Small World Networks

Alok Advin K.
04CS1005 {alokadvin@gmail.com}

14 April 2008

1 Introduction

A network is said to exhibit the small-world phenomenon when any two nodes in the network are likely to be connected through a short sequence of links or hops. This property is synonymous with the idea of "six degrees of separation". Social networks, the connectivity of the Internet, and gene networks all exhibit small-world network characteristics. In social networks it is a basic statement about the abundance of short paths in a graph whose nodes are people, with links joining pairs who know one another.

1.1 Milgrams experiment

In the 1950's, Ithiel de Sola Pool (MIT) and Manfred Kochen (IBM) set out to prove the six degrees of separation theory mathematically. Although they were able to phrase the question *given a set N of people, what is the probability that each member of N is connected to another member via* k_1*,* k_2*,* $k_3 ... k_n$ *links?*, after twenty years they were still unable to solve the problem to their own satisfaction.

In 1967, American sociologist Stanley Milgram devised a new way to test the theory, which he called "the small-world problem". He randomly selected people in the mid-West to send packages to a stranger located in Massachusetts. The senders knew the recipient's name, occupation, and general location. They were instructed to send the package to a person they knew on a first-name basis who they thought was most likely, out of all their friends, to know the target personally. That person would do the same, and so on, until the package was personally delivered to its target recipient. Although the participants expected the chain to include at least a hundred intermediaries, on average it only took between five and seven intermediaries to get each package delivered.

There were various drawbacks to this experiment like it suffers from selection and nonresponse bias due to the way participants were recruited and high non-completion rates. In many cases, the participant may be unsure which of their friends is the most likely to know the target. Thus they will send the package

further away from the target rather than sending it along the shortest path creating wrong estimates of the average number of ties needed for two random people.

1.2 Properties of small-world networks

Small-world networks have high representation of cliques, and subgraphs that are a few edges shy of being cliques, i.e. small-world networks have sub-networks that are characterized by the presence of connections between almost any two nodes within them. Secondly, most pairs of nodes will be connected by at least one short path. Thus to qualify as small world network the graph got to have large clustering coefficient and small average shortest path. Additionally, there are several properties that are commonly associated with small-world networks even though they are not required for that classification. Typically there is an over-abundance of hubs - nodes in the network with a high number of connections.

1.3 Examples of small-world networks

Networks such as the electric power grid for Southern California, the network of movie-actor collaborations, social network and the neuronal network of the worm Caenorhabditis elegans seem to be small-world networks. The Web is also a small-world network where nodes are web pages and edges are the links between them. Other examples of Small world networks include Protein networks, scientific coauthorship graphs, airport network etc.

2 Construction of Small world Networks

There are various methods to model the small world network experimentally. Few are discussed here with their properties.

2.1 Watts-Strogatz model

One of the first mathematical models built specifically for small-world networks was the Watts-Strogatz model. It is a brilliantly simple model that provides the ability to calibrate between a ring lattice (a completely structured regular graph) and a random graph. There are three parameters to the Watts-Strogatz model. N is the number of nodes in the graph you want to simulate. K is the degree of each node at the initial step of the algorithm and p is the probability of randomly rewiring each edge in the second step of the algorithm. Here $0 \leq p \leq 1$ and $N \gg K \gg ln(N) \gg 1$. The average degree of the graph remains unaltered at the end of step 2. See Figure 1.

