An Analytic Model for Peer to Peer File Sharing Networks

Krishna Kant
Enterprise Architecture Lab, Intel Corporation, OR

Abstract—**In this paper we introduce a random-graph based model for studying the evolution of ad hoc peer-to-peer (P2P) communities such as in Gnutella or Freenet. The proposed random graph model generates a non-uniform graph and provides control over the nodal degree distribution. We study basic properties such as reachability from a given node in the network using an analytical approach. The model can be used in conjunction with a simulation model to study detailed performance tradeoffs in a P2P file-sharing network.**

Keywords: peer-to-peer computing, random graphs, probability generating function.

I. Background and Basic Model

The peer-to-peer music file sharing revolution started by Napster, Gnutella, Freenet and the like has led to considerable interest in ad hoc networking issues. In particular, these applications have spurred the use of peer to peer ideas for a very diverse set of problems such as distribution of streaming media, hosting of arbitrary web services, etc. In reference [3], we have discussed a large number of P2P applications and platforms and introduced a comprehensive taxonomy for them. In this paper, we introduce a random-graph model that attempts to capture some basic properties of these networks, and analytically study their reachability properties and resource utilization at each node. Estimation of queuing delays in file searching and retrieval in such networks is of great interest; however, analytic calculation of these parameters is usually intractable. We have also developed a simulation tool to study queuing behavior under various file retrieval/propagation policies [11]. Due to lack of space, this paper does not include any simulation results. The analytic results discussed here provide a convenient way of engineering the network prior to the simulation.

A. Modeling of Peer-to-Peer File Sharing Networks

A P2P file-sharing network is basically an ad-hoc network with arbitrary topology and size that change as existing nodes leave the network and new ones enter. It is possible to model such network directly, however, the very large degrees of freedom generally allow only approximate or asymptotic results. For example, reference [9] presents a model for building low-diameter dynamic P2P networks where nodes arrive according to a Poisson process and stay in the network for an exponential amount of time. The authors assume a particular connection strategy that ensures that the network diameter increases only logarithmically with the number of nodes. They prove a number of interesting asymptotic results concering the connectivity and the diameter of the largest connected component.

In order to obtain more concrete results, we instead disregard dynamic changes to the network. The justification for this approach follows from the observations concerning real Gnutella networks in [5], [1], [7]. It has been observed in these and several other papers that Gnutella is not "democratic" in the sense that all nodes don't behave in a similar way. Instead, it is possible to identify roughly 3 types"tiers" of nodes as discussed below:

tier1: These are globally known nodes that stay in the network almost indefinitely, have high-speed connectivity and are always available. Effectively these nodes act like "servers" in traditional networks. We call these "distinguished nodes".

tier2: These are non-permanent nodes that are not free-riding [1], i.e., they contribute significant number of files to the community, are reasonably resource rich, and stay on far longer than an average request-response time. We call these as "undistinguished nodes".

tier3: These are transient nodes that join the network for short periods of time, connect to only a few other nodes, and typically do not contribute much to the community. We call these as "transient nodes".

For the purposes of the analysis in this paper, we shall ignore the transient nodes completely. Instead, we assume that the traffic generated by these nodes essentially originates at other nodes to which they connect to. This assumption has little impact on the results but simplifies the problem quite considerably. In particular, we can now decompose the request-response process in the network from the network topology modification process (due to undistinguished nodes joining/leaving the network). This allows us to work with a network instance with some given number of nodes, and completely avoid simultaneous consideration of the network change dynamics.

In view of the above, we are primarily interested in the performance of a file-sharing network assuming that it has grown to a certain size. Thus, although the mechanism of community establishment is not directly relevant, it provides a convenient way to characterize and control the topology of the eventual network.

Because of lack of global knowledge in an uncoordinated P2P network such as Gnutella, requests for documents invariably result in a spanning-tree like search through the network up to a certain number of "hops". If each message is stamped with a globally unique ID, duplicate responses for the same request from a given node can be easily avoided (as in Gnutella). A more sophisticated approach is for each node to glean information about the documents stored at other nodes, either by snooping the responses to the document requests or by a separate protocol (e.g., each node periodically exchanges information with its neighbors about what documents it has). Freenet [2] uses the former method. It shares storage, rather than files, and stores a small number of files and location data about a larger number of files. Also, the response (which contains the file) winds its way to the source node along the reverse search path and thus can be cached by intermediate nodes. We shall consider situations with both direct and reverse path responses.

Recently there have been several other proposals to reduce the cost of the searches. In particular, the idea of "distributed hash tables" (DHT) has been used in several recent projects including PASTRY, CAN, Chord and Tapestry. Here, each node is made responsible for a subset of the entire key space by assigning an Id from the hash over the key space. A search message is directed to the node whose Id is closest to the key in the hash space. Thus, a resource can usually be located in a small number of hops by redirection towards the node whose Id has a better match with the key than the current node. References to these and some discussion of how to exploit network proximity in the search can be found in [10].

