The rigorous solution for the average distance of a Sierpinski network

Zhongzhi Zhang\textsuperscript{1,2}, Lichao Chen\textsuperscript{1,2}, Lujun Fang\textsuperscript{3}, Shuigeng Zhou\textsuperscript{1,2}, Yichao Zhang\textsuperscript{4} and Jihong Guan\textsuperscript{4}

\textsuperscript{1} Department of Computer Science and Engineering, Fudan University, Shanghai 200433, People’s Republic of China
\textsuperscript{2} Shanghai Key Laboratory of Intelligent Information Processing, Fudan University, Shanghai 200433, People’s Republic of China
\textsuperscript{3} Electrical Engineering and Computer Science Department, University of Michigan, 2260 Hayward Avenue, Ann Arbor, MI 48109, USA
\textsuperscript{4} Department of Computer Science and Technology, Tongji University, 4800 Cao’an Road, Shanghai 201804, People’s Republic of China
E-mail: zhangzz@fudan.edu.cn, chenlichao@gmail.com, ljfang@umich.edu, sgzhou@fudan.edu.cn, achilles_talk@126.com and jhguan@tongji.edu.cn

Received 29 October 2008
Accepted 3 January 2009
Published 12 February 2009

Abstract. The closed-form solution for the average distance of a deterministic network—the Sierpinski network—is found. This important quantity is calculated exactly with the help of recursion relations, which are based on the self-similar network structure and enable one to derive the precise formula analytically. The rigorous solution obtained confirms our previous numerical result, which shows that the average distance grows logarithmically with the number of network nodes. The result is at variance with that derived from random networks.

Keywords: exact results, network dynamics, random graphs, networks

ArXiv ePrint: 0810.5172
1. Introduction

Structural properties [1], such as degree distribution [2], average distance [3], degree correlations [4], community [5], motifs [6], fractality [7], and symmetry [8], have received much attention in the field of complex networks, since these features play significant roles in characterizing and understanding complex networked systems in Nature and society. Among these important features, average distance characterizes the small-world behavior commonly observed in various disparate real networks [3]. It has been established that average distance is related to other structural properties, such as degree distribution [9,10], fractality [11,12] and symmetry [13]. On the other hand, average distance has an important consequence for dynamical processes taking place on networks, including disease spreading [3], routing [14,15], robustness [11,12], percolation [16], and so on. Thus far, average distance has been a focus of attention for the scientific community [17]–[22].

Using above-mentioned structural properties, extensive empirical studies on diverse real systems have been made in an attempt to uncover and understand the generic features and complexity of these systems, and various network models have been proposed for reproducing or explaining the common characteristics of real-life networks [23]–[26]. Recently, inspired by the well-known Sierpinski gasket, we proposed a novel network, called the Sierpinski network [27]. The Sierpinski network belongs to a class of deterministically growing networks that have attracted considerable attention and have turned out to constitute a useful tool [28]–[34]. Many relevant topological properties of Sierpinski networks such as degree distribution, clustering coefficient, and strength distribution have been determined analytically [27]. Also, the average distance of the Sierpinski network has been investigated numerically; it was shown to exhibit a logarithmic scaling with the number of network nodes (vertices) [27].

In view of the importance and usefulness of the quantity—average distance—here we derive a closed-form formula for the average distance characterizing the Sierpinski network.
The rigorous solution for the average distance of a Sierpinski network

Figure 1. The first two stages in the construction for a variant of the Sierpinski gasket.

The analytic method is based on the recursive construction and self-similar structure of the Sierpinski network. Our precise result shows that the average distance of the Sierpinski network increases logarithmically with the number of nodes. This scaling behaves differently from that of random networks [9,10]. Our rigorous solution confirms the scaling relation between average distance and number of network nodes that was previously obtained numerically in [27].

2. A brief introduction to the Sierpinski network

The Sierpinski network is derived from a variation of the Sierpinski gasket [35]. The Sierpinski gasket variant, shown in figure 1, is constructed as follows [36]. We start with an equilateral triangle, and we denote this initial configuration as that of generation $t=0$. Then in the first generation $t=1$, we divide the three sides of the equilateral triangle into three, then join these points and remove the three downward pointing triangles. This forms six copies of the original triangle, and the procedure is repeated indefinitely for all the new copies. In the limit of infinite number $t$ of generations, we get a fractal variant of the Sierpinski gasket. The Hausdorff dimension of the fractal obtained is $d_f = 1 + \frac{\ln 2}{\ln 3}$ [37]. From the fractal, one can define the Sierpinski network [27], where vertices correspond to the removed triangles and two vertices are connected if the boundaries of the corresponding triangles contact each other. Note that for uniformity, the three sides of the initial equilateral triangle at step 0 also correspond to three different vertices. Figure 2 shows the network construction process.

