from *SIAM News*, Volume 32, Number 10

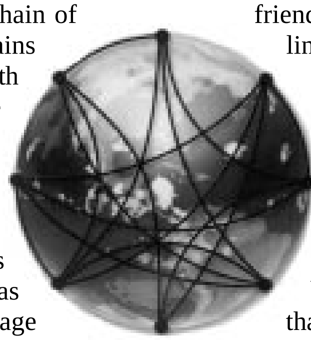
Random Shortcuts Make It A Small World Indeed

Small Worlds: The Dynamics of Networks between Order and Randomness. *By Duncan J. Watts, Princeton University Press, 1999, xv + 262 pages, \$39.50.*

BOOK REVIEW

By Gilbert Strang

"Small world" refers to a familiar experience: You meet someone new, say on an airplane. You explore connections with this new person, usually by mentioning friends. Sure enough, a short chain of friendships is discovered between you. And the "shortcut" established by this conversation makes the chains linking a lot of other people that much shorter. You both agree that it is a small world. In that path-length sense, it is.



The natural problem for a mathematician is to understand why these paths are, on average, so short. If everyone in the world has a thousand friends, then in principle three steps could connect you to a billion people. This is a very wrong model. The parameter that measures clustering—the strong probability that many of your friends are friends of one another—is obviously crucial. Somehow the graph with people as nodes, and friendships as edges, is highly clustered but still small in diameter. "Six Degrees of Separation" has become a popular measure—although six is probably closer to the average than to the maximum distance (and everything depends on the meaning of friendship). Well, this book review can't go into a question as deep as that.

There are huge graphs, like the Web, that we can probe locally. The Web sites are nodes (approaching a billion) and the links are edges. We have just the sketchiest idea of the global properties of this network. (Average distance is estimated as 18, and the distribution of links appears close to a power law). The graph with all telephones as nodes, and phone calls on a given day as edges, has been discussed at SIAM conferences. The graph of all actors is known in great detail (joint movies yield the edges, and the distance to Kevin Bacon is computed the most). In mathematics we have Erdös numbers, and the graph of all our joint papers could be fun. The average number of edges per node is probably increasing quickly.

Duncan Watts has added important examples from social science, electric power networks, and neuroscience. His highly interesting book analyzes the key properties of the graphs. The well-developed theory of random graphs (which have low clustering) does not directly apply. But there is a *partly* random element to the enormous graphs that we see in applications. This observation was the core of a brief paper by Watts and Steve Strogatz that appeared in *Nature* (June 4, 1998). Those few paragraphs caught the attention of science journalists everywhere. By January 1999 even *The New Yorker* had joined in!

The book is essentially the Cornell PhD thesis that Watts wrote with encouragement and advice from Strogatz—a great combination of student and adviser. They studied graphs that combine structure with randomness. The structure might be simply a circle of N nodes (two friends each, left and right). Its diameter is $D = N/2$, and the average path length is $L = N/4$. If we add M random shortcuts, how do the properties of the graph depend on N and M? The cal-culations of Watts and Strogatz showed a rapid decrease in the average distance as M increases, always with $M << N$.

Taking this problem as our model, Watts ended his preface with a correct prediction: "By the time this book is actually printed, multiple additions and refinements will no doubt have been made." This has indeed happened, mostly from his fruitful collaboration at the Santa Fe Institute with Mark Newman. I have to report that their key insights came from physics (or rather from the mathematics of physicists). A continuum mean field model led them to a formula for L that is asymptotically correct in the range $1 << M << N$.

With only one shortcut ($M = 1$), this reviewer contributed a completely elementary calculation. Suppose the shortcut leaves a fraction p of the N nodes on one side, and a fraction $1 - p$ on the other side. Effectively, the two ends of the shortcut become a single node (since we can ignore the one shortcut edge in computing path lengths). So our graph is now two cycles, of length pN and $(1 - p)N$, meeting at one point. A random pair of nodes is on one circle or the other, or split between them, with probabilities p^2, $(1 - p)^2$, and $2p(1/ - p)$. The average distances in those three cases are $pN/4$, $(1-p)N/4$, and $N/4$. Combining those possibilities and averaging over $0 < p < 1$ gives $L = 5N/24$. Thus, a single shortcut reduces L by one sixth, from $N/4 = 6N/24$.

With two shortcuts and major help from Henrik Eriksson and Maple, the average distance was calculated as $L = 131N/720$. This triple integral gave me a new and more respectful view of both freshmen and the limits of integration. (The first question was whether the shortcuts cross. I innocently thought that the chances were 50–50. Matlab thought otherwise; they cross only one third of the time.) Meanwhile, Newman and Watts established that the asymptotic relation must be $L = f(M)$, by a renor-malization group argument, and computer experiments produced a graph of f. I did not expect an explicit function, but from joint work with Chris Moore (http://www.lanl.gov/abs/cond-mat/9909165) they now have one:

$$f(M) = \log[(M/2)^{1/2} + (M/2 + 1)^{1/2}] / 2(M^2 + 2M)^{1/2}.$$

Recently, Jon Kleinberg (of the Cornell computer science department) studied a further key question: Can good paths created by shortcuts be found quickly? The rule is that only local information is available (looking out from nodes already reached). His answer is “no” when all shortcuts are equally likely, as above, but “yes” when the probability of a shortcut is proportional to $1/(\text{length})^2$. For all powers other than this one, you can’t find the good paths in a decentralized way. (See also *Scientific American* for June 1999, where a group from IBM describes work with Kleinberg on locating the best Web sites for any given topic—a decentralized way to find the sites that others link to. The algorithm is a power method!)

All this can be seen as a little niche in the larger and deeper theory of graphs. But the idea of “small worlds” can be very useful. The dynamic evolution of these networks is the subject of Part II of the book—and there is a lot still to be understood. I believe we will see more good mathematics (and real applications) for these graphs.

Gilbert Strang, the president of SIAM, is a professor of mathematics at MIT.