EPJ manuscript No.
(will be inserted by the editor)

Planar unclustered graphs to model technological and biological networks

Alicia Miralles[1 a], Lichao Chen[2,3], Zhongzhi Zhang[2,3 b], and Francesc Comellas[1c]

[1] Departament de Matemàtica Aplicada IV, Universitat Politècnica de Catalunya, 08680 Castelldefels, Catalonia, Spain

[2] School of Computer Science, Fudan University, Shanghai 200433, China

[3] Shanghai Key Lab of Intelligent Information Processing, Fudan University, Shanghai 200433, China

Received: date / Revised version: date

Abstract. Many real life networks present an average path length logarithmic with the number of nodes and a degree distribution which follows a power law. Often these networks have also a modular and self-similar structure and, in some cases - usually associated with topological restrictions- their clustering is low and they are almost planar. In this paper we introduce a family of graphs which share all these properties and are defined by two parameters. As their construction is deterministic, we obtain exact analytic expressions for relevant properties of the graphs including the degree distribution, degree correlation, diameter, and average distance, as a function of the two defining parameters. Thus, the graphs are useful to model some complex networks, in particular technological and biological networks.

PACS. 89.75.Hc Networks and genealogical trees – 05.45.Df Fractals – 89.75.Fb Structures and organization in complex systems

1 Introduction

Ten years have past since the publication of the groundbreaking papers by Watts and Strogatz [1] on small-world networks and Barabasi and Albert [2] on scale-free networks. Their works led researches to the design of new network models to describe complex systems in nature and society like the Internet, protein-protein interactions, transportation systems or social and economic networks. Their models try to match observational studies which have identified at least three important common characteristics for real-life networks: They exhibit a small aver-

[a] e-mail: almirall@ma4.upc.edu
[b] e-mail: zhangzz@fudan.edu.cn
[c] Corresponding author. e-mail: comellas@ma4.upc.edu

age distance and diameter (compared to a random network with the same number of nodes and links); the number of links attached to the nodes obeys a power-law distribution (the networks are scale-free); and recently it has been discovered that, often, real networks are self-similar, see [3] sometimes showing a degree hierarchy related to the modularity of the system.

Many of the proposed models are stochastic as this is the case for the now classical preferential attachment method [2]. Thus, the use of probabilistic techniques is required to estimate the main parameters of a network [4]. However, a deterministic approach has proven useful to complement and enhance the probabilistic and simulation techniques. Deterministic models have a clear advantage, as they allow an analytical exact determination of relevant network parameters, which then may be compared with experimental data coming from real and simulated networks

Among the different techniques to generate deterministic models are of particular interest those based in recursive or iterative methods, where at each generation step nodes are connected to a known substructure of the network. This is the case of the pseudo-fractal networks [5] where, at each step, all existing links are considered and a new vertex is added to each of them. This construction can be generalized if cliques of a given size are used instead of links (which are 2-cliques), see [6]. Similar rules give the interesting Apollonian networks [7–9]. A related technique produces networks by duplication of certain substructures, see [10].

A generalization of these former methods introduces at each iteration a more complex substructure than a single node which is added to the network in a deterministic way. Substructures considered are triangles [11], cycles [12] and paths [13].

In this paper we go an step further in the generalization by considering the introduction of d parallel paths. The result is a family of planar, modular, hierarchical and self-similar networks, with small-world scale-free characteristics and with clustering coefficient zero, and all these parameter are determined by d as well as by the iteration step t. We note that some important real life networks, for example those associated to electronic circuits, Internet and some biological systems [14,15], have these characteristics as they are modular, almost planar and with a reduced clustering coefficient and have small-world and scale-free properties. Thus, these networks can be modeled by our construction which can be considered as a new tool to study their associated complex systems. In the next section we introduce the family of graphs object of study and in Section 3 we calculate analytically some relevant properties for the graphs, namely, the degree distribution, degree correlations, the diameter and the average distance. The last section provides some conclusions.

