

SCIENTIFIC REPORTS



OPEN

On the Fractality of Complex Networks: Covering Problem, Algorithms and Ahlfors Regularity

Lihong Wang^{1,2}, Qin Wang³, Lifeng Xi², Jin Chen⁴, Songjing Wang², Liulu Bao¹, Zhouyu Yu⁵ & Luming Zhao²

Received: 22 August 2016
Accepted: 20 December 2016
Published: 27 January 2017

In this paper, we revisit the fractality of complex network by investigating three dimensions with respect to minimum box-covering, minimum ball-covering and average volume of balls. The first two dimensions are calculated through the minimum box-covering problem and minimum ball-covering problem. For minimum ball-covering problem, we prove its NP-completeness and propose several heuristic algorithms on its feasible solution, and we also compare the performance of these algorithms. For the third dimension, we introduce the random ball-volume algorithm. We introduce the notion of Ahlfors regularity of networks and prove that above three dimensions are the same if networks are Ahlfors regular. We also provide a class of networks satisfying Ahlfors regularity.

Complex networks arise from natural and social phenomena such as the Internet, the protein interactions, the collaborations in research, and the social relationships. Readers are referred to Watts-Strogatz's¹ small-world network model and Barabási-Albert's² scale-free network model, and Newman's review³ and book⁴, etc.

In this paper, we revisit the fractality of complex network by investigating three dimensions d_B^5 , d_{ball}^6 and d_f^7 with respect to minimum box-covering, minimum ball-covering and average volume of balls. The compact box burning algorithm (CBB)^{8,9} and random ball-covering algorithm⁶ are proposed to calculate d_B and d_{ball} respectively. However the minimum box-covering problem and minimum ball-covering problem are NP-complete, which are proved rigorously in Theorem 1 and Proposition 2 respectively. The NP-completeness implies that the CBB algorithm and the random ball-covering algorithm do not have high performance, then we suggest some algorithms to improve the random ball-covering algorithm. For the third dimension d_f , we obtain an efficient algorithm: the random ball-volume algorithm. When do the three dimensions coincide? To answer this question, we introduce the notion of Ahlfors regularity of networks and prove that $d_B = d_{ball} = d_f$ (Theorem 2) if networks are Ahlfors regular. Then for Ahlfors regular networks, the random ball-volume algorithm is efficient to obtain the above three fractal dimensions.

Fractal dimensions and covering problems

Song, Havlin and Makse⁵ reveal that many real networks have self-similarity and fractality, and Gallos, Song, Havlin and Makse give a review of fractality of complex networks¹⁰. The algorithms to numerically calculate the fractal dimension of complex networks have been proposed: For example, the CBB algorithm^{8,9} is applied to calculate the fractal dimension of complex networks through the *minimum box-covering*; Kim, Goh, Kahng and Kim¹¹ improve the CBB algorithm to investigate the fractal scaling property in scale-free networks; Zhou, Jing and Sornette¹² propose the edge-covering box algorithm; Gao, Hu and Di⁶ give the *minimum ball-covering* approach to calculate the fractal dimension of complex networks.

Recall some notation. Considering a network as a graph $G = (V, E)$ equipped with geodesic distance d , we let an l -**box** A denote a subset of V such that the geodesic distance of any two points in the subset is less than l , an l -**ball** centered at x_0 the subset $\{y : d(y, x_0) < l\}$. Let N_l be the smallest number of l -boxes needed to cover V , and B_l the smallest number of l -balls needed to cover V . Suppose that

¹Faculty of Mechanical Engineering and Mechanics, Ningbo University, Ningbo 315211, P. R. China. ²Department of Mathematics, Ningbo University, Ningbo 315211, P. R. China. ³Department of Software Engineering, Zhejiang Wanli University, Ningbo 315100, P. R. China. ⁴College of Science, Huazhong Agricultural University, 430070, Wuhan, P. R. China. ⁵Faculty of Mechanical Engineering and Mechanics, Ningbo University, 315211, P. R. China. Correspondence and requests for materials should be addressed to L.X. (email: xilifengningbo@yahoo.com)

$$\#V/N_l \sim l^{d_B} \text{ and } \#V/B_l \sim l^{d_{ball}},$$

where d_B is the fractal dimension defined by Song, Havlin and Makse⁵, and d_{ball} is defined by Gao, Hu and Di⁶.

For **box-covering**, Song, Gallos, Havlin and Makse⁹ point out that the minimum l -box-covering problem is NP-complete for any $l \geq 2$. On the other hand, for **ball-covering**, which is *far from box-covering* in graph theory, we have

Theorem 1. *The minimum l -ball covering problem is NP-complete for any $l \geq 2$.*

Ball-covering algorithms

Due to the NP-completeness, for finding the feasible solution of minimum ball-covering problem, we can apply the usual *random ball-covering algorithm (RBC)*⁶: when l is fixed, in each time t , we randomly choose one node x_t in the vertex set V_{t-1} remained in time $(t-1)$, and obtain V_t by cutting all nodes in $B(x_t, l) = \{y : d(y, x_t) < l\}$.

In the RBC algorithm we give a random sorting for nodes in V_{t-1} and take the first node. Moreover, given some function $f : V \rightarrow \mathbb{N}$, we can sort these nodes according to the values of function f .

Given a function $f : V \rightarrow \mathbb{N}$, suppose we sort nodes according to values of f in nondecreasing order: If f is the degree function, we can obtain *degree-order ball-covering algorithm (DOBC)*; If $f(x) = \#B(x, l)$ and, we obtain *volume-order ball-covering algorithm (VOBC)*.

For a function $g : V \rightarrow \mathbb{N}$, assume we sort nodes according to values of g in nonincreasing order, we propose the following greedy algorithm:

- (1) Assume that $V_{t-1} = \{x_{t-1}^1, x_{t-1}^2, \dots\}$ such that $g(x_{t-1}^1) \geq g(x_{t-1}^2) \geq \dots$
- (2) Set $V_t = V_{t-1} \setminus B(x_{t-1}^1, l)$ and the sorting of nodes in V_t inherits from $V_0 = V$.

When $g(x) = \#B(x, l)$, we obtain the *volume-greedy ball-covering algorithm (VGBC)*. Let $g(x) = \deg(x)$, we have the *degree-greedy ball-covering algorithm (DGBC)*.

In the point of view on fractal geometry, the box dimension is independent of the geometric shapes of covering, such as ball or box. It is easy to check that $B_l \leq N_l \leq B_{l/2}$, hence $|d_B - d_{ball}| \leq \left| \frac{\log N_l}{\log l} - \frac{\log B_l}{\log l} \right| \leq \left| \frac{\log B_{l/2}}{\log l} - \frac{\log B_l}{\log l} \right| \approx \left| \frac{\log B_{l/2}}{\log 2 + \log(l/2)} - \frac{\log B_{l/2}}{\log(l/2)} \right| \approx \frac{\log 2}{\log l} \cdot d_{ball}$. By the above estimate, when the diameter of network is large enough to insure that l can be taken large enough, we have

Proposition 1. *The fractal dimensions d_B and d_{ball} w.r.t. the box covering and ball covering respectively are the same.*

However, for real networks with small-world effect, we can not take l large enough, and the upper bound $\frac{\log 2}{\log l} \cdot d_{ball}$ of error is not small enough. On the other hand, we only find the feasible solutions of minimum covering problems due to their NP-completeness. See the following example.

Example 1. *Through above 5 algorithms (Fig. 1), we calculate d_{ball} for the WWW network (Table 1).*

In Table 1, the value of the RBC algorithm is exactly the value $d_{ball} = 4.2$ by Gao, Hu and Di⁶. Note that Song, Havlin, and Makse⁵ obtain that $d_B = 4.1$.

For the WWW network, we also compare the above 5 algorithms (Fig. 2). It seems that the VGBC algorithm is the best and the performance of the RBC is the worst and close to the VOBC.

Random ball-volume algorithm

Based on Shanker's work¹³, Guo and Cai⁷ investigate the power law between the average volume of balls and the their radii. Given a network, let $p(l)$ be the average cardinality of nodes in a ball with radius l , suppose that

$$p(l) \sim l^{d_f}.$$

We call d_f the volume dimension. Please also see generalized volume dimension¹⁴ by Wei *et al.*

We will discuss the volume dimension d_f related to average ball-volume and propose the random ball-volume algorithm for networks. Compared with the minimum box-covering algorithm and the minimum ball-covering algorithm, we have the following algorithm to calculate the average volume of ball with size l approximately.