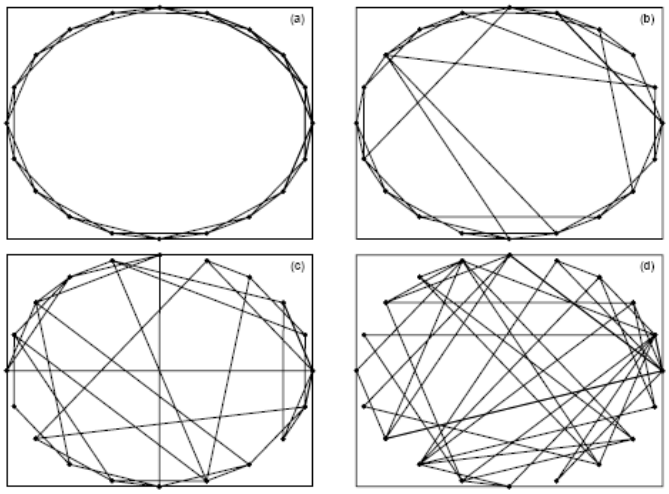

Figure 1: Examples of networks obtained by the procedure described for Watts-Strogatz model, for k = 2, N = 20. (a): p = 0, regular networks; (b), (c): intermediate values of p; (d): p = 1.

Step 1. Construct a regular ring lattice, a graph with N nodes each connected to K neighbors, K/2 on each side. That is, if the nodes are labeled labeled n_0, n_1 ...n_{N-1}, there is an edge (n_i,n_j) if and only if $|i-j| \equiv k \pmod N$ for some $|k| \in \left(1, \frac{K}{2}\right)$

Step 2. For every node n_i where $0 \leq i \leq N-1$ take every edge (n_i,n_j) with $i < j$, and rewire it with probability p. Rewiring is done by replacing edge (n_i,n_j) with edge (n_i,n_t) where t is chosen with uniform probability from all possible values that avoid loops ($t \neq i$) and link duplication (there is no edge (n_i, n_t') with $t' = t$ at this point in the algorithm).

2.2 Newman-Watts model

There are various drawbacks in Watts-Strogatz model for the purpose of analytical treatment. One problem is that the distribution of shortcuts(total number of rewired bonds) is not completely uniform; not all choices of the positions of the rewired bonds are equally probable. This nonuniformity of the distribution makes an average over different realizations of the randomness hard to perform. The other problem is that the average distance between pairs of vertices on the graph, is poorly defined. The reason is that there is a finite probability of a portion of the lattice becoming detached from the rest in this model. Formally, we can represent this by saying that the distance from such a portion to a vertex elsewhere on the lattice is infinite. However, this means that the average

vertex-vertex distance on the lattice is then itself infinite, and hence that the vertex-vertex distance averaged over all realizations is also infinite.

In Newman-Watts model both of these problems are circumvented by a slight modification of the SW model. In this version of the small-world model start again with a regular one-dimensional lattice, but now instead of rewiring each bond with probability p, add extra edge(long links) between pairs of vertices chosen uniformly at random but do not remove any bonds from the regular lattice. More than one bond between any two vertices, or a bond which connects a vertex to itself is also allowed. Add with probability p one shortcut for each bond on the original lattice. The average degree of the graph increases after addition of edges, $K' = K(1 + p)$.

2.3 Scaling of small world networks

We can also build the above network models out of higher-dimensional lattices, such as square or cubic lattices. Consider a hypercube of side L and dimension d. Here number of vertices $N = L^d$ and the average vertex–vertex distance increases as L, or equivalently as $N^{1/d}$. For small values of d this does not give us small-world behavior. If we allow the dimension d of the lattice to become large, then $N^{1/d}$ becomes a slowly increasing function of N, and so the lattice does show the small-world effect. The small world property is also observed when mean coordination number z (average degree) of the vertices is much higher than twice the dimension d of the lattice.

The average vertex-vertex distance l obeys the scaling form $l = \xi G(L/\xi)$, where $G(x)$ is a universal scaling function of its argument x and ξ is a characteristic length-scale for the model which is assumed to diverge in the limit of small p according to $\xi \sim p^{-\tau}$. Numerical results show that τ is 1. The small-world model has only one non-trivial length-scale other than the lattice spacing, which we can equate with the variable ξ above, and for one dimensional grid it is given by

$$\xi = 1/pz \tag{1}$$

and for d-dimensional model

$$\xi = \frac{1}{(pzd)^{1/d}} \tag{2}$$

τ is indeed 1 for d = 1, or $\tau = 1/d$ for general d and, since there are no other length-scales present, l must be of the form

$$l = \frac{L}{2dz} F(pzL^d) \tag{3}$$

where $F(x)$ is another universal scaling function. The divergence of ξ as $p \to 0$ gives something akin to a critical point in this limit. Both the scaling function

$F(x)$ and the scaling variable $x \equiv pzL^d$ have simple physical interpretations. The variable x is two times the average number of shortcuts on the graph for the given value of p, and $F(x)$ is the average fraction by which the vertex–vertex distance on the graph is reduced for the given value of x.

