Journal of Graph Algorithms and Applications
http://jgaa.info/ vol. 15, no. 5, pp. 661–682 (2011)

A New Model for a Scale-Free Hierarchical Structure of Isolated Cliques

Takeya Shigezumi[1] *Yushi Uno*[2] *Osamu Watanabe*[3]

[1] Aoyama Gakuin University, Sagamihara, 252-5258, Japan
[2] Osaka Prefecture University, Sakai 599-8531, Japan
[3] Tokyo Institute of Technology, Tokyo 152-8552, Japan

Abstract

Scale-free networks are usually defined as the ones that have power-law degree distributions. Since many of real world networks such as the World Wide Web, the Internet, citation networks, biological networks, and so on, have this property in common, scale-free networks have attracted interests of researchers so far. They also revealed that such networks have some typical properties such as high cluster coefficient and small diameter as well, and a lot of network models have been proposed to explain those properties. Recently, it is reported that the following new properties about self-similar structures of a real world network are observed [Uno and Oguri, FAW and AAIM, 2011]. For a special kind of cliques in a network, 1. the size distributions of these cliques show a power-law, 2. the degree distribution of the network after contracting these cliques show a power-law, and 3. by regarding the contracted network as the original, 1 and 2 are observed repeatedly. In this paper, we propose a new network model constructed by a 'clique expansion' procedure, and show that it can explain this 'hierarchical structure of cliques'.

Submitted:	Revised:	Revised:	Accepted:
April 2010	Jan 2011	July 2011	July 2011
	Final:	Published:	
	July 2011	October 2011	

Article type:	Communicated by:
Regular paper	M. S. Rahman and S. Fujita

Research partly supported by JSPS Global COE program "Computationism as a Foundation for the Sciences" and also partly supported by Grant-in-Aid for Scientific Research (KAKENHI), No. 23500022. A preliminary version of this paper appeared in the Proceedings of WALCOM 2010 [14].

E-mail addresses: shigezumi@ise.aoyama.ac.jp (Takeya Shigezumi) uno@mi.s.osakafu-u.ac.jp (Yushi Uno) watanabe@is.titech.ac.jp (Osamu Watanabe)

1 Introduction

Cluster structures have been observed on many real world networks. A community structure that is often seen in large web networks is one of the typical examples of such cluster structures, but it seems to have some specific structural property. In order to analyze this property, Uno et al. [17] adopted "isolated cliques" and investigated the distribution and the structure of isolated cliques in some large web networks. An *isolated clique* (of size k) [9] is a clique consisting of k nodes that does not have more than k edges to its outside (see the next section for the precise definition). That is, an isolated clique is, while it is maximally dense in its inside, sparsely connected to its outside. Furthermore, there is an efficient algorithm [9] that can extract all of isolated cliques from a given graph. Uno et al. used this algorithm to analyze an undirected graph (which we call a "webgraph" here) representing some web network links, and they found some interesting properties that are summarized as follows.

Observation 1. The size distribution of isolated cliques in the webgraph follows a power-law distribution with an exponent that is larger than the exponent for the degree distribution.

Observation 2. Contract each isolated clique to one node and obtain a reduced graph. Then the degree distribution of this reduced graph follows the power-law with almost the same exponent as the degree distribution of the original graph. Furthermore, the reduced graph has again many isolated cliques whose size distribution follows almost the same power-law as the isolated clique size distribution of the original graph.

Observation 3. This contraction can be conducted for several times (at least five times) until the number of isolated cliques becomes very small. Then in these reduced graphs, more or less almost the same degree distribution and isolated clique size distribution can be observed (Figure 1).

We may call this observed structure *hierarchical clique structure*. Let us also call the final reduced graph that has almost no isolated cliques a *prime network*. Although many scale-free network models have been proposed to explain networks in the real world, e.g., [2, 11], most of them can only generate graphs without large cliques (not to mention, isolated cliques). There some clique based models [7, 5, 19], these models can generate k-trees which only contains size $k+1$ cliques for some fixed parameter k, and cannot explain the size distribution of cliques. Up until now, no models have been proposed for the hierarchical clique structure. Recently, a different type of some hierarchical structure, called a fractal property, has been also studied by Song et al. [15]. They observed the power-law degree distribution on the reduced graph obtained by contracting randomly and greedily chosen connected subgraph. They also proposed a model to represent this fractal property [16], which generates a tree so has neither cliques nor hierarchical clique structure. On the other hand, it may be possible that this hierarchical clique structure and the structure of a

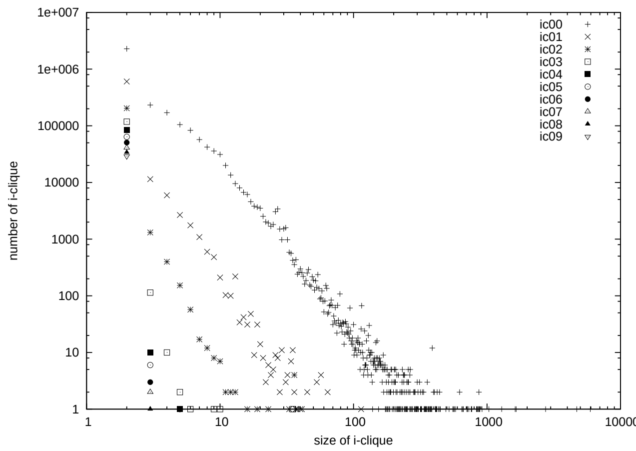

Figure 1: The size distribution of isolated cliques on the reduced graph. ic0i shows the size distribution of 1-isolated cliques on the ith graphs obtained by the contraction procedure.

prime network are independent. The purpose of this paper is to provide some model or method for adding the hierarchical clique structure to any given scale-free network. Thus, for example, we may use the BA model by Barabási and Albert [4] as a prime network model, and based on it a network with the hierarchical clique structure can be constructed by our method.

For explaining some of the features of our method, we introduce some basic notations (see the next section for their precise definitions). For a given graph W, its reduced graph $\mathcal{C}(W)$ is a graph obtained by contracting all isolated cliques of W into one vertex, where the contraction is made as shown in Figure 2.

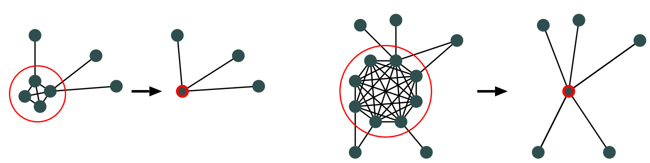

Figure 2: Examples of the contraction of an isolated clique.

Let W^0 denote the original webgraph and define $W^1 = \mathcal{C}(W^0)$, $W^2 = \mathcal{C}(W^1)$, ..., and so on. Uno et al. [17] observed that W^i follows almost the same power-law degree and isolated clique size distributions as W^0 for several times (at least three times).