The Sierpinski network can be generated using an iterative algorithm [27]. We denote the network after $t$ iterations by $F_t$ with $t \geq 0$. Then the network is constructed as follows. For $t = 0$, $F_0$ is a triangle. Next, three nodes are added into the original triangle. These three new nodes are connected to each other forming a new triangle, and both ends of each edge of the new triangle are linked to a node of the original triangle. Thus $F_1$ is obtained; see figure 3. For $t \geq 1$, we can get $F_t$ from $F_{t-1}$. Each of the existing triangles of $F_{t-1}$ that does not consist of three simultaneously emerging nodes and has never generated a node before we define as an active triangle. We replace each of the active triangles in $F_{t-1}$ by the connected cluster on the right-hand side of figure 3 to obtain $F_t$. 

doi:10.1088/1742-5468/2009/02/P02034
The rigorous solution for the average distance of a Sierpinski network

The resulting network presents the typical characteristics of real-life networks in Nature and society [27]. It has power-law distributions of degree and strength, with exponents $\gamma_k = 2 + (\ln 2/\ln 3)$ and $\gamma_s = \frac{3}{2} + (\ln 2/2 \ln 3)$, respectively. There also exists a power-law scaling relation between the strength $s$ and degree $k$ of an individual node, i.e. $s \sim k^2$. On the other hand, for any individual vertex, its clustering coefficient is $C(k) = (4/k) - (1/(k - 1))$; so when $k$ is large, $C(k)$ is approximately inversely proportional to degree $k$. The mean value $C$ of clustering coefficients for all vertices is very large; it asymptotically reaches a constant value 0.598. Moreover, the network is a maximal planar graph.

3. Rigorous derivation of the average distance

After introducing the Sierpinski network, we now derive analytically the average distance. We represent all the shortest path lengths of network $F_t$ as a matrix in which the entry $d_{ij}$ is the distance between node $i$ and $j$ that is the length of a shortest path joining $i$ and $j$. A measure of the typical separation between two nodes in $F_t$ is given by the average distance $d_t$ defined as the mean of distances over all pairs of nodes:

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

where $D_t$ is the sum of all distances in the matrix, $N_t$ is the number of nodes in the network at time $t$. This formula represents the average distance in the network.
The rigorous solution for the average distance of a Sierpinski network

**Figure 4.** Schematic illustration of the second means of construction of the Sierpinski network. $F_{t+1}$ may be obtained by joining six copies of $F_t$ denoted as $F_t^{(\eta)} (\eta = 1, \ldots, 6)$, which are connected to one another at the edge nodes.

where

$$ D_t = \sum_{u \in F_t, v \in F_t, i \neq j} d_{ij} \quad (2) $$

denotes the sum of the distances between two nodes over all couples.

### 3.1. A recursive equation for total distances

We continue by exhibiting the procedure for determining the total distance and present the recurrence formula which allows us to obtain $D_{t+1}$ for the $t+1$ generation from $D_t$ for the $t$ generation. The Sierpinski network $F_t$ has a self-similar structure that allows one to calculate $D_t$ analytically [38]. As shown in figure 4, network $F_{t+1}$ may be obtained by joining at six edge nodes (i.e., $A, B, C, X, Y,$ and $Z$) six copies of $F_t$ that are labeled as $F_t^{(1)}, \ldots, F_t^{(6)}$ [39]. From this we can obtain the recursion relation

$$ N_{t+1} = 6 N_t - 12, \quad (3) $$

for the network order $N_t$, which is the number of nodes in the graph of generation $t$. This recursion, coupled with $N_0 = 3$, yields

$$ N_t = \frac{3 \times 6^t + 12}{5}, \quad (4) $$

as previously obtained in [27].

According to the second construction method, the total distance $D_{t+1}$ satisfies the recursion relation

$$ D_{t+1} = 6 D_t + \Delta_t - 6, \quad (5) $$

where $\Delta_t$ is the sum over all shortest path length whose end-points are not in the same $F_t^{(\eta)}$ branch. The last term, $-6$, on the right-hand side of equation (5) compensates for
the overcounting of certain paths: the shortest path between $A$ and $X$, with length 1, is included in both $F^{(1)}_t$ and $F^{(2)}_t$. Similarly the shortest path between $A$ and $Y$, the shortest path between $B$ and $X$, the shortest path between $B$ and $Z$, the shortest path between $C$ and $Y$, and the shortest path between $C$ and $Z$ are all computed twice. To determine $D_t$, all that is left is to calculate $\Delta_t$.

