Dynamics of Continuous, Discrete and Impulsive Systems
Series **B:** *Applications & Algorithms* **13 (2006)** 505-510

SIMPLEX TRIANGULATION INDUCED SCALE-FREE NETWORKS

Zhi-Ming Gu[1], Tao Zhou[2,3,*], Bing-Hong Wang[2], Gang Yan[3], Chen-Ping Zhu[1] and Zhong-Qian Fu[3]

[1]College of Science, Nanjing University of Aeronautics and Astronautics
Nanjing Jiangsu, 210016, P. R. China

[2]Nonlinear Science Center and Department of Modern Physics
University of Science and Technology of China, Hefei Anhui, 230026, P. R. China

[3]Department of Electronic Science and Technology
University of Science and Technology of China, Hefei Anhui, 230026, P. R. China

Abstract. We propose a simple rule that generates scale-free networks with very large clustering coefficients and very small average distances. These networks are called simplex triangulation networks (STNs) as they can be considered as a kind of network representations of simplex triangulations. We obtain the analytic results of the power-law exponent $\gamma = 2 + \frac{1}{d-1}$ for d-dimensional STNs, and the clustering coefficient C. We prove that the increasing tendency of the average distances of STNs is a little slower than the logarithm of the number of their nodes. In addition, the STNs possess hierarchical structures as $C(k) \sim k^{-1}$ when $k \gg d$, which is in accordance with the observations of many real-life networks.

Keywords. complex network, simplex triangulation, scale-free network, small-world network, clustering coefficient, average distance

AMS (MOS) subject classification: 05C75, 05C80

1 Introduction

Recently, empirical studies indicate that networks in various fields have some common characteristics, which inspires scientists to construct a general model [1-3]. The most important characteristics are the scale-free property and the small-world effect. The former means that the degree distribution obeys a power law as $p(k) \propto k^{-\gamma}$, where k is the degree, $p(k)$ is the probability density function for the degree distribution, and γ is called the power-law exponent, which is usually between 2 and 3 in real-life networks. The latter involves two factors: small average distance as $L \sim \ln N$ and great clustering coefficient C, where L is the average distance, N is the number of nodes in the network, and C is the probability that a randomly selected node's two randomly picked neighbors are neighbors. One of the most well-known

[1]* corresponding author, email: zhutou@ustc.edu

models is Watts and Strogatz's small-world network (WS network), which can be constructed by starting with a regular network and randomly moving one endpoint of each edge with probability p [4]. Another significant one is Barabási and Albert's scale-free network model (BA network) [5]. The BA model suggests that two main ingredients of self-organization of a network in a scale-free structure are growth and preferential attachment. However, both WS and BA networks fail to mimic the reality in some aspects. Therefore, a significant question is how to generate networks displaying both scale-free property and small-world effect [6-8].

In this paper, we propose a simple rule that generates scale-free networks with very large clustering coefficients and very small average distances. These networks are called simplex triangulation networks (STNs) as they can be considered as a kind of network representations of simplex triangulations [9]. Strictly speaking, if $\vec{a_0}, \vec{a_1}, \cdots, \vec{a_d}$ are linearly independent points in $\mathbb{R}^d$, then the set $\{\sum_{i=0}^{d} \lambda_i \vec{a_i} | \lambda_i \geq 0, \sum_{i=0}^{d} \lambda_i = 1\}$ is the d-simplex. For instance, 0-simplex is a single vertex, 1-simplex is a line segment, 2-simplex is a triangle, 3-simplex is a tetrahedron, and so on. Informally speaking, simplex triangulation is the process to triangulate an original simplex into several sub-simplices. For example, choosing an arbitrary point inside a d-simplex and then linking this point to all the vertices of the simplex, this simplex will be triangulated into $d+1$ sub-simplices.

Our model starts with a d-simplex, where $d \geq 2$ is an integer. The network representation of this simplex contains $d+1$ nodes and $\frac{d(d+1)}{2}$ edges. Then, at each time step, a simplex is randomly selected, and a new node is added inside this simplex and linked to all its $d+1$ vertices. During each time step, the numbers of simplices, nodes and edges increase by d, 1, $d+1$, respectively. Repeating this simple rule, one can get a d-STN of arbitrary order.

2 The degree distribution

As mentioned above, the degree distribution is one of the most important statistical characteristics of networks. Since many real-life networks are scale-free networks, whether the networks are of power-law degree distribution is a criterion to judge the validity of a model.

