M.E.J. Newman: Models of the Small World

A Review

Tiina Lindh-Knuutila

Adaptive Informatics Research Centre
Helsinki University of Technology

November 7, 2007

Vocabulary

- N number of nodes of the graph
- ℓ average distance between nodes
- D diameter of the graph
- d is the number of dimensions of the lattice
- z number of connections each node has – also called the coordination number of a graph
- C clustering coefficient of the graph
- ξ characteristic length-scale of a small-world model
- p is the probability of creating a link between two vertices in small-world model

Small-world phenomenon

Definition

A network is a small world network when two arbitrary nodes of the network are connected with a short chain of intermediate links.

Study of the distribution of path lengths in a social network (Milgram 1967)

- Letters addressed to a stockbroker in Boston, Mass. divided to random people in Nebraska
- To be passed along to a first-name acquaintance possibly nearer to the recipient in social sense
- The letters reached the recipient on average six steps

Random graph model

- First model of a small world
- Very simple model of social network (Erdös and Rènyi, 1959)
- N nodes, $\frac{1}{2Nz}$ edges between randomly drawn pairs of nodes
- Shows small-world effect
- Diameter of the graph increases slowly with the system size N: $D = \frac{\log(N)}{\log(z)}$

Why a random graph is not enough?

Clustering

- Real-world graphs show clustering effects
- Your friends are usually friends with each other as well
- Random graph does not have clustering properties

Network	C	C_{rand}
movie actors	0.79	0.00027
neural network	0.28	0.05
power grid	0.08	0.0005

Table: (Excerpt of Table 1 in Newman 2000) The clustering coefficients C for three real-world networks and the value for C in a random graph with the same parameters.

Motivation and background

How to balance between

- the small-world properties i.e. the slow increase of path length with system size and
- the clustering effect?

Opposite of a random graph: a completely ordered lattice

- $1 \ldots n$ dimensions
- Clustering coefficient $C = \frac{3(z-2d)}{4(z-d)}$ tends to $\frac{3}{4}$ for $z \gg 2d$
- No small-world effect – in $1D$ case, average distance grows linearly with system size

Model of Watts and Strogatz

- Watts and Strogatz model from 1998
- Balances between the clustering property of a regular lattice and small-world properties of a random graph

Creation of a small-world graph

- Begin with a low-dimensional regular lattice
- Randomly rewire some of the links with a probability p
- For small p a mostly regular graph is produced
- Small-world properties are obtained through the randomly wired links

Model of Watts and Strogatz

(a) (b) (c)

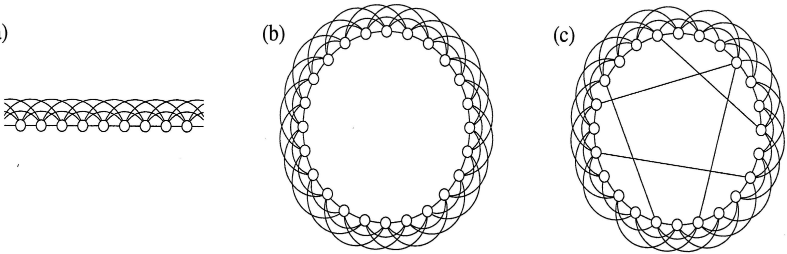


Figure 1: (a) A one-dimensional lattice with each site connected to its z nearest neighbors, where in this case $z = 6$. (b) The same lattice with periodic boundary conditions, so that the system becomes a ring. (c) The Watts–Strogatz model is created by rewiring a small fraction of the links (in this case five of them) to new sites chosen at random.

Analysis of the small-world model

- Numerical analysis indicates that the small-world model shows the log-increase of average path length and clustering properties simultaneously.
- A summary of some analytical results follows.
- We want to measure the average path length (or vertex-vertex distance) ℓ and find out what is the shape of its distribution.
- When we balance between ordered and random graph how does the transition from large-world to small-world occur?

Variation by Newman and Watts (1999)

Most analytical work has been done using a variation of Watts and Strogatz model

Usually both models are referred as small-world models

Differences to Watts-Strogatz model

- Added shortcuts
- No links removed from the underlying lattice
- No disconnected parts of the graph
- Easier to analyze as no distance is infinite

Findings by Barthémély and Amaral

- Average vertex-vertex distance obeys $\ell = \xi G(L/\xi)$, where ξ is the length-scale for the model and $G(x)$ a universal scaling function.
- ξ is assumed to diverge in the limit of small p according to $\xi \sim p^{-\tau}$
- Based on numerical simulations, they assumed that $\tau = \frac{2}{3}$
- Barrat (1999) disproved the numerical result and concluded that τ cannot be less than 1

A single length-scale of a small-world graph

Results from Newman and Watts 1999

Newman and Watts showed using numerical simulation and series expansion that

- there is a single, non-trivial length-scale in the small-world model that depends on the probability p,
- given by $\xi = \frac{1}{(pzd)^{1/d}}$ in general case

Definition

The average vertex-vertex distance scales with the system size according to
$\ell = \frac{L}{2dz} F(pzL^d)$, where $F(x)$ is a universal scaling function.
ξ diverges as $p \longrightarrow 0$

Interpretation of the average path length

- The average path length, ℓ, is defined by a single scalar function of a single scalar variable, if $\xi \gg 1$
- If we know the form of this function, we know everything.
- True only for small p, i.e. when most person's connections are local.
- In the limit $p \to 0$ model is a 'large-world' and typical path length tends to $\ell = \frac{L}{2z}$
- Scaling form shows that we can go from large-world to small-world either by increasing p or increasing system size

Interpretation of x and $F(x)$

Average path length equation again

$$\ell = \frac{L}{2dz} F(pzL^d)$$

- x is twice the average number of shortcuts for a given value of p
- $F(x)$ is the average fraction by which the vertex-vertex distance is reduced for a given value of x
- It takes about $5\frac{1}{2}$ shortcuts to reduce the average vertex-vertex distance by a factor of two, and 56 to reduce it by a factor of ten.

