Sharp dimension transition in a shortcut model

2008 J. Phys. A: Math. Theor. 41 285001

(http://iopscience.iop.org/1751-8121/41/28/285001)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 285001 (7pp) doi:10.1088/1751-8113/41/28/285001

Sharp dimension transition in a shortcut model

O Shanker[1]

Hewlett Packard Company, 16 399 W Bernardo Dr., San Diego, CA 92130, USA

E-mail: oshanker@gmail.com

Received 6 April 2008, in final form 28 May 2008
Published 17 June 2008
Online at stacks.iop.org/JPhysA/41/285001

Abstract
Earlier studies of a parametrized class of models, the shortcut models, whose fractal dimension transitions between integer values, indicated that the transition occurs infinitely sharply at the parameter value $p = 0$, as the system size increases to infinity. In this work, we explain the cause as being due to the combinatorially large increase in the available number of paths connecting a pair of points as the path length is increased.

PACS numbers: 05.45.Df, 89.75.HC, 05.50.+q

1. Introduction

The shortcut model was introduced [1, 2] while studying the dimension [3] of complex networks (graphs) [4–13]. The model interpolates between discrete regular lattices. It was observed that the fractal dimension apparently transitions sharply [14, 15] from one to two dimensions at the probability of shortcuts $p = 0$. In this work we provide an explanation of this property. The sharp transition occurs because of the combinatorially large increase in the available number of paths connecting a pair of points as the path length is increased. We also study the mean path length and consider some generalizations of the model.

The model is defined by starting with a regular discrete lattice of dimension d with periodic boundary conditions, and adding shortcuts between remote vertices in the lattice. If the shortcuts connect vertices a constant distance apart in one of the d dimensions, then the resulting complex network transitions from a d-dimensional regular lattice to a $(d + 1)$-dimensional regular lattice as the number of shortcuts is increased. The case of the transition from a one-dimensional to a two-dimensional lattice has been well studied.

Section 2 reviews the shortcut model and proposes a possible generalization. It also reviews the complex network zeta function. Section 3 studies the mean path length. Section 4 presents the proof that the dimension transitions infinitely sharply in the model as p increases from zero. Finally the conclusions are presented.

[1] http://www.geocities.com/oshanker/graphdimension.htm.

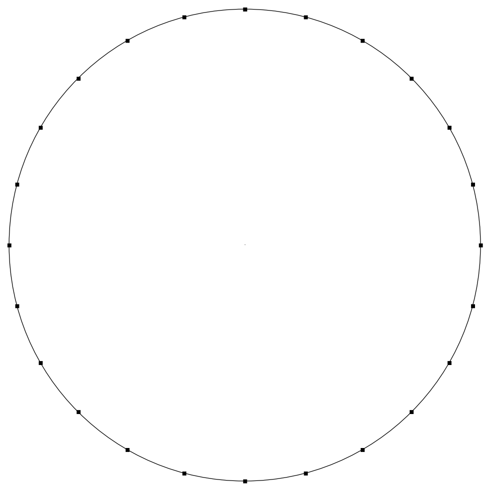

Figure 1. Shortcut model, $p = 0$.

2. Shortcut model

In this section we give a brief definition of the shortcut model [1, 2], and we specify a possible generalization. The model has fractal dimension as defined by the complex network zeta function [2], and transitions from a one-dimensional system to a two-dimensional system. The starting network is a one-dimensional lattice of N vertices with periodic boundary conditions. Each vertex is joined to its neighbors on either side, which results in a system with N edges. The network is extended by taking each node in turn and, with probability p, adding an edge to a new location m nodes ahead. We require that $N \gg m \gg 1$, say $m = \sqrt{N}$. The graphs are parametrized by

$$\text{size} = N, \tag{1}$$

$$\text{shortcut distance} = m \tag{2}$$

and

$$\text{shortcut probability} = p. \tag{3}$$

When the shortcut probability $p = 0$ (figure 1), we have a one-dimensional regular lattice of size N. The nodes are connected by edges represented by the arcs of the circle. When $p = 1$ (figure 2), every node is connected by a shortcut edge to a new location. Each node now has edges in two directions, the first along the original direction (on the circle), and the second along the shortcut edges. The graph is essentially a two-dimensional graph with m and N/m nodes in each direction. For p between 0 and 1 (figure 3), we have a graph which interpolates between the one- and two-dimensional systems.