3.2. Definition of the crossing distance

In order to compute $\Delta_t$, we classify the nodes in $F_{t+1}$ into two categories: the six edge nodes (such as $A$, $B$, $C$, $X$, $Y$, and $Z$ in figure 4) are called hub nodes, while the other nodes are named non-hub nodes. Thus $\Delta_t$, named the crossing distance, can be obtained by summing the following path lengths that are not included in the distance of node pairs in $F^{(n)}$: the length of the shortest paths between non-hub and non-hub nodes, the length of the shortest paths between hub and non-hub nodes, and the length of the shortest paths between hub nodes (i.e., $d_{AZ}$, $d_{BY}$, and $d_{CZ}$).

Denote as $\Delta_{t}^{\alpha,\beta}$ the sum of all shortest paths between non-hub nodes, whose endpoints are in $F^{(\alpha)}_t$ and $F^{(\beta)}_t$, respectively. That is to say, $\Delta_{t}^{\alpha,\beta}$ rules out the paths with end-point at the hub nodes belonging to $F^{(\alpha)}_t$ or $F^{(\beta)}_t$. For example, each path contributed to $\Delta_{1}^{1,2}$ does not end at node $A$, $B$, $X$ or $Y$. On the other hand, let $\Omega^{(n)}_t$ be the set of non-hub nodes in $F^{(n)}_t$. Then the total sum $\Delta_t$ is given by

$$
\Delta_t = \Delta_{t}^{1,2} + \Delta_{t}^{1,3} + \Delta_{t}^{1,4} + \Delta_{t}^{1,5} + \Delta_{t}^{1,6} + \Delta_{t}^{2,3} + \Delta_{t}^{2,4} + \Delta_{t}^{2,5} + \Delta_{t}^{2,6} + \Delta_{t}^{3,4} + \Delta_{t}^{3,5} + \Delta_{t}^{3,6} + \Delta_{t}^{4,5} + \Delta_{t}^{4,6} + \Delta_{t}^{5,6} + \sum_{j \in \Omega^{(n)}_t} d_{Aj} + \sum_{j \in \Omega^{(n)}_t} d_{Aj} + \sum_{j \in \Omega^{(n)}_t} d_{Bj} + \sum_{j \in \Omega^{(n)}_t} d_{Bj} + \sum_{j \in \Omega^{(n)}_t} d_{Cj} + \sum_{j \in \Omega^{(n)}_t} d_{Cj} + \sum_{j \in \Omega^{(n)}_t} d_{Xj} + \sum_{j \in \Omega^{(n)}_t} d_{Xj} + \sum_{j \in \Omega^{(n)}_t} d_{Yj} + \sum_{j \in \Omega^{(n)}_t} d_{Yj} + \sum_{j \in \Omega^{(n)}_t} d_{Zj} + \sum_{j \in \Omega^{(n)}_t} d_{Zj} + \sum_{j \in \Omega^{(n)}_t} d_{AZ} + d_{BY} + d_{CX}.
$$

(6)

By symmetry, $\Delta_{t}^{1,2} = \Delta_{t}^{1,6} = \Delta_{t}^{2,3} + \Delta_{t}^{3,4} = \Delta_{t}^{4,5} = \Delta_{t}^{5,6} = \Delta_{t}^{1,5} = \Delta_{t}^{2,4} = \Delta_{t}^{2,6} = \Delta_{t}^{3,5} = \Delta_{t}^{4,6}$, $\Delta_{t}^{1,3} = \Delta_{t}^{2,5} = \Delta_{t}^{3,6}$, $\sum_{j \in \Omega^{(n)}_t} d_{Aj} = \sum_{j \in \Omega^{(n)}_t} d_{Aj} = \sum_{j \in \Omega^{(n)}_t} d_{Bj} = \sum_{j \in \Omega^{(n)}_t} d_{Bj} = \sum_{j \in \Omega^{(n)}_t} d_{Cj} = \sum_{j \in \Omega^{(n)}_t} d_{Cj} = \sum_{j \in \Omega^{(n)}_t} d_{Xj} = \sum_{j \in \Omega^{(n)}_t} d_{Xj} = \sum_{j \in \Omega^{(n)}_t} d_{Yj} = \sum_{j \in \Omega^{(n)}_t} d_{Yj} = \sum_{j \in \Omega^{(n)}_t} d_{Zj} = \sum_{j \in \Omega^{(n)}_t} d_{Zj}$, and $d_{AZ} = d_{BY} = d_{CX} = 2$, so equation (6) can be simplified to

$$
\Delta_t = 6\Delta_{t}^{1,2} + 6\Delta_{t}^{1,3} + 3\Delta_{t}^{1,4} + 18 \sum_{j \in \Omega^{(n)}_t} d_{Aj} + 6.
$$