2 Generation of the graphs $M_d(t)$

In this section we introduce a family of modular, self-similar and planar graphs which have the small-world property and are scale-free. The family depends on an adjustable parameter d and the iteration number t. We provide an iterative algorithm, and also a recursive method, for its construction. The construction methods allow a direct determination of the order and size of the graph.

Iterative construction.– We give here an iterative formal definition of the proposed family of graphs, $M_d(t)$, characterized by $t \geq 0$, the number of iterations and a parameter d associated with the self-repeating modular structure.

First, we call *generating edge* the only edge of $M_d(0)$ and all edges of $M_d(t)$ whose endvertices have been introduced at different iteration steps t. All other edges of $M_d(t)$ will be known as *passive edges*. A generating edge becomes passive after its use in the construction.

The graph $M_d(t)$ is constructed as follows:

For $t = 0$, $M_d(0)$ has two vertices and a generating edge connecting them.

For $t \geq 1$, $M_d(t)$ is obtained from $M_d(t-1)$ by adding, to every generating edge in $M_d(t-1)$, d parallel paths P_4 of length three by identifying the two final vertices of each path with the endvertices of the generating edge.

The process is repeated until the desired graph order is reached, see Fig. 1. We note that the graph order can be also adjusted with the parameter d (number of parallel paths that are attached to each generating edge).

Recursive modular construction.– The graph $M_d(t)$ is also defined as follows:

For $t = 0$, $M_d(0)$ has two vertices and a generating edge connecting them.

For $t = 1$, $M_d(1)$ is obtained from $M_d(0)$ by adding to its only edge d parallel paths P_4 of length three by identifying the two final vertices of each path with the endvertices of the initial edge.

For $t \geq 2$, $M_d(t)$ is made from $2d$ copies of $M_d(t-1)$, by identifying, vertex to vertex, the initial edge of each $M_d(t-1)$ with the generating edges of $M_d(1)$, see Fig. 1.

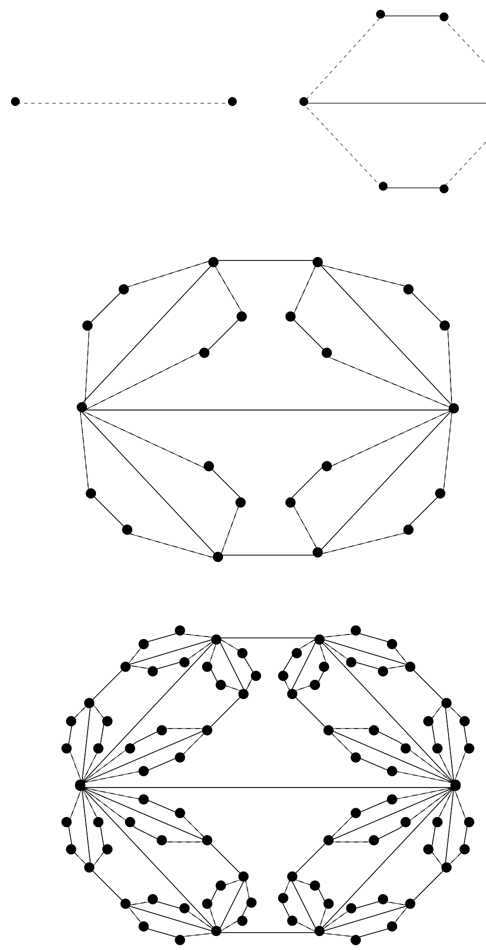

Fig. 1. Graphs $M_d(t)$ produced at iterations $t = 0, 1, 2$ and 3 for $d = 2$.

Order and size of $M_d(t)$.– We use the following notation: $\tilde{V}(t)$, $\tilde{E}(t)$ and $\tilde{E}_g(t)$ denote, respectively, the set of vertices, edges and generating edges introduced at step t, while $V(t)$ and $E(t)$ denote the set of vertices and edges of the graph $M_d(t)$.