It is interesting to note the difference between the shortcut model and the 'small-world model' of Watts and Strogatz [16–18]. In the small-world model also one starts with a regular lattice and adds shortcuts with probability p. However, the shortcuts are not constrained to connect to a node a fixed distance ahead. Instead, the other end of the shortcut can connect to any randomly chosen node. As a result, the small-world model tends to a random graph rather than a two-dimensional graph as the shortcut probability is increased.

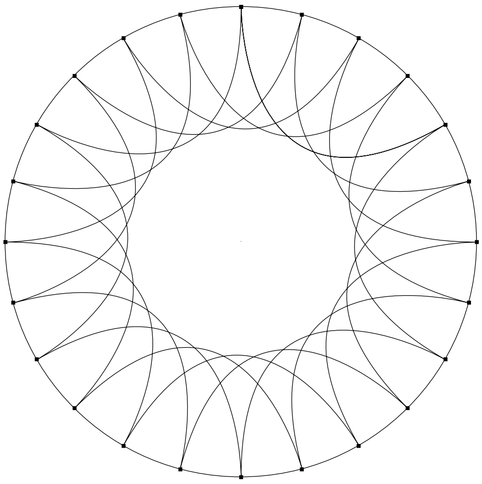

Figure 2. Shortcut model, $p = 1$.

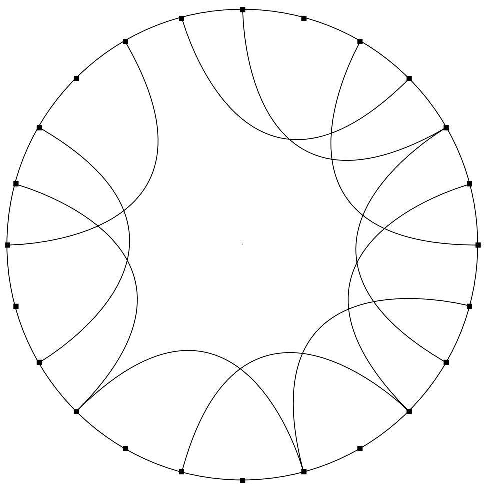

Figure 3. Shortcut model, $0 < p < 1$.

One possible generalization of the shortcut model is to have a hierarchy of shortcuts. For example, the network can be extended by taking each node in turn and, with probability p_1, adding an edge to a new location m_1 nodes ahead, and with probability p_2, adding an edge to a new location m_2 nodes ahead. We require that $N \gg m_1 \gg m_2 \gg 1$. The dimension of the extended model would lie between d and $d+2$, where d is the dimension of the starting lattice.

The dimension of a complex network is defined as the value at which the complex network zeta function transitions from non-convergence to convergence in the infinite system limit. For completeness, the complex network zeta function is defined below. The complex network zeta function $\zeta_G(\alpha)$ is defined as

$$\zeta_G(\alpha) := \frac{1}{N} \sum_i \sum_{j \neq i} r_{ij}^{-\alpha}, \tag{4}$$

where N is the graph size measured by the number of nodes and r_{ij} is the distance between nodes i and j. The distance is measured by the number of edges along the shortest path connecting two nodes. One could generalize this by considering weighted edges. The definition equation (4) can be expressed as a weighted sum over the node distances. The graph surface function, $S(r)$, is defined as the number of nodes which are exactly at a distance r from a given node, averaged over all nodes of the network. This gives the Dirichlet series expression for the complex network zeta function:

$$\zeta_G(\alpha) = \sum_r S(r)/r^{\alpha}. \tag{5}$$

When the exponent α tends to infinity, the sum in equation (4) gets contributions only from the nearest neighbors of a node. The other terms tend to zero. Thus, $\zeta_G(\alpha)$ tends to the average vertex degree for the complex network. When α is zero the sum in equation (4) gets a contribution of one from each node. This means that $\zeta_G(\alpha)$ is $N - 1$, and hence tends to infinity as the system size increases.