(7)

Having $\Delta_t$ in terms of the quantities of $\Delta_{t}^{1,2}$, $\Delta_{t}^{1,3}$, $\Delta_{t}^{1,4}$, and $\sum_{j \in \Omega^{(n)}_t} d_{Aj}$, the next step is to explicitly determine these quantities.
The rigorous solution for the average distance of a Sierpinski network

3.3. Classification of interior nodes

To calculate the crossing distances $\Delta_{t}^{1,2}$, $\Delta_{t}^{1,3}$, $\Delta_{t}^{1,4}$, and $\sum_{j\in\Omega_{t}^{1}} d_{Aj}$, we classify interior nodes in network $F_{t+1}$ into seven different parts according to their shortest path lengths to each of the three hub nodes (i.e. A, B, C) of the peripheral triangle $\triangle ABC$. Notice that nodes A, B, C themselves are not partitioned into any of the seven parts represented as $P_1$, $P_2$, $P_3$, $P_4$, $P_5$, $P_6$, and $P_7$, respectively. The classification of nodes is shown in figure 5. For any interior node $v$, we denote the shortest path lengths from $v$ to A, B, C as $a$, $b$, and $c$, respectively. By construction, $a$, $b$, $c$ can differ by at most 1 since vertices A, B, C are adjacent. Then the classification function $\text{class}(v)$ of node $v$ is defined to be

$$
\text{class}(v) = \begin{cases} 
    P_1 & \text{for } a < b = c, \\
    P_2 & \text{for } b < a = c, \\
    P_3 & \text{for } c < a = b, \\
    P_4 & \text{for } a = c < b, \\
    P_5 & \text{for } a = b < c, \\
    P_6 & \text{for } b = c < a, \\
    P_7 & \text{for } a = b = c.
\end{cases}
$$

(8)

It should be mentioned that the definition of node classification is recursive. For instance, classes $P_1$ and $P_4$ in $F_{t}^{(1)}$ belong to class $P_1$ in $F_{t+1}$, classes $P_3$ and $P_5$ in $F_{t}^{(1)}$ belong to class $P_2$ in $F_{t+1}$, classes $P_2$, $P_6$, and $P_7$ in $F_{t}^{(1)}$ belong to class $P_1$ in $F_{t+1}$. Since the three nodes A, B, and C are symmetrical, in the Sierpinski network we have the following equivalent relations from the viewpoint of class cardinality: classes $P_1$, $P_2$, and $P_3$ are equivalent to one another, and it is the same with classes $P_4$, $P_5$, and $P_6$. We denote the number of nodes in network $F_{t}$ that belong to class $P_1$ as $N_{t,P_1}$, the number

Figure 5. Illustration of the classification of interior nodes in $F_{t}^{(\eta)} (\eta = 1, \ldots, 6)$, from which we can derive recursively the classification of interior nodes in network $F_{t+1}$.
of nodes in class $P_2$ as $N_{t,P_2}$, and so on. By symmetry, we have $N_{t,P_1} = N_{t,P_2} = N_{t,P_5}$ and $N_{t,P_2} = N_{t,P_3} = N_{t,P_5}$. Therefore in the following computation we will only consider $N_{t,P_1}$, $N_{t,P_4}$, and $N_{t,P_7}$. It is easy to conclude that

$$
N_{t} = N_{t,P_1} + N_{t,P_2} + N_{t,P_3} + N_{t,P_4} + N_{t,P_5} + N_{t,P_7} + 3
= 3N_{t,P_1} + 3N_{t,P_4} + N_{t,P_7} + 3.
$$

(9)