In this paper, we shall stick to the simple Gnutella type of network. The "intelligent search" techniques discussed above essentially result in yet another layer of overlay network that is different for each search. Note that the P2P network itself is overlaid on the physical Internet; thus, even a direct modeling of Gnutella doesn't quite capture what happens in the physical network. The issues of efficient mapping of P2P networks onto physical network are well known but very difficult problems that we shall ignore in this paper. Our results still provide some insight into the performance wrt to the overlay network being considered.

B. A Proposed Graph Model

Following the discussion in the last section, we consider a 2-tier graph model with "distinguished" and "undistinguished" nodes. The distinguished nodes and connections between them exist right from the beginning and this subnetwork acts as the nucleus of the file sharing network. We assume an initial network with N_d distinguished nodes connected in a regular pattern with degree $M > 0$. The undistinguished nodes join the network sequentially. Each joining node, say X, connects to a total of $\mathcal{K}$ nodes, where $\mathcal{K}$ is a random variable with the mass function $\theta_k, k = 1 .. K$. Each connection attempt from joining node X is directed to an undistinguished node with probability q_u, and to a distinguished node with probability $(1-q_u)$. While the network contains less than K undistinguished nodes, any excess connections are directed to distinguished nodes. Given that the total number of nodes in the graph, henceforth denoted as N_t, is much larger than K, this initial perturbation has very little impact on the overall network.

One important property of such a graph construction is nonuniformity: distinguished nodes and the nodes that enter the graph early on have, on the average, higher connectivity than others. Moreover, the connectivity distribution has a heavy tail (unlike the exponentially decaying tail in a classical random graph). In fact, we explicitly introduce a parameter β which controls how heavy the tail would be. When the ℓ_2th node attempts to create a connection with an undistinguished node, it has the choice of connecting to up to $\ell_2 - 1$ of them. The function $\beta(\ell, \ell_2)$ gives the probability of selecting level ℓ node in this case. Note that if $\beta(.,.) = 1$, the average degree of a node entering at level ℓ is governed by the term $\sum_{i=\ell}^{\ell_2-1} 1/i$. For a given ℓ_2, this is a very slowly decaying function of ℓ. By choosing $\beta(\ell, \ell_2)/approx \ell^{-\mu}$ for some $\mu \geq 1$, it is possible to obtain a much faster decay and hence a less heavy tail in the degree distribution. It may be noted that heavy-tailed degree distribution is commonly observed in not only Gnutella [5] P2P networks but also in other contexts including web [4] and disease propagation networks [8].

In order to study the performance issues stated above, we need to characterize the following performance metrics.

1. Overall and *undistinguished degree* of a node at each level. The latter is needed since we assume that only undistinguished nodes generate requests.
2. Number of distinct nodes (overall and undistinguished) reachable from any given node in a given number of hops.
3. Given some request generation process at each node, the total request traffic arriving at each node.
4. Given some assumptions about response generation (e.g., response for every request, response only if the requested content exists, etc.), total response traffic passing through each node.
5. Queuing delays experienced by requests and responses at each node.
6. Response time (i.e., time between request transmission and response reception) from every node for requests issued by each node.

We attempt to do this initially with an analytical model in the next section. As stated earlier, the last 2 issues are best handled via simulation and are not addressed in this paper.

II. Analytic Model

Let us start with the characterization of nodal degree. It is convenient to introduce the notion of a "level" for each node. Since all distinguished nodes are treated identically, we consider all of them to lie at level 0. Each undistinguished node, however, introduces a new level. Thus, when the network has a total of N_t nodes, it has $N_t - N_d + 1$ levels, numbered $0 .. N_t - N_d$, with all level 0 nodes being distinguished, and others undistinguished. Henceforth, we let $L = N_t - N_d$ as the maximum level number.

Let $P_n(\ell_2, \ell)$ denote the probability that a node that entered the network at level ℓ has a degree n when a level ℓ_2 node enters the network. Note that following the inclusion of level ℓ_2 node, the total number of nodes in the network is $\ell_2 + N_d$. Also, ℓ_2 denotes the number of undistinguished nodes *after* the addition of the new node. In the following, we show how to obtain the probability generating function $\Phi(z|L, \ell)$ of $P_n(\ell_2, \ell)$, and hence compute $D_{at}(\ell|L)$, defined as the mean degree *at* a node that enters the network at level ℓ assuming a total of L levels.

A. Computation of Nodal Degree

In order to write a recurrence equation for $P_n(\ell_2, \ell)$, we introduce the connection probability $\alpha(\ell_2, \ell)$, defined as the probability that the addition of a new node to the network (at level ℓ_2) will result in one more connection to a level ℓ node.