Furthermore, $\zeta_G(\alpha)$ is a decreasing function of α. Thus, if it is finite for any value of α, it will remain finite for all higher values of α. If it is infinite for some value of α, it will remain infinite for all lower values of α. Thus, there is at most one value of α, $\alpha_{\text{transition}}$, at which $\zeta_G(\alpha)$ transitions from being infinite to being finite. This is reminiscent of the behavior of Hausdorff dimension [3]. We define the complex network dimension as the value of the exponent α at which $\zeta_G(\alpha)$ transitions from being infinite to being finite. To further motivate the definition of dimension using the complex network zeta function, we consider the regular discrete d-dimensional lattices $\mathbf{Z}^d$ as graphs with the edges being the links between the points along the coordinate axes. Since we are defining distance as the number of edges between two nodes along the shortest path, the graph distance r corresponds in this case to the distance defined by the L^1 norm:

$$r = \|\vec{n}\|_1 = \|n_1\| + \cdots + \|n_d\|, \tag{6}$$

where $\vec{n}$ is the vector joining the two points on the regular discrete lattice. For the regular lattices (considered as graphs) the transition of the complex network zeta function from non-convergence to convergence occurs when α is equal to the dimension d, as one would expect.

In section 3 we study the mean path length. The mean path length is given in terms of the graph surface function or the complex network zeta function by

$$\ell = \sum_r rS(r)/\sum_r S(r) = \lim_{N\to\infty} \zeta_G(-1)/\zeta_G(0). \tag{7}$$

In equation (7) we have normalized the mean path length using the number of nodes N in the denominator, rather than the maximum possible number of edges, $N(N-1)/2$. For the scaling behavior of the mean path length with N, this just results in the scaling exponent increasing by 1. We do this merely for the convenience of working with positive scaling exponents.

3. Mean path length

In this section we calculate the mean path length for the shortcut model. For a d-dimensional system we expect the mean path length ℓ to scale with size N as $N^{1/d}$. We show that the scaling exponent changes sharply as p increases from 0. This provides further evidence for the sharp transition of the dimension of the shortcut model at $p = 0$, a property which we explain in section 4.

When the shortcut probability $p = 0$ the mean path length scales linearly. We calculated the mean path lengths using equation (7) for different p for N varying from 1000 to 16 384.

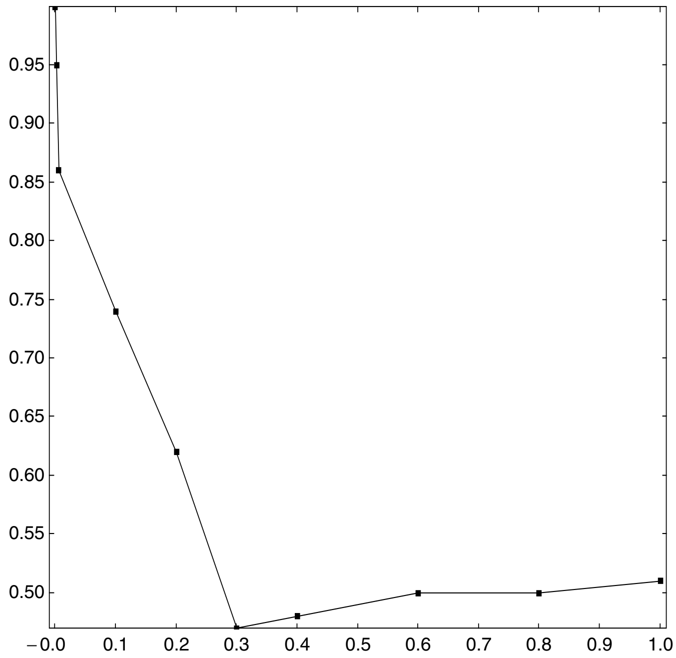

Figure 4. Scaling exponent for mean path length versus shortcut probability. The vertical axis is the scaling exponent for the mean path length and the horizontal axis is the shortcut probability.