Now our method is, roughly speaking, to use some randomized procedure to create G' from a given graph G so that (i) both G and G' follow the same degree distribution, and (ii) G' contains isolated cliques whose size distribution follows the power-law distribution with exponent that is about +1 larger than the one for the degree distribution (of G). We will give precise definition of this procedure, $\mathcal{E}()$, which satisfy the above properties. Let $\mathcal{E}(G)$ to denote the result of the procedure, applied to G. Consider a graph G^0 that is obtained by any model for scale-free networks (where we may assume that no isolated clique exists in G^0), and define $G^1 = \mathcal{E}(G)$, $G^2 = \mathcal{E}(G^1)$, ..., to G^t for some sufficiently large t. Then we show that the graph $W^0 \triangleq G^t$ has the desired property; that is, each W^i that is obtained from this W^0 by the contraction follows the same power-law degree and isolated clique size distributions as W^0.

Technically an interesting point in our analysis is that $\mathcal{C}(\cdot)$ is not necessarily the inverse of $\mathcal{E}(\cdot)$. Thus, the fact that W^i has the desired degree and isolated clique size distributions is not immediate from the above properties (i) and (ii) of $\mathcal{E}(\cdot)$.

The organization of this paper is as follows. In the rest of this section, we give some previous and related work. We give basic definitions of graphs, scale-free property and basic notations in Section 2. We explain our model precisely in Section 3, and give analysis in Section 4. Finally, we conclude the paper giving some future topics in Section 5.

Related Work

Various kinds of community structures have been introduced and investigated in the literature. Web mining using complete bipartite graph (CBG) has been investigated by Kleinberg [10]. They assumed that web communities contain at least one CBG which is called the core of the community. Reddy and Kitsuregawa [12] relaxed the criteria of existence of a community by defining a dense bipartite graph structure. They investigated a community hierarchy of the World Wide Web extracting all dense bipartite graphs found in the World Wide Web.

Many other models than mentioned above have been presented so far, there were only few mathematical analysis of the size distribution of communities for these models. Recently, a different type of some hierarchical structure, called the fractal property, has been also studied by Song et al. [15]. They observed that the power-law degree distribution on the reduced graph obtained by contracting randomly and greedily chosen connected subgraph. They also proposed a model to represent this fractal property [16] and they analysed a minimal model which generates a tree, thus it has neither cliques nor hierarchical clique structures.

Up until now, o models have been proposed for the hierarchical clique structure. Some clique based models had been presented [7, 5, 19]. All models in [7, 5, 19] generate k-trees for some fixed parameter k. A k-tree contains size $k+1$ cliques only, so these models cannot explain the hierarchical clique structure either.

2 Preliminaries

Throughout this paper, we consider only simple undirected graphs without multiple edges and self loops, and we denote a graph as $G = (V, E)$, where V is a set of vertices and E is a set of unordered pairs $e = \{u, v\}$ of V denoting edges. For any graph $G = (V, E)$, let $V[G] = V$ and $E[G] = E$ denoting the set of vertices and edges respectively. For any vertex $v \in V$, a vertex u is called *adjacent* to v if there is an edge $\{u, v\}$ in E. The *neighborhood* of a vertex v is a set $N_G(v) = \{u \in V[G] \mid \{u, v\} \in E[G]\}$, i.e., the set of adjacent vertices of v in G. The *degree* of v is $|N_G(v)|$, which is denoted by $d_G(v)$ and the *maximum degree* of G is $\max_{v \in V} d_G(v)$ and denoted by Δ. A graph $G' = (V', E')$ is called a subgraph of $G = (V, E)$ if $V' \subseteq V$ and $E' \subseteq (V' \times V') \cap E$. A subgraph of G is called a *clique* if every pair of vertices in this subgraph has an edge between them. A clique C is called *c-isolated* if the number of *outgoing* edges from $V(C)$ to $V \backslash V(C)$ is less than $c|V(C)|$. Although finding large cliques in a graph is intractable, finding isolated cliques is not so hard. Furthermore, 1-isolated clique can be enumerated in linear time [9], and it is investigated in [17].

In the following analysis, we assume that the graph is connected. We consider contraction and expansion procedures, and both procedures do not change the connectivity of the graph. Thus, if a given graph is not connected, we can apply our model separately to each of the connected component. By this assumption, we can also assume that all 1-isolated cliques are disjoint. Two 1-isolated cliques overlap only when they share 1 or $k - 1$ vertices. In both of these cases, there is no edge which connects vertex in those cliques and vertex on the outside of those cliques. If there are overlaps among two or more 1-isolated cliques, these overlaps can exist as an isolated component consisting themselves. In Figure 3, we present an example of two size k cliques shares $k - 1$ vertices (the case of $k = 4$). Thus, we can assume that 1-isolated cliques are disjoint without loss of generality and the contraction procedure can be uniquely defined. We consider a process of *contracting an isolated clique* of G into one vertex. We use $\mathcal{C}(G)$ to denote a reduced graph obtained from G by contracting all isolated cliques in G.

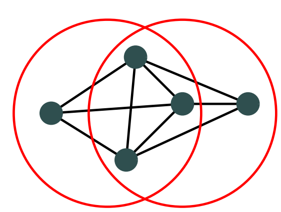

Figure 3: Overlapping two isolated cliques of size 4.

The scale-free property is considered as one of the basic properties characterizing real world large graphs. We say that G is 'scale-free' if its degree distribution follows power-law, i.e., a distribution proportional to $k^{-\gamma}$ for some constant γ. Let us make these notions more precise for our discussion. The *degree distribution* of G is a sequence $\{\frac{n_k}{n}\}_{k\geq 1}$, where n_k is the number of vertices with degree k and $\frac{n_k}{n}$ is the ratio of them among all vertices in G. Then we say that G's degree distribution follows a *power-law* if $n_k/n = \Theta(k^{-\gamma})$ for some γ, that is, there are some constants c_1 and c_2 such that $c_1 k^{-\gamma} \leq n_k/n \leq c_2 k^{-\gamma}$ for all $1 \leq k \leq \Delta$. In this paper, we extend this notion to isolated clique size distributions. The *isolated clique size distribution* of G is a sequence $\{\frac{m_s}{m}\}_{s\geq 1}$, where m_s is the number of isolated cliques of s vertices and m is the total number of isolated cliques. We say that G's isolated clique size follows a *power-law* if the sequence $\{\frac{m_s}{m}\}_{s\geq 1}$ satisfies $m_s/m = \Theta(s^{-\gamma})$ for some γ.

It does not make sense for discussing the above properties for any fixed finite graph G. Thus, in this paper, we will consider a family of graphs consisting of infinite number of graphs defined in a certain way and discuss power-law properties with constants c_1 and c_2 that are independent from k and the choice of a graph in the family. Thus, when claiming for example that G's degree distribution follows a power-law with some exponent γ, we formally imply that its degree sequence $\{n_k/n\}_{k\geq 1}$ satisfies $n_k/n = \Theta(k^{-\gamma})$ under some fixed constants c_1 and c_2 for all graphs in our assumed graph family.

In this paper, we consider a random process to generate graphs. To deal with degree distribution of such random graphs, we consider the *expected degree distribution*. We consider a sequence of $\left\{\frac{\mathrm{E}[N_k]}{\mathrm{E}[N]}\right\}_{k\geq 1}$ instead of $\left\{\frac{n_k}{n}\right\}_{k\geq 1}$, where $\mathrm{E}[N_k]$ is the expected number of vertices with degree k and $\mathrm{E}[N]$ is the expected number of vertices in G. In other words, it is the ratio of the expected number of vertices with degree k in G.