Let us first suppose that $\ell > 0$, i.e., we are considering how the degree of a nondistinguished node changes with the new node addition. Suppose that the level ℓ_2 node attempts to connect to k nodes, of which i are undistinguished nodes. When $\ell_2 = 1$, $\alpha(\ell_2, \ell) = 0$ for $\ell > 0$ since there are no undistinguished nodes to connect. Otherwise, if $\ell_2 - 1 < i$, only $\ell_2 - 1$ connections are actually possible with undistinguished nodes. It follows that for $\ell_2 > 1$, $\ell > 0$,

$$\alpha(\ell_2,\ell) = \sum_{k=0}^{K} \theta_k \left[\sum_{i=1}^{\min(\ell_2-1,k)} \beta(\ell,\ell_2) \frac{i\, B_{k,i}(q_u)}{\ell_2 - 1} + \sum_{i=\ell_2}^{k} B_{k,i}(q_u) \right]$$

where $B_{k,i}(q_u)$ is the binomial mass function $B_{k,i}(q_u) = \binom{k}{i} q_u^i (1-q_u)^{k-i}$ with $B_{0,0}(q_u) = 1$. The equation follows from the fact that the probability of connection attempt to i undistinguished nodes is $B_{k,i}(q_u)$, and a given node will be selected with probability $\frac{\beta(\ell,\ell_2)i}{(\ell_2-1)}$ if $i \leq \ell_2 - 1$. For $i > \ell_2 - 1$, every undistinguished node must be targeted for connection.

For $\ell_2 > K$, the entire expression collapses to the following intuitively obvious expression:

$$\alpha(\ell_2,\ell) = q_u E[\mathcal{K}] \frac{\beta(\ell,\ell_2)}{(\ell_2-1)}, \qquad \ell_2 > K \tag{1}$$

Note that only the mean value of $\mathcal{K}$ is relevant in this case; the rest of the distribution doesn't matter.

Next, let us evaluate $\alpha(\ell_2, 0)$. Note that if the new node connects to i undistinguished nodes, it will connect to $k - i$ distinguished nodes. Since there are a total of $N_d \geq K$ distinguished nodes, we have $\alpha(\ell_2, 0) = \sum_{k=0}^{K} \theta_k \zeta(k, \ell_2)$ where

$$\zeta(k,\ell_2) = \sum_{i=0}^{\min(\ell_2-1,k)} \frac{k-i}{N_d} B_{k,i}(q_u) + \sum_{i=\ell_2}^{k} \frac{k-\ell_2+1}{N_d} B_{k,i}(q_u)$$

If $\ell_2 > K$, the entire expression again collapses to the obvious equation:

$$\alpha(\ell_2,0) = (1-q_u)E[\mathcal{K}]/N_d, \qquad \ell_2 > K \tag{2}$$

Assuming that $P_n(.,.) = 0$ for $n < 0$, we can now write the following recurrence equation for $P_n(\ell_2, \ell)$ as:

$$\begin{aligned} P_n(\ell_2,\ell) &= \alpha(\ell_2,\ell) P_{n-1}(\ell_2-1,\ell) \\ &+ [1-\alpha(\ell_2,\ell)] P_n(\ell_2-1,\ell) \end{aligned} \tag{3}$$

For $\ell > 0$, this equation applies for $\ell < \ell_2$, i.e., when level ℓ node in consideration is not the same as the node being added to the network. For $\ell = \ell_2$, the degree of the added node is simply $\mathcal{K}$. That is,

$$P_k(\ell,\ell) = \theta_k, \qquad 0 \leq k \leq K \tag{4}$$

For $\ell = 0$, the above recurrence equation fails to apply only initially, i.e., when the network consists of only distinguished nodes. Since we are assuming a regular M-ary connectivity initially, we have:

$$P_n(0,0) = \begin{cases} 1 & n = M \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Let $\Phi(z|L,\ell)$ denote probability generating function of the degree distribution. We have:

$$\begin{aligned} \Phi(z|\ell_2,\ell) &= \sum_{n=0}^{\ell_2-\ell+K} z^n P_n(\ell_2,\ell) \\ &= [1-\alpha(\ell_2,\ell)(1-z)]\Phi(z|\ell_2-1,\ell), \qquad \ell < L \end{aligned}$$

where we have used the fact that $P_{\ell_2-\ell+K}(\ell_2-1,\ell) = 0$. Using this relationship recursively, we get

$$\Phi(z|L,\ell) = \Phi(z|\ell,\ell) \prod_{i=\ell+1}^{L} [1-\alpha(i,\ell)(1-z)] \tag{6}$$

