CS599: Structure and Dynamics of Networked Information (Spring 2005)
03/09/2005: Power Laws via Optimization
Scribes: Ranjit Raveendran and Farnoush Banaei-Kashani

In the past lecture, we investigated the preferential attachment model for the generation of power law degree distributions. While "rich-get-richer" models may be considered close to the truth for WWW-like graphs, they do not seem appropriate for graphs such as the Internet, as there appears no natural reason why nodes would choose to attach to high-degree nodes (or be more likely to find out about them). Hence, this class of models is not very useful for explaining the power laws observed in the Internet at the AS level [4].

Fabrikant et al. [3] argue that the reason power laws evolve in the Internet is a heuristic optimization performed by the owners of machines. Indeed, optimization as a cause for power laws has been investigated before. Mandelbrot [5] and Zipf [7] argue that power laws in word lengths are caused by the "Principle of Least Effort": languages evolve so as to optimize the average information transmitted per character or phoneme. (Notice, however, that the same result about word lengths can be obtained by assuming that characters, including the space bar, are pressed completely randomly [6].) Carlson and Doyle [2] extend the argument for file sizes and other parameters, and Fabrikant et al. [3] apply it to Internet-like graphs.

They posit the following graph growth model. A communication tree is built as nodes arrive uniformly at random in a unit square. Let O be the first node that arrives, and assume that it arrives at the center of the square. Each node i arriving subsequently connects to a node j that had arrived earlier. The issue is which node j should i connect to. [3] argues that nodes want to be central in the network, i.e., few hops from O. At the same time, they want to have low "last mile" cost for their connection to j. The tradeoff is accomplished by considering the objective function $d_{ij} + \beta h_j$ where d_{ij} is the Euclidean distance between nodes i and j, and h_j is the number of hops from j to O in the communication tree, and β is a given constant. That is, node i connects to the node j minimizing $d_{ij} + \beta h_j$, and consequently has $h_i = h_j + 1$.

Depending on the value of the parameter β, the graph evolves in very different ways.

- If β is large (e.g., $\beta \geq \frac{1}{\sqrt{2}}$), then the hop count is more important than any possible distance (all distances to O are at most $\frac{1}{\sqrt{2}}$), so the graph will evolve to be a star with node O as the center. In particular, the degree distribution will not be heavy-tailed or power-law.

- If $\beta = 0$, then the Euclidean distances become the only criterion for minimizing the objective function. Thus, nodes connect to their closest neighbors. We will analyze this process briefly.

  If node i has two neighbors j, j', such that $d_{ij}, d_{ij'} \geq r$, then we also have that $d_{jj'} \geq r$ (else j, j' would have been neighbors instead of both connecting to i). It follows that each neighbor j of i at distance r has a circle of influence with area at least $\Omega(r^2)$ around it, which does not contain any other neighbors of i.

  Now consider the $O(\log n)$ rings around i of the form $R_k := \{j \mid 2^{-k+1} < d_{i,j} \leq 2^{-k}\}, k = 0, \ldots, 3 \log n$. Note that with high probability, each node other than i lies inside one of the R_k. Within each R_k, node i can have at most a constant number of neighbors, because the area of each R_k is at most $O(2^{-2k})$, while each neighbor of i in R_k has an area of at least $\Omega(2^{-2(k+1)})$ in which no other neighbor can lie. Hence, there is at most a constant number of neighbors in each such ring, and a total of at most $O(\log n)$.

  Thus, all degrees are bounded by $O(\log n)$, and the distribution of degrees is not power law, or even heavy-tailed.

- If $\beta > 0$ is a small enough constant, then Fabrikant et al. [3] show that the degree distribution is heavy-tailed: there are polynomially many nodes with polynomially large degree. We will prove this fact in the next lecture.

  [3] claims that the distribution is in fact power-law, and this claim was made in class as well. However, Berger et al. [1] show that the distribution is in fact not power-law. It has a heavy tail, but does not exhibit all intermediate degrees to the corresponding extent.

References

[1] N. Berger, B. Bollobás, C. Borgs, J. Chayes, and O. Riordan. Degree distribution of the FKP network model. In *Proc. 30th Intl. Colloq. on Automata, Languages and Programming*, pages 725–738, 2003.

[2] J. Carlson and J. Doyle. Highly optimized tolerance: A mechanism for power laws in designed systems. *Physics Review E*, 60:1412–1427, 1999.

[3] A. Fabrikant, E. Koutsoupias, and C. Papadimitriou. Heuristically optimized trade-offs: A new paradigm for power laws in the internet. In *Proc. 29th Intl. Colloq. on Automata, Languages and Programming*, pages 110–122, 2002.

[4] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *Proc. ACM SIGCOMM Conference*, pages 251–262, 1999.

[5] B. Mandelbrot. An informational theory of the statistical structure of languages. In W. Jackson, editor, *Communication Theory*, pages 486–502. 1953.

[6] G. Miller. Some effects of intermittent silence. *American Journal of Psychology*, 70:311–314, 1957.

[7] G. Zipf. *Selective studies and the principle of relative frequency in language.* Harvard University Press, 1932.