Random ball-volume algorithm (RBV) (for fixed size l):

- Step2.** Randomly take a node x in the network.
- Step2.** Calculate the volume $\nu(B(x, l))$.
- Step3.** Repeat the steps 1–2 and obtain average volume of random l -balls.

For the WWW network, using the RBV algorithm we obtain $d_f = 5.833$ (Fig. 3).

Ahlfors regularity of networks

Fractal geometry and fractal network have deep connection. We can generate complex network models from self-similar fractals. For example Andrade *et al.*¹⁵ and Zhou *et al.*¹⁶ discuss Apollonian networks generated from Apollonian fractal, Zhang *et al.*^{17–19} construct evolving networks modeled from Sierpinski gasket by taking the line segments as nodes. Besides Zhang *et al.*²⁰ construct the networks produced from Vicsek fractals, Liu *et al.*²¹ and Chen *et al.*²² explore some Koch networks related to Koch curves, Song *et al.*²³ study complex networks modeled on Platonic solids, Chen *et al.*²⁴ investigate networks generated by Sierpinski tetrahedron.

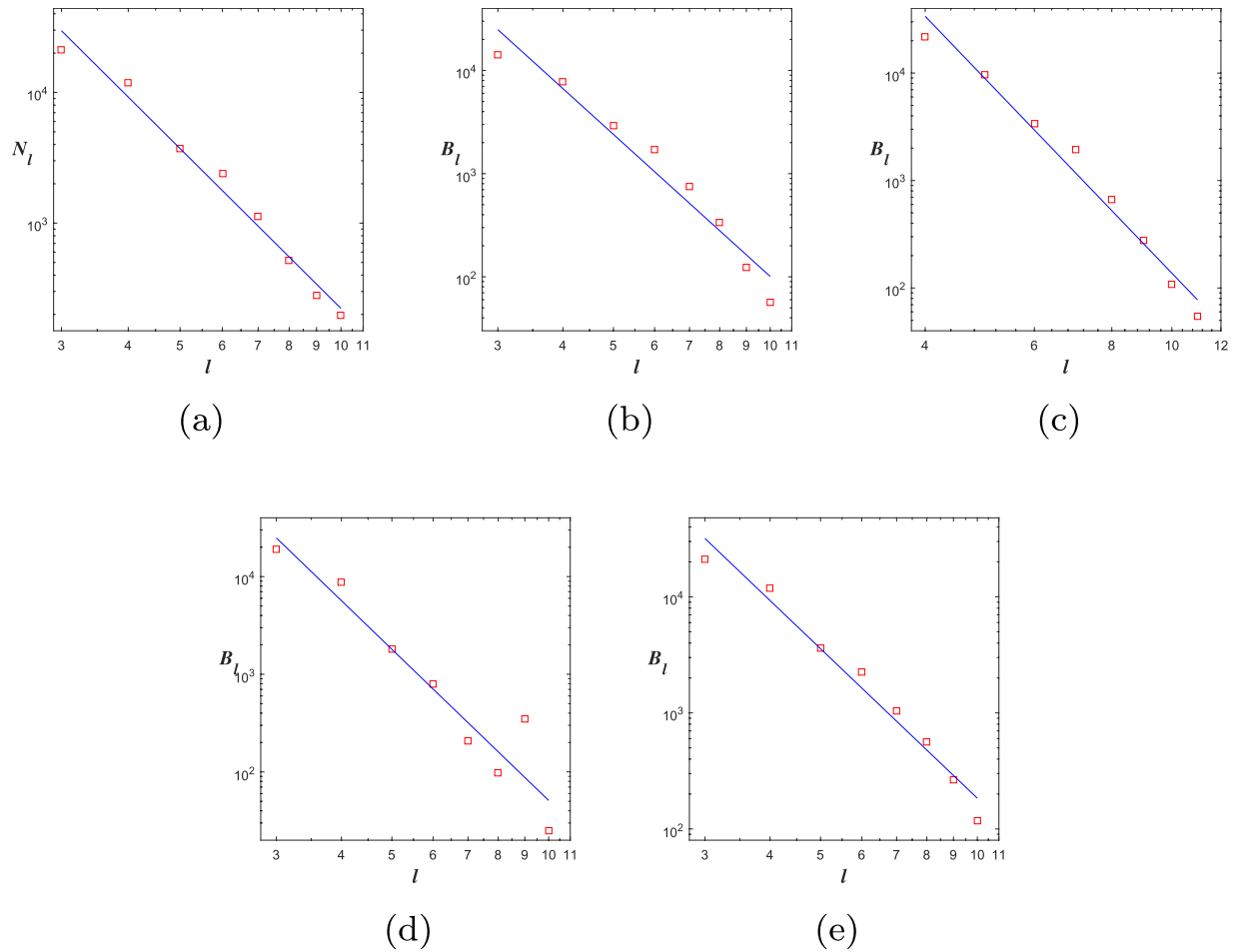


Figure 1. Slopes exist w.r.t. 5 algorithms for the WWW network: (a) RBC, (b) DGBC, (c) DOBC, (d) VGBC, (e) VOBC.

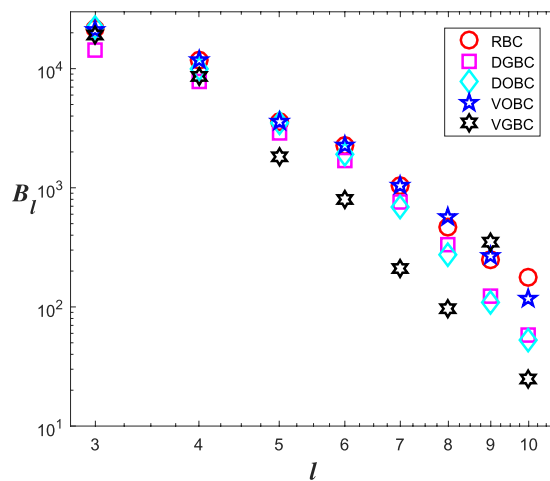


Figure 2. Comparison of 5 algorithms for the WWW network.

Algorithm	RBC	DGBC	DOBC	VGBC	VOBC
d_{ball}	4.1811	4.5693	5.0805	5.0950	4.2680

Table 1. d_{ball} w.r.t. 5 algorithms for the WWW network.

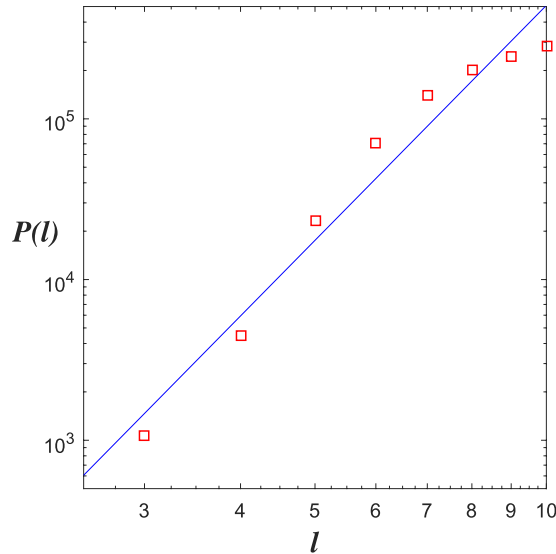


Figure 3. RBV for the WWW network.



Figure 4. Rule 1.