Considering the self-similar structure of the Sierpinski network, we can easily see that at time $t+1$, the quantities $N_{t+1,P_1}$, $N_{t+1,P_4}$, and $N_{t+1,P_7}$ evolve according to the following recursive equations:

$$
\begin{align*}
N_{t+1,P_1} &= 3N_{t,P_1} + 4N_{t,P_4} + N_{t,P_7}, \\
N_{t+1,P_4} &= 3N_{t,P_1} + N_{t,P_4} + N_{t,P_7} + 1, \\
N_{t+1,P_7} &= 3N_{t,P_4},
\end{align*}
$$

(10)

where we have used the equivalent relations $N_{t,P_1} = N_{t,P_2} = N_{t,P_5}$ and $N_{t,P_2} = N_{t,P_3} = N_{t,P_5}$. With the initial condition $N_{2,P_1} = 4$, $N_{2,P_4} = 2$, and $N_{2,P_7} = 3$, we can solve the recursive equation (10) to obtain

$$
\begin{align*}
N_{t,P_1} &= \frac{1}{240} \left[ -112 + 25 \times (-2)^t + 27 \times 6^t \right], \\
N_{t,P_4} &= \frac{1}{120} \left[ 16 - 25 \times (-2)^t + 9 \times 6^t \right], \\
N_{t,P_7} &= \frac{1}{80} \left[ 32 + 25 \times (-2)^t + 3 \times 6^t \right].
\end{align*}
$$

(11)

For a node $v$ in network $F_{t+1}$, we are also interested in the smallest value of the shortest path length from $v$ any of the three peripheral hub nodes $A$, $B$, and $C$. We denote the shortest distance as $f_v$, which can be defined to be

$$
f_v = \min(a, b, c).
$$

(12)

Let $d_{t,P_1}$ denote the sum of the $f_v$ for all nodes belonging to class $P_1$ in network $F_t$. Analogously, we can also define the quantities $d_{t,P_2}$, $d_{t,P_3}$, $d_{t,P_4}$, $d_{t,P_5}$, $d_{t,P_6}$, and $d_{t,P_7}$. Again by symmetry, we have $d_{t,P_1} = d_{t,P_2} = d_{t,P_3}$, $d_{t,P_4} = d_{t,P_5} = d_{t,P_6}$, and $d_{t,P_1}$, $d_{t,P_4}$, $d_{t,P_7}$ can be written recursively as follows:

$$
\begin{align*}
d_{t+1,P_1} &= 3d_{t,P_1} + 4d_{t,P_4} + d_{t,P_7}, \\
d_{t+1,P_4} &= 3d_{t,P_1} + d_{t,P_4} + d_{t,P_7} + 3N_{t,P_5} + 1, \\
d_{t+1,P_7} &= 3(d_{t,P_4} + N_{t,P_4}).
\end{align*}
$$

(13)

Substituting equation (11) into equation (13), and considering the initial condition $d_{2,P_1} = 4$, $d_{2,P_4} = 2$, and $d_{2,P_7} = 6$, equation (13) is solved inductively:

$$
\begin{align*}
d_{t,P_1} &= \frac{1}{25600} \left[ 2048 + 327 \times 6^t + 990t \times 6^t + 25(33 + 70t)2^t e^{i\pi t} \right], \\
d_{t,P_4} &= \frac{1}{12800} \left[ -2075(-2)^t - 1024(-1)^{2t} + 699 \times 6^t + 10t(-175 \times (-2)^t + 33 \times 6^t) \right], \\
d_{t,P_7} &= \frac{1}{25600} \left[ 4096 + 1329 \times 6^t + 330t \times 6^t + 25(359 + 210t)2^t e^{i\pi t} \right].
\end{align*}
$$

(14)

doi:10.1088/1742-5468/2009/02/P02034
3.4. Calculation of crossing distances

Having obtained the quantities $N_{t,P_i}$ and $d_{t,P_i}$ ($i = 1, 2, \ldots, 7$), we now begin to determine the crossing distance $\Delta_{t,2}^1$, $\Delta_{t,3}^1$, $\Delta_{t,4}^1$, and $\sum_{j \in Q_i^t} d_{Aj}$ expressed as a function of $N_{t,P_i}$ and $d_{t,P_i}$. Here we only give the computation details for $\Delta_{t,2}^1$, while the processes for computing $\Delta_{t,3}^1$, $\Delta_{t,4}^1$, and $\sum_{j \in Q_i^t} d_{Aj}$ are similar. For convenience of computation, we use $\Gamma_{t,i}^n$ to denote the set of interior nodes belonging to class $P_i$ in $F_t^{(n)}$. Then $\Delta_{t,2}^1$ can be written as