2.4 Ising model

The Ising model, named after the physicist Ernst Ising, is a mathematical model in statistical mechanics. It is defined on a discrete collection of variables called spins, which can take on the value 1 or -1. The model consists of a chain of N Ising spins s_i, with nearest-neighbor interactions J (chain bonds) and shortcut interactions I (long range bonds), both J and I being positive, so that the Hamiltonian reads

$$H = -J \sum_{i=0}^{N-1} \sigma_i \sigma_{i+1} - I \sum_{(ij) \in S} \sigma_i \sigma_j - h \sum_{i=0}^{N-1} \sigma_i \tag{4}$$

with $\sigma_i = \pm 1$ and periodic boundary conditions: $\sigma_0 = \sigma_N$. The set S contains $N_b = pN$ shortcut pairs of spins, and the last term accounts for the effect of external magnetic field. Consider a transformation from site variables to bond variables, which is formulated for a general Ising model in zero-magnetic field,

$$H = -\sum_{(ij)} J_{ij} \sigma_i \sigma_{i+1} \tag{5}$$

where J_{ij} can be chosen arbitrary (though in what follows they are only 0, J, or I).

3 Network Geometrical Properties

3.1 Connectivity

In Watts and Strogatz model for $p = 0$ each vertex has the same connectivity K. On the other hand, a non-zero value of p introduces disorder into the network, in the form of a non-uniform connectivity, while maintaining a fixed average connectivity $\bar{c} = K$. Let us denote $P_p(c)$ the probability distribution of the connectivities and $k = K/2$, number of neighbors of a node in either direction (clockwise or anticlockwise).

Since k of the initial $2k$ connections of each vertex are left untouched by the construction, the connectivity of a vertex i can be written $c_i = k + n_i$, with $K_i \geq 0$. n_i can then again be divided in two parts: $n_i^1 \leq k$ links have been left in place (each one with probability $1 - p$), the other $n_i^2 = n_i - n_i^1$ links have

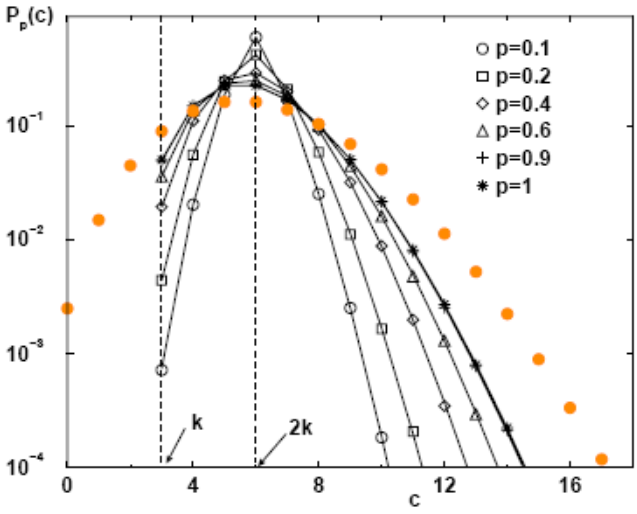

Figure 2: Probability distributions of the connectivity c for $k = 3$ and various values of p: $c \geq k$, and the mean connectivity is $\overline{c} = 2k = 6$. $N = 1000$.

been reconnected towards i, each one with probability p/N.

$$P_1(n_i^1) = \binom{k}{n_i^1}(1-p)^{n_i^1}p^{k-n_i^1} \tag{6}$$

$$P_2(n_i^2) = \frac{(kp)^{n_i^2}}{n_i^2!}\exp(-pk) \tag{7}$$

$$P_p(c) = \sum_{n=0}^{min(c-k,k)} \binom{k}{n}(1-p)^n p^{k-n} \times \frac{(kp)^{c-k-n}}{(c-k-n)!}\exp(-pk), c \geq 0 \tag{8}$$

in Figure 2 the probability distributions for k =3 and various values of p: as p grows, the distribution becomes broader.