In this paper, we try to find out the farther connection between the fractal networks and fractal geometry. Recall some classical result on fractal dimension. We find out that many dimension results have measure versions. Suppose μ is a Borel (finite) measure supported on a compact subset E , denoted by $\text{spt } \mu \subset E$. For any $x \in E$, let the lower local dimension of μ at point x be defined by $\underline{\dim}_{loc} \mu(x) = \underline{\lim}_{r \rightarrow 0} \frac{\log \mu(B(x, r))}{\log r}$. A classical result^{25,26} on Hausdorff dimension $\dim_H(\cdot)$ is

$$\dim_H E = \sup \{s : \text{there exists a Borel measure } \mu \text{ with } \text{spt } \mu \subset E \text{ such that } \underline{\dim}_{loc} \mu(x) \geq s \text{ for } \mu\text{-almost all } x \in E\}.$$

That means for Hausdorff dimension, we have the corresponding measure version. When replacing $\underline{\dim}_{loc} \mu(x)$ by $\overline{\dim}_{loc} \mu(x) = \overline{\lim}_{r \rightarrow 0} \frac{\log \mu(B(x, r))}{\log r}$, we obtain packing dimension $\dim_p(\cdot)$ ^{25,26}. We always have $\dim_H E \leq \dim_p E \leq \overline{\dim}_B E$, where $\overline{\dim}_B$ is upper box dimension. A reasonable case is $\dim_H E = \dim_p E = \overline{\dim}_B E = s$ and there is a suitable measure μ such that $\lim_{r \rightarrow 0} \frac{\log \mu(B(x, r))}{\log r} = s$, or we can pose the **Ahlfors regularity** assumption on the measure

$$c^{-1}r^s \leq \mu(B(x, r)) \leq cr^s \text{ for all } x \in E \text{ and } r < \text{diam}(E),$$

where c is an independent constant.

We give a *natural measure* on a graph $G = (V, E)$. For $\Omega \subset V$, we let $\nu(\Omega)$ be the cardinality of Ω , which is called the *volume* of Ω . We say that $\{G_t\}_t$ is a family of growing networks, i.e., $G_t \subset G_{t+1}$, which means the node set of G_{t+1} contains node set of G_t , and neighbors of G_t are still neighbors of G_{t+1} . When $\{G_t\}_t$ is growing, we let $\nu_t(\Omega)$ denote cardinality of $\Omega \cap V_t$, where V_t is the node set of G_t .

Remark 1. When taking $\bar{\nu}(\Omega)$ as the sum of degrees of nodes in Ω , Wei et al.¹⁴ obtain the generalized volume dimension.

Definition 1. Given $s > 0$, if $\nu(B(x, r)) \sim r^s$, we call the network an Ahlfors s -regular network. When $\{G_t\}_t$ is growing, we call $\{G_t\}_t$ Ahlfors s -regular networks, if there is an independent constant c such that for all $x \in V_t$, $r < \text{diam}(V_t)$ and t ,

$$c^{-1}r^s \leq \nu_t(B(x, r)) \leq cr^s.$$

When the diameter of network is large enough, we have

Theorem 2. $d_f = d_{ball} = d_B = s$ if the network or growing networks are Ahlfors s -regular.

When the networks are regular, we can use RBV algorithm to obtain their fractal dimensions efficiently.

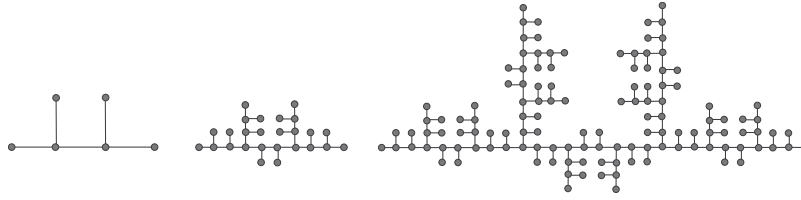


Figure 5. G_1, G_2, G_3 of growing trees w.r.t. rule 1.

Ahlfors regular trees

Now, we obtain a rule (rule 1) of generating Ahlfors s -regular networks and growing trees in Figs 4 and 5. We have $d_B = d_{ball} = d_f = 1.4\cdots$ in Fig. 6. By embedding the self-similar tree into the self-similar fractal in \mathbb{R}^2 , we find that the volume of ball in the tree is comparable with the (self-similar) measure of ball in plane, then we can obtain

Theorem 3. *The growing self-similar trees defined above are Ahlfors s -regular with $s = \log 5/\log 3$. Therefore, we have*

$$d_B = d_{ball} = d_f = \log 5/\log 3 = 1.4649\cdots.$$

We also have rule 2 and growing trees in Figs 7 and 8. For this self-similar tree with respect to rule 2, we have $d_B = d_{ball} = d_f = \log 4/\log 3$.

Fix an infinite sequence x_1, x_2, \cdots of 1 and 2 such that $\lim_{t \rightarrow \infty} \frac{\#\{x_i = 1 : i \leq t\}}{t} := p$ exists. We can construct a family of growing networks as follows by induction: for time t , we take rule 1 if $x_t = 1$, else take rule 2. For example, if the sequence is 211 \cdots , we obtain our growing networks G_1, G_2, G_3 as in Fig. 9. This is a family of deterministic growing networks.

Then we can generate a **Moran tree** with mixed rules. For this Moran tree without self-similarity, we have $d_B = d_{ball} = d_f = \frac{p \log 5 + (1-p) \log 4}{\log 3}$. We also obtain **random growing networks**, for each time t , we can choose rule 1 in probability p and rule 2 in probability $1-p$.

The rest of paper is organized as follows. Section 2 is devoted to the rigorous proofs on the NP-completeness of minimum ball-covering problem (Theorem 2) and minimum box-covering problem (Proposition 2). Section 3 is the preliminary on the Ahlfors regularity of fractal geometry, including covering inequality and self-similar fractal. In this section, we also recall the fact that the open set condition of self-similar fractal implies the Ahlfors regularity of fractal measure. Replacing the fractal measures by the cardinalities of subsets of networks, we obtain the Ahlfors regularity of networks. In Section 4, we prove Theorem 2 by using covering inequality shown in Section 2, and obtain Ahlfors regularity of a class of self-similar network (Theorem 3) by constructing bilipschitz mappings from a self-similar fractal, satisfying the open set condition, to self-similar networks, and estimating the cardinalities of balls of graph from the Ahlfors regularity of the fractal measure.

NP-completeness of minimum covering problems

Recall some notation of computer science. For an alphabet Σ , let Σ^* be the set of finite strings of elements of Σ , and Π the class of functions from Σ^* into Σ^* defined by one-tape Turing machine which operate in polynomial time.

Definition 2. *Let L and M be languages. Then $L \propto M$ (L is reduced to M) if there is a function $f \in \Pi$ such that $f(x) \in M \Leftrightarrow x \in L$. We say that some language $M \in NP$ is **NP-complete**, if $L \propto M$ for all $L \in NP$.*

The concept of NP-completeness was introduced in 1971 by Cook²⁷. In Cook's theorem, he proved that the **Boolean satisfiability problem** is NP-complete.

In 1972, Karp²⁸ proved that several other problems were also NP-complete. For example, we give the following two in Karp's 21 NP-complete problems.

(1) **Clique covering problem**

Input: graph $G = (V, E)$, positive integer k

Property: V is the union of k or fewer cliques, where a *clique* is a subset of vertices of G such that its induced subgraph is complete.

(2) **Set covering problem**

Input: universe U and a family S of subsets of U , positive integer k

Property: there is a set covering of size k or less, where a *set covering* is a subfamily $C \subseteq S$ of sets whose union is U . In 1992, Kann²⁹ proved that the set covering problem, which is NP-complete, can be reduced to the following dominating set problem (hence it's also NP-complete).

(3) **Dominating set problem**

Input: graph $G = (V, E)$, positive integer k

Property: there is a dominating set of k or fewer nodes, where a *dominating set* is a subset D of V such that every vertex not in D is adjacent to at least one member of D .

In this section, we will show the following two problems are NP-completes.

(4) **l -ball-covering problem**

Input: graph $G = (V, E)$, positive integer k

Property: V is the union of k or fewer l -balls.

