

Navigation in a small world

It is easier to find short chains between points in some networks than others.

The small-world phenomenon — the principle that most of us are linked by short chains of acquaintances — was first investigated as a question in sociology^{1,2} and is a feature of a range of networks arising in nature and technology^{3–5}. Experimental study of the phenomenon¹ revealed that it has two fundamental components: first, such short chains are ubiquitous, and second, individuals operating with purely local information are very adept at finding these chains. The first issue has been analysed^{2–4}, and here I investigate the second by modelling how individuals can find short chains in a large social network.

I have found that the cues needed for discovering short chains emerge in a very simple network model. This model is based on early experiments¹, in which source individuals in Nebraska attempted to transmit a letter to a target in Massachusetts, with the letter being forwarded at each step to someone the holder knew on a first-name basis. The networks underlying the model follow the ‘small-world’ paradigm³: they are rich in structured short-range connections and have a few random long-range connections.

Long-range connections are added to a two-dimensional lattice controlled by a clustering exponent, α , that determines the probability of a connection between two nodes as a function of their lattice distance (Fig. 1a). Decentralized algorithms are studied for transmitting a message: at each step, the holder of the message must pass it across one of its short- or long-range connections; crucially, this current holder does not know the long-range connections of nodes that have not touched the message. The primary figure of merit for such an algorithm is its expected delivery time T , which represents the expected number of steps needed to forward a message between a random source and target in a network generated according to the model. It is crucial to constrain the algorithm to use only local information — with global knowledge of all connections in the network, the shortest chain can be found very simply⁶.

A characteristic feature of small-world networks is that their diameter is exponentially smaller than their size, being bounded by a polynomial in $\log N$, where N is the number of nodes. In other words, there is always a very short path between any two nodes. This does not imply, however, that a decentralized algorithm will be able to discover such short paths. My central finding is that there is in fact a unique value of the exponent α at which this is possible.

When $\alpha = 2$, so that long-range connections

follow an inverse-square distribution, there is a decentralized algorithm that achieves a very rapid delivery time; T is bounded by a function proportional to $(\log N)^2$. The algorithm achieving this bound is a ‘greedy’ heuristic: each message holder forwards the message across a con-

nection that brings it as close as possible to the target in lattice distance. Moreover, $\alpha = 2$ is the only exponent at which any decentralized algorithm can achieve a delivery time bounded by any polynomial in $\log N$; for every other exponent, an asymptotically much larger delivery time is required, regardless of the algorithm employed (Fig. 1b).

These results indicate that efficient navigability is a fundamental property of only some small-world structures. The results also generalize to d -dimensional lattices for any value of $d \geq 1$, with the critical value of the clustering exponent becoming $\alpha = d$. Simulations of the greedy algorithm yield results that are qualitatively consistent with the asymptotic analytical bounds (Fig. 1c).

In the areas of communication networks⁷ and neuroanatomy⁸, the issue of routing without a global network organization has been considered; also in social psychology and information foraging some of the cues that individuals use to construct paths through a social network or hyper-linked environment have been discovered^{9,10}. Although I have focused on a very clean model, I believe that a more general conclusion can be drawn for small-world networks — namely that the correlation between local structure and long-range connections provides critical cues for finding paths through the network.

When this correlation is near a critical threshold, the structure of the long-range connections forms a type of gradient that allows individuals to guide a message efficiently towards a target. As the correlation drops below this critical value and the social network becomes more homogeneous, these cues begin to disappear; in the limit, when long-range connections are generated uniformly at random, the result is a world in which short chains exist but individuals, faced with a disorienting array of social contacts, are unable to find them.

Jon M. Kleinberg

Department of Computer Science, Cornell University, Ithaca, New York 14853, USA

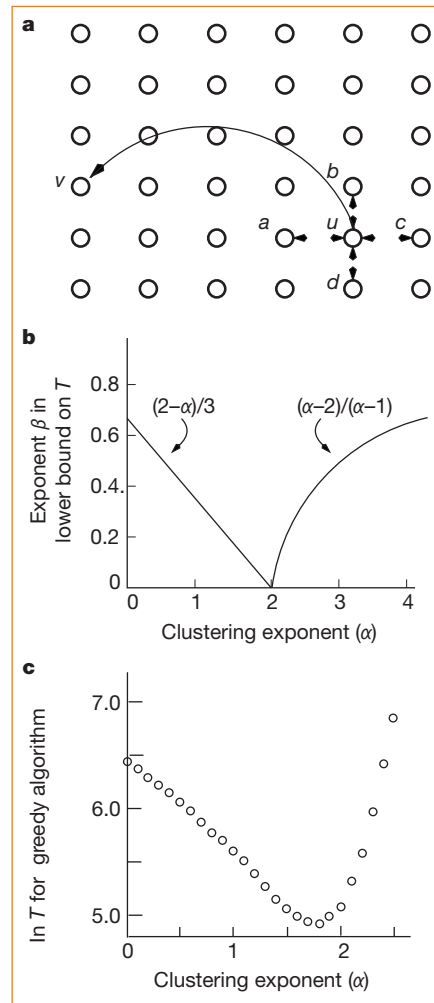


Figure 1 The navigability of small-world networks. **a**, The network model is derived from an $n \times n$ lattice. Each node, u , has a short-range connection to its nearest neighbours (a , b , c and d) and a long-range connection to a randomly chosen node, where node v is selected with probability proportional to $r^{-\alpha}$, where r is the lattice (‘Manhattan’) distance between u and v , and $\alpha \geq 0$ is a fixed clustering exponent. More generally, for $p, q \geq 1$, each node u has a short-range connection to all nodes within p lattice steps, and q long-range connections generated independently from a distribution with clustering exponent α . **b**, Lower bound from my characterization theorem: when $\alpha \neq 2$, the expected delivery time T of any decentralized algorithm satisfies $T \geq cn^\beta$, where $\beta = (2 - \alpha)/3$ for $0 \leq \alpha < 2$ and $\beta = (\alpha - 2)/(\alpha - 1)$ for $\alpha > 2$, and where c depends on α , p and q , but not on n . **c**, Simulation of the greedy algorithm on a $20,000 \times 20,000$ toroidal lattice, with random long-range connections as in **a**. Each data point is the average of 1,000 runs.

- Milgram, S. *Psychol. Today* **1**, 61–67 (1967).
- Kochen, M. (ed.) *The Small World* (Ablex, Norwood, NJ, 1989).
- Watts, D. & Strogatz, S. *Nature* **393**, 440–442 (1998).
- Albert, R. et al. *Nature* **401**, 130–131 (1999).
- Adamic, L. in *Proc. 3rd European Conference on Digital Libraries* (eds Abiteboul, S. & Vercoustre, A.-M.) 443–452 (Springer Lecture Notes in Computer Science, Vol. 1696, Berlin, 1999).
- Cormen, T., Leiserson, C. & Rivest, R. *Introduction to Algorithms* (McGraw-Hill, Boston, 1990).
- Peleg, D. & Upfal, E. *J. Assoc. Comput. Machinery* **36**, 510–530 (1989).
- Braitenberg, V. & Schüz, A. *Anatomy of the Cortex* (Springer, Berlin, 1991).
- Killworth, P. & Bernard, H. *Social Networks* **1**, 159–192 (1978).
- Pirollo, P. & Card, S. *Psychol. Rev.* **106**, 643–675 (1999).