$$
\Delta_{t,2}^1 = \sum_{u \in \Gamma_{t,1}^{1,1}, v \in F_t^{(2)}} d_{uv}^+ \sum_{u \in \Gamma_{t,1}^{1,2}, v \in F_t^{(2)}} d_{uv}^+ \sum_{u \in \Gamma_{t,1}^{1,3}, v \in F_t^{(2)}} d_{uv}^+ \sum_{u \in \Gamma_{t,1}^{1,4}, v \in F_t^{(2)}} d_{uv}^+
+ \sum_{u \in \Gamma_{t,2,1}} d_{uv}^+ \sum_{u \in \Gamma_{t,2,3}} d_{uv}^+ \sum_{u \in \Gamma_{t,2,5}} d_{uv}^+. \tag{15}
$$

The seven terms on the right-hand side of equation (15) are represented consecutively as $\delta_i^t$ ($i = 1, 2, \ldots, 7$). Next we will calculate the quantities $\delta_i^t$. By symmetry, $\delta_i^t = \delta_i^t$, $\delta_i^t = \delta_i^t$. Therefore, we need only compute $\delta_i^t$, $\delta_i^t$, $\delta_i^t$, $\delta_i^t$ and $\delta_i^t$. Firstly, we evaluate $\delta_i^t$. By definition,

$$
\delta_i^t = \sum_{u \in \Gamma_{t,1}^{1,1}, v \in F_t^{(2)}} d_{uv}^+ \sum_{u \in \Gamma_{t,1}^{1,2}, v \in F_t^{(2)}} (d_{uv}^+ + d_{Av})
+ \sum_{u \in \Gamma_{t,1}^{1,3}, v \in F_t^{(2)}} (d_{uv}^+ + d_{AY}) + \sum_{u \in \Gamma_{t,1}^{1,4}, v \in F_t^{(2)}} (d_{uv}^+ + d_{AX} + d_{Xv})
= N_{t,P_i}(3d_{t,P_i} + 3d_{t,P_i} + d_{t,P_i} + 2N_{t,P_i} + N_{t,P_i})
+ d_{t,P_i}(3N_{t,P_i} + 3N_{t,P_i} + N_{t,P_i}). \tag{16}
$$

Proceeding similarly, we obtain

$$
\delta_i^t = N_{t,P_i}(3d_{t,P_i} + 3d_{t,P_i} + d_{t,P_i} + 4N_{t,P_i} + 3N_{t,P_i} + N_{t,P_i})
+ d_{t,P_i}(3N_{t,P_i} + 3N_{t,P_i} + N_{t,P_i}), \tag{17}
$$

$$
\delta_i^t = N_{t,P_i}(3d_{t,P_i} + 3d_{t,P_i} + d_{t,P_i} + N_{t,P_i} + d_{t,P_i}(3N_{t,P_i} + 3N_{t,P_i} + N_{t,P_i}), \tag{18}
$$

$$
\delta_i^t = N_{t,P_i}(3d_{t,P_i} + 3d_{t,P_i} + d_{t,P_i} + 2N_{t,P_i} + N_{t,P_i}) + d_{t,P_i}(3N_{t,P_i} + 3N_{t,P_i} + N_{t,P_i}), \tag{19}
$$

and

$$
\delta_i^t = N_{t,P_i}(3d_{t,P_i} + 3d_{t,P_i} + d_{t,P_i} + N_{t,P_i}) + d_{t,P_i}(3N_{t,P_i} + 3N_{t,P_i} + N_{t,P_i}). \tag{20}
$$

With the results obtained for $\delta_i^t$, we have

$$
\Delta_{t,2}^1 = 2(3d_{t,P_i} + 3d_{t,P_i} + d_{t,P_i})(3N_{t,P_i} + 3N_{t,P_i} + N_{t,P_i}) + N_{t,P_i}(3N_{t,P_i} + 3N_{t,P_i} + N_{t,P_i})
+ 2(N_{t,P_i} + N_{t,P_i})(2N_{t,P_i} + N_{t,P_i}) + N_{t,P_i}(N_{t,P_i} + N_{t,P_i}) + (N_{t,P_i})^2. \tag{21}
$$