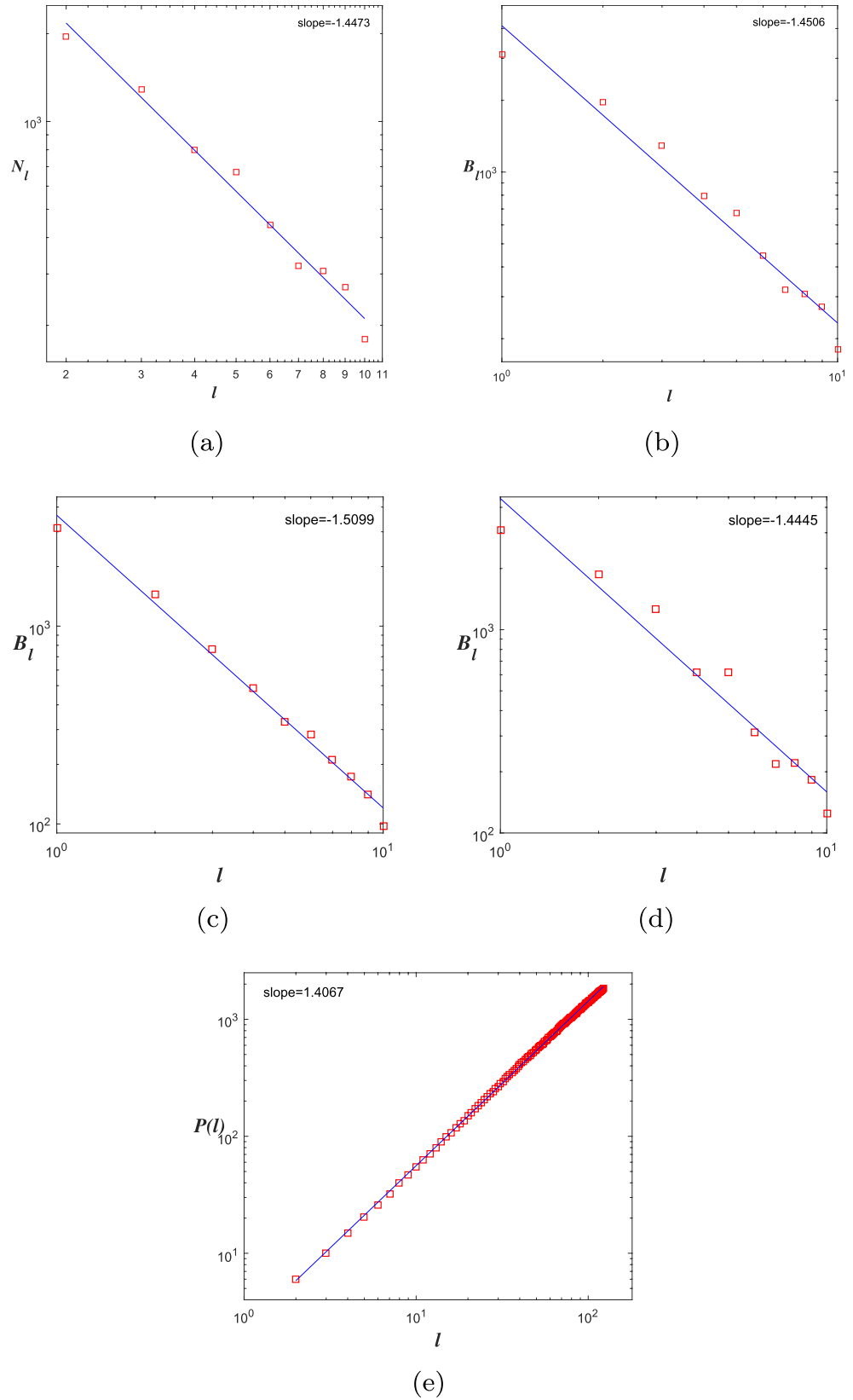


Figure 6. Fractal dimensions of G_5 : (a) CBB, (b) RBC, (c) DGBC, (d) DOBC, (e) RBV.



Figure 7. Rule 2.

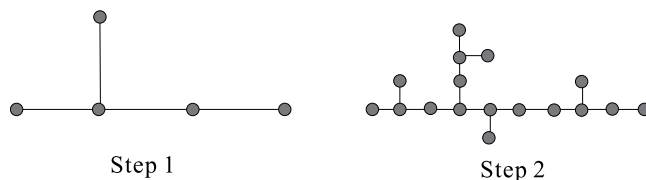


Figure 8. G_1, G_2 of growing trees w.r.t. rule 2.

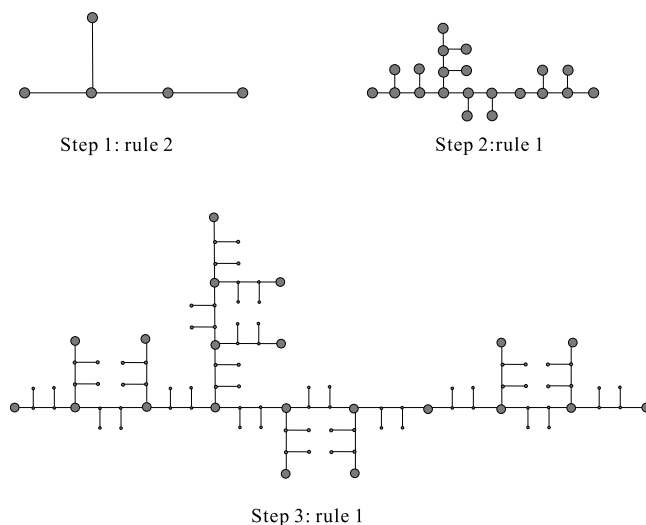


Figure 9. The first three steps according to an infinite sequence 211...

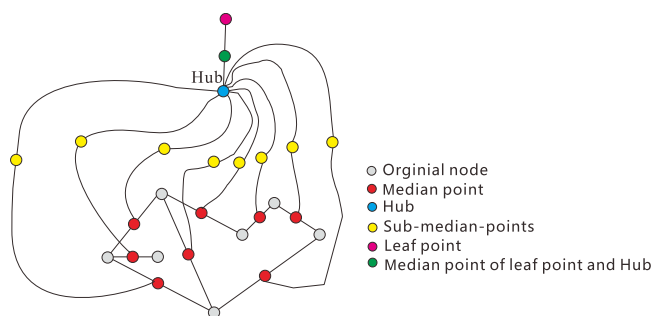


Figure 10. The reduction process for $l=3$.

(5) l -box-covering problem

Input: graph $G=(V, E)$, positive integer k
 Property: V is the union of k or fewer l -boxes.

Proof of Theorem 1. If $l=2$, then l -ball-covering problem is exactly the dominating set problem, which is NP-complete. If $l=3$, given a undirected graph $G=(V, E)$ as in Fig. 10, we construct a new graph $\bar{G}=(\bar{V}, \bar{E})$ in polynomial time w.r.t. the size of G .

Step I. For any $x \sim y \in V$, we insert a median point z (in red) in the edge $(x, y) \in E$ with degree 2 in \bar{G} , i.e., in \bar{G} we have $x \sim z, z \sim y$ and x, y are not neighbors in \bar{G} .

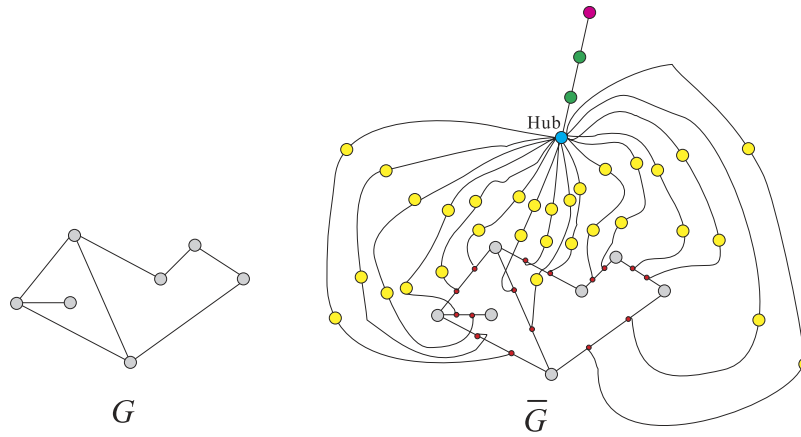


Figure 11. The reduction process for $l=4$.

- Step II.** We add a Hub (in blue) to connect all median points.
- Step III.** Insert sub-median-point (in yellow) for every edge between one median point (in Step I) and Hub.
- Step IV.** We construct a leaf node (in pink) and the median point (in green) between leaf node and the Hub.

We have the following

Claim 1. *There is a dominating set of k or fewer nodes in G if and only if \bar{V} is the union of $(k+1)$ or fewer 3-balls.*

To verify this claim, we notice the following facts.

