Deterministic hierarchical networks *

Lali Barrière, Francesc Comellas, Cristina Dalfó

Dept. Matemàtica Aplicada IV, EPSC, Universitat Politècnica de Catalunya

Av. Canal Olímpic s/n, 08860 Castelldefels, Barcelona, Catalonia (Spain)

e-mails: {lali,comellas,cdalfo}@ma4.upc.edu

Abstract

It has recently been shown that many networks associated with complex systems are small-world (they have both a large local clustering and a small average distance and diameter) and they are also scale-free (the degrees are distributed according to a power-law). Moreover, these networks are very often hierarchical, as they describe the modularity of the systems which are modeled. While most of the studies for complex networks are based on stochastic methods, a deterministic approach, with an exact determination of the main relevant parameters of the networks, has proven useful to complement and enhance the probabilistic and simulation techniques and therefore to provide a better understanding of the systems modeled.

In this paper we find the diameter, clustering and degree distribution of a generic family of deterministic hierarchical small-world scale-free networks which has been considered for modeling real life complex systems.

*Research supported by the Secretaria de Estado de Universidades e Investigación (Ministerio de Educación y Ciencia), Spain, and the European Regional Development Fund (ERDF) under project TEC2005-03575

1 Introduction

With the publication in 1998 and 1999 of papers by Watts and Strogatz on small-world networks [21] and Barabási and Albert on scale-free networks [3], there has been a renewed interest in the study of networks associated to complex systems which has received a considerable boost as an interdisciplinary subject.

Many real life networks, transportation and communication systems (including the power distribution and telephone networks), the Internet [10], the World Wide Web [2] , and several social and biological networks [11, 12, 14] , belong to a class of networks known as small-world scale-free networks. All these networks exhibit both strong local clustering (nodes have many mutual neighbors) and a small average distance and diameter. Another important characteristic is that the number of links attached to the nodes usually obeys a power-law distribution (it is scale-free). Several authors also noticed that the modular structure of a system can be identified in the network as a specific clustering distribution which depends on the degree. The network is then called hierarchical [18, 4, 22, 20]. Moreover, with the introduction of a new measuring technique for graphs it has been discovered that many real networks can also be categorized as self-similar, see [19].

Along with these observational studies, researchers have developed different models [1, 9, 15], most of them stochastic, which should help to understand and predict the behavior and characteristics of the systems. However, new deterministic models constructed by recursive methods, based on the existence of cliques, have also been introduced [6, 8, 13, 7, 23]. These deterministic models have the advantage that they allow one to compute analytically relevant properties and parameters, which may be compared with data from real and simulated networks. In [6], Barabási et al. introduced a simple hierarchical family of deterministic networks and showed it had a small-world scale-free nature. However, their clustering is zero, in contrast with many real networks which have a high clustering. Another family of hierarchical networks is proposed in [18] to combine a modular structure with a scale-free topology and to model the metabolic networks of living organisms and networks associated with generic system-level cellular organizations. A simple variation of this hierarchical network is considered in [17] , where the authors study other modular networks as the WWW, the actor network, the Internet at the domain level, etc. The model is further generalized in [16].

In this paper, we study a family of hierarchical networks recursively and deterministically defined from an initial complete graph K_n. We find some of the main properties for this family: diameter, degree distribution and clustering distribution.

2 Hierarchical graphs $H_{n,k}$

Deterministic hierarchical graphs can be constructed from a complete graph K_n by connecting to a selected root vertex $n-1$ replicas of K_n. Next, $n-1$ replicas of the new whole structure are added, again to the same root. At this step the graph will have n^3 vertices. The process continues until we reach the desired graph order. There are many variations for these hierarchical graphs, depending on the order of initial graph, the introduction of extra edges among the different copies of subgraphs, etc. However, given the starting complete graph and the number of iterations they have no more parameters to adjust and the main characteristics of the graph become fixed.

In this section we introduce a family of networks defined by parameters n (order of the initial complete graph) and k (number of iterations) which generalizes the deterministic hierarchical network introduced in [18], see also [5] (which corresponds to $H_{4,k}$). The deterministic hierarchicals networks introduced in [17] and generalized in [16], constitute a subgraph of $H_{5,k}$ (some edges are not present). Our model enhances the modularity and self-similarity of the network, and allows the exact determination of the diameter, degree distribution and clustering distribution.