Table 1. Scaling exponent for mean path length.

Probability	Scaling exponent	Dimension
0.002	0.95	1.05
0.006	0.86	1.16
0.1	0.74	1.35
0.2	0.62	1.61
0.4	0.48	2.08
0.6	0.50	2.00
0.8	0.50	2.00
1.0	0.51	1.96

Table 1 shows the scaling exponent for ℓ. Figure 4 shows a plot of the scaling exponent versus shortcut probability. Note, in particular, the rather sharp drop in the exponent from the value 1 as p increases from 0. The scaling exponent shows the change in behavior from a one-dimensional system to a two-dimensional system.

In the following section, we provide an explanation for the sharp transition of the dimension at $p = 0$.

4. Theory for sharp transition

In this section we show that the dimension of the shortcut model transitions sharply at $p = 0$. The sharp transition is related to the combinatorially large increase in the available number of paths connecting a pair of points as the path length is increased, as explained below.

For a one-dimensional regular lattice the graph surface function $S_1(r)$ in equation (5) is exactly two for all values of r. This is because there are two nearest neighbors, two next-nearest neighbors, etc. Thus, the complex network zeta function $\zeta_G(\alpha)$ of equation (4) is equal to $2\zeta(\alpha)$, where $\zeta(\alpha)$ is the usual Riemann zeta function. Thus, the transition from non-convergence to convergence occurs at $\alpha = 1$. The graph surface function $S_d(r)$ for a lattice of dimension d is a polynomial in r of order $d-1$ for a given d. From combinatorics it can be written as

$$S_d(r) = \sum_{i=0}^{d-1} (-1)^i 2^{d-i} \binom{d}{i} \binom{d+r-i-1}{d-i-1}. \tag{8}$$

It scales aymptotically as $S_d(r) \rightarrow 2^d r^{d-1}/\Gamma(d)$ for large r. $r \rightarrow \infty$ corresponds to $\alpha \rightarrow \alpha_{\text{transition}}$. Thus, $\zeta_G(\alpha) \rightarrow 2^d \zeta(\alpha - d + 1)/\Gamma(d)$ as $\alpha \rightarrow \alpha_{\text{transition}}$. The largest pole of $\zeta_G(\alpha)$ occurs for $\alpha = d$.

Now consider the shortcut model which starts with a one-dimensional regular lattice. When $p = 0$ the graph surface function is constant as the path length is varied. When $p = 1$ the graph surface function increases linearly with the path length r, i.e., $S(r) = 4r$. The increase in the graph surface function arises because points which were far apart when $p = 0$ are brought closer together by the shortcut edges when $p = 1$. This much is fairly straightforward. The interesting question is the behavior when p lies between 0 and 1. Is there a smooth change in the complexity measure as p is increased? Let us consider a pair of points which are separated by i shortcut edges and $r - i$ normal edges when $p = 1$. We show that with very high probability, the distance between this pair of nodes will be the same as it is in the two-dimensional limit, for all non-zero values of p. Thus, the transition is sharp, like in a phase transition. There are

$$\binom{r}{i}$$

possible different paths connecting the two points, corresponding to all the possible ways of choosing the i shortcut edges in the total path of length r. This is a very large number when r and i are large. Consider a particular path from among this large set. When p lies between 0 and 1 the probability that this particular path will have all the i shortcut edges present is given by p^i. The probability that the particular path does not have all the required shortcut edges is $(1 - p^i)$. Thus, any one particular path will have a small probability of having all the required shortcut edges present. However, we need only one path from among the huge number of available paths to have all the required shortcuts. We first estimate the probability $p_{\text{no-paths}}$ that none of the possible paths have the required shortcuts. Since different paths share edges, the probabilities for different paths are not independent. This complicates the analysis. To get an order of magnitude estimate, we ignore the correlations between the probabilities for the different paths. This is similar in spirit to the mean field theories. Ignoring the correlations, $p_{\text{no-paths}}$ is given by