It thus follows:

$$\Phi(z|L,\ell) = \Psi(z) \prod_{i=\ell+1}^{L} [1-\alpha(i,\ell)(1-z)] \tag{7}$$

where the function $\Psi(z)$ is defined as:

$$\Psi(z) = \begin{cases} \sum_{k=0}^{K} z^k \theta_k & \ell > 0 \\ z^M & \ell = 0 \end{cases} \tag{8}$$

This completes the characterization of the degree distribution. Let $D_{at}(\ell|L)$ denote the mean degree *at* a node that enters the network at level ℓ, given a total of L levels. By definition,

$$D_{at}(\ell|L) = \Phi'(z|L,\ell)\Big|_{z=1} = \Psi'(1) + \sum_{j=\ell+1}^{L} \alpha(j,\ell) \tag{9}$$

Now we separately consider the $\ell > 0$ and $\ell = 0$ cases. In the former case, if $\ell \leq K$, the above summation can be divided into two parts, one for the range $\ell+1 \mathinner{..} K$ and the other from $K+1 \mathinner{..} L$. Then, using equation (1), for $0 < \ell < K$, we get:

$$D_{at}(\ell|L) = E[\mathcal{K}] + q_u E[\mathcal{K}] \sum_{i=K}^{L-1} \frac{\beta(\ell,i)}{i} + \sum_{j=\ell+1}^{K} \alpha(j,\ell) \tag{10}$$

where the last term drops out for $\ell > K$. Now, for $\ell = 0$, we have two ranges, one $0 \mathinner{..} K$, and the other $K+1 \mathinner{..} L$. Using equation (2), we have:

$$D_{at}(0|L) = M + E[\mathcal{K}](1-q_u)\frac{L-K}{N_d} + \sum_{j=1}^{K} \alpha(j,0) \tag{11}$$

If required, the overall average degree of a node could then be computed as: $D_{at}(L) = (N_d D_{at}(L,0) + \sum_{\ell=1}^{L} D_{at}(\ell|L))/N_t$.

Let us now adapt the above analysis to obtain the undistinguished degree of a node, defined as the number of undistinguished nodes to which a given node is directly connected. Let us denote this as $P_n^{(u)}(\ell_2,\ell)$, defined just like the overall degree $P_n(\ell_2,\ell)$. Since every added node in the network construction process is a undistinguished node, equation (3) holds for $P^{(u)}$ as well, except that the initial conditions are different. In particular, let k_u denote the number of undistinguished nodes that a newly arriving level $\ell > 0$ node connects to. Obviously, $k_u \in 0 \mathinner{..} \min(K, \ell-1)$; therefore, $\theta_u(k_u,\ell)$, the probability of a newly arriving level ℓ node connecting to k_u undistinguished nodes is:

$$\theta_u(k_u,\ell) = \sum_{k=k_u}^{K} \theta_k B_{k,k_u}(q_u) \tag{12}$$

Note that even if $\theta_0 = 0$ (i.e., an incoming node always connects to at least one node), k_u could still be 0. Also, $\theta_u(k_u,\ell)$

depends on ℓ for $\ell \leq K$ only. Now, equation (4) becomes $P^{(u)}(k_u|\ell,\ell) = \theta_u(k_u,\ell)$. Since, a distinguished node is initially not connected to any undistinguished node, eqn(5) reduces to:

$$P_n^{(u)}(0,0) = \begin{cases} 1 & n=0 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

With this, the PGF of $P^{(u)}$ is given by an equation like (7), but with Ψ function suitably modified for the changed initial conditions, i.e.,

$$\Psi_u(z) = \begin{cases} \sum_{k=0}^{K} z^k \theta_u(i,\ell) & \ell > 0 \\ \\ 1 & \ell = 0 \end{cases} \tag{14}$$

Let $D_{at,u}$ be defined like D_{at} except that we are now interested in only the undistinguished degree of a node. Equations (10–11) then apply to $D_{at,u}$ as well except for the first term modified to correspond to $\Psi'_u(1)$ instead of $\Psi'(1)$. Note that for $\ell > K$,

$$\Psi'_u(1) = \sum_{k_u=0}^{K} k_u \theta_u(k_u,\ell) = \sum_{k=0}^{K} \theta_k \sum_{k_u=0}^{k} k_u B_{k,k_u}(q_u)$$

which evaluates to $q_u E[\mathcal{K}]$, as expected.