- (a) For any nodes $x \sim y \in V$, in \bar{G} their geodesic distance $d_{\bar{G}}(x, y) = 2$.
- (b) The subset of all nodes not in V is a 3-ball centered at the Hub.
- (c) The geodesic distance between the pink node and any node in V is 5, that means any 3-ball can not contain the pink node and any node of V simultaneously.
- (d) For any 3-ball D with $D \cap V \neq \emptyset$, we can find a node u in $D \cap V$ such that

$$D \cap V \subset \{v \in V : d_G(u, v) \leq 1\}.$$

Suppose $\{x_i\}_{i=1}^s$ is the minimum dominating set of G and there is a minimum 3-ball covering $\{D_i\}_{i=1}^t$ of \bar{G} . We only need to show that

$$t = s + 1. \tag{1}$$

In fact, we have $V \subset \cup_{i=1}^s \{y \in V : d_G(x_i, y) \leq 1\}$. It follows from the fact (a) that

$$\{y \in V : d_G(x_i, y) \leq 1\} \subset \{u \in \bar{V} : d_{\bar{G}}(x_i, u) \leq 2\}$$

for any $i \leq s$. Applying the fact (b), we see that there exists a 3-ball covering with $(s+1)$ balls. Hence

$$t \leq s + 1. \tag{2}$$

On the other hand, considering the minimum 3-ball covering $\{D_i\}_{i=1}^t$, by fact (d), we obtain a dominating set $\{u_i\}_{i \in \bar{\Lambda}}$ of G , where $\bar{\Lambda} = \{i : D_i \cap V \neq \emptyset\}$. Therefore, $\#\bar{\Lambda} \geq s$. Since the pink point must belong to some ball D_{i_0} , by fact (c), we have $i_0 \notin \bar{\Lambda}$. Therefore we have

$$t \geq \#\bar{\Lambda} + 1 \geq s + 1. \tag{3}$$

Then (1) follows from (2) and (3). Then Theorem 1 is proved for $l=3$.

For $l \geq 4$, we have the similar construction during reduction. In fact, we insert $(l-2)$ median points into each edge of G , add a Hub to connect all median points, insert $(l-2)$ sub-median-point for every edge between one median point and Hub. Finally, we construct the leaf node and connect it to the Hub, insert $(l-2)$ the median point between leaf node and the Hub. See Fig. 11 for $l=4$.

Remark 2. *To prove one problem is NP-complete, we always find a reduction from a known NP-complete problem to our problem. On the other hand, we can always construct a reduction from our (NP) problem to a known NP-complete problem due to the definition of NP-completeness.*

We give a proof of the following fact which is pointed out by Song, Gallos, Havlin and Makse⁹.

Proposition 2. *For any fixed size l , the l -box-covering problem is NP-complete.*

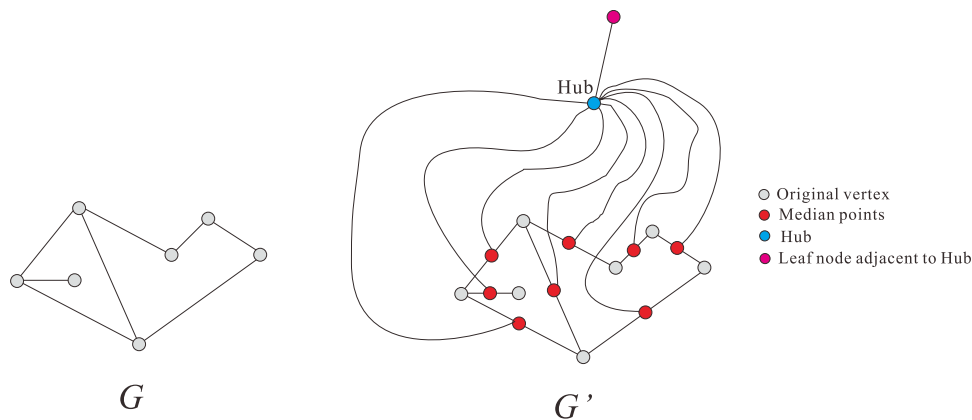


Figure 12. The reduction process for $l=3$.

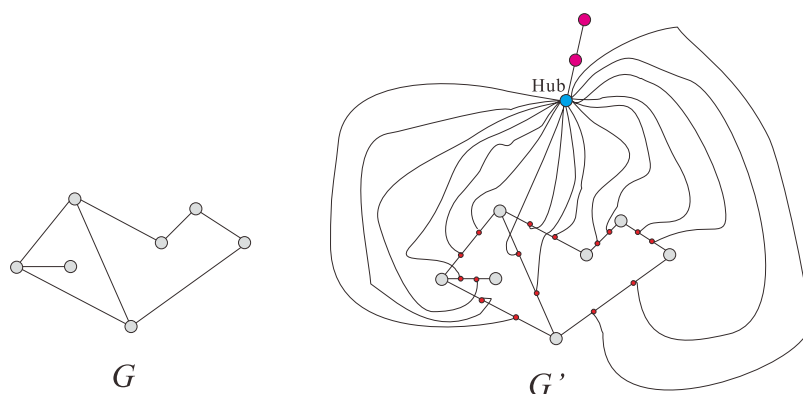


Figure 13. The reduction process for $l=4$.

Proof. If $l=2$, l -box-covering problem is exactly the clique covering problem, which is NP-complete.

If $l=3$, given a undirected graph $G=(V, E)$, as in Fig. 12, we construct a new graph $G'=(V', E')$ in polynomial time with respect to the size of G .

- Step 1.** For any $x \sim y \in V$, we insert a median point z (in red) in the edge (x, y) with degree 2, i.e., $x \sim z, z \sim y$ and x, y are not neighbors in G' .
- Step 2.** We add a Hub (in blue) to connect all median points.
- Step 3.** We construct a leaf node (in pink) adjacent to the Hub.

We have the following

Claim 2. V is the union of k or fewer cliques if and only if V' is the union of $(k+1)$ or fewer 3-boxes.

To verify this claim, we notice the following facts.

1. For any nodes $x \sim y \in V$, in G' their geodesic distance $d_{G'}(x, y) = 2$.
2. The subset of nodes not in V is a 3-box.
3. The geodesic distance between leaf node (in pink) and any node in V is 3.

Suppose $\{A_i\}_{i=1}^s$ is a family of cliques of G such that s is minimal one. Suppose there is a minimum 3-box covering $\{B_i\}_{i=1}^t$ of G' . We only need to show that

$$t = s + 1. \tag{4}$$

In fact, we have $V \subset \cup_{i=1}^s A_i$. It follows from the fact (i) that A_i is a 3-box in G' for any $i \leq s$. Applying the fact (ii), we see that there exists a 3-box covering with $(s+1)$ boxes. Hence

$$t \leq s + 1. \tag{5}$$

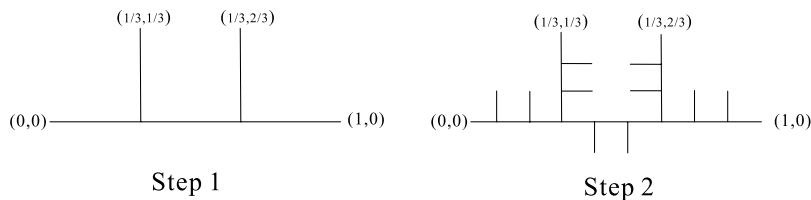


Figure 14. The first two steps of self-similar fractal (model 1).

On the other hand, it follows from fact (i) that $\{B_i \cap V\}_{i \in \Lambda}$ is a family of cliques in G where $\Lambda = \{i : B_i \cap V \neq \emptyset\}$. Therefore, $\#\Lambda \geq s$. We also notice that if the pink point belongs to some B_{i_0} , by fact (iii), we have $i_0 \notin \Lambda$. Therefore we have

$$t \geq \#\Lambda + 1 \geq s + 1. \tag{6}$$

Then (4) follows from (5) and (6).

