

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSI*net*

Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

A Swirl-shaped Deterministic Network with High Clustering Coefficient

¹Chong Li, ²Shi-Ze Guo, ³Zhe-Ming Lu and ¹Yu-Long Qiao

¹College of Information and Communication, Harbin Engineering University,
Harbin 150001, China

²North Electronic Systems Engineering Corporation, Beijing 100083, China

³School of Aeronautics and Astronautics, Zhejiang University,
Hangzhou 310027, China

Abstract: Complex network is a young and booming research area, which makes people know more about the characteristic of complex systems and get deeper acquaintanceship about nature. As randomness is a common characteristic of many real-life complex systems, people take a long time to research it and constructed many stochastic models. However, these models still cannot make people clear how the complex network is formed vividly step by step. Therefore, deterministic models have attracted considerable interests for their topological features can be analytically obtained. Swirl is a common natural phenomenon, which has attracted much attention in recent years. In this study, inspired by the swirl phenomenon, a deterministic model is first constructed to describe swirl-shaped networks and then both analytical solutions and experimental results are presented for the topological characteristics of the proposed model. The results demonstrate that the proposed model is a constant-degree centrosymmetric network with a high clustering coefficient.

Key words: Complex network, deterministic model, swirl-shaped network, centrosymmetric network, clustering coefficient

INTRODUCTION

The study of networks pervades all kinds of science, from neurobiology to statistical physics. The greatest challenge today has been the accurate and complete description of complex systems. The booming complex network theory is a useful tool that can help us to learn how the world compose and work. As a new branch of network science, it comes from the random graph theory created by two Hungary mathematicians, Erdos and Renyi (Albert and Barabasi, 2002; Watts, 2004).

With the development of complex network theory, many kinds of models are proposed and researched broadly (Asad, 2009). In general, a network is characterized by three widely-used topological features, i.e., degree distribution, clustering coefficient and average path length. Regular networks and random networks are two kinds of traditional models. The former possess large clustering coefficients, while the latter bear short average path lengths. In the past several decades, researchers have found that many real-life networks are in possession of non-trivial topological features. Some of these features do not occur in traditional regular lattices or random graphs. These networks are usually called complex

networks (Benzargua *et al.*, 2006; Li and Fang, 2009; Liu *et al.*, 2009; Tian *et al.*, 2011; Wu *et al.*, 2010). Two most famous and extensively studied kinds of complex networks are scale-free networks and small-world networks. The former can be characterized by power-law degree distributions, while the latter possess short path lengths and high clustering coefficients. Specially, small-world properties are ubiquitous in real-life networks, such as the World Wide Web and the electric power grid (Kwon *et al.*, 2006; Babainejad *et al.*, 2010; Edriss *et al.*, 2008). Therefore, small-world networks become a recent research focus of complexity theory.

From the view point of the randomness in model construction, network models can be categorized into probabilistic models and deterministic models. For a long time, researchers have focused on several famous stochastic models, such as the WS small-world model (Watts and Strogatz, 1998), the NW small-world model (Newman and Watts, 1999) and the BA scale-free network (Barabasi and Albert, 1999). Randomness is a common characteristic of many real-life networks, but it will be hard for people to understand how these networks are formed and how nodes in the network have interactions with each other (Barabasi *et al.*, 2001). In general, it is hard for

people to get analytical solutions to topological features from probabilistic network models. Therefore, many researchers turn their steps to deterministic network models in order to obtain analytical results. In recent years, the research on deterministic models has made significant progress. The first deterministic model that is a variant of uniform recursive tree (Smythe and Mahmoud, 1995) known as deterministic uniform recursive tree. Based on graph theories, the first deterministic small-world model was proposed by Comellas *et al.* (2000). This pioneer model attracts a lot of interest from researchers who become concern about deterministic small-world networks. For example, a special type of deterministic small-world networks named “Hanoi Networks” based on the famous tower of Hanoi puzzle was proposed by Boettcher *et al.* (2008). A deterministic small-world network model based on the Clayey graph was established by Xiao and Parhami (2006). Through studying the prime factor decomposition, Corso (2001) constructed a deterministic network whose each node stands for a natural number. In addition, Chandra and Dasgupta (2005) constructed a prime-number-based small-world network using a similar idea to Corson’s work.

