CS599: Structure and Dynamics of Networked Information (Spring 2005)
03/21/2005: Heavy-tailed degree distributions in Internet-like graphs
Scribes: Fang Bian and Dustin Reishus

From last time, consider the following random graph model for networks in which nodes trade off between number of hops and "last mile" costs. Node O is initially placed at the center of the unit square. From then on, nodes fall randomly in the unit square with uniform distribution. When a node i falls, it makes one connection to another node j that has already been placed. Assume each node i minimizes $\beta h_j + d_{i,j}$ over j, where β is a small constant, h_j is the number of hops from node j to O, and $d_{i,j}$ is the Euclidean distance from i to j. We show that the degree distribution resulting from this process is heavy-tailed:

Theorem 1 *There is some $\varepsilon > 0$ (depending on β), such that there are at least $\Omega(n^{\varepsilon/6})$ nodes of degree at least $n^{1-\varepsilon}$.*

Proof. We first show that each neighbor of the first node O carves out a region of influence so that all nodes landing in the region must link to it. This gives those nodes high degrees. Then, we show that there will be enough such neighbors of O. Let S denote the set of all neighbors of O.

Consider a node i that connects to O, and has $\beta \leq d_{i,O} \leq 2\beta$. Let $r_i = d_{i,O} - \beta$. At the time node i arrived, there are no other nodes within distance r_i of it, because it linked to the root (otherwise, it would connect to such a node). No node j arriving after i within distance $d_{i,j} \leq r_i/2$ connects to node O, because it would be cheaper to connect to node i. So there are no other nodes of S within distance $r_i/2$ of i. Every node j that lands within $r_i/4$ of i links to i, because the cost to connect to node i is at most $\beta + r_i/4$, while the cost to connect to another node i' within distance $d_{i',i} \leq r_i/2$ of i is at least 2β, and the cost to connect to a node i' with distance $d_{i',i} > r_i/2$ of i is at least $\beta + d_{i',j} \geq \beta + r_i/4$ by the Triangle Inequality.

So node i has a region of influence with area $\Omega(r_i^2)$ around it in which all nodes link to it, and thus receives at least an r_i^2 fraction of all edges.[1]

So we know that any such node i has high degree (as long as r_i is large enough). We still need to show that there are enough such nodes i. Among others, we want to ensure that not too many nodes will link to i. To address this concern, consider a node j arriving after i. The line of indifference between connecting to i and connecting to O is given by $d_{j,O} = \beta + d_{j,i}$. This curve is a hyperbola, i.e., the nodes linking to i are a subset of the interior of that hyperbola.

Consider the following rings (Figure 1) around O: Ring I has inner radius β and outer radius $\beta + n^{-\varepsilon/2}$, Ring II has inner radius $\beta + n^{-\varepsilon/2}$ and outer radius $\beta + \frac{1}{2}n^{-\varepsilon/3}$, and Ring III has inner radius $\beta + \frac{1}{2}n^{-\varepsilon/3}$ and outer radius $\beta + n^{-\varepsilon/3}$. (Chose ε small enough that $\beta + n^{-\varepsilon/3} \leq 2\beta$.) Note that the area of Ring II is a constant fraction of the total area of Rings I, II, and III.

No node will ever link to a node $j \in S$ with $d_{j,O} < \beta$, because by Triangle Inequality, it would be cheaper to link to O directly. Nor will any node that falls in these rings ever connect to any node more than one hop from node O, because the cost would be at least 2β, so it would be cheaper to connect to node O directly. Similarly, nodes in Ring II will not connect to a node j outside the outer circle, because the cost is at least $\beta + \frac{1}{2}n^{-\varepsilon/3}$, while a direct connection to node O would cost at most $\beta + \frac{1}{2}n^{-\varepsilon/3}$.

In summary, each node i with $\beta + n^{-\varepsilon/2} < d_{i,O} < \beta + \frac{1}{2}n^{-\varepsilon/3}$ links either to O or to some node $j \in S$ with $\beta \leq d_{j,O} \leq \beta + n^{-\varepsilon/3}$. Each such node that links to O carves out a hyperbola, so that subsequent nodes that fall in the hyperbola do not link to node O.

If a new node i from Ring II arrives outside all existing hyperbolas of nodes j with $\beta \leq d_{j,O} \leq 2\beta$, then i links to node O. We want to find a lower bound on the number of hyperbolas defined by such i. Some

---

[1] Note: to show that i will have degree $\Omega(r_i^2 n)$, we need to be sure that enough other nodes will arrive after i to give it high degree. We can solve this problem by only considering the degrees for nodes that are among the first $n/2$ to arrive. Then, at least $n/2$ nodes arrive subsequently, giving us the desired high degree for those nodes. We only lose a factor of 2.

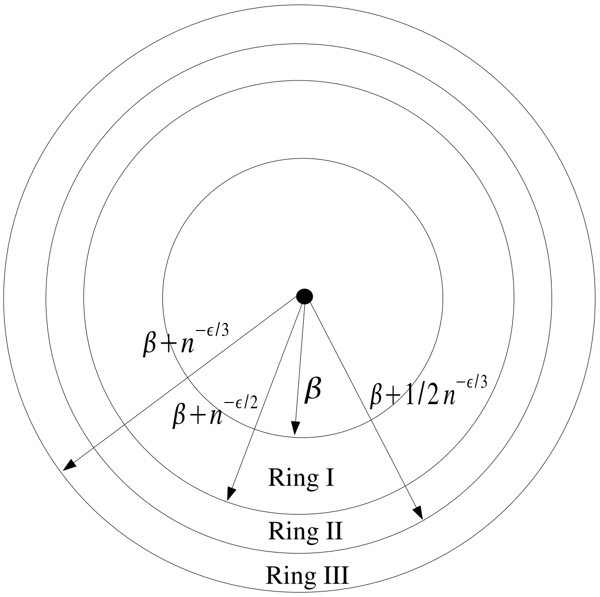

Figure 1: Rings

analytic geometry shows that the largest angle θ of each hyperbola is $O\left(\sqrt{r_i/\beta}\right)$. Because $r_i < n^{-\varepsilon/3}$ in Ring II, and β is a constant, θ is $O\left(\sqrt{n^{-\varepsilon/3}}\right) = O(n^{-\varepsilon/6})$. So there is room for $\Omega(n^{\varepsilon/6})$ disjoint hyperbolas.

Each node i that carves out a hyperbola has degree at least $\Omega(nr_i^2)$. Because $r_i = \Omega(n^{-\varepsilon/2})$ in Ring II, the degree of these nodes is at least $\Omega(n \cdot (n^{-\varepsilon/2})^2) = \Omega(n^{1-\varepsilon})$. Among all the nodes arriving in any of the rings (and thus claiming hyperbolas), at most a constant fraction lie outside of Ring II, so at least $\Omega(n^{\varepsilon/6})$ nodes will be in Ring II, claim hyperbolas, and thus have degree at least $\Omega(n^{1-\varepsilon})$.

Notice that while we only talk about the expected number of such nodes (and their expected degree), standard occupancy bounds can be used to show concentration, i.e., that the actual outcome will be close to the expectation.

This completes the proof. ■