The structural properties of the generalized Koch network

Jingyuan Zhang and Weigang Sun
School of Science, Hangzhou Dianzi University, Hangzhou 310018, People’s Republic of China
E-mail: zjy.anhui@sina.com and wgsun999@yahoo.com.cn

Received 11 May 2010
Accepted 16 June 2010
Published 9 July 2010

Online at stacks.iop.org/JSTAT/2010/P07011
doi:10.1088/1742-5468/2010/07/P07011

Abstract. Real-world systems in nature can be described by complex networks. It is of current interest for researchers to construct network models depicting complex systems. In this paper, the generalized Koch network is proposed, which exhibits some properties of Koch networks. Analytical expressions involving the degree distribution, clustering coefficient and average path length for the generalized Koch network are given. Our models can provide valuable insights into evolution systems.

Keywords: fractal growth (theory), exact results, random graphs, networks
1. Introduction

Complex networks have been viewed as a useful tool for describing real-world systems in nature and society [1]–[3]. Prompted by the computerization of data acquisition and the increasing computational power of computers, researchers have done many empirical studies on diverse real networked systems and discovered some basic facts: a power-law degree distribution \( P(k) \sim k^{-\gamma} \) [4], a power-law relation between the number of boxes needed to cover the network and the size of the box [5, 6], the small-world effect with large clustering coefficient and small average path length (APL) [7], some variables obeying simple scaling laws under renormalization [8, 9], and degree correlations [10].

To reproduce or explain the striking common features of real-life systems, some network models have been constructed. In addition to the seminal Watts–Strogatz’s small-world network model [7] and Barabási–Albert’s scale-free network model [4], a considerable number of models have been developed to mimic natural systems, including initial attractiveness [11], aging and cost [12], accelerating growth [13], coevolution [14], weight or traffic driven evolution [15, 16], fitness models [17], fractal models based on the branching tree [18, 19], and duplication [20], to name but a few. Although significant progress has been made in the field of network modeling, it is still a fundamental task and of current interest to put forward models and reproduce their generic properties from different angles. Apart from the various proposed models, the dynamical properties, e.g., synchronization of complex networks, have received much attention in the field of complex networks (see [21, 22] and the many references cited therein).

Among the studies on various network models, the deterministic complex network models have attracted increasing attention. The first successful attempt was the pseudofractal model [23], which started the interest and research on deterministic models (see [6, 18, 19, 24, 25] and many references cited therein). On the basis of the well-known Koch fractal, Zhang et al firstly proposed a deterministic network, named a Koch network, incorporating some key properties, such as a power-law distribution, a high clustering coefficient, small average path length and degree correlations [24, 25]. Let \( K(t) \) be a Koch
network of \( t \) generation. A algorithm to create a Koch network is as follows (see [24]): initially \( K(0) \) consists of three nodes forming a triangle, then each of the three nodes of the initial triangle gives birth to two nodes and these two new nodes and its mother node are linked to each other shaping a new triangle, thus \( K(1) \) is obtained. For \( t \geq 1 \), \( K(t) \) is obtained from \( K(t-1) \) when each of existing triangle of \( K(t-1) \) is replaced with \( K(1) \). When \( t \) is large enough, the cumulative degree distribution \( P_{\text{cum}}(k) \) of \( K(t) \) is \( 4 \times k^{-2} \), the clustering coefficient \( C(k) \) for a node of degree \( k \) of \( K(t) \) is \( 1/(k-1) \), the average path distance \( d_t \) for Koch network \( K(t) \) is

\[
d_t = \frac{4 + 14 \times 4^t + 12t \times 4^t}{3(4^t + 1)}.
\]

The quantity \( k_{nn}(k) \), related to degree correlations, is the average degree of the nearest neighbors for nodes with degree \( k \) of \( K(t) \), \( k_{nn}(k) \) is \( t + 2 \). In this paper, inspired by the Koch network, we propose a family of deterministic networks, called generalized Koch networks. Generalized Koch networks not only exhibit some properties of the Koch network, but also provide valuable insights into evolution systems.

