CS599: Structure and Dynamics of Networked Information (Spring 2005)
03/28/2005: The Watts-Strogatz Model
Scribes: David Kempe

In the past lecture, we began our quest for graphs that would model the empirically observed "small world phenomenon": the fact that most (or all) pairs of people tend to be connected through short paths of mutual acquaintances. We observed that expander graphs tend to have small diameter, i.e., short paths, and that random d-regular graphs for $d \geq 3$ are with high probability expander graphs.

While this in principle will give us "small worlds", it is unsatisfactory in that real social networks do not at all look like random graphs. They exhibit a lot of "clustering" in the form of triangles or other short cycles. Indeed, two people sharing a common acquaintance are more likely to know each other as well. This observation has led to the definition of the *clustering coefficient* of node v as

$$\frac{|\{(u, u') \mid (u, u') \in E, (u, v) \in E, (u', v) \in E\}|}{\binom{d_v}{2}},$$

i.e., the fraction of actual triangles among pairs of neighbors of v. Random graphs have low clustering coefficient (roughly d/n for d-regular graphs), while real social or collaboration networks have much higher clustering.

In order to form a more realistic model giving both small diameter and high clustering coefficient, Watts and Strogatz [3] propose the following small-world model. Start with a (structured) graph H (for instance, the 2-dimensional grid, or a ring), and add one or a few random edges from each node (alternately, "rewire" one or a few edges randomly). The result will be that the edges from graph H will lead to high clustering coefficient, while the random edges lead to small diameter, as per our analysis of random graphs. The resulting graph will still "resemble" H a lot. Intuitively, the graph H is supposed to model the "traditional" social network, based perhaps on physical proximity or related employment, while the random edges model "random" friends.

An interesting additional outcome of Milgram's experiment [1] was not only the existence of short paths in social networks, but also the fact that individuals, who do not have a map of the graph, were able to actually find short paths. Their decisions were only based on local and qualitative information. Hence, an interesting question is what properties of a graph allow efficient decentralized routing based solely on local information. In addition to a better understanding of social networks, this question is also relevant for designing networks permitting simple routing.

To make this question more concrete, we can work with the Watts-Strogatz Model, and ask about the effect that different distributions of long-range links have on the ability to route with a decentralized algorithm. Watts and Strogatz considered the case of uniformly random links, i.e., each node links to one other uniformly randomly chosen one. We will show that in the case of the 2-dimensional grid of size $n \times n$, with this distribution, there is no local routing protocol reaching its destination in $o(n^{2/3})$ steps in expectation. Here, a *local protocol* is one in which a node only knows the history of the message in the network so far (including which nodes it has visited), and its own links.

For the proof, consider a ball B of radius $n^{2/3}$ around the destination node t, and assume that the source s of the message lies outside this ball. By the Principle of Deferred Decisions[1], we can assume that each node only generates its random outgoing link once it receives the message. Consider the first $\delta n^{2/3}$ steps of the protocol (for some constant $\delta \ll 1$). Because each node's outgoing link hits B with probability at most $O(\frac{(n^{2/3})^2}{n^2}) = O(n^{-2/3})$, only $O(\delta) \ll 1$ long-range links hit B in expectation, and with constant

[1] The Principle of Deferred Decisions [2] states the intuitively obvious fact that we can defer coin flips or other random decisions until their outcome is actually needed/referenced.

probability (by Markov's Inequality), none do. Thus, every step ending in B must have been a step on the grid. As a result, with constant probability, t was not reached in $\delta n^{2/3}$ steps, completing the proof. (Notice that the analysis is tight, in that in $O(n^{2/3})$ rounds, with constant probability, B is reached, and within at most another $O(n^{2/3})$ grid steps, the message makes it to t). In the next lecture, we will continue a more fine-grained analysis of the effect of distributions on the routing speeds.

References

[1] S. Milgram. The small world problem. *Psychology Today*, 1:61–67, 1967.

[2] R. Motwani and P. Raghavan. *Randomized Algorithms.* Cambridge University Press, 1990.

[3] D. Watts and S. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998.