3.2 Chemical distances

The chemical distance between the vertices of a graph is the minimal number of links between two vertices. We note d_{ij} the chemical distance between vertices i and j, and

$$\ell(N,p) = \frac{1}{N(N-1)}\overline{\sum_{i\neq j} d_{ij}} \tag{9}$$

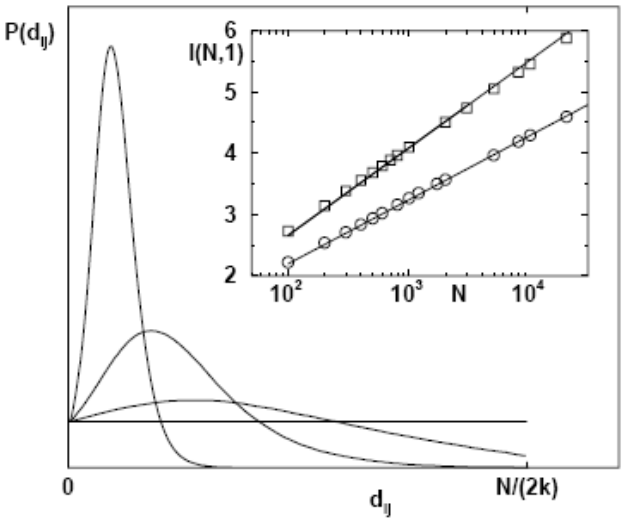

Figure 3: Probability distribution of the distance d_{ij} between two vertices i and j of small-world graphs, for $k = 3$, $N = 2000$, $p = 2^{-20}$(flat distribution), $p = 2^{-12}$, $p = 2^{-10}$, $p = 2^{-8}$ (curves becoming more and more peaked as p grows), averaged over 500 samples for each p. The maximum value of d_{ij} is of course $\frac{N}{2k}$. Inset: $\ell(N, 1)$ versus N for $k = 3$ and $k = 5$, together with the $ln(N)/ln(2k - 1)$ straight lines.

the mean chemical distance, averaged over all pairs of vertices and over the disorder induced by the rewiring procedure. Watts and Strogatz have shown that the mean distance between vertices $\ell(N, p)$ decreases very rapidly as soon as p is non-zero. For p = 0, we have a linear chain of sites, so that we easily find

$$\ell(N, p) = \frac{N(N + 2k - 2)}{4k(N - 1)} \sim \frac{N}{4k} \tag{10}$$

growing like N. On the other hand, for p = 1 $\ell(N, 1)$ grows like $\frac{\ln N}{\ln 2k-1}$, the graph is then random. Besides, the distribution of lengths, being uniform between 1 (shortest possible distance) and $N/(2k)$ for the linear chain, becomes more and more peaked around its mean value as p grows. (see Figure 3)

In Figure 4, we plot $\ell(N, p)/\ell(N, 0)$ for various values of N and $k = 2$. It is clear that $\ell(N, p)$ decreases very fast already for small p (note the logarithmic scale for p): from this point of view, the network is very soon similar to a random network.In particular, as N becomes larger, the drop in the curve occurs for smaller and smaller values of p, showing that no finite critical value of p can be determined this way: in the thermodynamic limit, $\ell(N, p)/\ell(N, 0)$ goes to 0 for all $p > 0$.