doi:10.1088/1742-5468/2009/02/P02034
Analogously, we find
\[
\Delta_{t,1} = 2(3d_{t,P_1} + 3d_{t,P_4} + d_{t,P_7})(3N_{t,P_1} + 3N_{t,P_4} + N_{t,P_7}) + 2(N_{t,P_1})^2 \\
+ 2N_{t,P_1}(3N_{t,P_1} + 3N_{t,P_4} + N_{t,P_7}) + N_{t,P_1}(3N_{t,P_1} + 3N_{t,P_4} + N_{t,P_7}) + (N_{t,P_1})^2 \\
+ 2N_{t,P_1} + N_{t,P_1})(2N_{t,P_1} + N_{t,P_1}),
\]
(22)
\[
\Delta_{t,1} = 2(3d_{t,P_1} + 3d_{t,P_4} + d_{t,P_7})(3N_{t,P_1} + 3N_{t,P_4} + N_{t,P_7}) + (3N_{t,P_1} + 3N_{t,P_4} + N_{t,P_7})^2 \\
+ 3(N_{t,P_1})^2,
\]
(23)
and
\[
\sum_{j \in \Omega_2} d_{Aj} = (3d_{t,P_1} + 3d_{t,P_4} + d_{t,P_7}) + (3N_{t,P_1} + 3N_{t,P_4} + N_{t,P_7}) + N_{t,P_1}.
\]
(24)

Substituting equations (21), (22), (23), and (24) into equation (7), we get the final expression for crossing distances \(\Delta_t\):
\[
\Delta_t = \frac{1}{192000}[-17920 + 95040 \times 6^t + 160380 \times 6^{2t} + 50 \times 2^t(-32 + 27 \times 6^t) e^{i\pi t} \\
+ 44550t \times 6^{2t} + e^{2i\pi t}(6875 \times 4^t - 82944 \times 6^t + 8019 \times 6^{2t} \\
+ 26730t \times 6^{3t})].
\]
(25)

### 3.5. The rigorous result for the average distance

With the above-obtained results and recursion relations, we now readily calculate the sum of the shortest path lengths between all pairs of nodes. Inserting equation (25) into equation (5) and using the initial condition \(D_2 = 555\), equation (5) is solved inductively:
\[
D_t = \frac{1}{192000}[266240 - 34375 \times 4^t + 2000 \times (-2)^t + 225264 \times 6^t - 750(-12)^t \\
+ 27621 \times 6^{2t} + 20160t \times 6^t + 23760t \times 6^{2t}].
\]
(26)

Substituting equation (26) into equation (1) yields the exactly analytic expression for the average distance:
\[
d_t = \frac{1}{34560(84 + 57 \times 6^t + 9 \times 6^{2t})}[266240 - 34375 \times 4^t + 2000 \times (-2)^t + 225264 \times 6^t \\
- 750(-12)^t + 27621 \times 6^{2t} + 20160t \times 6^t + 23760t \times 6^{2t}].
\]
(27)

In the large \(t\) limit, \(d_t \sim t\), while the network order \(N_t \sim 6^t\), which is obvious from equation (4). Thus, the average distance grows logarithmically with increasing order of the network. This scaling is consistent with the speculation in [27] based on computer simulations. We have also checked our analytic result provided by equation (27) against numerical calculations for different network orders up to \(t = 8\) which corresponds to \(N_8 = 100772\). In all the cases we obtain a complete agreement between our theoretical formula and the results of numerical investigation; see figure 6.

Recently, it has been suggested that for random scale-free networks with degree exponent \(\gamma_k < 3\) and network order \(N\), their average distance \(d(N)\) behaves as a double-logarithmic scaling with \(N\): \(d(N) \sim \ln \ln N\) [9, 10]. However, for a deterministic Sierpinski network, in spite of the fact that its degree exponent \(\gamma_k = 2 + (\ln 2/\ln 3) < 3\), its average distance scales as a logarithmic scaling with network order, showing an obvious departure from that for the stochastic scale-free counterparts.
4. Conclusion

Average distance plays an important role in the characterization of the internal structure of a network, and has a profound impact on a variety of dynamical processes on the network. In this article, we have obtained rigorously the solution for the average distance of a deterministic Sierpinski network. We have explicitly shown that in the limit of infinite network order, the average distance of the Sierpinski network scales logarithmically with the number of network nodes, verifying our previously suggested scaling obtained through simulations [27]. Our findings display that the scaling of the average distance for the deterministic Sierpinski network is strikingly distinct from the counterpart for stochastic scale-free networks [9,10]. This disparity of the scaling for average distance between the deterministic Sierpinski network and random scale-free networks merits study in the future.