$$p_{\text{no-paths}} = (1 - p^i)^{\binom{r}{i}}. \tag{9}$$

For large r and i, equation (9) can be written as

$$p_{\text{no-paths}} = \exp\left(-\left(\frac{p^f}{f^f\left(1-f\right)^{1-f}}\right)^r \frac{1}{\sqrt{2\pi r f\left(1-f\right)}}\right), \tag{10}$$

where $f = i/r$. $p_{\text{no-paths}}$ goes to zero when $f < f_{\text{critical}}$ where f_{critical} is the solution of the transcendental equation

$$p = f(1 - f)^{\frac{1-f}{f}}. \tag{11}$$

For a finite fraction of nodes $p_{\text{no-paths}}$ rapidly goes to zero when p lies strictly between 0 and 1 because of the combinatorially large value of the exponent. Thus, the probability $p_{\text{path-exists}}$ that there is at least one path which has all the i required shortcuts

$$p_{\text{path-exists}} = 1 - (1 - p^i)^{\binom{r}{i}}, \tag{12}$$

is very close to 1. When p lies strictly between 0 and 1, the dependence of the graph surface function on distance becomes essentially the same as for the case $p = 1$, i.e., it becomes linear with slope greater than $4f_{\text{critical}}$. This accounts for the sudden transition observed empirically in earlier studies for the dimension, when p increases from 0.

5. Conclusions

In this work we showed that the dimension of the shortcut model shows a sharp transition at $p = 0$. The sharp transition had been conjectured earlier, based on the study of the complex network zeta function, and on processes like the random walk. The property is now explained as being due to the large number of potential paths available between a given pair of points.

References

[1] Shanker O 2007 Defining dimension of a complex network *Mod. Phys. Lett.* B **21** 321–6
[2] Shanker O 2007 Zeta function and dimension of complex network *Mod. Phys. Lett.* B **21** 639–44
[3] Falconer K 2003 *Fractal Geometry: Mathematical Foundations and Applications* 2nd edn (New York: Wiley)
[4] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D-U 2006 Complex networks: structure and dynamics *Phys. Rep.* **424** 175–308
[5] Shanker O and Motta Giovanni 2007 Use of word relationship graphs to analyze texts in different languages Technical Report http://www.geocities.com/oshanker/TextAnalysis.pdf
[6] Shanker O 2007 Complex network dimension and path counts Technical Report http://www.geocities.com/oshanker/complexNetworkZetaFunction.pdf
[7] Newman M E J 2003 The structure and function of complex networks *SIAM Rev.* **45** 167–256
[8] Albert R and Barabási A-L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47–97
[9] Dorogovtsev S N and Mendes J F F 2003 *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford: Oxford University Press)
[10] Newman M E J, Barabási A-L and Watts D J 2006 *The Structure and Dynamics of Networks* (Princeton, NJ: Princeton University Press)
[11] Girvan M and Newman M E J 2002 Community structure in social and biological networks *Proc. Natl Acad. Sci. USA* **99** 7821–6
[12] Newman M E J 2004 Detecting community structure in networks *Eur. Phys. J.* B **38** 321–30
[13] Lee C-Y and Jung S 2006 Statistical self-similar properties of complex networks *Phys. Rev.* E **73** 066102
[14] Shanker O 2008 Algorithms for fractal dimension calculation *Mod. Phys. Lett.* B **22** 459–66
[15] Shanker O 2008 Random walk in shortcut models *Mod. Phys. Lett.* B **22** 727–33
[16] Watts D J and Strogatz S H 1998 Collective dynamics of 'small-world' networks *Nature* **393** 440–2
[17] Watts D J 1999 *Small Worlds* (Princeton, NJ: Princeton University Press)
[18] Watts D J 1999 Networks, dynamics, and the small world phenomenon *Am. J. Sociol.* **105** 493–592