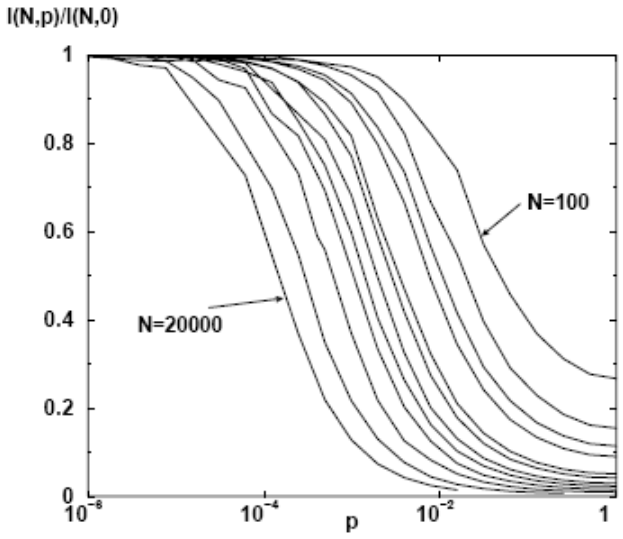

Figure 4: Mean chemical length $\ell(N,p)$ normalized by $\ell(N,0)$, versus p, for k = 2, and N from 100 to 20000: the drop in the curve occurs at lower and lower values of p as N grows.

3.3 Clustering coefficient

Clustering coefficient $C(p)$ quantifies the "cliquishness" of the small world network. It is defined as follows: if c_i is the number of neighbors of a vertex i, there are a priori $c_i(c_i-1)/2$ possible links between these neighbours. Denoting C_i the fraction of these links that are really present in the graph, $C(p)$ is the average of C_i over all vertices. On a linear-log plot, $C(p)/C(0)$ is close to 1 for a wide range of values of p, and its drop occurs around $p \approx 0.01$.

For $p = 0$, each vertex has $2k$ neighbours; the number of links between these neighbours is $\aleph_0 = 3k(k-1)/2$. Then $C(0) = \frac{3(k-1)}{2(2k-1)}$. For $p > 0$, two neighbours of i that were connected at $p = 0$ are still neighbours of i and linked together with probability $(1-p)^3$, up to terms of order $1/N$. The mean number of links between the neighbours of a vertex is then clearly $\aleph_0(1-p)^3 + O(1/N)$. The clustering coefficient $C(p)$ is defined as the mean of the ratio $C_i = \frac{\aleph_i}{c_i(c_i-1)/2}$. If instead we define $\overline{C}(p)$ as the ratio of the mean number of links between the neighbours of a vertex and the mean number of possible links between the neighbours of a vertex, we obtain

$$\overline{C}(p) = \frac{3(k-1)}{2(2k-1)}(1-p)^3 \tag{11}$$

We see in Figure 5 that the difference between $C(p)$ and $\overline{C}(p)$ is very small. The behavior of $C(p)$ is therefore very simply described by $C(p) \approx C(0)(1-p)^3$, and

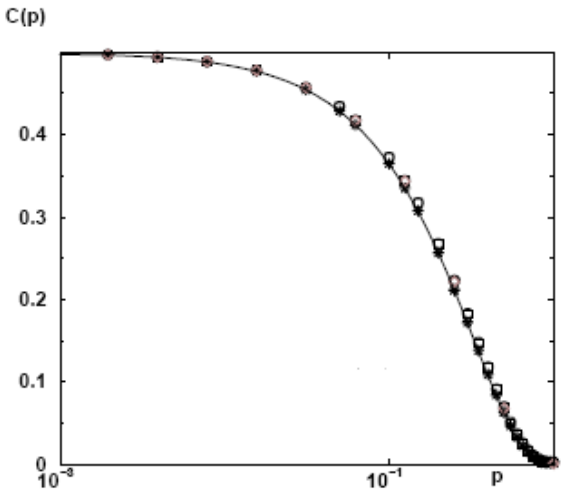

Figure 5: $C(p)$ and $\overline{C}(p)$ versus p, for $k = 2$ ($C(0) = \overline{C}(0) = 0.5$), N = 1000,2000,5000: open symbols are for $C(p)$, and the crosses are for $\overline{C}(p)$; the line is $C(0)(1-p)^3$.

the dependence on N is very small.