2.1 Definition

A hierarchical graph $H_{n,k}$ with defining parameter n and at level k is defined recursively as follows (see Figure 1):

- $H_{n,1}$ is the complete graph K_n. One of its vertices is distinguished and called *root*. All other vertices are called *peripheral*.
- For $k > 1$, $H_{n,k}$ is obtained by adding some edges to the union of n disjoint copies of $H_{n,k-1}$, denoted by $H^0_{n,k-1}, \ldots, H^{n-1}_{n,k-1}$.
- The edge set of $H_{n,k}$ contains:
  - the edges of each $H^j_{n,k-1}$;
  - the edges connecting the root of $H^0_{n,k-1}$ with every peripheral vertex of $H^j_{n,k-1}$, for $j = 1, \ldots, n-1$; and
  - all possible edges among the roots of $H^j_{n,k-1}$, for $j = 1, \ldots, n-1$.
- The root of $H_{n,k}$ is the root of $H^0_{n,k-1}$.
- The set of peripheral vertices of $H_{n,k}$ is the union of the peripheral vertices of $H^j_{n,k-1}$, for $j = 0, \ldots, n-1$.

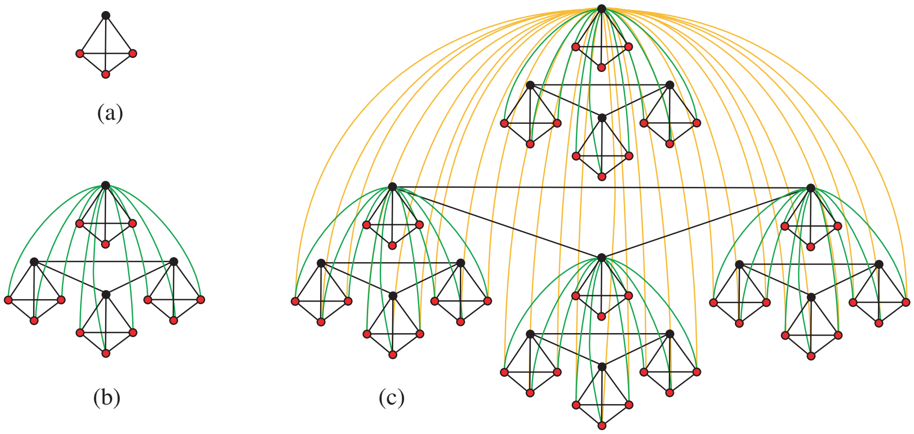

Figure 1: Hierarchical graphs with initial order 4: (a) $H_{4,1}$, (b) $H_{4,2}$, (c) $H_{4,3}$

It can be easily proved that the order and the size of $H_{n,k}$ are $|V_{n,k}| = n^k$ and $|E_{n,k}| = \frac{3}{2}n^{k+1} - 2n^k - (n-1)^{k+1} - \frac{n-2}{2}$.

2.2 Hierarchical properties

The hierarchical properties of these graphs can be summarized by the following two facts:

- For every $i = 0, \ldots, k$, the graph $H_{n,k}$ can be decomposed into n^{k-i} subgraphs each of them isomorphic to $H_{n,i}$.
- In $H_{n,k}$, by collapsing every subgraph isomorphic to $H_{n,i}$ into a node, and all multiple edges into one, we obtain a graph isomorphic to $H_{n,k-i}$.

2.3 Labeling

With this definition, we can assign labels to the nodes of $H_{n,k}$, as follows:

- The vertices of $H_{n,1}$ are labeled $0, \ldots, n-1$.
- A vertex labeled ω in $H^j_{n,k-1}$ is labeled $j \cdot \omega$ in $H_{n,k}$.

Labeling the vertices in this way we have, for instance, that the root of $H_{n,k}$ is labeled $00\ldots0$ (word of length k, all zeros), and the peripheral vertices of $H_{n,k}$ are precisely those vertices which have no zeros in their labels.

2.4 Diameter

Notation. Every vertex is identified with its label. Then, the vertex x of $H_{n,k}$ is $x = j \cdot x'$ if x is the same vertex as x' in $H^j_{n,k-1}$. In particular, that means that x' has length $k-1$. Moreover, we use the following notation:

- d_k denotes the distance in $H_{n,k}$,
- r_j denotes the root of $H^j_{n,k-1}$,

- P denotes the set of peripheral vertices of $H_{n,k}$, and
- P_j denotes the set of peripheral vertices of $H^j_{n,k-1}$.

