LETTER

A Tree-Structured Deterministic Small-World Network

Shi-Ze GUO[†], **Zhe-Ming LU**[††a)], **Guang-Yu KANG**[†], **Zhe CHEN**[†], ***Nonmembers*****,** ***and*** **Hao LUO**[††], ***Member***

SUMMARY Small-world is a common property existing in many real-life social, technological and biological networks. Small-world networks distinguish themselves from others by their high clustering coefficient and short average path length. In the past dozen years, many probabilistic small-world networks and some deterministic small-world networks have been proposed utilizing various mechanisms. In this Letter, we propose a new deterministic small-world network model by first constructing a binary-tree structure and then adding links between each pair of brother nodes and links between each grandfather node and its four grandson nodes. Furthermore, we give the analytic solutions to several topological characteristics, which shows that the proposed model is a small-world network.

key words: *deterministic small-world models, tree-structured network, Interconnection network, average path length, clustering coefficient*

1. Introduction

Small-world properties have been found in a number of real-life social, technological and biological networks. People define small-world networks as those having clustering coefficients much larger than random networks, and at the same time their average shortest path length grows proportionally to the logarithm of the number of nodes. In the past dozen years, many probabilistic models have been proposed to characterize real-life small-world networks. In 1998, Watts and Strogatz [1] proposed the pioneering small-world network named WS model, which promoted a great deal of research on the characteristics of small-world networks. One year later, Newman and Watts [2], [3] presented another intensively studied small-world network named NW model. These two classical small-world models are both derived from the same regular ring lattice. However, with a certain probability, the WS model rewires each link while the NW model adds links between each pair of unlinked nodes. Kasturirangan [4] presented a small-world model called R + T network, which is actually a regular network coupled with a tree structure. Subsequently, Kleinberg [5] proposed a generalized WS model based on a two-dimensional lattice. Furthermore, Ozik et al. [6] introduced a simple evolution model to construct small-world networks based on geographical attachment preference. The common feature of the above models lies in that they construct networks in a probabilistic manner, and thus they cannot provide us with a vivid and concrete explanation of how networks are formed link by link.

To calculate topological features analytically, many researchers have turned their steps to constructing scale-free or small-world networks in some deterministic manners. Based on graph-theoretic methods, Comellas et al. [7] presented the first deterministic small-world network. Based on the famous tower of the Hanoi puzzle, Boettcher et al. [8] proposed a kind of deterministic small-world networks named "Hanoi Networks". However, in above two models, the number of nodes is fixed during the generation process. To obtain growing small-world networks, Zhang et al. presented a deterministic small-world network generated by edge iterations [9]. Besides above models, many other types of deterministic small-world models have been proposed, such as those with specific degree distributions [10], those based on prime numbers [11], [12] and those based on Cayley graphs [13]. In this Letter, we aim to derive a deterministic small-world network from the binary tree by adding some edges in each iteration with a simple mechanism, resulting in a high clustering coefficient.

2. Proposed Deterministic Small-World Network

In most real-life networks, the number of nodes often grows exponentially with time. Therefore, we adopt a binary tree whose number of nodes increases exponentially with layer. However, the clustering coefficient of the binary tree is zero since there is not any triangles in the tree. To get a high clustering coefficient, here we propose a generation algorithm by adding some links based on a simple mechanism at each iteration. Assume the obtained network after t iterations is SW_t that has N_t nodes and E_t edges, where $t = 0, 1, 2, \ldots, T-1$, and T is the number of iterations performed. Assume each node is labeled with a nature number increasing with the generation time, then the proposed generation process can be illustrated as follows:

Step 0. Initialization: Set $t = 0$, SW_0 contains one node labeled as "1". Obviously, $N_0 = 1$ and $E_0 = 0$, and the number of layers is 1.

Step 1. Generation of SW_1 from SW_0: Two child nodes labeled as "2" and "3" branch from Node "1", and an extra link is constructed between the two child nodes. Thus, $N_1 = 3$ and $E_1 = 3$, and the number of layers is 2.

Step 2. Generation of SW_{t+1} from SW_t for $t > 0$: This step includes following three substeps.