B. Reachability Distribution

Next we consider the computation of the number of nodes reached starting from a given node. For this, we first characterize the reachability matrix $\mathcal{R}_h$ whose (i,j)th element gives the probability of reaching level i from level j in exactly h hops (i.e., via paths with exactly $h-1$ unique intermediate nodes different from nodes i and j). This reachability matrix, along with the nodal degree characterization, yields the average number of nodes reached in h hops. We henceforth let $G(\ell_2,h)$ denote the average number of nodes reached *at* hth hop, starting from level ℓ_2. Similarly, let $G_u(\ell_2,h)$ denote the number of undistinguished nodes reached in h hops starting from level ℓ_2. The computation of $\mathcal{R}_h$, $G(\ell_2,h)$, and $G_u(\ell_2,h)$ is discussed below.

We start with the reachability of all nodes from a given node. For this, we arbitrarily number the distinguished nodes as $1..N_d$, and subsequently a level ℓ node as $\ell + N_d$. Let $R_1(n_1,n_2)$ denote the probability of reaching node n_1 from node n_2 in one hop. Let R_1 denote the corresponding $N_t \times N_t$ matrix. To specify R_1, we first compute a related matrix R_0 that enumerates the probability of finding an arc between any given pair of nodes. Since R_0 is a symmetric matrix and for any node n_1, $R_0[n_1,n_1] = 0$; therefore, it suffices to assume $n_2 > n_1$ in the following:

$$R_0[n_1,n_2] = \begin{cases} M/(N_d-1) & n_1 \leq N_d, n_2 \leq N_d \\ \alpha(n_2 - N_d, 0) & n_1 \leq N_d, n_2 > N_d \\ \alpha(n_2 - N_d, n_1 - N_d) & n_1 > N_d, n_2 > n_1 \end{cases} \tag{15}$$

Given R_0, we construct R_1 by normalizing the columns to 1. This ensures that the total probability of reaching all other nodes from any node is 1. Note that R_1 is not a symmetric matrix, but it is possible to exploit its special structure.

Let $R_h[n_1,n_2]$ denote the probability of reaching node n_1 from node n_2 in h hops. In order to compute the matrix R_h, it is necessary to enumerate all unique paths of length h that do not use any intermediate node more than once. The latter requirement increases the computational cost, since a straightforward recursion on h cannot avoid node reuse. We also need to renormalize the matrix at each step to ensure that the columns sum to 1. We denote the unnormalized matrices as R'_h where we specifically set $R'_h[n,n] = 0$ for all n. Then, $R_h = R'_h \times \text{diag}(\mathbf{e}.R'_h)$ where $\mathbf{e}$ denotes a row vector of all 1's and the function diag converts a row vector into a diagonal matrix by putting the vector elements along the diagonal. Now, for $n_1 \neq n_2$,

$$R'_2[n_1,n_2] = \sum_{n_3=1}^{N_t} R_1[n_1,n_3] R_1[n_3,n_2] \tag{16}$$

where it is not necessary to specifically avoid $n_3 = n_1$ or $n_3 = n_2$ since $R_1(n,n) =$ for all n. However, for R'_3, we have:

$$R'_3[n_1,n_2] = \sum_{\substack{n_4=1 \\ n_4 \neq n_2}}^{N_t} \sum_{\substack{n_3=1 \\ n_3 \neq n_1}}^{N_t} R_1[n_1,n_4] R_1[n_4,n_3] R_1[n_3,n_2]$$

Similar equations can be written for R'_i, $i > 3$. We now define a modified $(L+1) \times (L+1)$ matrix $\mathcal{R}_h$ where the rows and columns indicate level, rather than individual nodes:

$$\mathcal{R}_h[\ell_1,\ell_2] = \begin{cases} (N_d - 1)\, R_h[1,2] & \ell_1 = 0, \ell_2 = 0 \\ R_h[\ell_1 + N_d, 1] & \ell_1 > 0, \ell_2 = 0 \\ N_d\, R_h[1, \ell_2 + N_d] & \ell_1 = 0, \ell_2 > 0 \\ R_h[\ell_1 + N_d, \ell_2 + N_d] & \ell_1 > 0, \ell_2 > 0 \end{cases}$$

This matrix suffices for further computation. The sole reason for defining R matrices in terms of node numbers earlier was to ensure that no node gets used more than once on a path.

Let $D_{fr}(\ell_2,h|L)$ denote the average degree of the nodes reached in h hops starting *from* a node at level ℓ_2 when the maximum level is L. It is convenient to denote row vectors of $D_{at}(\ell|L)$'s and $D_{fr}(\ell,h|L)$'s for $\ell = 0..L$ as $\mathbf{D}_{at}(L)$ and $\mathbf{D}_{fr}(h|L)$ respectively. Then, $\mathbf{D}_{fr}(1|L) = \mathbf{D}_{at}(L)$, and for $h > 1$

$$\mathbf{D}_{fr}(h|L) = \mathbf{D}_{at}(L)\mathcal{R}_{h-1} \tag{17}$$

In the analysis presented here, the last equation is the only place where the matrix $\mathcal{R}_h$ is used; therefore, explicit computation of $\mathcal{R}_h$ is unnecessary (and very expensive). Instead, it is desirable to compute $\mathbf{D}_{fr}(h|L)$ directly for each h.