3 Model

The main idea of our model is as follows. Let G^0 be a prime scale-free graph generated by a certain scale-free model, e.g., BA model, which cannot generate graphs with cliques. Consider that a vertex in G^0 is either a "node" that represents a contracted 1-isolated clique or a "(simple) vertex", otherwise. We decide whether a vertex in G^0 is a "node" or a simple vertex randomly. We replace each node by an isolated clique whose size is the same as the degree of the original node as shown in the Figure 4. We call this replacement *expansion* and call this isolated clique *expanded clique*. Then we regard these new vertices in the isolated clique could be "nodes" or vertices, so, we decide them recursively. In order to technically simplify our analyses and discussions, we here change the definition of c-isolated cliques. A clique C is called *c-isolated* if the number of *outgoing* edges from $V(C)$ to $V \backslash V(C)$ is less than *or equals to* $c|V(C)|$. In this paper, we consider only 1-isolated cliques, so we simply call them *isolated cliques*. Note that we can obtain almost similar results even if we used the original definition of the isolated clique.

In our model, all vertices in expanded clique has one outgoing edge, which also implies the number of outgoing edges from the expanded clique equals to the size of the number of vertices in the clique. However, the requirement of the isolated clique is the number of outgoing edges is less than or equals to the number of vertices in the clique. We adopt simpler model since all vertices in expanded clique have one outgoing edge, those vertices have the same degree as the original node, that makes the analysis much simpler.

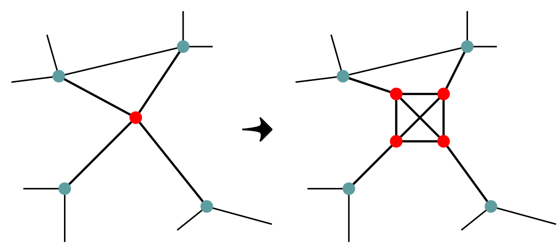

Figure 4: Replacing a "node" of degree 4 by an isolated clique of 4 nodes.

We now explain this idea precisely. Let $G^0 = (V^0, E^0)$ and a parameter μ_0 be inputs of our model. Let us assume G^0 is a *prime scale-free graph*, i.e., G^0 contains no isolated cliques and its degree distribution follows a power-law. From a given graph G^0, we expand it to G^i recursively and randomly. For $G^i = (V^i, E^i)$ $(i \geq 0)$, consider two subsets U^i and A^i of V^i such that $A^i \subseteq U^i \subseteq V^i$, where A^i denotes a set of "nodes" which are regarded as contracted isolated cliques, and U^i denotes a set of candidates of being "nodes". At the first step, all vertices of G^0 are candidate, i.e., $U^0 = V^0$. First, decide a set of "nodes" $A^i \subseteq U^i$ randomly. Consider a vertex v in U^i with degree k. We choose v into A^i with probability $p_k = \frac{\mu_0}{k}$ where μ_0 is a parameter in the input. It is independent to the choice of other vertices. We choose $p_k = \frac{\mu_0}{k}$ since it makes the expected number of vertices in one expansion $(p_k k = \mu_0)$ constant, independent of k. We also discuss about the case that we set $p_k = \frac{\mu_0}{k^a}, (a > 1)$ in Section 4.

Second, for each $v \in A^i$, let C_v be a clique of size $k = d_{G^i}(v)$. Let us define $G^{i+1} = (V^{i+1}, E^{i+1})$ and U^{i+1} as follows.

$$
\begin{aligned}
V^{i+1} &= V^i - A^i + \bigcup_{v \in A^i} V[C_v], \\
E^{i+1} &= \big\{\{u, v\} \in E^i \mid u, v \in (V^i - A^i)\big\} \\
&\quad + \bigcup_{v \in A^i} \left(E[C_v] \cup \{\{u_j, v_j\} \mid (*)\}\right), \\
((*) &: \{u_1, \ldots, u_{d_{G^i}(v)}\} = N_{G^i}(v) \text{ and } \{v_1, \ldots, v_{d_{G^i}(v)}\} = V[C_v].) \\
U^{i+1} &= \bigcup_{v \in A^i} V[C_v].
\end{aligned}
$$

Let us denote the above expansion procedure by a function $\mathcal{E}(\cdot)$, i.e., $(G^{i+1}, U^{i+1}) = \mathcal{E}(G^i, U^i)$ for any $i \geq 0$. In this paper, we always set $U^0 = V^0$, so the obtained

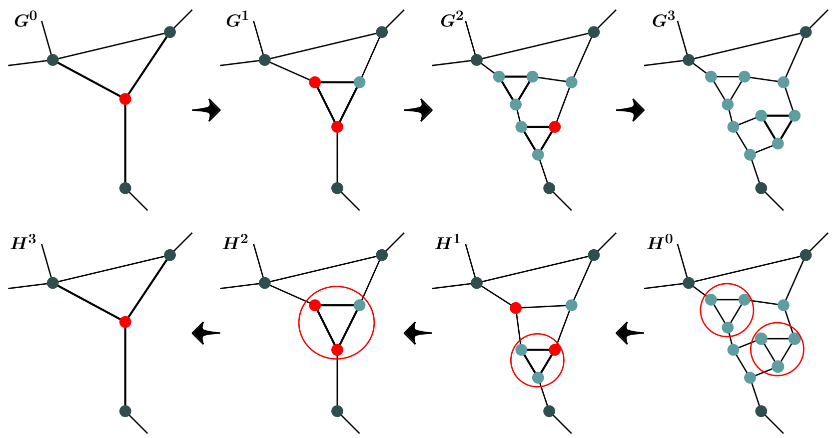

Figure 5: Expansion procedure is not necessarily inverse of the contraction.

$(G^1, U^1), (G^2, U^2), \ldots$ is a sequence of random graphs. We omit U^i and simply write them as $G^i = \mathcal{E}(G^{i-1})$ if no confusion arises.

As shown in Figure5, $\mathcal{C}()$ may not be an inverse function of $\mathcal{E}()$. In Figure5, $G^{i+1} = \mathcal{E}(G^i)$ for $i = 0, 1, 2$, and let $H^0 = G^3$, $H^{i+1} = \mathcal{C}(H^i)$ for $i = 0, 1, 2$. We here have $G^2 \neq H^1 (= \mathcal{C}(\mathcal{E}(G^2)))$.

When $A^t = \emptyset$ for some t, we let $H = G^t$ be an output of our model. We choose the parameter μ_0 as $\mu_0 < 1$, since otherwise, t may become infinite with positive probability. (The recursive procedure will not stop with positive probability.) This can be obtained by the classical analysis of the branching processes. (See next section or a literature e.g. [3].)

4 Analysis

In the following analysis, we focus on a vertex with degree k. The number of vertices expanded from one vertex obeys the following branching process (as known as Galton-Watson process) starting with one node. Many detailed analysis has been done for the branching process in the literature (see e.g. [3, 6]).Our expansion procedure can be expressed as the following branching process. (i) start from a single node that is set open; (ii) at each step, on each open node, the decision of "expansion" is made with probability p_k independently; (iii) those decided not to expand are set closed, and those decided to expand are also set closed after adding new k children that are set open; and (iv) repeat (ii) and (iii) until no open node exists. Let T denote a tree generated by this expansion process. We also call T a *tree representation of the expansion*. We present an example of a tree representation of an expansion in Figure 6. Note that we should consider the forest $\{T_v\}_{v \in V^0}$, a set of trees starting from each node $v \in V^0$, for the analysis of the number of nodes or the number of isolated cliques. However, we will focus on one tree since each tree is created independently at random and the number of total nodes or isolated cliques are the sum of them in each tree. In our expansion procedure, we say that vertices in C_v are *expanded from* v. When u is expanded from w and w is expanded from v, we say u is expanded from v. In this case, v is an ancestral node of u in the tree representation of the expansion.