Manuscript received November 9, 2011.
Manuscript revised January 9, 2012.
[†]The authors are with North Electronic Systems Engineering Corporation, Beijing 100083, China.
[††]The authors are with School of Aeronautics and Astronautics, Zhejiang University, Hangzhou 310027, China.
a) E-mail: zheminglu@zju.edu.cn (Corresponding Author)
DOI: 10.1587/transinf.E95.D.1536

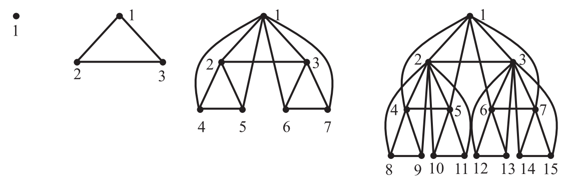

Fig. 1 The first four iterations of the growth of the proposed network.

Step 2.1. Two new nodes branch from each node in the last layer of SW_t, resulting in a new layer. In other words, Node "$(N_t+1)/2+p$" is linked to the newly-generated Node "$N_t + 1 + 2p$" and Node "$N_t + 2p + 1$" respectively, $p = 0, 1, \ldots, (N_t + 1)/2 - 1$.

Step 2.2. Construct a link between each pair of full-brother nodes in the newly-generated layer. In other words, Node "$N_t + 1 + 2p$" and Node "$N_t + 2p + 1$" are linked, $p = 0, 1, \ldots, (N_t + 1)/2 - 1$.

Step 2.3. For each newly-generated node, we link it to its grandfather node. In other words, each node in the t-th layer labeled as "$(N_t+1)/4+p$" is linked to its four grandson nodes in the $t + 2$-th layer labeled as "$N_t + 4p + 1$", "$N_t + 4p + 2$", "$N_t + 4p + 3$" and "$N_t + 4p + 4$" respectively, $p = 0, 1, \ldots, (N_t + 1)/4 - 1$.

Obviously, after above three substeps, we have $N_{t+1} = 2N_t + 1$ and $E_{t+1} = 3N_t + 4N_{t-1}$.

Step 3: If $t < T - 1$, set $t = t + 1$ and go to Step 2. Otherwise, the algorithm is terminated.

Above iterative process is repeated for $T - 1$ times, and then we can obtain a deterministic network with a high clustering coefficient as shown below. In fact, in the above generation process, we mimic the family relationship to some extent, for example, the pair of full-brother nodes branching from the same father is linked and the grandfather is linked to its four grandsons. Figure 1 shows the obtained network after the first four iterations. According to the relationships $N_{t+1} = 2N_t + 1$ and $E_{t+1} = 3N_t + 4N_{t-1}$ together with the initial conditions $N_0 = 1$, we can easily prove that $N_t = 2^{t+1} - 1$ and $E_t = 5 \times 2^t - 7$ ($t > 0$), thus we can obtain the average node degree as follows

$$\langle k \rangle_t = \frac{2E_t}{N_t} = \frac{2(5 \cdot 2^t - 7)}{2^{t+1} - 1} = 5\left(1 - \frac{2.8}{2^{t+1} - 1}\right) \tag{1}$$

Since $\lim_{t\to\infty} \langle k \rangle_t = 5$, we can see that the proposed model is a sparse network with much fewer links than possible. In comparison with our model, the average node degree of the deterministic model presented in [9] approaches 4.

3. Topological Properties

3.1 Degree Distribution

Degree distribution is one of the most important topological features of a network. The degree of Node i is defined as the number of connections it has to other nodes, and degree distribution $P(k)$ is defined as the fraction of nodes in the network with degree k. According to the iteration algorithm, for $t > 3$, we can easily prove that the possible degree values in the proposed network are 6, 8, 9, 5, 3, which correspond to the nodes in the first layer, the second layer, the middle layers, the second last layer and the last layer respectively. Thus, we can easily obtain

$$\begin{aligned} P(k) = {} & \frac{1}{2^{t+1} - 1}\delta(k - 6) + \frac{2}{2^{t+1} - 1}\delta(k - 8) \\ & + \frac{2^{t-1} - 4}{2^{t+1} - 1}\delta(k - 9) + \frac{2^{t-1}}{2^{t+1} - 1}\delta(k - 5) \\ & + \frac{2^t}{2^{t+1} - 1}\delta(k - 3) \end{aligned} \tag{2}$$