Swirl is an important research branch of hydromechanics. To make up a swirl, three factors are required, i.e., viscosity, baroclinicity and non-conservative forces. Swirl is a complex movement form and its cause of formation is not clear now. In general, there are four causes that can be used to explain the source of swirl, i.e., naturally formed from the nonpolar energy, driven by existing swirls, separated or released from existing swirls and formed by coupling a number of exiting swirls. Inspired by the swirl phenomenon, the aim of this study is to construct a deterministic swirl-shaped model with several non-trivial topological features.

PROPOSED SWIRL-SHAPED DETERMINISTIC MODEL

The proposed idea comes from the swirl shape. Four directions are considered during the swirl-shaped model generation. When the swirl revolves round the centroid by one cycle, four new nodes are linked to the network with a specific manner. In other words, the proposed model grows by four nodes every time. Assume the obtained network after *t* iterations is *U_t* with *N_t* nodes and *E_t* edges, where *t* = 0,1,2,...,T-1 and T is the number of iterations performed, then the proposed network generation process can be illustrated as follows:

Step 0: Initialization. Set *t*=0, *U₀* is a complete network with four nodes. Obviously, *N₀* = 4 and *E₀* = 6

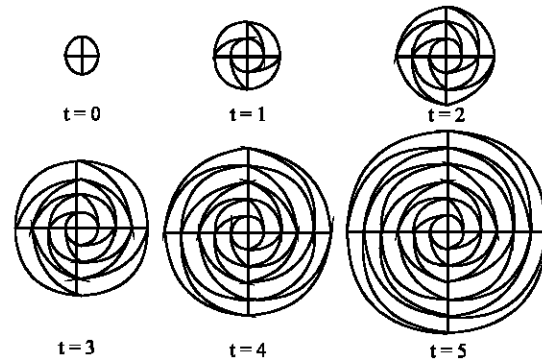


Fig. 1: The first six iterations of the growth process of the proposed network

Step 1: Generation of *U_{t+1}* from *U_t*. This step includes following two substeps

Step 1.1: Four new nodes are added to the network, each linked to its nearest old node and next nearest old node on the right side

Step 1.2: Each new node is linked to two nearest new nodes

Obviously, after above two substeps, *N_{t+1}* = *N_t*+4 and *E_{t+1}* = *E_t*+12 can be obtained

Step 2: If *t*<T-1, set *t* = *t*+1 and go to Step 1. Otherwise, the algorithm is terminated

Above iterative process is repeated for T-1 times and then a deterministic network with a high clustering coefficient can be obtained as shown below. Figure 1 shows the obtained network after the first six iterations. According to the relationships *N_{t+1}* = *N_t*+4 and *E_{t+1}* = *E_t*+12 together with the initial conditions *N₀* = 4 and *E₀* = 6, it can be easily proved that *N_t* = 4 *t*+4 and *E_t* = 12*t*+6, thus the average node degree can be obtained as follows:

$$\langle k \rangle_t = \frac{2E_t}{N_t} = \frac{24t+12}{4t+4} = 6 - \frac{3}{t+1} \tag{1}$$

Since, $\lim_{t \rightarrow \infty} \langle k \rangle_t = 6$, a conclusion can be drawn that the proposed network is a sparse graph whose nodes have much fewer links than possible.