It is well known that if $p_k k < 1$, then T is finite with probability 1. We defined $\mu_0 < 1$ and thus $p_k k = \mu_0 < 1$ in our model, so our expansion procedure generates a finite tree with probability 1.

The initial node is called a *root* node and a node with no child node is called a *leaf* node. For each node v of T, we define its height $h(v)$ and level $l(v)$ inductively as follows.

$$h(v) = \begin{cases} 0, & \text{if } v \text{ is a root node, and} \\ h(v') + 1 & \text{where } v' \text{ is the parent node of } v; \end{cases}$$

$$l(v) = \begin{cases} 0, & \text{if } v \text{ is a leaf node, and} \\ \max\{l(v_1), \ldots, l(v_k)\} + 1 & \text{where } v_1, \ldots, v_k \text{ are child nodes of } v. \end{cases}$$

The height of a tree is the maximum height of nodes in T and note that the height of a tree equals the level of the root node of the tree.

An example of a tree representation of an expansion procedure and corresponding height and level of nodes are shown in Figure 6. Let $H^0 = H$ and $H^1 = \mathcal{C}(H^0)$, $H^2 = \mathcal{C}(H^1), \ldots$, and so on. As shown in Figure 6, we can easily obtain the following observation.

Observation 4. Consider any node v in G^0, and consider a subgraph of G^t which is expanded from v. On the tree representation of the expansion from v, the number of leaves (which has level 0) is the number of nodes in a subgraph of $G^t(= H^0)$ expanded from v. For $l \geq 1$, the number of nodes in the tree with level l represents the number of isolated cliques in a subgraph of H^{l-1} expanded from v. The number of nodes in the tree with height i represents the number of nodes in a subgraph of G^i expanded from v.

So, we will analyse the number of nodes with level l for any $l \geq 0$ in this section. For any $l \geq 0$, define the following values:

$$\begin{aligned} M(l) &= \text{the expected number of level } l \text{ nodes in } T, \\ q(l) &= \Pr[T \text{ has a node of level } l] = \Pr[\text{the height of } T \geq l] \\ P(l) &= \Pr[\text{the height of } T \text{ is } l] = \Pr[\text{the level of the root of } T \text{ is } l] \end{aligned}$$

4.1 Degree distribution of H

Let V_k be a set of vertices with degree k in G^0 and let $n_k = |V_k|$. N_k denotes the number of vertices with degree k in $H(= G^t)$ and for any $v \in V[G^0]$, L_v denotes the number of vertices in the subgraph of H, expanded from v. Since for any vertices with degree k in H, there exists $v \in V_k$ in G, such that it is expanded

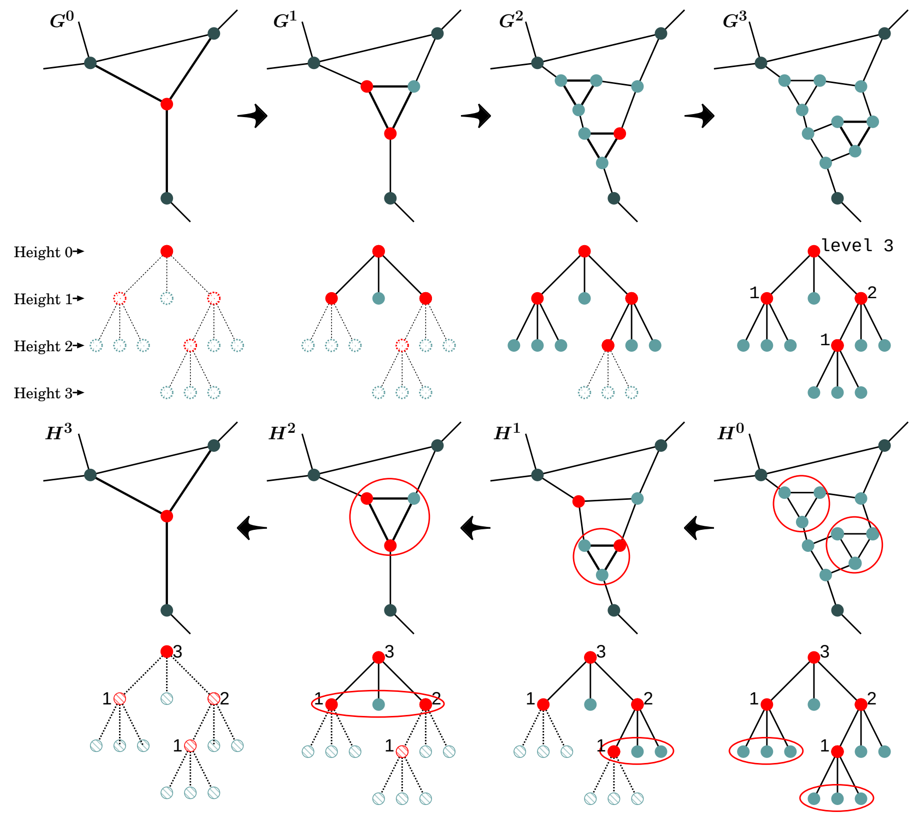

Figure 6: A tree representation of an expansion, the height and level, and the contraction.

from $v \in V_k$. Thus, we have $N_k = \sum_{v \in V_k} L_v$. Note that H is created by a random expansion process, N_k and L_v can be considered as random variables.

The distribution of L_v is well studied in the literature, e.g., [8, 6]. We can obtain the probability generating function (p.g.f.) of L_v as follows. Let $g(z)$ be the p.g.f. of the number of children of one node; $g(z) = 1 - p_k + p_k z^k$. Let $g_0(z) = z$, $g_1(z) = g(z)$ and $g_i(z) = g(g_{i-1}(z))$ for $i > 1$. Then, we have [8, 6];

Theorem 1. The p.g.f. of the number of nodes with height i on T is $g_i(z)$ for any $i \geq 0$.

Proof: Let Z_i be the number of nodes with height i on T, $Z_0 = 1$, and let $g_{(i)}(z)$ be the p.g.f. of Z_i for $i = 0, 1, \ldots$. Firstly, $g_{(0)}(z) = z$ and $g_{(1)}(z) = g(z)$. Under the condition of $Z_i = n$, the distribution of Z_{i+1} can be represented as the sum of the number of children of n nodes. So it has the p.g.f. $(g(z))^n$ for any $n = 0, 1, \ldots$. Accordingly, the p.g.f. of Z_{i+1} is;

$$g_{(i+1)}(z) = \sum_{n=0}^{\infty} \Pr[Z_i = n] \left(g_{(i)}(z)\right)^n = g_{(i)}(g(z)) \quad i = 0, 1, \ldots.$$

Since $g_{(1)}(z) = g(z) = g_1(z)$, we can obtain $g_{(i)}(z) = g_i(z)$ for any $i = 1, \ldots$ by induction. □

It is hard to obtain the closed-form of $g_i(z)$, however, we can obtain the expected value of Z_i, i.e. $g'_i(1)$.