Since after a new node is added to the network, the number of simplices increases by d, we can immediately get that when the network is of order N, the number of simplices is

$$N_s = d(N - d - 1) + 1 = dN - d^2 - d + 1. \tag{1}$$

Let N_s^i be the number of simplices containing node i. The probability that a newly added node will link to the ith node is N_s^i/N_s. Apparently, N_s^i will increase by $d-1$ while k_i increases by one, and when the node is newly added, it is of degree $d+1$ and contained by $d+1$ different simplices. Thus,

$$N_s(k) = k(d-1) - (d+1)(d-2), \tag{2}$$

where $N_s(k)$ denotes the number of simplices containing a node of degree k. Let $n(N,k)$ be the number of nodes with degree k when N nodes are present. Now, we add a new node to the network. Then, $n(N,k)$ evolves according to the following rate equation [10]:

$$n(N+1,k+1) = n(N,k)\frac{N_s(k)}{N_s} + n(N,k+1)\left(1 - \frac{N_s(k+1)}{N_s}\right). \tag{3}$$

When N is sufficient large, $n(N,k)$ can be approximated as $Np(k)$, where $p(k)$ is the probability density function for the degree distribution. In terms of $p(k)$, the above equation can be rewritten as

$$p(k+1) = \frac{N}{N_s}[p(k)N_s(k) - p(k+1)N_s(k+1)]. \tag{4}$$

Using Eqs.(1) and (2) and the expression $p(k+1) - p(k) = \frac{dp}{dk}$, we can get the continuous form of Eq. (4):

$$\left(2d - 1 - \frac{d^2+d-1}{N}\right)p(k+1) + [k(d-1) - (d+1)(d-2)]\frac{dp}{dk} = 0. \tag{5}$$

For large k ($p(k+1) \approx p(k)$, $k \gg d$) and $N \gg d$, we have

$$(2d-1)p(k) + (d-1)k\frac{dp}{dk} = 0. \tag{6}$$

This leads to $p(k) \propto k^{-\gamma}$ with $\gamma = 2 + \frac{1}{d-1} \in (2,3]$, in accordance with many real-life networks having the power exponent between 2 and 3.

3 The clustering coefficient

The clustering coefficient of a node is the ratio of the total number of existing edges between all its neighbors and the number of all possible edges between them. The clustering coefficient C of the whole network is defined as the average of all nodes' clustering coefficients. At first, let us derive the analytical expression of $C(k)$ denoting the clustering coefficient of a node with degree k. When a node is newly added into the network, its degree and clustering coefficient are $d+1$ and 1. After that, if its degree increases by one, then its new neighbor must link to its d existing neighbors. Hence, we have

$$C(k) = \frac{\frac{d(d+1)}{2} + d(k-d-1)}{\frac{k(k-1)}{2}} = \frac{d(2k-d-1)}{k(k-1)}. \tag{7}$$

The clustering coefficient of the whole network can be obtained as the mean value of $C(k)$ with respect to the degree distribution $p(k)$:

$$C = \int_{k_{\mathtt{min}}}^{k_{\mathtt{max}}} C(k)p(k)dk, \tag{8}$$

where $k_{\mathtt{min}} = d+1$ is the minimal degree and $k_{\mathtt{max}} \gg k_{\mathtt{min}}$ is the maximal degree. Combining Eq. (7) and Eq. (8), and noting that $p(k) = Ak^{\frac{2d-1}{d-1}}$ and $\int_{k_{\mathtt{min}}}^{k_{\mathtt{max}}} Ap(k)dk = 1$, we can get the analytical result of C by approximately treating $k_{\mathtt{max}}$ as $+\infty$. For example, when $d = 2$, the clustering coefficient is

$$C = \frac{46}{3} - 36\mathtt{ln}\frac{3}{2} \approx 0.7366, \tag{9}$$

and, when $d = 3$, it is

$$C = 18 + 36\sqrt{2}\mathtt{arctan}\sqrt{2} + \frac{9}{2}\pi - 18\sqrt{2}\pi \approx 0.8021. \tag{10}$$

For larger d, the expression is too long, thus it will not be shown here. The integral values for $d = 4, 5, 6, 7, 8, 9, 10$ are 0.8406, 0.8660, 0.8842, 0.8978, 0.9085, 0.9171, and 0.9241, respectively.

It is remarkable that, the clustering coefficients of BA networks are very small and decrease with the increasing of network orders, following approximately $C \sim \frac{\mathtt{ln}^2 N}{N}$ [11]. Since the data-flow patterns show a large amount of clustering in interconnection networks, the STNs may perform better than BA networks. In addition, the demonstration exhibits that most real-life networks have large clustering coefficient no matter how many nodes they have. That agrees with the case of STNs but conflicts with that of BA networks. Furthermore, many real-life networks are characterized by the existence of hierarchical structures[12], which can usually be detected by the negative correlation between the clustering coefficient and the degree. The BA network, which does not possess a hierarchical structure, is known to have the clustering coefficient $C(x)$ of node x independent of its degree $k(x)$, while the STN has been shown to have $C(k) \sim k^{-1}$, in accordance with the observations of many real networks [12].