Further analysis of the results

- In the limit of large p, the small-world models becomes a nearly random graph
- ℓ should scale logarithmically with system size L when p is large and also when L is large
- When small L or p, ℓ should scale linearly with L
- Cross-over from small- and large-x in the vicinity of $L = \xi$

Limiting forms for $F(x)$

$$F(x) = \begin{cases} 1 & \text{for } x \ll 1 \\ (\log x)/x & \text{for } x \gg 1 \end{cases}$$

Open questions in the small-world models

- Actual distribution of path lengths in the small-world model
- The calculation of the exact average path length ℓ
- Exact analytical calculations very hard for the small-world model

Some attempts towards the distribution and average path length

- The form of the scaling function calculated for $d = 1$ and small or large x but not for $x \simeq 1$ (Newman et al. 2000) $F(x) = \frac{4}{\sqrt{x^2+4x}} \tanh^{-1} \frac{x}{\sqrt{x^2+4x}}$
- In addition, a mean-field approximation was used to solve the distribution
- Can be used as a simple model of a spread of disease

Dynamical systems defined on small-world graphs

Several studies use small-world structures instead of regular lattices in dynamical systems problems:

- Cellular automata: density classification becomes easier
- In simple games: e.g. multi-player Prisoner's dilemma is more difficult
- In oscillators: Small-world topology helps oscillators to synchronize

Other applications

- Solution for the ferromagnetic Ising model for d=1 with a phase transition in a finite temperature
- Small-world graph as a model of a neural network: able to produce fast responses to external stimuli and coherent oscillation.
- Model of species coevolution

Disease spread in small-world graphs

Small-world graphs are suitable for modeling spread of disease (or information) in a population

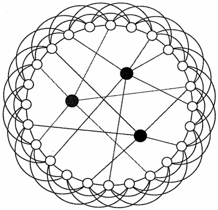
- First idea: use the approximate distribution of ℓ as a simple model.
  - Disease spreads from neighborhoods of the infected people
  - Number of people n infected after t time steps: those t steps away from the initial carrier.
- More complex idea: Only a certain fraction q is susceptible
  - What does the fraction q need to be to make disease an epidemic?

Multiple scales in small world graphs, (Kasturirangan, 1999)

Definition

The small-world phenomenon arises because there are few 'hubs' in the network that have unusually high number of neighbors, not because a few long-range connections.

- Shows small-world effects even with one sufficiently-connected hub
- For a graph with one single central hub, it is possible to calculate the scaling function exactly



The Albert and Barabási model

Network model based on their observations on the World Wide Web

- Small-world models operate only on sparse graphs.
- Highly connected sites dominate the Web.
- Distribution of the coordination numbers of sites is not bimodal but follows power-law.
- Does not show clustering which is present in the Web as well.

Creating an Albert and Barabási network

Network creation algorithm

- Start with a random network
- Take two vertex at random and add a link if it brings the distribution of z nearer to power-law
- Continue until correct coordination numbers reached
- Still otherwise as random graph

The network could also be created by generating N vertices with lines out of them according to power-law distribution and joining lines randomly until none are left.

The Kleinberg model

This model was discussed in detail in the previous lecture

- Comment on Watts-Strogatz model: No simple algorithm for finding the path using only local information
- For Kleinberg's model there is
  - a simple algorithm for finding a short path using only local information
  - for those structures for which the exponent of the power law is the dimension of the grid

Comments from the article:

- For other values of r than $r = d$ path-finding becomes a hard task.
- There is more to the small-world effect than the existence of short paths.

Conclusions

- Overview on some theoretical work on the small-world phenomenon
- Analytic and numerical results for Watts-Strogatz model and its variants
- Continuing research to determine the exact structure

Most important points

- Small-world network behavior different from either regular graph or a random one
- Transition from large-world to small-world implication: disease or information spreads first as a power of time, then changes to exponential increase and flattens off when the graph becomes saturated
- Dynamical systems behave differently on small-world graphs than on regular lattices
- There are other characteristics in addition to small-world effect: e.g. scale-free distribution

Further reading

Kasturirangan, R.
Multiple scales in small-world graphs.
Massachusetts Institute of Technology AI Lab Memo 1663. 1999. Also cond-mat/9904055

Newman, M.E.J.
The Structure and Function of Complex Networks,
SIAM Reviews, 45(2): 167-256, 2003. Also cond-mat/0303516

Watts, D.J. and Strogatz, S.H.
Collective dynamics of “small-world” networks.
Nature, 393, 440–442. 1998.