Let $G(\ell_2,h)$ denote the average number of nodes reached *at* hth hop, and $G^{(t)}(\ell_2,h)$ the total number of nodes reached in h hops (ℓ_2 is the starting level in both cases). Again, it is convenient to work with the corresponding row vectors henceforth denoted as $\mathbf{G}(h)$ and $\mathbf{G}^{(t)}(h)$. Then, $\mathbf{G}(1) = \mathbf{D}_{at}(L)$ (i.e., the degree of the starting node) and $\mathbf{G}^{(t)}(0) = \mathbf{e}$ (the starting or "root" node). For $h > 0$, we have:

$$\mathbf{G}^{(t)}(h) = \mathbf{G}^{(t)}(h-1) + \mathbf{G}(h), \quad 1 \leq h \leq H \tag{18}$$

where H denotes the highest hop count of interest. Now, to obtain an expression for $\mathbf{G}(h)$ for $h > 1$, we imagine that the nodes at $(h-1)^{st}$ hop are expanded sequentially in order to count the unique nodes reachable from each of them. Let $g_m(\ell_2, h)$ denote the number of unique nodes accounted for at the mth step of the expansion. Then, in the $G(\ell_2, h-1)$th step, we would have accounted for all $G(\ell_2, h)$ nodes. With $g_0(\ell_2, h) = 0$ for all ℓ_2 and h, we can write the following recurrence equation for $g_m(\ell_2, h)$, $h > 1$

$$g_m(\ell_2, h) = g_{m-1}(\ell_2, h) + D_{fr}(\ell_2, h|L) \frac{N_t - G^{(t)}(\ell_2, h-1) - g_{m-1}(\ell_2, h)}{N_t} \quad (19)$$

This equation is obtained as follows: At mth step, $N_t - G^{(t)}(\ell_2, h-1) - g_{m-1}(\ell_2, h)$ nodes out of a total of N_t nodes have already been accounted for; therefore, the number of new nodes added is simply the average node degree multiplied by the fraction of unreached nodes. This equation can be rewritten as

$$g_m(\ell_2, h) = a.g_{m-1}(\ell_2, h) + b \quad (20)$$

where the quantities a and b are independent of the index m and are given by $a = 1 - D_{fr}(\ell_2, h|L)/N_t$ and $b = [N_t - G^{(t)}(\ell_2, h)](1-a)$. A repeated expansion of equation (20) along with $g_0(\ell_2, h) = 0$, yields the following explicit recurrence for $G(\ell_2, h)$.

$$\begin{aligned} G(\ell_2, h) &= g_{G(\ell_2,h-1)}(\ell_2, h) = b\frac{1 - a^{G(\ell_2,h-1)}}{1-a} \\ &= [1 - a^{G(\ell_2,h-1)}][N_t - G^{(t)}(\ell_2, h-1)] \end{aligned} \quad (21)$$

Thus, $\mathbf{G}(h)$, and hence $\mathbf{G}^{(t)}(h)$ can be computed.

Next we adapt the above equations to obtain reachability to only undistinguished nodes. For this, we first need to compute the average undistinguished degree of nodes reached in h hops from a given level ℓ_2, given a total of L levels. We henceforth denote this as $D_{fr,u}(\ell_2, h|L)$ and also define the corresponding row vector $\mathbf{D}_{fr,u}(h|L)$. As with $\mathbf{D}_{fr}(h|L)$, we have $\mathbf{D}_{fr,u}(1|L) = \mathbf{D}_{at,u}(L)$, and for $h > 1$

$$\mathbf{D}_{fr,u}(h|L) = \mathbf{D}_{at,u}(L)\mathcal{R}_{h-1} \quad (22)$$

Next, we define the quantities $\mathbf{G}_u(h)$ and $\mathbf{G}_u^{(t)}(h)$ respectively, which have the same meaning as $\mathbf{G}(h)$ and $\mathbf{G}^{(t)}(h)$ except that reachability to only undistinguished nodes is being considered. These quantities can be determined by a slight tweak to the analysis presented above. Note that the paths to undistinguished nodes could well pass through distinguished nodes; therefore, if we are interested in paths of length h, we still proceed as in the last subsection up to length $h-1$ and then estimate the number of reachable undistinguished nodes in the last step. Let $g_{m,u}(\ell_2, h)$ denote the number of undistinguished nodes reached in the mth step for length h paths. Note specifically that the upper bound on m is $G^{(t)}(\ell_2, h-1)$ [instead of $G_u^{(t)}(\ell_2, h-1)$]. With $g_{0,u}(\ell_2, h) = 0$, equation (19) can be adapted as follows:

$$g_{m,u}(\ell_2, h) = g_{m-1,u}(\ell_2, h) + D_{fr,u}(\ell_2, h|L) \frac{L - G_u^{(t)}(\ell_2, h-1) - g_{m-1,u}(\ell_2, h)}{L} \quad (23)$$