RELEVANT TOPOLOGICAL CHARACTERISTICS

From the network generation process, it is obvious that the proposed model is symmetric such that there are strong similarities among the nodes. If the model is divided into four parts, then any of three parts can be obtained by rotating the fourth part. The nodes in the same iteration layer have the same properties. Therefore,

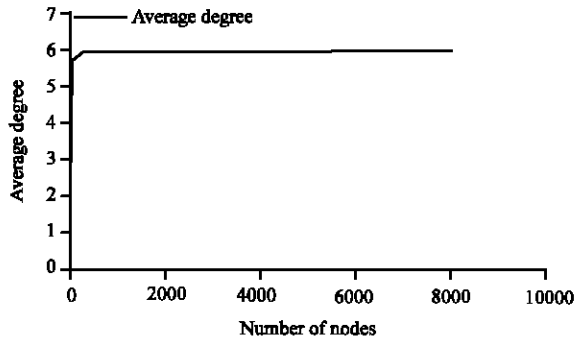


Fig. 2: The evolvement of average clustering coefficient with the number of nodes

in this section, the topological characteristics can be calculated based on the layer-by-layer analysis method.

Degree distribution: Degree distribution is one of the most important topological features to characterize a network. The degree of a node is defined as the number of edges directly connected with it. Degree distribution $P(k)$ is defined as the probability that a randomly selected node has exactly k edges directly connected with it. In the proposed model, for the first layer, all four nodes are with the same degree value 5. From the second layer to the t -th layer, all nodes are with the same degree value 6. For the last layer, i.e., the $t+1$ -th layer, all four nodes are with the same degree value 4. On the whole, for $t > 2$, the proposed model has a discrete degree distribution on three values $\{4, 5, 6\}$ as follows:

$$P(k) = \frac{\delta(k-4)}{t+4} + \frac{(t-1)\delta(k-5)}{t+4} + \frac{\delta(k-6)}{t+4} \quad (2)$$

When $t \rightarrow \infty$, the following degree distribution can be easily obtained

$$\lim_{t \rightarrow \infty} P(k) = \delta(k-5) \quad (3)$$

Therefore, the degree distribution of the proposed network focuses mainly on the degree value 5.

Clustering coefficient: The clustering coefficient is an important measure of degree to which nodes in a network incline to clustering together. The clustering coefficient can be obtained by averaging the local clustering coefficients over all the nodes. The local clustering coefficient C_i for Node i with degree k_i is defined as the number of links L_i that actually exist between its nearest neighbors divided by the number of links that could possibly exist between them, i.e., $C_i = 2L_i/[k_i(k_i-1)]$. In the proposed model, the nodes can be grouped into three classes according to their degree values. For the first layer, when $t > 1$, the clustering coefficient for all nodes is

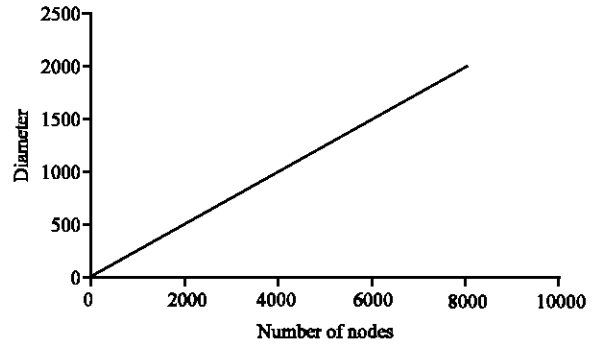


Fig. 3: The evolvement of diameter with the number of nodes

$3/5$. From the second layer to the t -th layer, the clustering coefficient of all nodes is $2/5$. For the last layer, the clustering coefficient of all four nodes is $1/2$. Based on above results, the average clustering coefficient can be easily computed as follows:

$$C(t) = \frac{\frac{3}{5} \times 4 + \frac{2}{5} \times 4 \times (t-1) + \frac{1}{2} \times 4}{4t+4} = \frac{\frac{8}{5} \times (t-1) + \frac{11}{10} \times 4}{4t+4} \quad (4)$$

When t approaches infinity, the average clustering coefficient approaches a constant value $2/5 = 0.40$ that is much bigger than the average clustering coefficient of a random network with the same size. Figure 2 shows how the average clustering coefficient changes as the number of nodes increases.