Proposition 1 *The diameter of $H_{n,k}$ is $2k-1$.*

The proof of this proposition uses the following three Lemmas.

Lemma 1 *Let $k > 1$ and x, y be two arbitrary vertices in $H_{n,k}$. Then, we have one of the following three cases:*

1. *There exists j, $0 \leq j \leq n$, such that $x = j \cdot x'$, $y = j \cdot y'$, and $d_k(x, y) = d_{k-1}(x', y')$.*
2. *There exists j, $1 \leq j \leq n$, such that $x = 0 \cdot x'$, $y = j \cdot y'$, and $d_k(x, y) = d_{k-1}(x', r_0) + 1 + d_{k-1}(P_j, y')$.*
3. *There exist i, j, $1 \leq i < j \leq n$, such that $x = i \cdot x'$, $y = j \cdot y'$, and $d_k(x, y) = \min\{d_{k-1}(x', P_i) + 2 + d_{k-1}(P_j, y'),\ d_{k-1}(x', r_i) + 1 + d_{k-1}(r_j, y')\}$.*

Proof. By construction of $H_{n,k}$. ■

Lemma 2 *For all x in $H_{n,k}$:*

$$d_k(x, r_0) \leq \begin{cases} k-1 & \text{if } x = 0 \cdot x' \\ k & \text{otherwise} \end{cases} \quad \text{and} \quad d_k(x, P) \leq \begin{cases} k & \text{if } x = 0 \cdot x' \\ k-1 & \text{otherwise.} \end{cases}$$

Proof. By induction on k.

Case $k = 1$: If $x = 0 = r_0$, then $d_1(x, r_0) = 0$ and $d_1(x, P) = 1$; otherwise, $x \in P$, then $d_1(x, r_0) = 1$ and $d_1(x, P) = 0$.

Case $k > 1$: We only need to observe that, by construction of $H_{n,k}$,

$$d_k(x, r_0) = \begin{cases} d_{k-1}(x', r_0) & \text{if } x = 0 \cdot x' \\ d_{k-1}(x', P_j) + 1 & \text{if } x = j \cdot x',\ j \neq 0 \end{cases}$$

and

$$d_k(x, P) = \begin{cases} d_{k-1}(x', r_0) + 1 & \text{if } x = 0 \cdot x' \\ d_{k-1}(x', P_j) & \text{if } x = j \cdot x',\ j \neq 0. \end{cases}$$

By the induction hypothesis, the Lemma holds. ■

In the next Lemma, $0101\ldots$ denotes the vertex with label $\ell = \ell_1 \ldots \ell_k$, where $\ell_i = i-1 \pmod 2$ and $1010\ldots$ denotes the vertex with label $\ell = \ell_1 \ldots \ell_k$, where $\ell_i = i \pmod 2$.

Lemma 3 *$d_k(0101\ldots, r_0) = k - 1$, $d_k(0101\cdots, P) = k$, $d_k(1010\cdots, r_0) = k$, and $d_k(1010\cdots, P) = k - 1$.*

Proof. By induction on k.

Case $k = 1$: $0101\cdots = 0$ and $1010\cdots = 1$.

Case $k > 1$: We only need to observe that $0101\cdots = 0\cdot 1010\cdots$ and $1010\ldots = 1\cdot 0101\cdots$. This, and Lemma 1, imply that

- $d_k(0101\cdots, r_0) = d_{k-1}(1010\cdots, r_0) = k - 1$,
- $d_k(0101\cdots, P) = d_{k-1}(1010\cdots, r_0) + 1 = k - 1 + 1 = k$,
- $d_k(1010\cdots, r_0) = d_{k-1}(0101\cdots, P_1) + 1 = k$,
- $d_k(1010\cdots, P) = d_{k-1}(0101\cdots, P_0) = k - 1$.

■

Proof of Proposition 1. First we prove by induction on k that, for any given pair of vertices of $H_{n,k}$, x and y, we have $d_k(x, y) \leq 2k - 1$.

Case $k = 1$: $2k - 1 = 1$ and $H_{n,1} = K_n$.

Case $k > 1$: We distinguish the three cases in Lemma 1.