### 2. Generation algorithm of a generalized Koch network

Let \( GK(t) \) be generalized Koch network of the \( t \)th generation. An iterative algorithm to create a generalized Koch network is as follows. Initially, \( GK(0) \) consists of \( m \) nodes forming a \( m \)-polygon, where \( m \geq 3 \). Then, each of the \( m \) vertices of the initial \( m \)-polygon gives birth to \((m-1) \) vertices. These \((m-1) \) new vertices and their mother vertex are so linked as to form new \( m \)-polygon. Thus we get \( GK(1) \). For \( t \geq 1 \), \( GK(t) \) is obtained from \( GK(t-1) \) as follows, we replace each of the existing \( m \)-polygons of \( GK(t-1) \) with \( GK(1) \) to obtain \( GK(t) \). The growing process is repeated until the network reaches a desired order. In figure 1, \( GK(t) \) at \( t = 1 \) with \( m = 4 \) is shown. We compute the order and the size (number of all edges) of \( GK(t) \). To this end, we first calculate the total number of \( m \)-polygons existing at step \( t \), which we denote as \( L_v(t) \). By construction, this quantity increases by a factor of \( m+1 \), i.e., \( L_v(t) = (m+1)L_v(t-1) \). Considering the initial condition \( L_v(0) = 1 \), it follows that \( L_v(t) = (m+1)^t \). Let \( L_v(t) \) and \( L_e(t) \) be the respective number of nodes and edges created at step \( t \). Noting that each \( m \)-polygon in \( GK(t-1) \) will lead to the addition of \( m \times (m-1) \) new nodes and \( m \times m \) new edges at step \( t \), then one can easily obtain the following relations: \( L_v(t) = m(m-1)L_v(t-1) = m(m-1)(m+1)^{t-1} \) and \( L_e(t) = m^2L_e(t-1) = m^2(m+1)^{t-1} \) for arbitrary \( t > 0 \). According to these results, we then compute the order and the size of the generalized Koch network. The total number of vertices \( N_t \) and edges \( E_t \) present at step \( t \) is

\[
N_t = \sum_{k=0}^{t} L_v(k) = (m-1)(m+1)^t + 1
\]

and

\[
E_t = \sum_{k=0}^{t} L_e(k) = m(m+1)^t,
\]
Figure 1. $GK(t)$ at $t = 1$ with $m = 4$.

respectively. Thus, the average degree is

$$
\langle k \rangle = \frac{2E_t}{N_t} = \frac{2m(m+1)^t}{(m-1)(m+1)^t+1},
$$

which is approximately $2m/(m-1)$ for large $t$.

3. Structural properties of the generalized Koch network

Now we study some relevant characteristics of the generalized Koch network $GK(t)$, mainly focusing on degree distribution, clustering coefficient and average path length.

3.1. Degree distribution

We define $k_i(t)$ as the degree of a node $i$ at time $t$. When node $i$ is added to the network at step $t_i$ ($t_i \geq 0$), it has a degree of two, namely, $k_i(t_i) = 2$. To determine $k_i(t)$, we firstly calculate the number of $m$-polygons involving node $i$ at step $t$, which is represented by $L_o(i, t)$. These $m$-polygons will create nodes connected to node $i$ at step $t + 1$. Then at step $t_i$, $L_o(i, t_i) = 1$, by construction, $L_o(i, t) = 2L_o(i, t - 1)$. One can easily derive $L_o(i, t) = 2^{t - t_i}$. The relation between $k_i(t)$ and $L_o(i, t)$ is

$$
k_i(t) = 2L_o(i, t) = 2^{t + 1 - t_i}.
$$

Equation (5) shows that the degree spectrum of a generalized Koch network is discrete. It follows that the cumulative degree distribution [2] is given by

$$
P_{\text{cum}}(k) = \frac{1}{N_t} \sum_{\tau \leq t_i} L_v(\tau) = \frac{(m - 1) \times (m + 1)^{t_i} + 1}{(m - 1) \times (m + 1)^t + 1}.
$$

doi:10.1088/1742-5468/2010/07/P07011
Substituting $t_i = t + 1 - \ln k / \ln 2$ in equation (7), we obtain

$$P_{\text{cum}}(k) = \frac{(m - 1) \times (m + 1)^t \times (m + 1) \times k^{-\ln(m+1)/\ln 2} + 1}{(m - 1) \times (m + 1)^t + 1}. \tag{8}$$

When $t$ is large enough, one can obtain

$$P_{\text{cum}}(k) = (m + 1) \times k^{-\ln(m+1)/\ln 2}. \tag{9}$$

So the degree distribution follows a power-law form with the exponent $\gamma = \ln(m+1)/\ln 2 + 1$; in particular $\gamma = 3$, as in [24], when $m$ is 3.

