan (envían un PONG, que es lo mismo que publicar el contenido $k=0$) son automáticamente agregados a la tabla de vecinos. Cuando no se responde es necesario tener un criterio para eliminarlos luego de un tiempo, es decir un criterio de expiración de conexión de nodo. El caso más simple es considerar que si en las últimas e rondas no se recibió un PONG entonces el nodo ya no es vecino, esto es:

$$E_{i,j,0}^t = \sum_l \left(P_{l,i,j,0}^{t-e} + \dots + P_{l,i,j,0}^t \right).$$

4. CONCLUSIONES

En este trabajo se presenta un modelo detallado sobre el comportamiento dinámico de las redes de contenido. Para mostrar la generalidad, se aplica el modelo a dos redes reales muy distintas: Napster[5][6] y Gnutella[2][3][4] (en [13] se presentan otras aplicaciones, por ejemplo al sistema de nombres DNS[7][8]). Algunas hipótesis del modelo incluyen el considerar que la red está en estado estacionario y es libre de congestión (no presenta pérdida de paquetes).

La generalidad de este modelo tiene un costo en cantidad de cómputo necesario para su utilización. En particular, el mismo incluye el empleo de una cantidad muy grande de variables del sistema, con eventos modelados a nivel muy fino, lo que dificulta su potencial empleo en modelos con grandes cantidades de nodos (las actuales redes de pares para compartir archivos pueden presentar millones de usuarios simultáneos). Esto también lo evidencian los simuladores de redes P2P implementados, que al momento de escrito este documento escalan solo a miles de nodos. Ejemplos son: Peersim[9], P2psim[10], Simpastry[11], Anthill[1], etc. Por tanto el modelo sería útil para la simulación o evaluación de una red de contenido con una cantidad limitada de nodos, sin embargo sus requerimientos computacionales lo hacen inadecuado para emplear un método de optimización al diseño de una red así modelada.

5. REFERENCIAS

- [1] Babaoglu O., Meling H., and Montresor A.. Anthill: A framework for the development of agent-based peer-to-peer systems. Technical Report UBLCS-2001-09, University of Bologna, Italy, 2001. <http://www.cs.unibo.it/projects/anthill/>.
- [2] Gnutella Homepage, <http://gnutella.wego.com/>.
- [3] Gnutella Protocol Specification v0.4 http://www9.limewire.com/developer/gnutella_protocol_0.4.pdf.
- [4] Gnutella hosts. <http://www.gnutellahosts.com>.
- [5] Napster Homepage, <http://www.napster.com/>.
- [6] Napster Messages. <http://opennap.sourceforge.net/napster.txt>.
- [7] Mockapetris, P.V. Internet Domain Name System Standard: Domain names - concepts and facilities. Rfc-Editor Home Page, <ftp://ftp.rfc-editor.org/in-notes/rfc1034.txt>, 1987.
- [8] Mockapetris, P.V. Internet Domain Name System Standard: Domain names - implementation and specification. Rfc-Editor Home Page, <ftp://ftp.rfc-editor.org/in-notes/rfc1035.txt>, 1987.
- [9] Peersim Peer-to-Peer Simulator Home Page. <http://peersim.sourceforge.net/>
- [10] P2psim. A Simulator for Peer-to-Peer protocols. <http://pdos.csail.mit.edu/p2psim/>
- [11] Pastry Home Page. <http://research.microsoft.com/~antr/Pastry/>
- [12] Rodríguez-Bocca P., Cancela Bosi H.. "Redes de contenido: un panorama de sus características y principales aplicaciones". AST 2004 (5th Argentine Symposium on Computing Technology), 33 JAIIO, Córdoba - Argentina. September 20-24, 2004.
- [13] Rodríguez-Bocca P.. "Redes de Contenido: Taxonomía y Modelos de evaluación y diseño de los mecanismos de descubrimiento de contenido" Tesis de Maestría en Informática, PEDECIBA / Facultad de Ingeniería, UDELAR. 2005.