For $l \geq 4$, we have the similar construction during reduction. See Fig. 13 for $l = 4$. □

Covering inequality, self-similar fractal and Moran fractal

Covering and packing on metric space. Given a compact metric space (X, d) , let a δ -ball centered at x_0 be an open ball $B(x, \delta) = \{y : d(y, x_0) < \delta\}$, and a δ -cube a cube of Euclidean space with side length δ , a δ -box B is a subspace of X such that its diameter less than δ , i.e., $d(x, y) < \delta$ for all $x, y \in B$. Denote

- B_δ : the smallest number of δ -balls needed to cover X ,
- N_δ : the smallest number of δ -boxes needed to cover X ,
- M_δ : the smallest number of δ -cubes needed to cover X ($\subset \mathbb{R}^n$),
- P_δ : the maximal number of δ -balls pairwise disjoint,

We recall an elementary inequality²⁶ which is important in this paper. We give the proof for the self-containedness of this paper.

Lemma 1. $B_{2\delta} \leq P_\delta \leq B_{\delta/2}$.

Proof. Suppose $\{B(x_i, \delta)\}_{i=1}^{P_\delta}$ is a packing family of δ -balls, we conclude that $X \subset \cup_{i=1}^{P_\delta} B(x_i, 2\delta)$. Otherwise, suppose $y \notin \cup_{i=1}^{P_\delta} B(x_i, 2\delta)$, for any $y_i \in B(x_i, \delta)$, we have $d(y, y_i) \geq d(y, x_i) - d(y_i, x_i) \geq 2\delta - \delta = \delta$. That means $B(y, \delta) \cap B(x_i, \delta)$ is empty for any i . Now, we obtain a new packing family of $\{B(y, \delta)\} \cup \{B(x_i, \delta)\}_{i=1}^{P_\delta}$, which is a contradiction. Therefore, we have $X \subset \cup_{i=1}^{P_\delta} B(x_i, 2\delta)$, and thus we have $P_\delta \geq B_{2\delta}$.

Assume $\{B(x_i, \delta)\}_{i=1}^{P_\delta}$ is a packing family of δ -balls, then $d(x_i, x_j) \geq \delta$ for all $i \neq j$. Notice that on Euclidean space, we have $d(x_i, x_j) \geq 2\delta$ for all $i \neq j$. Suppose there is a minimum covering of $\delta/2$ -balls $\{B(y_j, \delta/2)\}_{j=1}^{B_{\delta/2}}$. Now, every $\delta/2$ -ball contains at most one points in $\{x_i\}_{i=1}^{P_\delta}$ since the diameter of a $\delta/2$ -ball is less than δ and $d(x_i, x_j) \geq \delta$ for all $i \neq j$. On the other hand, every x_i must be contained in some $\delta/2$ -ball. Therefore, we obtain $P_\delta \leq B_{\delta/2}$. □

We also have

$$B_\delta \leq N_\delta \leq B_{\delta/2} \text{ and } M_\delta \geq B_{\sqrt{n}\delta} \geq M_{2\sqrt{n}\delta} \text{ (on } \mathbb{R}^n \text{)}.$$

By the above inequalities, the classical result^{25,26} on box dimension is that

$$\dim_B X = \lim_{\delta \rightarrow 0} \frac{\log N_\delta}{-\log \delta} = \lim_{\delta \rightarrow 0} \frac{\log P_\delta}{-\log \delta} = \lim_{\delta \rightarrow 0} \frac{\log B_\delta}{-\log \delta} = \lim_{\delta \rightarrow 0} \frac{\log M_\delta}{-\log \delta} \text{ (on } \mathbb{R}^n \text{)}.$$

In fact, in the above formula, we take upper box dimension $\overline{\dim}$ or lower box dimension $\underline{\dim}$ when the limit does not exist.

Self-similar set on Euclidean space. Let $K = \cup_{i=1}^m S_i(K)$ be a self-similar set³⁰ on a Euclidean space \mathbb{R}^n , where S_i is a similarity with ratio r_i , i.e., $|S_i x - S_i y| = r_i |x - y|$ for all $x, y \in \mathbb{R}^n$. In fact, $S_i(x) = r_i R_i x + b_i$ where $r_i \in (0, 1)$, $b_i \in \mathbb{R}^n$ and R_i is orthogonal. That means any similarity is the compositions of homothety, translation and orthogonal transformation.

We say that the open set condition (OSC) holds if there exists a non-empty open set V such that

$$\cup_{i=1}^m S_i(V) \subset V \text{ and } S_i(V) \cap S_j(V) = \emptyset \text{ for all } i \neq j. \tag{7}$$

Let $(r_1)^s + \dots + (r_m)^s = 1$ and the probability vector $(p_1, \dots, p_m) = (r_1^s, \dots, r_m^s)$. According to ref. 30, there is a unique Borel measure μ (self-similar measure) satisfying $\mu = \sum_{i=1}^m p_i (\mu \circ S_i^{-1})$. When the OSC holds, Hutchinson³⁰ obtained that $\dim_H K = \dim_B K = s$, and there is a constant $C \geq 1$ such that for all $x \in K$ and $r \leq |K|$ (the diameter of K),

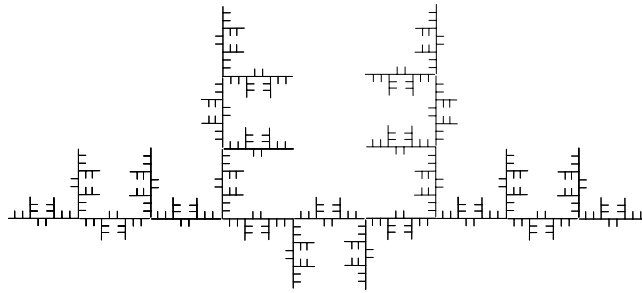


Figure 15. Step 4 of self-similar fractal (model 1).

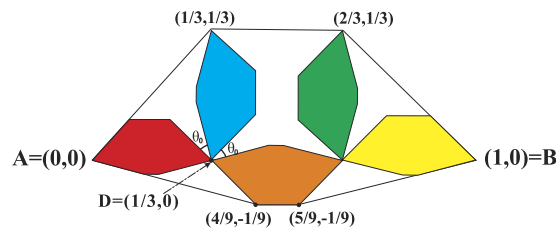


Figure 16. OSC holds.

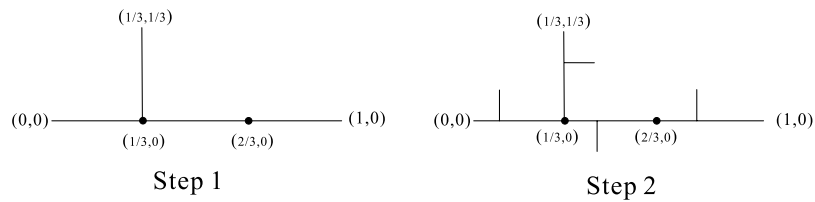


Figure 17. The first two steps of self-similar fractal of model 2.

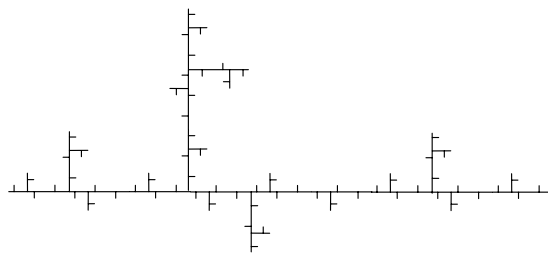


Figure 18. Step 4 of self-similar fractal of model 2.

$$C^{-1}r^s \leq \mu(B(x, r)) \leq Cr^s. \tag{8}$$

A compact set E is said to be Ahlfors s -regular²⁶, if there is a Borel measure μ supported on E satisfying (8). That means the self-similar set satisfying the OSC is Ahlfors regular.

Self-similar fractals. We introduce a special self-similar fractal on \mathbb{R}^2 (Figs 14 and 15). Let

$$\begin{aligned} S_1\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x/3 \\ y/3 \end{pmatrix}, & S_2\begin{pmatrix} x \\ y \end{pmatrix} &= T_2\begin{pmatrix} x/3 \\ y/3 \end{pmatrix} + \begin{pmatrix} 1/3 \\ 0 \end{pmatrix}, \\ S_3\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x/3 \\ y/3 \end{pmatrix} + \begin{pmatrix} 2/3 \\ 0 \end{pmatrix}, & S_4\begin{pmatrix} x \\ y \end{pmatrix} &= T_4\begin{pmatrix} x/3 \\ y/3 \end{pmatrix} + \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}, \\ S_5\begin{pmatrix} x \\ y \end{pmatrix} &= T_5\begin{pmatrix} x/3 \\ y/3 \end{pmatrix} + \begin{pmatrix} 2/3 \\ 0 \end{pmatrix}, \end{aligned}$$

where orthogonal matrixes $T_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, T_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, T_5 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

Let V be the interior of polygon with vertexes $(0, 0), (1/3, 1/3), (2/3, 1/3), (1, 0), (4/9, -1/9)$ and $(5/9, -1/9)$. Then (7) holds for $m = 5$ (Fig. 16).