### 3.2. Clustering coefficient

The clustering coefficient [7] of a node $i$ with a degree $k_i$ is given by $C_i = 2e_i/(k_i(k_i-1))$, where $e_i$ is the number of existing triangles attached to node $i$, and $k_i(k_i-1)/2$ is the total number of possible triangles including node $i$. For a single node with degree $k$ in $GK(t)$, the clustering coefficient to this node is denoted by $C(k)$. When $m$ is 3, the number of existing triangles attached to this node is $k/2$, we have

$$C(k) = \frac{1}{k-1}, \tag{10}$$

while when $m$ is greater than 3, since no triangle exists attached to this node, we get

$$C(k) = 0. \tag{11}$$

### 3.3. Average path length

Let $d_t$ denote the APL of the generalized Koch network $GK(t)$. Since a generalized Koch network is self-similar, the APL can be computed analytically to obtain an explicit formula by using a method similar to those in [24]. We represent all the shortest path lengths of $GK(t)$ as a matrix in which the entries $d_{ij}$ represent the shortest distance from node $i$ to node $j$, then

$$d_t = \frac{D_t}{N_t(N_t - 1)/2} \tag{12}$$

is defined as the mean of $d_{ij}$ over all pairs of nodes, where

$$D_t = \sum_{i \in GK(t), j \in GK(t)} d_{ij} \tag{13}$$

denotes the sum of the shortest path length between two nodes over all pairs. We must mention that in equation (13), for a pair of nodes $i$ and $j$ ($i \neq j$), we only count $d_{ij}$ or $d_{ji}$, not both. Since the generalized Koch network is a self-similar structure, to obtain $GK(t+1)$, we can make $m+1$ copies of $GK(t)$ and join them at the hub nodes. As shown in figure 2, network $GK(t+1)$ may be obtained by the juxtaposition of $m+1$ copies of $GK(t)$, which are labeled as $GK_1^t$, $GK_2^t$, $\ldots$, $GK_{m+1}^t$, respectively. Then we can write the sum $D_{t+1}$ as

$$D_{t+1} = (m + 1)D_t + \Delta_t, \tag{14}$$

doi:10.1088/1742-5468/2010/07/P07011
where $\Delta_t$ is the sum over all shortest paths whose end points are not in the same $GK(t)$ branch. The solution of equation (14) is

$$D_t = (m + 1)^{t-1} D_1 + \sum_{x=1}^{t-1} m^{t-1-x} \Delta_x.$$  

(15)

The paths that contribute to $\Delta_t$ must all go through at least one of the $m$ edge nodes at which the different $GK(t)$ branches are connected. The analytical expression for $\Delta_t$, called the length of crossing paths, is found below. Denote $\Delta_t^{\alpha,\beta}$ as the sum of the lengths of all shortest paths with end points in $GK_t^\alpha$ and $GK_t^\beta$, respectively. If $GK_t^\alpha$ and $GK_t^\beta$ meet at an edge node, $\Delta_t^{\alpha,\beta}$ rules out the paths where either end point is that shared edge node. If $GK_t^\alpha$ and $GK_t^\beta$ do not meet, $\Delta_t^{\alpha,\beta}$ excludes the paths where either end point is any edge node. Then the total sum $\Delta_t$ is

$$\Delta_t = \sum_{1 \leq i < j \leq m+1} \Delta^{i,j}_t.$$  

(16)