Acknowledgments

This research was supported by the National Basic Research Program of China under grant No. 2007CB310806, the National Natural Science Foundation of China under Grants Nos. 60704044, 60873040 and 60873070, Shanghai Leading Academic Discipline Project No. B114, and the Program for New Century Excellent Talents in University of China (NCET-06-0376).

References

[1] Costa L da F, Rodrigues F A, Travieso G and Boas P R V, 2007 Adv. Phys. 56 167
[2] Barabási A-L and Albert R, 1999 Science 286 509
[3] Watts D J and Strogatz H, 1998 Nature 393 440
[4] Newman M E J, 2002 Phys. Rev. Lett. 89 208701
[5] Girvan M and Newman M E J, 2002 Proc. Nat. Acad. Sci. 99 7821
[6] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D and Alon U, 2002 Science 298 824
[7] Song C, Havlin S and Makse H A, 2005 Nature 433 392

doi:10.1088/1742-5468/2009/02/P02034
The rigorous solution for the average distance of a Sierpinski network

[8] Xiao Y, Xiong M, Wang W and Wang H, 2008 Phys. Rev. E 77 066108
[9] Chung F and Lu L, 2002 Proc. Nat. Acad. Sci. 99 15879
[10] Cohen R and Havlin S, 2003 Phys. Rev. Lett. 90 058701
[11] Song C, Havlin S and Makse H A, 2006 Nat. Phys. 2 275
[12] Zhang Z Z, Zhou S G and Zou T, 2007 Eur. Phys. J. B 56 259
[13] Xiao Y, MacArthur B D, Wang H, Xiong M and Wang W, 2008 Phys. Rev. E 78 046102
[14] Yan G, Zhou T, Hu B, Fu Z Q and Wang B H, 2006 Phys. Rev. E 73 046108
[15] Zhang Z Z, Comellas F, Rrapaj A, Rong L L and Zhou S G, 2008 J. Phys. A: Math. Theor. 41 035004
[16] Zhang Z Z, Zhou S G, Zou T and Chen G S, 2008 J. Stat. Mech. P09008
[17] Dorogovtsev S N, Mendes J F F and Samukhin A N, 2003 Nucl. Phys. 653 307
[18] Fronczak A, Fronczak P and Holyst J A, 2004 Phys. Rev. E 70 056110
[19] Lovejoy W S and Loch C H, 2003 Soc. Netw. 25 333
[20] Holyst J A, Sienkiewicz J, Fronczak A, Fronczak P and Suchecki K, 2005 Phys. Rev. E 72 026108
[21] Dorogovtsev S N, Mendes J F F and Oliveira J G, 2006 Phys. Rev. E 73 056122
[22] Zhang Z Z, Zhou S G, Chen L C, Yin M and Guan J H, 2008 J. Phys. A: Math. Theor. 41 485102
[23] Albert R and Barabási A-L, 2002 Rev. Mod. Phys. 74 47
[24] Dorogovtsev S N and Mendes J F F, 2002 Adv. Phys. 51 1079
[25] Newman M E J, 2003 SIAM Rev. 45 167
[26] Boccaletti S, Latore V, Moreno Y, Chavez M and Hwanga D-U, 2006 Phys. Rep. 424 175
[27] Zhang Z Z, Zhou S G, Fang L J, Guan J H and Zhang Y C, 2007 Europhys. Lett. 79 38007
[28] Barabási A-L, Ravasz E and Vicsek T, 2001 Physica A 299 559
[29] Dorogovtsev S N, Goltsev A V and Mendes J F F, 2002 Phys. Rev. E 65 066122
[30] Jung S, Kim S and Kahng B, 2002 Phys. Rev. E 65 056101
[31] Ravasz E, Somera A L, Mongru D A, Oltvai Z N and Barabási A-L, 2002 Science 297 1551
[32] Andrade J S Jr, Herrmann H J, Andrade R F S and Silva L R da, 2005 Phys. Rev. Lett. 94 018702
[33] Hinczewski M, 2007 Phys. Rev. E 75 061104
[34] Boettcher S, Gonçalves B and Guclu H, 2008 J. Phys. A: Math. Theor. 41 252001
[35] Sierpinski W, 1915 Comptes Rendus (Paris) 160 302
[36] Hambly B M, 1992 Probab. Theory Relat. Fields 94 1
[37] Hutchinson S, 1981 Indiana Univ. Math. J. 30 713
[38] Hinczewski M and Berker A N, 2006 Phys. Rev. E 73 066126
[39] Bollt E M and ben-Avraham D, 2005 New J. Phys. 7 26

doi:10.1088/1742-5468/2009/02/P02034