Lemma 1. Let $g_1(z) = g(z) = 1 - p_k + p_k z^k$ and $g_i(z) = g(g_{i-1}(z))$ for $i \geq 1$. The expected number of nodes with height i, $\mathrm{E}[Z_i]$, is;

$$\mathrm{E}[Z_i] = g'_i(1) = \mu_0^i$$

where $\mu_0 = p_k k$.

Proof: At first, we have $g'(1) = p_k k = \mu_0^1$. By induction, we have

$$g'_i(1) = g'(g_{i-1}(1))g'_{i-1}(1) = g'(1)g'_{i-1}(1) = \mu_0 \mu_0^{i-1} = \mu_0^i.$$

□

So, the expected total number of nodes on T is $\sum_{i \geq 0} \mu_0^i = \frac{1}{1-\mu_0}$. Since T is a full k-ary tree (such that every inner node has exactly k children), the expected number of leaves is $\left(1 + \frac{\mu_0}{1-\mu_0}\left(1 - \frac{1}{k}\right)\right)$. See Appendix for the derivation.

From above, we can obtain the following Theorem.

Theorem 2. The expected number of vertices with degree k in H is;

$$\mathrm{E}[N_k] = \left(1 + \frac{\mu_0}{1-\mu_0}\left(1 - \frac{1}{k}\right)\right) n_k.$$

Since we assumed that the tree is finite, we can obtain another simple proof for the expected number of leaves. For further analysis in the later of this section, we show the another proof here.

Proof: By observation 4, the expected number of leaves expanded from v is $M(0)$, i.e. $\mathrm{E}[L_v] = M(0)$. If v_0 is not expanded, then the number of leaves is 1, and this occurs with probability $1 - p_k$. Otherwise, the number of leaf nodes is the sum of the number of leaf nodes in subtrees under k child nodes. Thus, we have

$$M(0) = p_k \cdot k M(0) + (1 - p_k) \cdot 1.$$

Hence

$$M(0) = \frac{1 - p_k}{1 - p_k k} = \left(1 + \frac{\mu_0}{1-\mu_0}\left(1 - \frac{1}{k}\right)\right).$$

Since $\mathrm{E}[N_k] = \sum_{v \in V_k} \mathrm{E}[L_v] = |V_k| M(0)$, the expected number of vertices with degree k in H is;

$$\mathrm{E}[N_k] = \left(1 + \frac{\mu_0}{1-\mu_0}\left(1 - \frac{1}{k}\right)\right) n_k.$$

□

Here, let N denote the total number of nodes in H to consider the degree distribution of H. Since N is the sum of the N_k for all k,

$$\mathrm{E}[N] = \sum_{k=1}^{\Delta} \mathrm{E}[N_k] = n + \frac{\mu_0}{1-\mu_0}\left(n - \sum_{k=1}^{\Delta} \frac{n_k}{k}\right) = \left(1 + \frac{\mu_0 C}{1-\mu_0}\right) n,$$

where C is a constant satisfying $C = 1 - \frac{\sum_{k=1}^{\Delta} \frac{n_k}{k}}{n}$.

Since the number of vertices in H is proportional to n, Theorem 2 gives the following ratio of the expectation of number of nodes with degree k;

$$\frac{\mathrm{E}[N_k]}{\mathrm{E}[N]} = \frac{\left(1 + \frac{\mu_0}{1-\mu_0}\left(1 - \frac{1}{k}\right)\right) n_k}{\left(1 + \frac{\mu_0 C}{1-\mu_0}\right) n}.$$

Let c_1 and c_2 be

$$c_1 = \frac{1 + \frac{\mu_0}{2(1-\mu_0)}}{1 + \frac{\mu_0 C}{1-\mu_0}}, \quad c_2 = \frac{1 + \frac{\mu_0}{(1-\mu_0)}}{1 + \frac{\mu_0 C}{1-\mu_0}}.$$

Then we obtained

$$c_1 \frac{n_k}{n} \leq \frac{\mathrm{E}[N_k]}{\mathrm{E}[N]} \leq c_2 \frac{n_k}{n}.$$

Corollary 1. If the input graph G^0 has the power-law degree distribution with exponent γ, $n_k/n = \Theta(k^{-\gamma})$, the expected degree distribution of H also follows the power-law distribution, i.e., $\mathrm{E}[N_k]/\mathrm{E}[N] = \Theta(k^{-\gamma})$.

4.2 Degree and isolated clique size distributions of H^i

In this section, we analyze the expected degree distribution and the expected number of isolated cliques in H^i. We must note that the contraction procedure $\mathcal{C}(\cdot)$ is not an inverse procedure of the expansion $\mathcal{E}(\cdot)$. It is easy to observe the fact by an example of the Figure 6.

Let us denote the number of isolated cliques of size k in H^i by $M_k(H^i)$, and the number of vertices with degree k in H^i by $N_k(H^i)$. First, we have the following obvious bound.

Theorem 3. Let $C'_1 = 1$ and $C'_2 = 1 + \frac{\mu_0}{1-\mu_0}$. Then, for any i,

$$C'_1 n_k \leq \mathrm{E}[N_k(H^i)] \leq C'_2 n_k.$$

Proof: It is clear that $C'_1 n_k = N_k(G^0) \leq \mathrm{E}[N_k(H^i)] \leq \mathrm{E}[N_k(H^0)] \leq C'_2 n_k$. □

Corollary 2. Consider an input graph G^0 has a power-law degree distribution with exponent γ, $\frac{n_k}{n} = \Theta(k^{-\gamma})$, and G^0 has no isolated cliques. Then the expected degree distribution of H^i also follows the power-law distribution with exponent γ.

For the expected number of isolated cliques of size k in H^i, we have the following bounds.

Theorem 4. Let C_1 and C_2 be $C_1 = 1 - \mu_0$ and $C_2 = \frac{1}{1-\mu_0}$. Then for any $0 \leq i$,

$$C_1 \mu_0^{i+1} \frac{n_k}{k} < \mathrm{E}[M_k(H^i)] < C_2 \mu_0^{i+1} \frac{n_k}{k}.$$

Proof: As mentioned in Observation 4, we will consider the distribution of the number of nodes which has level i. In the literature, e.g., [6], the distribution of the number of nodes with height i is mentioned. However, the analysis of the distribution of the number of nodes which has level i has not been provided before.

Let us remind the reader some definitions for the analysis.

$$\begin{aligned} M(l) &= \text{the expected number of level } l \text{ nodes in } T, \\ q(l) &= \Pr[\text{the level of the roof of } T \ \geq l], \\ P(l) &= \Pr[\text{the level of the root of } T \text{ is } l]. \end{aligned}$$

The expected number of isolated cliques expanded from one vertex and on H^i equals to $M(i+1)$, so the total number of isolated cliques of size k is $\mathrm{E}[M_k(H^i)] = n_k M(i+1)$.