To see figure 2, by symmetry we obtain $\Delta_t^{m+1,1} = \Delta_t^{m+1,2} = \cdots = \Delta_t^{m+1,m}$, $\Delta_t^{1,2} = \Delta_t^{2,3} = \cdots = \Delta_t^{m,1}$, $\Delta_t^{1,3} = \Delta_t^{2,5} = \cdots = \Delta_t^{m,2}$, and $\Delta_t^{1,[m/2]} = \Delta_t^{2,[m/2]+1} = \cdots = \Delta_t^{m,[m/2]-1}$, thus we have

$$\Delta_t = \begin{cases} 
  m \Delta_t^{1,m+1} + m \Delta_t^{1,2} + \cdots + m \Delta_t^{1,1/2+1}, & \text{m is odd;} \\
  m \Delta_t^{1,m+1} + m \Delta_t^{1,2} + \cdots + m \Delta_t^{1,m/2} + \frac{m}{2} \Delta_t^{1,m/2+1}, & \text{m is even.}
\end{cases}$$  

(17)

In order to find $\Delta_t^{\alpha,\beta}$, we define

$$s_t = \sum_{i \in G_t \setminus X} d_i.$$  

(18)
According to the self-similar network structure, we can easily obtain the following recursive relation. When $m$ is odd, we have

$$s_{t+1} = 2s_t + 2[s_t + (N_t - 1)] + 2[s_t + 2(N_t - 1)] + \cdots + 2[s_t + \frac{m-1}{2}(N_t - 1)]$$

$$= (m + 1)s_t + \frac{(m-1)^2}{4}(m + 1)^{t+1};$$  \hspace{1cm} (19)

when $m$ is even, we have

$$s_{t+1} = 2s_t + 2[s_t + (N_t - 1)] + 2[s_t + \frac{m-2}{2}(N_t - 1)] + \cdots + 2[s_t + \frac{m-2}{2}(N_t - 1)] + \frac{m^2}{4}(m - 1)(m + 1)^t.$$  \hspace{1cm} (20)

Using $s_1 = m(m - 1)((m-1)/2 + 1)$ when $m$ is odd and $s_1 = 2m[m/2 + ((m - 2)/2)((m - 2)/2 + 1)] = m^3/2$ when $m$ is even, we have

$$s_{t+1} = (m + 1)^{t+1}m(m - 1) + \frac{(m-1)^2}{4}t(m + 1)^{t+1}$$  \hspace{1cm} (21)

when $m$ is odd and

$$s_{t+1} = (m + 1)^{t}m^3/2 + \frac{m^2}{4}(m - 1)t(m + 1)^t$$  \hspace{1cm} (22)

when $m$ is even. On the other hand, by the definition given above, we have

$$\Delta_{t}^{1,m+1} = \sum_{i \in G_{K}^{1}, j \in G_{K}^{m+1}} d_{i,j}$$

$$= \sum_{i \in G_{K}^{1}, j \in G_{K}^{m+1}} (d_{iX} + d_{jX})$$

$$= (N_t - 1) \sum_{i \in G_{K}^{1}, i \neq X} d_{iX} + (N_t - 1) \sum_{j \in G_{K}^{m+1}, j \neq X} d_{iX}$$

$$= 2(N_t - 1) \sum_{i \in G_{K}^{1}, i \neq X} d_{iX} = 2(N_t - 1)s_t$$  \hspace{1cm} (23)

and

$$\Delta_{t}^{\alpha,\beta} = \sum_{i \in G_{K}^{\alpha}, i \neq X} \sum_{j \in G_{K}^{\beta}, j \neq Y} d_{i,j} = \sum_{i \in G_{K}^{\alpha}, i \neq X} \sum_{j \in G_{K}^{\beta}, j \neq Y} (d_{iX} + d_{XY} + d_{jY})$$

$$= 2(N_t - 1)s_t + (N_t - 1)^2 d_{XY}. \hspace{1cm} (24)$$

\[\text{doi:10.1088/1742-5468/2010/07/P07011}\]
Thus, when \( m \) is odd, substituting equations (23) and (24) into (16), we obtain
\[
\Delta_t = 2m(N_t - 1)s_t + 2m\frac{m - 1}{2}(N_t - 1)s_t + m(N_t - 1)^2 \left[ 1 + 2 + \cdots + \frac{m - 1}{2} \right] \\
= (N_t - 1)s_t(m^2 + m) + (N_t - 1)^2m\frac{(m + 1)(m - 1)}{8} \\
= (m + 1)^2\left[ \frac{(m - 1)^3m(m + 1)}{4}t + \frac{(m-1)^2m^2(m+1)}{2} - \frac{(m-1)^3m(m+1)}{8} \right].
\]