Taking $\{S_i\}_{i=1}^4$, we give a self-similar fractal of model 2 (Figs 17 and 18).

$$\begin{aligned} S_1\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x/3 \\ y/3 \end{pmatrix}, & S_2\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x/3 \\ y/3 \end{pmatrix} + \begin{pmatrix} 1/3 \\ 0 \end{pmatrix}, \\ S_3\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x/3 \\ y/3 \end{pmatrix} + \begin{pmatrix} 2/3 \\ 0 \end{pmatrix}, & S_4\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} x/3 \\ y/3 \end{pmatrix} + \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}. \end{aligned}$$

Then the OSC also holds. Let E_1, E_2 be the self-tree of models 1 and 2 respectively. Then

$$\begin{aligned} \dim_H E_1 &= \dim_B E_1 = \log 5 / \log 3, \\ \dim_H E_2 &= \dim_B E_2 = \log 4 / \log 3. \end{aligned}$$

Moran fractal and random fractal. Fix an infinite sequence x_1, x_2, \dots of 1 and 2, we can generate a Moran fractal with mixed model. If $x_i = 1$ then we take model 1, else we take model 2. Let

$$p_t = \frac{\#\{x_i = 1 : i \leq t\}}{t}.$$

If $\lim_{t \rightarrow \infty} p_t := p$ exists, then the corresponding fractal has fractal dimension

$$\frac{p \log 5 + (1 - p) \log 4}{\log 3}.$$

An interesting fact is that this is a deterministic fractal without self-similarity. This is a Moran fractals³¹.

An alternative is a random fractal such that for each time t , we can choose model 1 in probability p and model 2 in probability $1 - p$. Then we obtain the above dimension almost surely.

Ahlfors regularity of networks

Proof of Theorem 2. By the definition of Ahlfors regularity, we have $d_f = s$.

Suppose $c^{-1}r^s \leq \nu(B(x, r)) \leq cr^s$. Since the network is covered by B_l balls of radius l , that means

$$\#V \leq \sum_{i=1}^{B_l} \nu(B(x_i, l)) \leq c(B_l \cdot l^s).$$

On the other hand, we have $P_{l/2}$ packing balls of radius $l/2$, which implies

$$\#V \geq \sum_{i=1}^{P_{l/2}} \nu(B(y_i, l/2)) \geq c^{-1} \left(P_{l/2} \cdot \left(\frac{l}{2} \right)^s \right).$$

That means

$$B_l \geq (c^{-1} \#V) l^{-s} \text{ and } B_l \leq P_{l/2} \leq (2^s c \#V) l^{-s},$$

here we use the inequality $B_{2\delta} \leq P_\delta \leq B_{\delta/2}$ in Lemma 1. Therefore,

$$(c^{-1} \#V) l^{-s} \leq B_l \leq (2^s c \#V) l^{-s},$$

which implies $\#V/B_l \sim l^s$, i.e., $d_{ball} = d_B = s$.

Proof of Theorem 3. Let $A = (0, 0)$ and $B = (1, 0)$. Let $V_t = \{S_{i_1} S_{i_2} \dots S_{i_t}(x) : 1 \leq i_j \leq 5 \text{ and } x \in \{A, B\}\}$.

Remark 3. One node may have distinct codings $(i_1 \dots i_t, x)$ and $(j_1 \dots j_t, x')$ if $S_{i_1} S_{i_2} \dots S_{i_t}(x) = S_{j_1} S_{j_2} \dots S_{j_t}(x')$. We also notice that each node has three codings at most.

Two different nodes $x, y \in V_t$ are neighbors if and only if there exists a word $i_1 \dots i_t$ such that

$$x = S_{i_1} S_{i_2} \cdots S_{i_t}(a) \text{ and } y = S_{i_1} S_{i_2} \cdots S_{i_t}(b) \text{ with } \{a, b\} = \{A, B\}.$$

Let d_t be the geodesic distance on G_t .

We denote $a_i \approx b_i$ if there is a constant $d > 0$ independent of the index i such that $d^{-1}b_i \leq a_i \leq db_i$.

Now, we will prove the following important

Lemma 2. *There is a constant $c > 0$ independent of t such that*

$$c^{-1}|y_1 - y_2| \leq \frac{d_t(y_1, y_2)}{3^t} \leq c|y_1 - y_2| \text{ for all } y_1, y_2 \in V_t. \tag{9}$$

Proof. Suppose

$$y_1 = S_{i_1} S_{i_2} \cdots S_{i_k} S_{i_{k+1}} \cdots S_{i_t}(a) \text{ and } y_2 = S_{i_1} S_{i_2} \cdots S_{i_k} S_{i'_{k+1}} \cdots S_{i_t}(a')$$

where $i_{k+1} \neq i'_{k+1}$. Notice that

$$d_t(y_1, y_2) = d_{t-k}(y'_1, y'_2)$$

where $y'_i = (S_{i_1} S_{i_2} \cdots S_{i_k})^{-1}(y_i)$. Without loss of generality, we assume that $i_1 \neq i'_1$.

Case 1. If $S_{i_1}(E) \cap S_{i'_1}(E)$ is empty, then $|y_1 - y_2| \leq \text{diam}(E) \leq \sqrt{2}$ and $|y_1 - y_2| \geq 1/9$, and

$$\frac{3^t}{3} = d_t((1/3, 0), (2/3, 0)) \leq d_t(y_1, y_2) \leq 3^t.$$

Then (9) follows in this case. □

Case 2. If $S_{i_1}(E) \cap S_{i'_1}(E)$ is non-empty, we may assume that $S_{i_1}(E) \cap S_{i'_1}(E) = (1/3, 0)$ without loss of generality.

For $D = (1/3, 0)$, let $\theta = \angle y_1 D y_2$. Then there exists $\theta_0 = \arcsin \frac{3}{\sqrt{34}} \in (0, \pi/2)$ (Fig. 16) such that $\theta \geq \theta_0 (> 0)$. Now,

$$\begin{aligned} |y_1 - y_2|^2 &= |y_1 - D|^2 + |y_2 - D|^2 - 2 \cos \theta (|y_1 - D| \cdot |y_2 - D|) \\ &\geq |y_1 - D|^2 + |y_2 - D|^2 - 2 \cos \theta_0 (|y_1 - D| \cdot |y_2 - D|) \\ &\geq \cos \theta_0 (|y_1 - D| - |y_2 - D|)^2 + (1 - \cos \theta_0) (|y_1 - D|^2 + |y_2 - D|^2) \\ &\geq \left(\frac{1 - \cos \theta_0}{2} \right) (|y_1 - D| + |y_2 - D|)^2 \\ &= \frac{1 - \frac{5}{\sqrt{34}}}{2} (|y_1 - D| + |y_2 - D|)^2, \end{aligned}$$

which implies

$$|y_1 - y_2| \geq 0.26(|y_1 - D| + |y_2 - D|).$$

We also have $|y_1 - y_2| \leq |y_1 - D| + |y_2 - D|$. Therefore, we have

$$|y_1 - y_2| \approx |y_1 - D| + |y_2 - D|. \tag{10}$$

On the other hand,

$$d_t(y_1, y_2) = d_t(y_1, D) + d_t(y_2, D) \tag{11}$$

by the tree structure. It follows from (10) and (11) that we only need to verify (9) for the pairs (y_1, D) and (D, y_2) . By the self-similarity, now we only need to prove the case when $y_1 \in \{A, B\}$.