1. There exists j, $0 \leq j \leq n$, such that $x = j \cdot x'$, $y = j \cdot y'$, and $d_k(x, y) = d_{k-1}(x', y')$. By the induction hypothesis, $d_{k-1}(x', y') \leq 2(k-1) - 1 = 2k - 3 < 2k - 1$.
2. There exists j, $1 \leq j \leq n$, such that $x = 0\cdot x'$, $y = j\cdot y'$, and $d_k(x, y) = d_{k-1}(x', r_0) + 1 + d_{k-1}(P_j, y')$.

By Lemma 2, $d_{k-1}(x', r_0) \leq k-1$ and $d_{k-1}(P_j, y') \leq k-1$. Then, $d_k(x,y) \leq 2(k-1)+1 = 2k-1$.

3. There exist i, j, $1 \leq i < j \leq n$, such that $x = i \cdot x'$, $y = j \cdot y'$, and $d_k(x,y) = \min\{d_{k-1}(x', P_i) + 2 + d_{k-1}(P_j, y'),\, d_{k-1}(x', r_i) + 1 + d_{k-1}(r_j, y')\}$.

   By Lemma 3, $d_{k-1}(x', r_i) \leq k-1$ and $d_{k-1}(r_j, y') \leq k-1$. Then, $d_k(x,y) \leq 2(k-1)+1 = 2k-1$.

Now, we have to prove that there exist two vertices in $H_{n,k}$ at distance exactly $2k-1$. Let $x = 0101\cdots$ and $y = 1010\cdots$. It follows by the Lemmas 1 and 3 that $d_k(x,y) = 2k-1$. That completes the proof. ■

Note that the diameter scales logarithmcally with the order $N = |V_{n,k}| = n^k$, since $d_k = \frac{2}{n-1} \log N - 1$.

3 Degree distribution and clustering distribution of $H_{n,k}$

Proposition 2 *The degree distribution of $H_{n,k}$ is as follows: the root of $H_{n,k}$ has degree $\frac{(n-1)^{k+1}-(n-1)}{n-2}$, the $(n-1)n^{i-1}$ roots of $H^j_{n,k-i}$ have degree $\frac{(n-1)^{k-i+1}-(n-1)}{n-2} + n - 2$ ($i = 1, 2, \ldots, k-1$), the $(n-1)^k$ peripheral vertices of $H_{n,k}$ have degree $n+k-2$, the $(n-*

$1)^{k-i}n^{i-1}$ *peripheral vertices of* $H^j_{n,k-i}$ *have degree* $n+k-i-2$ *(*$i = 1, 2, \ldots, k-1$*).*

Proof. The root of $H_{n,k}$ has degree $1 + (n-1) + (n-1)^2 + \cdots + (n-1)^k = \frac{(n-1)^{k+1}-1}{n-2}$. Each $H^j_{n,k-i}$ root has degree $1 + (n-1) + (n-1)^2 + \cdots + (n-1)^{k-i} = \frac{(n-1)^{k-i+1}-1}{n-2}$, plus $n-2$ corresponding to the edges which join this root to the others at the same level. The peripheral vertices of $H^j_{n,1}$ have degree $n-1$. Those of $H^j_{n,2}$ have degree n. Similarly, the peripheral vertices of $H^j_{n,k-i}$ for $i = 1, \ldots, k-2$ have degree $n+k-i-2$ (see Table 1). ∎

The average degree is $\frac{2|E_{n,k}|}{|V_{n,k}|} = \frac{3n^{k+1}-4n^k-2(n-1)^{k+1}-n+2}{n^k}$.

For large k, by looking at the degree distribution we see that the number of vertices with a given degree z, $N_{n,k}(z)$, decreases as a power of the degree z and therefore the graph is scale-free [3, 9, 7]. As the degree distribution of the graph is discrete, to relate the exponent of this discrete degree distribution to the standard γ exponent of a continuous degree distribution for random scale free networks we use a cumulative distribution $P_{cum}(z) \equiv \sum_{z' \geq z} \frac{|N_{n,k}(z')|}{|N_{n,k}(z)|} \sim z^{1-\gamma}$, where z and z' are points of the discrete degree spectrum. When

$$z \;=\; \frac{(n-1)^{k-i+1}-n+2}{n-2}$$

there are exactly $(n-1)n^{i-1}$ vertices with degree z. The number of vertices with this or a higher degree is

$$(n-1)n^{i-1} + \cdots + (n-1)n + (n-1) + 1 = 1 + (n-1)\sum_{j=0}^{i-1} n^j = n^i.$$