Where $\delta(k) = 1$ for $k = 0$ and $\delta(k) = 0$ for $k \neq 0$. When $t \to \infty$, we can easily obtain the following degree distribution

$$\begin{aligned} \lim_{t\to\infty} P(k) = {} & 0.25 \cdot \delta(k - 9) + 0.25 \cdot \delta(k - 5) \\ & + 0.5 \cdot \delta(k - 3) \end{aligned} \tag{3}$$

Therefore, the degree distribution of the proposed network is discrete and focuses mainly on three degree values, which is different from most small-world networks [9] that $P(k)$ is an exponential of a power of degree k.

3.2 Clustering Coefficient and Clustering-Degree Correlation

The clustering coefficient of a network is the average local clustering coefficient over all nodes in the network. The local clustering coefficient of a node in a network quantifies how close its neighbors are to be a clique. For Node i with degree k_i, its local clustering coefficient C_i is defined as $C_i = 2n_i/[k_i(k_i - 1)]$, where n_i is the number of links that actually exist between its nearest neighbors. Because of the symmetry of the proposed network, the nodes with the same degree have the same local clustering coefficient. At Iteration t, there are $t + 1$ layers. When $t > 3$, for the nodes in the first layer, we have $n_i = k_i + 1 = 7$. For the nodes in the second layer, we have $n_i = k_i + 2 = 10$. For the nodes in the t-th layer, we have $n_i = k_i + 1 = 6$. And for the nodes in the $t+1$-th layer, we have $n_i = k_i = 3$. For the remainder middle layers, we have $n_i = k_i + 3 = 12$. Thus, we can easily obtain the following results.

$$\begin{aligned} C(k) = {} & \frac{7}{15} \cdot \delta(k - 6) + \frac{5}{14} \cdot \delta(k - 8) + \frac{1}{3} \cdot \delta(k - 9) \\ & + \frac{3}{5} \cdot \delta(k - 5) + 1 \cdot \delta(k - 3) \end{aligned} \tag{4}$$

Furthermore, since $n_i \approx k_i$, we have $C_i = 2n_i/[k_i(k_i - 1)] \approx 2/k_i$, and thus the clustering-degree correlation can be simply described as $C(k) \propto k^{-1}$. That is to say, there is a negative correlation between the local clustering coefficient and the nodal degree, which means that the larger the degree of the node the lower its clustering coefficient.

Based on Eq. (4) and Eq. (2), for $t > 3$, the average clustering coefficient C for the whole network can be calculated as follows:

$$
\begin{aligned}
C(t) &= \frac{1}{2^{t+1}-1}\sum_{i=1}^{2^{t+1}-1} C_i(t) \\
&= \frac{1}{2^{t+1}-1}\Bigg[1\cdot\frac{7}{15}+2\cdot\frac{5}{14}+(2^{t-1}-4)\cdot\frac{1}{3} \\
&\quad +2^{t-1}\cdot\frac{3}{5}+2^{t}\cdot 1\Bigg] \\
&= \frac{11}{15}+\frac{61}{105\cdot(2^{t+1}-1)}
\end{aligned} \tag{5}
$$

Since $C(t+1) - C(t) < 0$, as the number of iterations t approaches infinity, C will monotonically decrease to the constant value $11/15 = 0.7333$, therefore the clustering coefficient of the proposed network is very high. In comparison with our model, the clustering coefficient of the deterministic model in [9] approaches 0.6931.

3.3 Diameter and Average Path Length

The most typical feature of a small-world network is that its average path length (APL) is short in comparison with its size. 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 of APL. In fact, if a network is with a small maximal distance, then this network is undoubtedly with a short APL. We call this maximal distance diameter, which characterizes the maximum communication delay in the network. Here, we denote the diameter at Iteration t as $D(t)$. Because of the tree structure, one can easily see that the diameter always lies between the non-brother nodes in the last layer. Take the node pair $(2^t, 2^{t+1} - 1)$ for example, while traveling from Node 2^t to Node $2^{t+1} - 1$, the shortest path includes t edges. Thus, we have the following simple formula

$$
D(t) = t = \frac{\ln(N_t + 1)}{\ln 2} - 1 \tag{6}
$$

Thus, the diameter D grows logarithmically with the number of nodes. In comparison with our model, the diameter of the model in [9] is $t + 1$. Because the average path length is smaller than D, thus the APL should increase more slowly. To show the relationship more clearly, we provide the simulation results in Fig. 2.