Same as the proof of Theorem 2, we will consider $M(l)$ as follows. We use $P(l)$ to denote the probability that the root has level l, i.e. the depth of T is l. Clearly, this contributes $P(l)$ to $M(l)$. Then consider the other case. Since $M(l)$ is 0 for $l \geq 1$ if the root was not expanded; thus, consider the situation that the root was expanded (which occurs with probability p_k). Let $v_1, \ldots, v_k$ denote the child nodes of the root and let $T_1, \ldots, T_k$ denote the trees rooted by these nodes. Each T_i follows the same probability distribution as T; thus, we may use $M(l)$ for the expected number of level l nodes of T_i. Since the number of nodes on the tree T is finite, hence we have

$$M(l) = P(l) + p_k k M(l)$$

and

$$M(l) = \frac{P(l)}{1 - p_k k} = \frac{P(l)}{1 - \mu_0}. \tag{1}$$

Before considering $P(l)$, we note some basic equations of $P(l)$ and $q(l)$.

$$\begin{aligned} P(l) &= q(l) - q(l+1) \quad (\text{for } l \geq 0) && (2) \\ q(l) &= p_k \left\{ 1 - (1 - q(l-1))^k \right\} \quad (\text{for } l \geq 1) && (3) \\ q(l) &< \mu_0 \, q(l-1) \quad (\text{for } l \geq 1). && (4) \end{aligned}$$

Equation (4) was derived from Equation (3) as follows;

$$q(l) = p_k \left\{ 1 - (1 - q(l-1))^k \right\} < p_k \left\{ 1 - (1 - k q(l-1)) \right\} = \mu_0 \ q(l-1).$$

From now on, we consider the upper and lower bound for $P(l)$.

Lemma 2. We have $P(0) = 1 - p_k$ and $P(1) = p_k(1-p_k)^k = \frac{\mu_0}{k}(1-\frac{\mu_0}{k})^k$. For any $l > 1$, we have

$$P(l) < \frac{\mu_0^l}{k}\left(1 - \frac{\mu_0}{k}\right)^k.$$

Proof: By definition, $P(0)$ and $P(1)$ are the probability that the root node has level 0 and 1 respectively, so we immediately have $P(0) = 1 - p_k$ and $P(1) = p_k(1-p_k)^k$. For any $0 < x < y < 1$, it is easy to show that

$$(1-x)^k - (1-kx) < (1-y)^k - (1-ky).$$

Equation (4) implies $q(l) < q(l-1)$, so we have

$$\begin{aligned} P(l) &= q(l) - q(l+1) = p_k\left[\left\{1-(1-q(l-1))^k\right\} - \left\{1-(1-q(l))^k\right\}\right] \\ &< p_k\left[\{1-(1-kq(l-1))\} - \{1-(1-kq(l))\}\right] \\ &= p_k k\left(q(l-1) - q(l)\right) = \mu_0 P(l-1). \end{aligned}$$

Hence we obtained $P(l) < \mu_0^{l-1}P(1) = \frac{\mu_0^l}{k}(1-\frac{\mu_0}{k})^k$. □

To analyse the lower bound of $P(l)$, we need to consider the upper and lower bound of $q(l)$.

Lemma 3. For $q(l)$, we have the following upper bound;

$$q(l) < \frac{\mu_0^l}{k}.$$

Proof:

First, $q(1) = p_k = \frac{\mu_0}{k}$. By equation (4) and induction hypothesis,

$$q(l) < \mu_0\ q(l-1) \leq \mu_0\left(\frac{\mu_0^{l-1}}{k}\right) = \frac{\mu_0^l}{k}.$$

□

The lower bound of $q(l)$ was well studied in the literature. The $q(l)$ satisfies the following relationships with the p.g.f. $g_l(z)$;

$$\begin{aligned} 1 - q(l) &= \Pr[\text{the level of the root node } < l] \\ &= \Pr[\text{the number of nodes with height } l \text{ is } 0] \\ &= \Pr[Z_l = 0] = g_l(0). \end{aligned}$$

However, as mentioned above, the closed-form of $g_l(z)$ is hard to obtain. In [1], Agresti used a fractional linear generating function (f.l.g.f.) to obtain a good upper/lower bound of $g_l(z)$. We can use their results and obtain the lower bound of $q(l)$.

Lemma 4. The lower bound of $q(l)$ is;

$$q(l) > \frac{\mu_0^l}{k}(1-\mu_0).$$

Proof: For any p.g.f. $g(z)$, let $U(z)$ be any p.g.f. satisfying $g(z) \leq U(z)$ for $0 \leq z \leq 1$. Then, Seneta[13] showed that;

Lemma 5. ([13], Lemma A)

$$g_l(z) \leq U_l(z), \quad \text{for any } 0 \leq z \leq 1,\ l \geq 1,$$

where $U_l(z) = U(U_{l-1}(z))$, $U_1(z) = U(z)$.

Proof: Since $U_l(z)$ is a p.g.f. so it is an increasing function and $g(z)$ is also a p.g.f. so it satisfies $0 \leq g(z) \leq 1$ for any $0 \leq z \leq 1$. Then we have;

$$\begin{aligned} g_l(z) &= g_{l-1}(g(z)) \\ &\leq U_{l-1}(g(z)) \quad \text{by induction;} \\ &\leq U_{l-1}(U(z)) \quad U_{l-1}(z) \text{ is increasing;} \\ &= U_l(z). \end{aligned}$$

□(Lemma 5)

So, if we have some $U(z)$, then we can obtain the lower bound of $q(l)$ as follows;

$$\begin{aligned} 1 - q(l) &= \Pr[\text{The level of the root } < l] \\ &= \Pr[\text{The number of nodes with height } l = 0] \\ &= g_l(0) \\ &\leq U_l(0). \end{aligned}$$

For $U(z)$, Agresti used the following fractional linear generating function;

Lemma 6. ([1], Lemma 3 (i)) Let $U(z)$ as

$$U(z) = 1 - p_k + \frac{p_k z}{k - (k-1)z}.$$

Then, $U(z)$ satisfies $g(z) \leq U(z)$ for any $0 \leq z \leq 1$.

Proof:

$$g(z) = 1 - p_k + p_k z^k \leq 1 - p_k + \frac{p_k z}{k - (k-1)z} \quad \text{for } 0 \leq z \leq 1,$$

which holds if and only if

$$t(z) = 1 - kz^{k-1} + (k-1)z^k = 1 - z^k - k(1-z)z^{k-1} \geq 0 \quad \text{for } 0 \leq z \leq 1.$$

Now $t(1) = 0$ and

$$\begin{aligned} t'(z) &= -kz^{k-1} - k(k-1)z^{k-2}(1-z) + kz^{k-1} \\ &= -k(k-1)z^{k-2}(1-z) \leq 0 \quad (\text{for } 0 \leq z \leq 1). \end{aligned}$$

So, $t(z) \geq 0$ and thus $g(z) \leq U(z)$ for $0 \leq z \leq 1$. □(Lemma 6)

Since $U(z)$ is a f.l.g.f., we can easily obtain the closed form of $U_l(z)$. The lth iterate of $U(z)$ is;

$$U_l(z) = 1 + \frac{\mu_0^l(1-\mu_0)(z-1)}{(k-1)\left(\mu_0^l - 1\right)z + \left(k - \mu_0 - (k-1)\mu_0^l\right)}.$$