Similarly, when \( m \) is even, we get
\[
\Delta_t = 2m(N_t - 1)s_t + (m/2 - 1) \times 2m(N_t - 1)s_t + \frac{m}{2} \times 2(N_t - 1)s_t \\
+ m \times (N_t - 1)^2\left[ 1 + 2 + \cdots + \left( \frac{m}{2} - 1 \right) \right] + \frac{m}{2} \times \frac{m}{2}(N_t - 1)^2 \\
= (m^2 + m)(N_t - 1)s_t + \frac{m^3}{8}(N_t - 1)^2 \\
= (m + 1)^{2t-1}\left[ \frac{(m - 1)^2m^3(m + 1)}{4}t + \frac{(m-1)m^4(m+1)}{2} \\
+ \frac{(m-1)^2m^3(m+1)}{8} \right].
\]

In particular, when \( m = 3 \), \( \Delta_t = 16t(24t + 60) \), as in (25). Denote \( C_1 \) as \((m - 1)^2m(m+1)/4 \) \( \) and \( C_2 \) as \((m - 1)^2m^2(m+1)/2 -(m - 1)^3m(m+1)/8 \), respectively. When \( m \) is odd, substituting equation (25) into (15), we have
\[
D_t = (m + 1)^{t-1}D_1 + C_1(m + 1)^t \frac{d}{dm}\left\{ \sum_{x=1}^{t-1}(m + 1)^x \right\} + C_2(m + 1)^t \frac{\sum_{x=1}^{t-1}(m + 1)^{x-1}}{m} \\
= (m + 1)^{t-1}D_1 + C_1(m + 1)^t \frac{mnt(m + 1)^{t-1} - (m + 1)^t + 1}{m^2} \\
+ C_2(m + 1)^t \left( \frac{m + 1}m \right)^{t-1} - 1,
\]
where
\[
D_1 = \frac{(m - 1)m(m + 1)(3m^2 - m)}{8}.
\]

Clearly when \( m = 3 \), \( D_1 \) is 72, as in [24]. Denote \( C_3 \) as \((m - 1)^2m^3(m+1)/4 \) and \( C_4 \) as \((m - 1)^2m^4(m+1)/2 + (m - 1)^2m^3(m+1)/8 -(m - 1)^2m^3(m+1)/4 \), respectively. When \( m \) is even, inserting equation (26) into (15), we have
\[
D_t = (m + 1)^{t-1}D_1 + C_3(m + 1)^{t-1} \frac{d}{dm}\left\{ \sum_{x=1}^{t-1}(m + 1)^x \right\} + C_4(m + 1)^{t-1} \frac{\sum_{x=1}^{t-1}(m + 1)^{x-1}}{m} \\
= (m + 1)^{t-1}D_1 + C_3(m + 1)^t \frac{mnt(m + 1)^{t-1} - (m + 1)^t + 1}{m^2} \\
+ C_4(m + 1)^{2t-2} - (m + 1)^{t-1},
\]

\ doi:10.1088/1742-5468/2010/07/P07011
where
\[ D_1 = (3m - 1) \frac{m^4}{8}. \]  
(30)

Using (12), it is easy to find an analytical result for \( d_t \) for a generalized Koch network. When \( t \to +\infty \), one can see \( d_t \sim t \).

4. Conclusion

The structural properties play an important role in the characterization of a network and greatly affect its dynamical processes. In this paper, based on the Koch network, we put forward the generalized Koch network, and obtain rigorously the solution for the average path length, clustering coefficient and degree distribution. Our explicit analyses show that the generalized Koch network has the properties of the Koch network. The proposed models will help us efficiently describe evolution systems and investigate their structural properties.

Acknowledgment

This work was supported by the Research Foundation of Hangzhou Dianzi University (No. KYS075608083).

References


doi:10.1088/1742-5468/2010/07/P07011