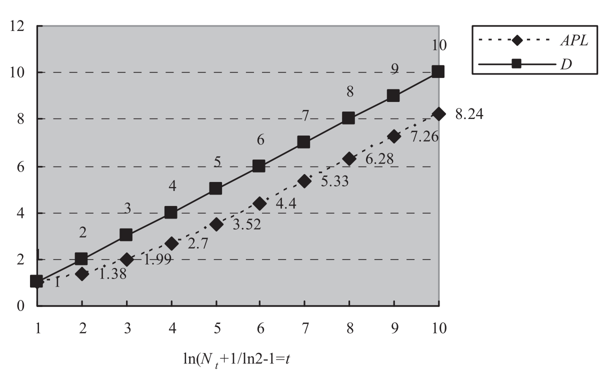

Fig. 2 The APL and D versus the logarithm of the number of nodes.

According to the above discussions, we can conclude that the proposed model is a deterministic small-world network, for it is sparse with a high clustering coefficient and a short average path length, satisfying the three main required properties for small-world networks.

4. Conclusions

In this Letter, we have presented a deterministic small-world model derived from the binary tree. By adding links between full-brother nodes and adding links between the grandfather node and its grandson nodes, we get a high clustering coefficient not less than 0.7333, which results in a small-world network. We have derived the analytic solutions for degree distribution, clustering coefficient, clustering-degree correlation and diameter of the deterministic model, and they are all close to those for existing random small-world networks. The proposed model provides a new way to generate a network with specific properties by modifying an existing network.

Acknowledgement

This work was supported by the National Natural Scientific Foundation of China under Grant No.61003255, No.61070208 and No.61071128.

References

[1] D.J. Watts and S.H. Strogatz, “Collective dynamics of ‘small-world’ networks,” Nature, vol.393, no.6684, pp.440–442, June 1998.
[2] M.E.J. Newman and D.J. Watts, “Renormalization group analysis of the small-world network model,” Phys. Lett. A, vol.263, no.4-6, pp.341–346, March 1999.
[3] M.E.J. Newman and D.J. Watts, “Scaling and percolation in the small-world network model,” Phys. Rev. E, vol.60, no.6, pp.7332–7342, May 1999.
[4] R. Kasturirangan, “Multiple scales in small-world networks,” cond-mat/9904055, 1999.
[5] J. Kleinberg, “Navigation in a small world,” Nature, vol.406, no.6798, p.845, Aug. 2000.
[6] J. Ozik, B.-R. Hunt, and E. Ott, “Growing networks with geographical attachment preference: emergence of small worlds,” Phys. Rev. E, vol.69, no.2, 026108, Feb. 2004.
[7] F. Comellas, J. Ozon, and J.G. Peters, “Deterministic small-world communication networks,” Inf. Process. Lett., vol.76, no.1-2, pp.83–90, Nov. 2000.
[8] S. Boettcher, B. Gongalves, and H. Guclu, “Hierarchical regular small-world networks,” J. Physics A: Mathematical and Theoretical, vol.41, no.25, 252001, June 2008.
[9] Z.Z. Zhang, L.L. Rong, and C.H. Guo, “A deterministic small-world network created by edge iterations,” Physica A, vol.363, no.2, pp.567–572, May 2006.
[10] F. Comellas and M. Sampels, “Deterministic small-world networks,” Physica A, vol.309, no.1-2, pp.231–235, June 2002.
[11] G. Corso, “Families and clustering in a natural numbers network,” Phys. Rev. E, vol.69, no.3, 036106, March 2004.
[12] A.K. Chandra and S. Dasgupta, “A small world network of prime numbers,” Physica A, vol.357, no.3-4, Nov. 2005.
[13] W.J. Xiao and B. Parhami, “Cayley graphs as models of deterministic small-world networks,” Inf. Process. Lett., vol.97, no.3, pp.115–117, March 2006.