Then, we have

$$z^{1-\gamma} = (\frac{(n-1)^{k-i+1} - n + 2}{n-2})^{1-\gamma} = \frac{n^i}{n^k} = n^{i-k}.$$

Therefore, for large k,

$$((n-1)^{k-i})^{1-\gamma} \sim n^{i-k}$$

and

$$\gamma \sim 1 + \frac{\log n}{\log(n-1)}.$$

Table 1: Degree and clustering distribution for $H_{n,k}$

Identification	Vertices	Degree	Clustering
$H_{n,k}$ root	1	$\frac{(n-1)^{k+1}-(n-1)}{n-2}$	$\frac{(n-2)^2}{(n-1)^{k+1}-2(n-1)+1}$
$H^j_{n,k-l}$ roots $j \neq 0$, $l = 1 \ldots k-1$	$(n-1)n^{l-1}$	$\frac{(n-1)^{k-i+1}-n+2}{n-2}$	$\frac{(n-2)^2}{(n-1)^{k-l+1}+(n-1)^2-3(n-1)+1}$
$H_{n,k}$ peripheral	$(n-1)^k$	$n+k-2$	$\frac{(n-1)^2+(2k-3)(n-1)+2-2k}{(n+k-2)(n+k-3)}$
$H^0_{n,k-l}$ peripheral $l = 1 \ldots k-1$	$(n-1)^{k-l}n^{l-1}$	$n+k-l-2$	$\frac{(n-1)^2+(2k-2l-3)(n-1)+2+2l-2k}{(n+k-l-2)(n+k-l-3)}$

Proposition 3 *The clustering distribution of $H_{n,k}$ is: the root of $H_{n,k}$ has clustering $\frac{(n-2)^2}{(n-1)^{k+1}-2n-+3}$, the $(n-1)n^{i-1}$ roots of $H^j_{n,k-i}$ have clustering $\frac{(n-2)^2}{(n-1)n^{k-i+1}+(n-1)^2-3n+4}$*

($i = 1, 2, \ldots, k-2$), the $(n-1)n^{k-2}$ roots of $H^j_{n,1}$ have clustering $\frac{n-2}{2n-3}$, the $(n-1)^k$ peripheral vertices of $H_{n,k}$ have clustering $\frac{(n-1)^2+(2k-3)(n-1)+2-2k}{(n+k-2)(n+k-3)}$, the $(n-1)^{k-i}n^{i-1}$ peripheral vertices of $H^j_{n,k-i}$ have clustering $\frac{(n-1)^2+(2k-2i-3)(n-1)+2+2i-2k}{(n+k-i-2)(n+k-i-3)}$ ($i = 1, 2, \ldots, k-2$) and the $(n-1)n^{k-2}$ peripheral vertices of $H^j_{n,1}$ have clustering 1.

Proof. The root of $H_{n,k}$ has clustering

$$\frac{\frac{(n-1)(n-2)}{2}(1+(n-1)+(n-1)^2+\cdots+(n-1)^{k-1})}{\frac{1}{2}\frac{(n-1)^{k+1}-(n-1)}{n-21}(\frac{(n-1)^{k+1}-n+1}{n-2}-1)} = \frac{(n-2)^2}{(n-1)^{k+1}-2n+3}.$$

The roots of $H^j_{n,k-i}$ ($i = 1, 2, \ldots, k-2$) have clustering

$$\frac{\frac{(n-1)n)}{2}\frac{(n-1)^{k-i}}{n-2}+\frac{(n-2)(n-3)}{2}}{\frac{1}{2}(\frac{(n-1)^{k-i+1}-n+1}{n-2}+n-2)(\frac{(n-1)^{k-i+1}-n+1}{n-2}+n-3)} = \frac{(n-2)^2}{(n-1)^{k-i+1}+(n-1)^2-3n+4}.$$

The clustering of the peripheral vertices of $H^j_{n,k-i}$ for $i = 1, \ldots, k-1$ is

$$\frac{\frac{(n-1)n-1}{2}+(n-2)(k-i-1)}{\frac{1}{2}(n+k-i-2)(n+k-i-3)} = \frac{(n-1)^2+(2k-2i-3)(n-1)+2+2i-2k}{(n+k-i-2)(n+k-i-3)}.$$

Note that for $i = k-1$, the peripheral vertices of $H^j_{n,1}$ have clustering $\frac{(n-1)^2-n+1}{(n-1)n} = 1$ (see Table 1). ∎

It is easy to check that, for each degree, the clustering of the corresponding vertices is inversely proportional to it. Then, the clustering of the graph is $C(z) \sim z^{-1}$. This is considered a signature for scale-free networks with high modularity (hierarchical), see [5].