This equation is again based on the observation that at mth step the number of new undistinguished nodes added is the undistinguished degree multiplied by the fraction of unvisited undistinguished nodes. The quantities $\mathbf{G}_u^{(t)}(h)$ and $\mathbf{G}_u(h)$ are again related as:

$$\mathbf{G}_u^{(t)}(h) = \mathbf{G}_u^{(t)}(h-1) + \mathbf{G}_u(h), \quad 1 \le h \le H \quad (24)$$

along with the initial conditions $\mathbf{G}_u(1) = \mathbf{D}_{at,u}(L)$ (i.e., the undistinguished degree of the starting node) and $\mathbf{G}_u^{(t)}(0) = [0, 1, \ldots, 1]$. As in the last section, equation (23) can be solved explicitly to give:

$$G_u(\ell_2, h) = [1 - a^{G(\ell_2,h-1)}][L - G_u^{(t)}(\ell_2, h-1)] \quad (25)$$

where $a = 1 - D_{fr,u}(\ell_2, h|L)/L$.

C. Performance Metrics

Given the above analysis, we are now in a position to estimate the total request and response traffic arriving at a given node X. This is given by a superposition of the traffic generated by all nodes from which X can be reached in H or fewer hops (including requests generated by the node itself). A precise characterization of the traffic process is intractable; here we only compute its average rate. Let $N_{req}(\ell, H)$ denote the total number of requests reaching *undistinguished nodes* in at most H hops from a level ℓ node. Then,

$$N_{req}(\ell, H) = 1 + \sum_{h=1}^{H} G_u(\ell, h) \quad (26)$$

Turning this around, $N_{req}(\ell, H)$ gives the total request traffic processed by a level ℓ node. Thus, if λ_0 denotes the request rate for each undistinguished node, the total request traffic arriving at a level ℓ node, denoted $\lambda_{req}(\ell)$, is simply $\lambda_0 N_{req}(\ell, H)$.

The response traffic can also be computed similarly. That is, $\lambda_{resp}(\ell)$, the total response traffic arriving at a level ℓ node, is $\lambda_0 N_{resp}(\ell, H)$ where $N_{resp}(\ell, H)$ denotes the total number of responses processed at level ℓ node for a single request. The estimation of $N_{resp}(\ell, H)$ depends on how the responses are generated and forwarded back to the requesting node. We consider the following possibilities in this regard:

1. Every node generates a response (found or not found), and sends it to the requester directly (without going through any intermediate nodes). Then, $N_{resp}(\ell, H) = N_{req}(\ell, H)$.

2. Every node generates a response which travels back along the request path. In this case, a request going to a node that is h hops away, will require response processing at each one of the h intermediate nodes (excluding the response generating node where request processing and response generation are considered a single atomic operation). It follows that:

$$N_{resp}(\ell, H) = \sum_{h=1}^{H} h\, G(\ell, h) \quad (27)$$

Let S_{req} and S_{resp} denote the average service times of requests and responses. Then, the level ℓ node utilization is given by:

$$U(\ell) = \lambda_{req}(\ell)S_{req} + \lambda_{resp}(\ell)S_{resp} \quad (28)$$

TABLE I

REACHABILITY AND TRAFFIC FOR A 100 NODE NETWORK

undist prob	no of hops	tot nodes reached	undist reached	responses per node	tot traf per node
0.05	1	5.9	3.3	4.9	6.1
	2	55.2	44.5	103.6	146.5
	3	99.1	85.8	235.2	320.5
	4	100	90	238.8	328.8
	5	100	90	238.8	328.8
0.50	1	5.9	4.3	4.9	8.4
	2	34.3	23.8	61.7	82.3
	3	91	73.9	231.7	304
	4	99.9	89.4	267.5	356.9
	5	100	89.6	267.7	357.3
0.95	1	5.9	5.3	4.9	10.6
	2	28.6	22.6	50.3	73.6
	3	76.7	63.8	194.6	258.4
	4	98.5	87.4	281.8	369.2
	5	99.7	89.3	287.8	377.2

TABLE II

REACHABILITY AND TRAFFIC FOR A 500 NODE NETWORK

undist prob	no of hops	tot nodes reached	undist reached	responses per node	tot traf per node
0.05	1	6	3.6	5	6.2
	2	243.7	232.7	480.5	711.5
	3	499.7	488.6	1248.4	1737
	4	500	490	1249.6	1739.6
0.50	1	6	4.7	5	8.5
	2	95.7	84.2	184.3	264.6
	3	483.5	465.1	1347.8	1812.4
	4	500	490	1413.9	1903.9
0.95	1	6	5.8	5	10.7
	2	35.1	29.1	63.2	91.7
	3	163.5	137.1	448.3	582.4
	4	405.7	367.7	1417.2	1782.7

As stated earlier, the queuing delays depend on the details of the arrival and service processes and are best determined via a simulation (which can also also issues such as limited queue depths, abandonments and retries). The above equation for utilization is useful for ensuring that the nodal utilization in the simulation model does not exceed a predetermined limit.