4 The average distance

Using symbol $d(i, j)$ to represent the distance between node i and j, the average distance of STN with order N, denoted by $L(N)$, is defined as $L(N) = \frac{2\sigma(N)}{N(N-1)}$, where the total distance is $\sigma(N) = \sum_{1 \le i < j \le N} d(i, j)$. Since a newly added node will not affect the distance between existing nodes, we have

$$\sigma(N+1) = \sigma(N) + \sum_{i=1}^{N} d(i, N+1). \tag{11}$$

Assume that node $N+1$ is added into the simplex $y_1y_2\cdots y_{d+1}$. Then Eq. (11) can be rewritten as

$$\sigma(N+1)=\sigma(N)+\sum_{i=1}^{N}(D(i,y)+1), \tag{12}$$

where $D(i,y)=\mathtt{min}\{d(i,y_j),j=1,2,\cdots,d+1\}$. Construct this simplex continuously into a single node y. Then, we have $D(i,y)=d(i,y)$. Since $d(y_j,y)=0$, Eq. (12) can be rewritten as

$$\sigma(N+1)=\sigma(N)+N+\sum_{i\in\Gamma}d(i,y), \tag{13}$$

where $\Gamma=\{1,2,\cdots,N\}-\{y_1,y_2,\cdots,y_{d+1}\}$ is a node set with cardinality $N-d-1$. The sum $\sum_{i\in\Gamma}d(i,y)$ can be considered as the total distance from one node y to all the other nodes in STN with order $N-d$. In a rough version, the sum $\sum_{i\in\Gamma}d(i,y)$ is approximated in terms of $L(N-d)$ as

$$\sum_{i\in\Gamma}d(i,y)\approx(N-d-1)L(N-d). \tag{14}$$

Apparently,

$$(N-d-1)L(N-d)=\frac{2\sigma(N-d)}{N-d}<\frac{2\sigma(N)}{N}. \tag{15}$$

Combining Eqs. (13), (14) and (15), one can obtain

$$\sigma(N+1)<\sigma(N)+N+\frac{2\sigma(N)}{N}. \tag{16}$$

Consider (15) as an equation. Then, we have

$$\frac{d\sigma(N)}{dN}=N+\frac{2\sigma(N)}{N}. \tag{17}$$

This equation leads to

$$\sigma(N)=N^2\mathtt{ln}N+H, \tag{18}$$

where H is a constant. As $\sigma(N)\sim N^2L(N)$, we have $L(N)\sim\mathtt{ln}N$, which should be paid attention to, since Eq. (16) is an inequality, indeed, the precise increasing tendency of L may be a little slower than $\mathtt{ln}N$.

5 Conclusion remarks

In conclusion, since simplex triangulation networks have very large clustering coefficients and very small average distances, they are not only scale-free networks but also small-world networks. Since many real-life networks are both scale-free and small-world, STNs may perform better in mimicking reality than BA and WS networks. In addition, STNs possess hierarchical structures in accordance with common observations of many real networks.

Furthermore, we have proposed an analytic approach to calculating the clustering coefficient, and an interesting technique to estimate the growing trend of average distances. Since in the earlier studies, only few analytic results about these two quantities of networks with randomicity are reported, we believe that our work may enlighten readers on this subject.

We also have done corresponding numerical simulations, where the simulation results agree with the analytical ones very well.

6 Acknowledgements

we acknowledge the support by the National Natural Science Foundation of China (NNSFC) under No. 70471084, 70271070, 10472116 and 70471033, the Specialized Research Fund for the Doctoral Program of Higher Education under No. 20020358009, and the Foundation for Graduate Students of the University of Science and Technology of China under No. KD200408.

7 References

[1] R. Albert and A. -L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.*, **74**, (2002) 47-97.

[2] M. E. J. Newman, The structure and function of complex networks, *SIAM Rev.*, **45**, (2003) 167-256.

[3] X. -F. Wang, Complex networks: topology, dynamics and synchronization, *Int. J. Bifurcation and Chaos*, **12**, (2002) 885-916.

[4] D. J. Watts and S. H. Strogatz, Collective dynamics of small-world networks, *Nature*, **393**, (1998) 440-442.

[5] A. -L. Barabási and R. Albert, Emergence of scaling in random networks, *Science*, **286**, (1999) 509-512.

[6] P. Holme and B. J. Kim, Growing scale-free networks with tunable clustering, *Phys. Rev. E*, **65**, (2002) 026107.

[7] J. S. Andrade Jr., H. J. Herrmann, R. F. S. Andrade and L. R. da Silva, Apollonian networks: Simultaneously scale-free, small world, Euclidean, space filling, and with matching graphs, *Phys. Rev. Lett.*, **94**, (2005) 018702.

[8] T. Zhou, G. Yan and B. -H. Wang, Maximal planar networks with large clustering coefficient and power-law degree distribution, *Phys. Rev. E*, **71**, (2005) 046141.

[9] J. R. Munkres, *Elements of Algebraic Topology*, Addison-Wesley, 1984.

[10] P. L. Krapivsky, S. Redner and F. Leyvraz, Connectivity of growing random networks, *Phys. Rev. Lett.*, **85**, (2000) 4629-4632.

[11] K. Klemm and V. M. Eguíluz, High clustered scale-free networks, *Phys. Rev. E*, **65**, (2002) 036123.

[12] E. Ravasz and A. -L. Barabási, Hierarchical organization in complex networks, *Phys. Rev. E*, **67**, (2003) 026112.