4 Conclusion

In this paper we have provided a graph model which generalizes the hierarchical network introduced in [18], which combines a modular structure with a scale-free topology in order to model modular structures associated to living organisms, social organizations and technical systems. We have calculated the diameter, the degree distribution and the clustering of the graphs and we have seen that they are scale-free with a power law exponent which depends on the initial complete graph, and that the clustering distribution $C(z)$ scales with the degree as z^{-1}, as in many networks associated to real systems [17].

References

[1] R. Albert and A.-L. Barabási, Statistical mechanics of complex networks, Rev Mod Phys 74 (2002), 47–97.

[2] R. Albert, H. Jeong, and A.-L. Barabási, Diameter of the world wide web, Nature 401 (1999), 130–131.

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

[4] A.-L. Barabási, Z. Dezso, E. Ravasz, S.-H. Yook, and Z. Oltvai, "Scale-free and

hierarchical structures in complex networks", Statistical Mechanics of Complex Networks, R. Pastor-Satorras, M. Rubi, A. Diaz-Guilera, Albert (Editors), Lecture Notes in Physics , Springer-Verlag, Berlin, 625 (2003).

[5] A.-L. Barabási and Z.N. Oltvai, Network biology: Understanding the cell's functional organization, Nature Rev Genetics 5 (2004), 101–113.

[6] A.-L. Barabási, E. Ravasz, and T. Vicsek, Deterministic scale-free networks, Physica A 299 (2001) 559–564.

[7] F. Comellas, G. Fertin, and A. Raspaud, Recursive graphs with small-world scale-free properties, Phys Rev E 69 (2004), 037104.

[8] S.N. Dorogovtsev, A.V. Goltsev, and J.F.F. Mendes, Pseudofractal scale-free web, Phys Rev E 65 (2002), 066122.

[9] S.N. Dorogvtsev and J.F.F. Mendes, Evolution of networks, Adv Phys 51 (2002), 1079–1187.

[10] M. Faloutsos, P. Faloutsos, and C. Faloutsos, On power-law relationships of the internet topology. Comput Commun Rev 29 (1999), 251–260.

[11] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, and A.-L. Barabási, The large-scale organization of metabolic networks, Nature 407 (2000), 651–654.

[12] H. Jeong, S. Mason, A.-L. Barabási, and Z.N. Oltvai, Lethality and centrality in protein networks, Nature 411 (2001) 41–42.

[13] S. Jung, S. Kim, and B. Kahng, Geometric fractal growth model for scale-free networks, Phys Rev E 65 (2002), 056101.

[14] M.E.J. Newman, The structure of scientific collaboration networks, Proc Natl Acad Sci USA 98 (2001), 404–409.

[15] M.E.J. Newman, The structure and function of complex networks, SIAM Rev 45 (2003), 167–256.

[16] J.D. Noh, Exact scaling properties of a hierarchical network model, Phys Rev E 67 (2003), 045103.

[17] E. Ravasz and A.-L. Barabási, Hierarchical organization in complex networks, Phys Rev E 67 (2003), 026112.

[18] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, Hierarchical organization of modularity in metabolic networks, Science 297 (2002), 1551–1555.

[19] C.M. Song, S. Havlin, and H.A. Makse, Self-similarity of complex networks Nature 433 (2005), 392–395.

[20] R. V. Solé and S. Valverde, "Information theory of complex networks: on evolution and architectural constraints", Complex Networks, E. Ben-Naim, H. Frauenfelder, and Z. Toroczkai (Editors), Springer-Verlag, Berlin, Lecture Notes in Phys 650 (2004), 189–207.

[21] D.J. Watts and S.H. Strogatz, Collective dynamics of ‘small-world’ networks, Nature 393 (1998), 440–442.

[22] S. Wuchty, E. Ravasz, and A.-L. Barabási, “The Architecture of Biological Networks”, Complex Systems in Biomedicine, T.S. Deisboeck, J. Yasha Kresh and T.B. Kepler (Editors), Kluwer Academic Publishing, New York, 2003.

[23] Z.Z. Zhang, F. Comellas, G. Fertin, and L.L. Rong, High dimensional Apollonian networks, J Phys A: Math Gen 39 (2006), 1811–1818 .