Now, we present some experimental results using the analytic model. We choose the parameters in this study as follows:

1. The network initially has 10 distinguished nodes, each of which is connected to 4 others.
2. Each newly arriving node connects from 1 to 4 nodes with equal probability. (Thus, the average connectivity is 2.5).
3. The β function is assumed to identically 1.
4. We consider two network sizes N_t of 100, and 500 nodes.
5. For each network, we consider 3 undistinguished node connection probabilities q_u of 5%, 50%, and 95%.

Tables I and II show overall averages on reachability, response traffic per node and total traffic per node for 100 and 500 node graphs respectively. For reachability, the tables list the average number of all nodes reached from an arbitrary node (column 3), and average number of undistinguished nodes reached from an undistinguished node (column 4). The response (and hence total) traffic computation assumes that the responses trace the request path backwards. These results point to a number of interesting conclusions about the network under consideration.

1. Even with a very limited immediate connectivity per node of 2.5, 4 hops (default for Gnutella and Freenet) can cover nearly all the nodes except when the undistinguished node connection probability is very high.
2. Distinguished nodes provide a "short cut" between nodes; therefore, with a low value of "undist prob" (or q_u), the overall reachability increases very fast.
3. As the size of the network increases, the percentage of nodes that can be covered with a given number of hops goes down (as expected), however, a small number of hops can still cover a lot of nodes.
4. Given the assumption of responses tracing back the request path, the response traffic per undistinguished node (column 5) increases very rapidly for the second and third hops (since these hops add an awful lot of nodes to the reachability tree). Beyond the third hop, the response traffic penalty is negligible.
5. An interesting traffic parameter is the total traffic divided by the network size. With 4 hops, this metric varies between 3.3 and 3.8. That is, if each request returns a response, the nodal capacity must increase 3-4 times as fast as the network size.

III. CONCLUSIONS

In this paper, we introduced a simple analytic model to study possible topologies of file-sharing P2P networks and its performance characteristics. The model results have been verified against the simulation results (not included here for lack of space). The results were also crucial in computing the node service times in the simulation model for target utilizations. It is possible to extend the model in many ways, including a more direct representation of network changes and considering the impact of mismatch between the P2P topology and that of the underlying network. These are currently underway and will be reported in the future.

REFERENCES

[1] E. Adar and B. A. Huberman, "Free riding in Gnutella", Xerox PARC report, Oct 2000, available at www.parc.xerox.com/istl/groups/iea/papers/gnutella.

[2] I. Clarke, "A Distributed Decentralized Information Storage and Retrieval system." M.S. Thesis, Division of Informatics, Univ of Edinburgh, UK, 1999.

[3] K. Kant, R. Iyer and V. Tewari, "On the Potential of Peer-to-Peer Computing: Classification and Evaluation", Proc. of CCGrid 2002, Berlin, Germany, available at http://kkant.ccwebhost.com/download.html

[4] R. Kumar, P. Raghvan, et. al., "The web as a graph", ACM POD conference, Dallas, TX, 2000.

[5] M.A. Jovanovic, F.S. Annexstein and K.A. Berman, "Scalability issues in large peer to peer networks – A case study of Gnutella", Tech. Report, Univ. of Cincinati, 2001.

[6] Matei Ripeanu, "Peer-to-Peer Architecture Case Study: Gnutella", Proc. of 2001 Intl Conf on Peer-to-peer Computing (P2P2001), Linkoping Sweden, 27-29 August 2001.

[7] A. Vahdat, J. Chase, et. al., "Self-organizing subsets: From each according to his abilities, to each according to his needs".

[8] C. Moore and M. Newman, "Epidemics and percolation in small-world networks".

[9] G. Pandurangan, P. Raghvan and E. Upfal, "Building low-diameter P2P networks".

[10] M. Castro, P. Druschel, et. al., "Exploiting Network Proximity in Distributed Hash Tables".

[11] K. Kant and R. Iyer, "SimP^2: A Performance Modeling Tool for Evaluating P2P File Sharing Networks", submitted for publication, available at kkant.ccwebhost.com/download.html.