4 Spectral properties

The set of all eigenvalues of a characteristic matrix of a graph, also referred to as the spectrum, is a well-known topology retrieval method. In other words, important information on the topological properties of a graph can be extracted from the eigenvalues of the associated adjacency, Laplacian or any other type of matrix.

4.1 Spectra of the Internet Graphs

Let G be a graph, and let $\aleph$ and denote the node set and the link set, consisting of $N = |\aleph|$ and $L = ||$, respectively. The Laplacian matrix of a graph G with N nodes is an $N \times N$ matrix $Q = \triangle - A$ where $\triangle = diag(D_i)$, D_i is the degree of the node $i \in N$ and A is the adjacency matrix of G. The eigenvalues of Q are called the Laplacian eigenvalues. The Laplacian eigenvalues are all real and nonnegative; they are contained in the interval $[0, \min(N, 2D_{max})]$, where D_{max} is the maximum degree of G.

The set of all N Laplacian eigenvalues $\lambda_N = 0 \leq \lambda_{N-1} \leq \ldots \leq \lambda_1$ is

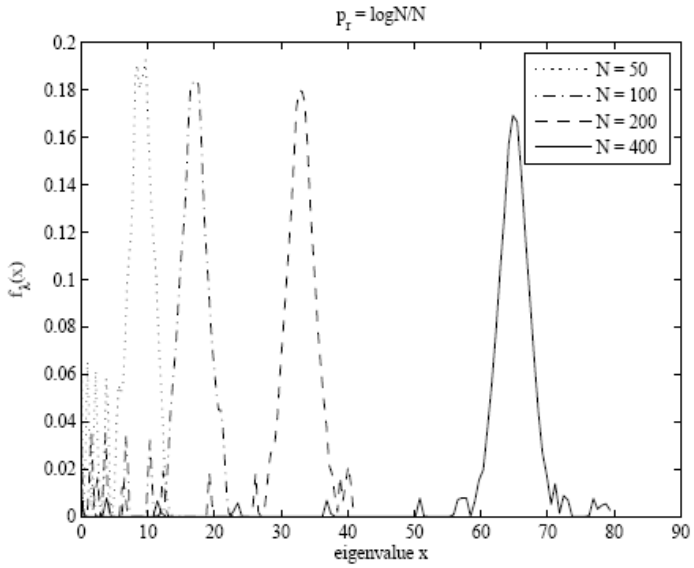

Figure 6: The Laplacian spectrum of the Watts and Strogatz small-world graph with $N = 50, 100, 200, 400$ and $p = \frac{\log N}{N}$.

called the Laplacian spectrum of a graph G. The second smallest eigenvalue is $\lambda_{N-1} \geq 0$, but equal to zero only if a graph is disconnected. Thus, the multiplicity of 0 as an eigenvalue of Q is equal to the number of components of G.

4.2 Spectra of the Small-World Graph of Watts and Strogatz

For the small rewiring probability $p = 0$ the Watts and Strogatz small-world graph is regular and also periodical. Because of the highly ordered structure, for small p the spectrum is highly skewed with the bulk towards the high eigenvalues, distributed around the mean nodal degree, which, irrespective of p equals $E[D] = 2k$. The spectrum of the two-dimensional lattice graph with $N \times N$ nodes aims to illuminate this effect. The Laplacian spectrum of the two dimensional lattice is the sum of two path graphs P_N whose eigenvalues are $\lambda_i(P_N) = 2 - 2\cos(\pi i/N),\ i = 1, 2, \ldots, N$. Consequently, the spectrum of the two-dimensional lattice converges to a pointy shape with a peak centered around the mean nodal degree.