Without loss of generality, let $y_1 = A$ and $y_2 = S_1 \cdots S_{i_k} S_{i'_k} \cdots S_{i_t}(a)$ where $i_k \neq i'_k$. Then

$$3^{t-k} \leq d_t(A, y_2) \leq 3^{t-k+1}$$

and

$$3^{-k} \leq |A - y_2| \leq 3^{-(k-1)} \text{diam}(E),$$

then (9) follows.

Since the OSC holds, then the self-similar measure μ with respect to the vector $(1/5, 1/5, 1/5, 1/5, 1/5)$ is Ahlfors s -regular for $s = \log 5 / \log 3$.

It follows from the above lemma and Remark 3 that

$$\frac{\#\{y' \in G_t : d_t(y, y') < l\}}{\#V_t} \approx \mu \left\{ \left\{ z \in \mathbb{R}^2 : |z - y| < \frac{l}{3^t} \right\} \right\} \approx \frac{l^s}{5^t},$$

where $\#V_t = 5^t + 1$. Therefore, we have

$$\#\{y' \in G_t : d_t(y, y') < l\} \approx l^s. \quad (12)$$

Notice that the constant in (12) is independent of t . Now, the growing networks $\{G_t\}$ are Ahlfors s -regular.

Conclusion

We focus on the NP-completeness of minimum ball-covering problem, propose some heuristic ball-covering algorithms such as GOBC, GDBC, VOBC and VGBC, and compare these algorithms with usual RBC algorithm. Inspired by the notion of measure on fractal, a natural measure on the finite graph is obtained such that the measure of every subset is the cardinality of subset. Based on this measure, we revisit the volume dimension d_f and propose the random ball-volume algorithm, which has performance better than the above five minimum covering algorithms due to the NP-completeness. Applying the notion of Ahlfors regularity from fractal geometry, we prove that $d_B = d_{ball} = d_f = s$ if the network is Ahlfors s -regular. Finally, we investigate the Ahlfors regularity of a class of self-similar trees and random trees which come from the self-similar fractals and Moran fractals respectively. Although we only prove Theorem 3 for self-similar tree of model 1, but our approach can be applied to many self-similar trees, Moran tree and random trees. Essentially, our approach is to embed our networks into a self-similar (or Moran) fractal (on Euclidean space) satisfying the open set condition, using the Ahlfors regularity of corresponding self-similar (or Moran) measure, we can estimate the volume of balls in networks.

References

- Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
- Barabási, A. L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
- Newman, M. E. J. The structure and function of complex networks. *Siam Review* **45**, 167–256 (2003).
- Newman, M. E. J. *Networks: An Introduction*. Oxford, Oxford University Press (2010).
- Song, C., Havlin, S. & Makse, H. A. Self-similarity of complex networks. *Nature* **433**, 392–395 (2005).
- Gao, L., Hu, Y. Q. & Di, Z. R. Accuracy of the ball-covering approach for fractal dimensions of complex networks and a rank-driven algorithm. *Physical Review E* **78**, 046109 (2008).
- Guo, L. & Cai, X. The fractal dimensions of complex networks. *Chin. Phys. Lett.* **26**, 088901 (2009).
- Song, C., Havlin, S. & Makse, H. A. Origins of fractality in the growth of complex networks. *Nature Physics* **2**, 275–281 (2006).
- Song, C., Gallos, L. K., Havlin, S. & Makse, H. A. How to calculate the fractal dimension of a complex network: the box covering algorithm. *J. Stat. Mech.: Theor. Exp.* **3**, 4673–4680 (2007).
- Gallos, L. K., Song, C. M., Havlin, S. & Makse, H. A. A review of fractality and self-similarity in complex networks. *Physica A* **386**, 686–691 (2007).
- Kim, J. S., Goh, K. I., Kahng, B. & Kim, D. A box-covering algorithm for fractal scaling in scale-free networks. *Chaos* **17**, 026116 (2007).
- Zhou, W. X., Jing, Z. Q. & Sornette, D. Exploring self-similarity of complex cellular networks: The edge-covering method with simulated annealing and log-periodic sampling. *Physica A* **375**, 741–752 (2007).
- Shanker, O. Defining dimension on a complex network. *Mod. Phys. Lett. B* **21**, 321–326 (2007).
- Wei, D., Wei, B., Zhang, H., Gao, C. & Deng, Y. A generalized volume dimension of complex networks. *J. Stat. Mech.: Theor. Exp.* **10**, P10039 (2014).
- Andrade, J. S. Jr., Herrmann, H. J., Andrade, R. F. & Da Silva, L. R. Apollonian networks: Simultaneously scale-free, small world, Euclidean, space filling, and with matching graphs. *Phys. Rev. Lett.* **94**, 018702 (2005).
- Zhou, T., Yan, G. & Wang, B. Maximal planar networks with large clustering coefficient and power-law degree distribution. *Phys. Rev. E* **71**, 046141 (2005).
- Zhang, Z., Zhou, S., Fang, L., Guan, J. & Zhang, Y. Maximal planar scale-free Sierpinski networks with small-world effect and power law strength-degree correlation. *Europhysics Letters* **79**, 38007 (2007).
- Zhang, Z., Zhou, S., Su, Z., Zou, T. & Guan, J. Random Sierpinski network with scale-free small-world and modular structure. *Eur. Phys. J. B* **65**, 141–147 (2008).
- Guan, J., Wu, Y., Zhang, Z., Zhou, S. & Wu, Y. A unified model for Sierpinski networks with scale-free scaling and small-world effect. *Physica A* **388**, 2571–2578 (2009).
- Zhang, Z., Zhou, S., Chen, L., Yin, M. & Guan, J. The exact solution of the mean geodesic distance for Vicsek fractals. *J. Phys. A: Math. Theor.* **41**, 485102 (2008).
- Liu, J. & Kong, X. Establishment and structure properties of the scale-free Koch network. *Acta Phys. Sinica* **59**, 2244–2249 (2010).
- Chen, R., Fu, X. & Wu, Q. On topological properties of the octahedral Koch network. *Physica A* **391**, 880–886 (2012).
- Song, W. M., Di Matteo, T. & Aste, T. Building complex networks with Platonic solids. *Phys. Rev. E* **85**, 046115 (2012).
- Chen, J., Gao, F., Le, A., Xi, L. & Yin, S. A small-world and scale-free network generated by Sierpinski tetrahedron. *Fractals* **24**, 1650001 (2016).
- Falconer, K. J. *Fractal geometry: mathematical foundations and applications*. Chichester, John Wiley & Sons Ltd. (1990).
- Mattila, P. *Geometry of sets and measures in Euclidean spaces*. Cambridge, Cambridge University Press (1995).
- Cook, S. A. The complexity of theorem proving procedures. In: *Proceedings, Third Annual ACM Symposium on the Theory of Computing* (Eds), ACM (1971).
- Karp, R. M. Reducibility among combinatorial problems. In: *Complexity of Computer Computations* (Eds), Plenum (1972).
- Kann, V. *On the approximability of NP-complete optimization problems*. PhD thesis, Department of Numerical Analysis and Computing Science. Stockholm, Royal Institute of Technology (1992).
- Hutchinson, J. E. Fractals and self-similarity. *Indiana University Mathematics Journal* **30**, 714–747 (1981).
- Wen, Z. Y. Moran sets and Moran classes. *Chinese Sci. Bull.* **46**, 1849–1856 (2001).

Acknowledgements

The authors wish to express their thanks to the anonymous referee for his/her patience and carefulness to improve the quality of the manuscript. The work is supported by National Natural Science Foundation of China (Nos 11371329, 11471124), NSF of Zhejiang Province (No. LR13A010001) and Scientific Research Fund of Zhejiang Provincial Education Department (No. Y201326678) and Philosophical and Social Science Planning of Zhejiang Province (No. 17NDJC108YB). The work is also supported by K.C. Wong Magna Fund in Ningbo University.

Author Contributions

L.X. designed the research. L.X., Qin W. and Lihong W. wrote the manuscript. Lihong W., J.C., Songjing W., L.B., Z.Y. and L.Z. collected the data, L.X. and Qin W. provided the proofs, Lihong W. prepared Figs 1, 2, 3 and 6. All authors discussed the results and reviewed the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Wang, L. *et al.* On the Fractality of Complex Networks: Covering Problem, Algorithms and Ahlfors Regularity. *Sci. Rep.* 7, 41385; doi: 10.1038/srep41385 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017