Diameter and average path length: In a network, the shortest distance between two nodes is defined as the minimum number of edges traveled from one node to another. The Average Path Length (APL) is defined as the average distance over all possible pairs of nodes in a connected network. In general, it is hard to obtain the analytic solution to APL. Diameter is defined as the largest or shortest distance between any two nodes in the network, which reflects the delay of the interaction between the network nodes. For the proposed network, when $t < 3$, it is easy to obtain that the diameter is $t+1$. When $t > 2$, via the swirl lines, each node in the last layer has the same distance from each node in the first layer. Therefore, the diameter of the network is the distance between a node in the last layer and a node in the first layer, which is just equal to the number of iterations t . Thus, the following equation can be easily obtained:

$$D(t) = \begin{cases} t+1 & t \leq 2 \\ t & t > 2 \end{cases} \quad (5)$$

From above equation, it is easy to conclude that the diameter increases linearly with the number of iterations

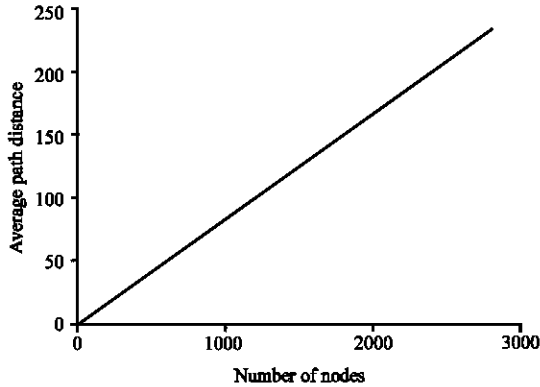


Fig. 4: The evolvement of average distance with the number of nodes

as well as the number of nodes. Figure 3 shows how the diameter changes as the number of nodes increases.

The next step is to obtain the analytic solution to the average path length, as well as its asymptote. According to the layer-by-layer analysis, one can see that the four nodes in the same layer share the same importance. Based on this consideration, the sum of the distances between all possible node pairs can be easily calculated as follows:

$$\sum_{i>j} d_{ij} = 38 + 68(t-1) + 16[3(t-2) + 4(t-3) + 5(t-4) + \dots + t(t-(t-1))] = \frac{8t^3}{3} + 8t^2 + \frac{76t}{3} + 2 \tag{6}$$

Dividing Eq. 6 by the number of possible node pairs, the average path length can be obtained as follows:

$$APL = \frac{\sum_{i>j} d_{ij}}{\frac{1}{2}(4t+4)(4t+3)} = \frac{1}{3}(t+1) + \frac{1}{3} \cdot \frac{1 + \frac{28}{t}}{4 + \frac{7}{t} + \frac{3}{t^2}} \tag{7}$$

Obviously, the APL approximates to $(t+1)/3$, which linearly increases with the number of iterations as well as the number of nodes. Figure 4 shows how the APL changes as the number of nodes increases.

DISCUSSION

From above two sections, one can see that the proposed network model is generated by a very simple deterministic mechanism but it is with some interesting topological features. Firstly, the proposed network is a sparse network, which complies with most real-life networks. Secondly, the proposed network under a large size nearly has a constant degree, which complies with

some regular networks. Thirdly, the proposed network is with a large average clustering coefficient, which complies with small-world networks and regular networks. Generally, it is difficult for people to calculate the average path length analytically, even for a regular lattice. However, because the proposed model has a symmetrical structure with a special layer-by-layer shape, its average path length can be fortunately calculated with an analytical expression. Based on Eq. 7, it can be concluded that the average path length increases linearly with the number of nodes, which complies with a regular network. At first sight, this result makes the authors disappointed for a small-world network is originally expected. In fact, the result is also inspiring for it provides an analytical solution to the average path length. Future work may concentrate on how to obtain a small-world network by improving the iterative mechanism.