The Laplacian spectrum of the ring lattice $C(N, k)$ with N nodes and $2k$ neighbors comprises the eigenvalues $\lambda_i(C(N, k)) = 2k - \left(\frac{\sin(\frac{\pi}{N}(i-1)(2k-1))}{\sin(\frac{\pi}{N}(i-1))} - 1\right)$, $i = 1, 2, \ldots, N$. Upon increasing k-regularity, the bulk part of the spectrum

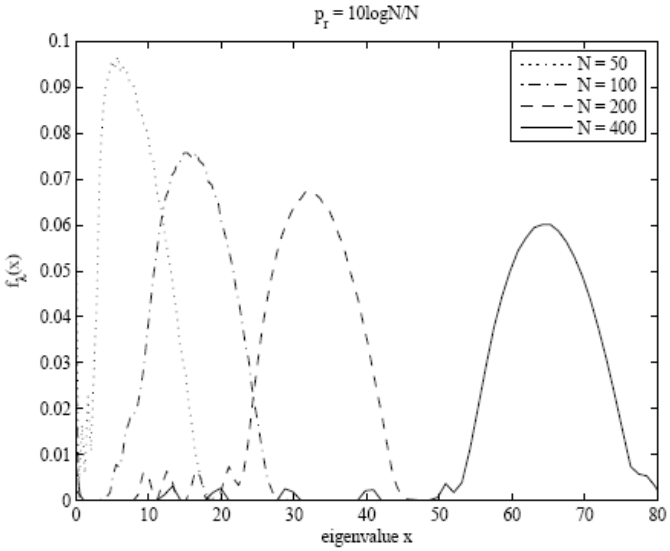

Figure 7: The Laplacian spectrum of the Watts and Strogatz small-world graph with $N = 50, 100, 200, 400$ and $p = \frac{10 \log N}{N}$.

shifts towards the mean nodal degree, similar to the Laplacian spectrum of the Erdos-Renyi random graph as shown in the figure 6 and 7.

5 Search in Small World

Its often required agents must search or navigate in a large decentralized small world network without knowledge of the global structure. The question is how to find links between vertices in small world with only local knowledge. Decentralized Algorithm of Kleinberg is studied here for that purpose. Consider a 2D grid lattice structure of $n \times n$ with long links added with some probability p for each original edge. What's given is set of all local contacts among all nodes, the location on the lattice of the target and the location and long range contact of all nodes that have come in contact of the message. The choice of which neighbor to contact is to be produced.

Delivery time is defined as expected number of steps, over random generation of graph and random start and target. The algorithm should have exponentially better delivery time, polynomial function of log n, not n (polylogarithmic). Kleinberg stated that There is a constant $c > 0$ such that the delivery time of any decentralized algorithm in the Watts-Strogatz model is at least $cn^{2/3}$. For two vertices v (i,j) and u (k,l), the lattice distance $d(v, u)$ is evaluated as $d(v, u) = |k - i| + |l - j|$.

Step 1. For each vertex v, add directed link to random long-range contact.

Step 2. Choose vertex w as the contact with probability proportional to $d(v,w)^{-r}$

There exist constants c_r ($r \geq 0$) such that for $r = 2$, there is a decentralized algorithm with delivery time $\leq c_2(\log n)^2$; and for any $r < 2$, the delivery time is $\geq c_r n^{(2-r)/3}$; and for any $r > 2$, the delivery time is $\geq c_r n^(r-2)/(r-1)$.

In other words rather than adding the long-range shortcuts uniformly at random, we add links between nodes of this network with a probability that decays like the d^{th} power of their distance (in d dimensions). This is the only link distribution of this form for which efficient search is possible. The intuition here is that a probability decaying like the d^{th} power of the distance is in fact uniform over all "distance scales" – a node is roughly as likely to form links at distances 1 to 10 as it is at distances 10 to 100, 100 to 1000, and so on.

The ability to construct a searchable network in this way, with long-range links whose probabilities decay with distance, has proved useful in the design of peer-to-peer file-sharing systems on the Internet, where content must be found by nodes consulting one another in a decentralized fashion. In other words, nodes executing these look-up protocols are behaving very much like participants in the Milgram experiments – a striking illustration of the way in which the computational and social sciences can inform one another, and the way in which mathematical models in the computational world turn into design principles with remarkable ease.