We give the derivation of this closed form in Appendix.

$$\begin{aligned} 1 - q(l) &\leq U_l(0) = 1 - \frac{\mu_0^l(1-\mu_0)}{k - \mu_0 - (k-1)\mu_0^l} \\ &< 1 - \frac{\mu_0^l(1-\mu_0)}{k}, \end{aligned}$$

and hence

$$q(l) > \frac{\mu_0^l}{k}(1-\mu_0).$$

□(Lemma 4)

By Lemma 3 and Lemma 4, $q(l)$ can be represented as

$$q(l) = \frac{\mu_0^l}{k}(1-\mu_0) + \epsilon_l,$$

where $0 < \epsilon_l < \frac{\mu_0^{l+1}}{k}$. Since $q(l) < \mu_0 q(l-1)$,

$$\frac{\mu_0^l}{k}(1-\mu_0) + \epsilon_l = q(l) < \mu_0 q(l-1) = \mu_0\left(\frac{\mu_0^{l-1}}{k}(1-\mu_0) + \epsilon_{l-1}\right).$$

So, $\epsilon_l < \mu_0\epsilon_{l-1} < \epsilon_{l-1}$. Thus, by $\epsilon_l - \epsilon_{l+1} > 0$,

$$\epsilon_l - \epsilon_{l+1} = q(l) - q(l+1) - \left\{\frac{\mu_0^l}{k}(1-\mu_0) - \frac{\mu_0^{l+1}}{k}(1-\mu_0)\right\} > 0. \quad (5)$$

By equation 5, we obtain the lower bound of $P(l)$;

$$P(l) = q(l) - q(l+1) > \left\{\frac{\mu_0^l}{k}(1-\mu_0) - \frac{\mu_0^{l+1}}{k}(1-\mu_0)\right\} = \frac{\mu_0^l}{k}(1-\mu_0)^2. \quad (6)$$

By equation (1), (6) and Lemma 2,

$$\frac{\mu_0^l}{k}(1-\mu_0) < M(l) < \frac{\mu_0^l}{k}\left(1 - \frac{\mu_0}{k}\right)^k \frac{1}{1-\mu_0} < \frac{\mu_0^l}{k}\frac{1}{1-\mu_0}.$$

Now letting $C_1 = 1 - \mu_0$ and $C_2 = \frac{1}{1-\mu_0}$,

$$C_1\mu_0^{i+1}\frac{n_k}{k} < \mathrm{E}[M_k(H^i)] < C_2\mu_0^{i+1}\frac{n_k}{k}.$$

This concludes the proof of Theorem 4. □

By Theorem 4, the expected number of isolated cliques in H^i is proportional to $\mu_0^{i+1}\frac{n_k}{k}$ for any size k. The total number of isolated cliques in H^i is also proportional to $\mu_0^{i+1}\sum_{k>1}\frac{n_k}{k}$. The ratio of the isolated clique of size k among all isolated cliques in H^i can be written as

$$\frac{C_1\mu_0^{i+1}\dfrac{n_k}{k}}{C_2\mu_0^{i+1}\sum_{k>1}\frac{n_k}{k}} < \frac{\mathrm{E}[M_k(H^i)]}{\mathrm{E}[\sum_{j\geq 1}M_k(H^j)]} < \frac{C_2\mu_0^{i+1}\dfrac{n_k}{k}}{C_1\mu_0^{i+1}\sum_{k>1}\frac{n_k}{k}}.$$

$\sum_{k>1}\frac{n_k}{k}$ can be considered as a constant independent from k, so let $M = \sum_{k>1}\frac{n_k}{k}$, $c_1' = \frac{C_1}{C_2 M}$ and $c_2' = \frac{C_2}{C_1 M}$. Finally, we have

$$c_1'\frac{n_k}{k} < \frac{\mathrm{E}[M_k(H^i)]}{\mathrm{E}[\sum_{j\geq 1}M_k(H^j)]} < c_2'\frac{n_k}{k}.$$

Corollary 3. Consider an input graph G^0 has a power-law degree distribution with exponent γ, $\frac{n_k}{n} = \Theta(k^{-\gamma})$, and G^0 has no isolated cliques. Then the expected size distribution of isolated cliques in H^i also follows the power-law distribution with exponent $\gamma+1$.

We here consider the case $p_k = \frac{\mu_0}{k^a}$ for some $a > 1$. Let $\nu = p_k k = \frac{\mu_0}{k^{a-1}}$, since $\nu < 1$, we can derive the same analysis as the above such that μ_0 is replaced by ν. Then, the equation in Theorem 4 becomes;

$$C_1\nu^{i+1}\frac{n_k}{k} < \mathrm{E}[M_k(H^i)] < C_2\nu^{i+1}\frac{n_k}{k}.$$

This equation is;

$$C_1\mu_0^{i+1}\frac{n_k}{k^{(a-1)i+a}} < \mathrm{E}[M_k(H^i)] < C_2\mu_0^{i+1}\frac{n_k}{k^{(a-1)i+a}}.$$

In this case, the power law exponent of the size distribution of isolated cliques are different for each i.

5 Concluding Remarks

In this paper, we proposed a new model to explain the hierarchical clique structure and its scale-free properties. Our model provides a graph with the similar properties to the ones that are observed in the World Wide Web.

However, our model generates a special kind of isolated cliques such that each member of the clique has exactly one outgoing edge. It is possible to consider some modifications of our model to this problem, randomly connect outgoing edges of the isolated cliques for example. In our model, we used some other model to generate a prime network (G^0). If we use a single vertex or a clique as a prime network, it generates a regular graph in our current model. We are trying to make more general model which can generate graphs with scale-free property and the hierarchical clique structure from one node or one clique. In our model, we set $\mu_0 < 1$ to let the output graph finite, we will try to study the distribution for the case of $\mu_0 \geq 1$.

Uno et al. also investigates the hierarchical structure of isolated stars [17, 18], we also apply our approach to them.

Acknowledgements

The authors wish to thank Elitza Maneva for helpful discussions and advice on the topic of branching process.

References

[1] A. Agresti. Bounds on the extinction time distribution of a branching process. *Advances in Applied Probability*, 6(2):322–335, 1974.

[2] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Review of Modern Physics*, 74:47–97, 2002.

[3] K. B. Athreya and P. E. Ney. *Branching Processes*. Springer-Verlag, 1972.

[4] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[5] C. Cooper and R. Uehara. Scale free properties of random k-trees. *Mathematics in Computer Science*, 3:489–496, 2010. 10.1007/s11786-010-0041-6.

[6] W. Feller. *An Introduction to Probability Theory and Its Applications*. Wiley, 3 edition, January 1968.

[7] Y. Gao. The degree distribution of random k-trees. *Theoretical Computer Science*, 410:688–695, March 2009.

[8] T. E. Harris. *The Theory of Branching Processes*. Springer-Verlag, Berlin, Heidelberg, 1963.

[9] H. Ito, K. Iwama, and T. Osumi. Linear-time enumeration of isolated cliques. In *Proceeding of ESA, Lecture Notes in Computer Science*, volume 3669, pages 119–130, 2005.

[10] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46(5):604–632, 1999.

[11] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45:167–256, 2003.