CONCLUSION

In this study, a deterministic network model has been proposed inspired by the shape of swirl. The proposed model is a growing network, where four new nodes are added to the network at each iteration. Topological features of the proposed network have been discussed both analytically and experimentally. The results have shown that the proposed network is a sparse network with constant degree distribution and high clustering coefficient, however the average path length increases linearly with the number of nodes. Therefore, the proposed network is actually a special kind of regular networks with an analytical solution to the average path length.

ACKNOWLEDGMENT

This study was supported by the National Natural Science Foundation of China under grant 61070208. This research project was conducted in China from January 2011 to December 2011.

REFERENCES

Albert, R. and A.L. Barabasi, 2002. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74: 47-97.
 Asad, J.H., 2009. Analysis of capacitance networks. *Asian J. Applied Sci.*, 2: 296-299.
 Babainejad, S., S. Babainejad, N. Bigdeli and K. Afshar, 2010. Setting up a virtual laboratory for evaluation of congestion control algorithms in TCP/IP networks. *Trends Applied Sci. Res.*, 5: 177-187.
 Barabasi, A.L. and R. Albert, 1999. Emergence of scaling in random networks. *J. Sci.*, 286: 509-512.

- Barabasi, A.L., E. Ravasz and T. Vicsek, 2001. Deterministic scale-free networks. *Phys. A: Stat. Mech. Appl.*, 299: 559-564.
- Benzargua, F., A. Chaker and N. Khalfallah, 2006. Optimization of reactive power in large electrical networks: Algerian network. *Inform. Technol. J.*, 5: 703-708.
- Boettcher, S., B. Goncalves and H. Guclu, 2008. Hierarchical regular small-world networks. *J. Phys. A: Math. Theor.*, 25: 252001-252007.
- Chandra, A.K. and S. Dasgupta, 2005. A small world network of prime numbers. *Phys. A: Stat. Mech. Appl.*, 357: 436-446.
- Comellas, F., J. Ozon and J.G. Peters, 2000. Deterministic small-world communication networks. *Inf. Process. Lett.*, 76: 83-90.
- Corso, G., 2004. Families and clustering in a natural numbers network. *Phys. Rev. E*, 69: 36106-36110.
- Edriss, M.A., P. Hosseinnia, M. Edrisi, H.R. Rahmani and M.A. Nilforooshan, 2008. Prediction of second parity milk performance of dairy cows from first parity information using artificial neural network and multiple linear regression methods. *Asian J. Anim. Vet. Adv.*, 3: 222-229.
- Kwon, I.W.G., D.R. Rickert and M.C.H. Chao, 2006. The use of the world wide web and internet in pharmacy practice: An exploratory study. *J. Pharmacol. Toxicol.*, 1: 1-11.
- Li, X.S. and H.J. Fang, 2009. Stability of continuous-time vehicles formations with time delays in undirected communication network. *Inform. Technol. J.*, 8: 165-172.
- Liu, P., H. Miao and J. Mei, 2009. Towards common acquaintance immunization strategy for complex network. *Inform. Technol. J.*, 8: 1129-1139.
- Newman, M.E.J. and D.J. Watts, 1999. Renormalization group analysis of the small-world network model. *Phys. Lett. A*, 263: 341-346.
- Smythe, R.T. and H. Mahmoud, 1995. A survey of recursive trees. *Theor. Probab. Math. Statist.*, 51: 1-27.
- Tian, L., C. Liu and G. Sun, 2011. A novel method to construct undirected network. *Inform. Technol. J.*, 10: 904-907.
- Watts, D.J. and S.H. Strogatz, 1998. Collective dynamics of small-world networks. *Nature*, 393: 440-442.
- Watts, D.J., 2004. The new science of networks. *Annu. Rev. Sociol.*, 30: 243-270.
- Wu, X., M. Zhou, L. Meng, J. Hua, K. Zhou and T. Wu, 2010. Connection stability analyze based on LEACH algorithm for mobile Ad Hoc network. *Inf. Technol. J.*, 9: 1212-1216.
- Xiao, W.J. and B. Parhami, 2006. Cayley graphs as models of deterministic small-world networks. *Inf. Process. Lett.*, 97: 115-117.