[12] P. K. Reddy and M. Kitsuregawa. Building a community hierarchy for the web based on bipartite graphs. *Proceedings of the 13-th IEICE Data Engineering Workshop*, pages C4–1, 2002.

[13] E. Seneta. On the transient behavior of a Poisson branching process. *Journal of the Australian Mathematical Society*, 7:465–480, 1967.

[14] T. Shigezumi, Y. Uno, and O. Watanabe. A new model for a scale-free hierarchical structure of isolated cliques. In M. Rahman and S. Fujita, editors, *WALCOM: Algorithms and Computation*, volume 5942 of *Lecture Notes in Computer Science*, pages 216–227. Springer Berlin / Heidelberg, 2010.

[15] C. Song, S. Havlin, and H. A. Makse. Self-similarity of complex networks. *Nature*, 433(7024):392–395, January 2005.

[16] C. Song, S. Havlin, and H. A. Makse. Origins of fractality in the growth of complex networks. *Nature Physics*, 2(4):275–281, April 2006.

[17] Y. Uno and F. Oguri. Contracted webgraphs: structure mining and scale-freeness. In *Proceedings of the Joint Conference of the 5th International Frontiers of Algorithmics Workshop (FAW) and the 7th International Conference on Algorithmic Aspects of Information Management (AAIM), Lecture Notes in Computer Science*, volume 6681, pages 287–299, 2011.

[18] Y. Uno, Y. Ota, and A. Uemichi. Web structure mining by isolated stars. *Proceedings of 4th WAW, Lecture Notes in Computer Science*, 4936:149–156, 2008.

[19] Z. Zhang, L. Rong, and F. Comellas. High-dimensional random apollonian networks. *Physica A: Statistical Mechanics and its Applications*, 364:610 – 618, 2006.

Appendix

Number of leaves in a full k-ary tree

Let T be a full k-ary tree. We here derive the relation between number of the leaf nodes and the number of total nodes of T.

Let N_{leaf} be a number of leaves of T, let N_{inner} be a number of inner nodes of T, and let N_{all} be a number of total nodes of T.

Each inner node has exactly k children, we have $1 + kN_{\text{inner}} = N_{\text{all}}$. Thus, $N_{\text{inner}} = \frac{N_{\text{all}}-1}{k}$. So we obtain

$$N_{\text{leaf}} = N_{\text{all}} - N_{\text{inner}} = 1 + (k-1)N_{\text{inner}} = 1 + (k-1)\frac{N_{\text{all}} - 1}{k}.$$

If the expected number of total nodes is $\frac{1}{1-\mu_0}$, the expected number of leaf node is;

$$1 + (k-1)\frac{\frac{1}{1-\mu_0} - 1}{k} = 1 + \frac{\mu_0}{1-\mu_0}\left(1 - \frac{1}{k}\right).$$

lth iteration of $U(z)$

We here derive the lth iteration of $U(z)$. Let us recall our definition of $U(z)$, that is,

$$U(z) = 1 - p_k + \frac{p_k z}{k - (k-1)z} = \frac{(k-1-\mu_0)z - (k-\mu_0)}{(k-1)z - k}.$$

Also recall that its lth iteration $U_l(z)$ is defined inductively by $U_l(z) = U(U_{l-1}(z))$ for $l > 1$ and $U_1(z) = U(z)$.

To derive $U_l(z)$, we use a linear function $L(z) = az + b$ and $f(z) = L^{-1}(U(L(z)))$. Due to the following lemma, for evaluating $U_l(z)$, it suffices to get good a and b such that $f_l(z)$ is easily calculated.

Lemma 7.

$$U_l(z) = L(f_l(L^{-1}(z))).$$

Proof: By $f(z) = L^{-1}(U(L(z)))$, we have $U(z) = L(f(L^{-1}(z)))$. Then we prove the lemma by induction. We already have it for $l = 1$. Let us assume that $U_l(z) = L(f_l(L^{-1}(z)))$. Then we have

$$\begin{aligned} U_{l+1}(z) &= U(U_l(z)) = L(f(L^{-1}(U_l(z)))) \\ &= L(f(L^{-1}(L(f_l(L^{-1}(z)))))) = L(f(f_l(L^{-1}(z)))) \\ &= L(f_{l+1}(L^{-1}(z))). \end{aligned}$$

□

Let $a = \frac{1-\mu_0}{k-1}$ and $b = 1$; then we have

$$\begin{aligned} f(z) = L^{-1}(U(L(z))) &= \frac{1}{a}(U(az+1)-1) \\ &= \frac{a(k-1-\mu_0)z + (k-1-\mu_0) - (k-\mu_0) - \{a(k-1)z + (k-1) - k\}}{a\{a(k-1)z + (k-1) - k\}} \\ &= \frac{a(k-1-\mu_0)z - 1 - a(k-1)z + 1}{a\{a(k-1)z - 1\}} \\ &= \frac{-\mu_0 z}{a(k-1)z - 1} = \frac{-\mu_0 z}{(1-\mu_0)z - 1} \qquad (\text{by } a = \tfrac{1-\mu_0}{k-1}) \\ &= \frac{z}{\left(1 - \frac{1}{\mu_0}\right) z + \frac{1}{\mu_0}}. \end{aligned}$$

Lemma 8. Let $K = \frac{1}{\mu_0}$. Then we have

$$f_l(z) = \frac{z}{K^l + (1-K^l)\,z}.$$

Proof: For $l = 1$, we have

$$f_1(z) = \frac{z}{\frac{1}{\mu_0} + \left(1 - \frac{1}{\mu_0}\right) z} = \frac{z}{K + (1-K)\,z},$$

and the lemma holds. For $l \geq 1$, we prove by induction as follows:

$$\begin{aligned} f_{l+1}(z) &= \frac{f_l(z)}{K + (1-K)\,f_l(z)} = \frac{\frac{z}{K^l + (1-K^l)z}}{K + (1-K)\,\frac{z}{K^l+(1-K^l)z}} \\ &= \frac{z}{K^{l+1} + (1-K^l)\,Kz + (1-K)z} = \frac{z}{K^{l+1} + (1-K^{l+1})\,z} \end{aligned}$$

□

We now have the closed form of $f_l(z)$. That is,

$$f_l(z) = \frac{z}{K^l + (1-K^l)\,z} = \frac{z}{\left(\frac{1}{\mu_0}\right)^l + \left(1 - \left(\frac{1}{\mu_0}\right)^l\right) z} = \frac{\mu_0^l}{\left(\frac{1-z}{z}\right) + \mu_0^l}.$$

By using Lemma 7, we obtain the closed form of $U_l(z)$ as follows:

$$\begin{aligned} U_l(z) &= L(f_l(L^{-1}(z))) = a\left(f_l\left(\frac{z-1}{a}\right)\right) + 1 \\ &= a\frac{\mu_0^l}{\left(\frac{a+1-z}{z-1}\right) + \mu_0^l} + 1 \\ &= 1 + \frac{a\mu_0^l z - a\mu_0^l}{\left(\mu_0^l - 1\right) z + \left(a + 1 - \mu_0^l\right)} \\ &= 1 + \frac{\mu_0^l(1-\mu_0)(z-1)}{(k-1)\left(\mu_0^l - 1\right) z + \left(k - \mu_0 - (k-1)\mu_0^l\right)}. \end{aligned}$$