Notice that, at each iteration, a generating edge is replaced by $2d$ new generating edges and d passive edges. Therefore: $|\tilde{E}_g(t+1)| = 2d \cdot |\tilde{E}_g(t)|$, and $|\tilde{E}_g(t)| = (2d)^t$. As each generating edge introduces at the next iteration $2d$ new vertices and $3d$ new edges we have $|\tilde{V}(t+1)| = 2d \cdot |\tilde{E}_g(t)| = (2d)^{t+1}$ and $|\tilde{E}(t+1)| = 3d \cdot |\tilde{E}_g(t)| = 3d \cdot (2d)^t$. As $|\tilde{V}(0)| = 2$ and $|\tilde{E}_g(0)| = 1$, the order and size of $M(t)$, $t \geq 0$, is:

$$|V(t)| = \sum_{i=0}^{t} |\tilde{V}(i)| = \frac{(2d)^{t+1} + 2d - 2}{2d-1},$$
$$|E(t)| = \sum_{i=0}^{t} |\tilde{E}(i)| = \frac{3d(2d)^t - d - 1}{2d-1}. \qquad (1)$$

Planarity.– A graph is planar if it can be drawn on the plane with no edges crossing. By construction of $M_d(t)$, the introduction at each iteration of d parallel paths connected to each generating edge, which afterwards becomes passive, adds $2d$ new vertices to the graph and they can be drawn without crossing edges. Planarity could also be proven from Kuratowski's theorem or from the known planarity test which states that a graph is planar if it has no cycles of length 3 and $|E| \leq 2|V| - 4, |V| > 3$, see [16].

3 Topological properties of $M_d(t)$

Thanks to the deterministic nature of the graphs $M_d(t)$, we can give exact values for the relevant topological properties of this graph family, namely, the degree distribution, degree correlations, the diameter and the average distance.

Degree distribution.– Initially, at $t = 0$, the graph has two vertices of degree one. When a new vertex i is added to the graph at iteration t_i, this vertex has degree 2 and it is connected to only one generating edge. We use the following notation: $k_g(i,t)$, $k_p(i,t)$ and $k(i,t)$ are, respectively, the number of generating edges, passive edges and total edges connected to vertex i, at step $t \geq t_i$. Therefore $k(i,t) = k_g(i,t) + k_p(i,t)$ is the degree of vertex i at this step.

From the construction process we can write,

$$\begin{cases} k_p(i,t+1) = k_p(i,t) + k_g(i,t) \\ k_g(i,t+1) = dk_g(i,t) \end{cases} \qquad (2)$$

with the initial conditions,

$$k_g(i,t_i) = 1 \quad \forall t_i \quad \text{and} \quad k_p(i,t_i) = \begin{cases} 0 & \text{if} \quad t_i = 0 \\ 1 & \text{otherwise} \end{cases} \qquad (3)$$

we have for $d > 1$,

$$k_g(i,t) = d^{t-t_i} \quad \forall t_i \quad \text{and}$$
$$k_p(i,t) = \begin{cases} 1 + \frac{d^t - d}{d-1} & \text{if} \quad t_i = 0 \\ 2 + \frac{d^{t-t_i} - d}{d-1} & \text{otherwise.} \end{cases} \qquad (4)$$

All the vertices that have been introduced at step t_i have the same degree at step t:

1. The two vertices introduced at step $t_i = 0$ have degree,

$$k(i,t) = k_g(i,t) + k_p(i,t) =$$
$$= d^t + 1 + \frac{d^t - d}{d-1} = \frac{d^{t+1} - 1}{d-1}. \qquad (5)$$

2. The $|\tilde{V}(t_i)| = (2d)^{t_i}$ vertices introduced at step $t_i > 0$ have degree,

$$k(i,t) = k_g(i,t) + k_p(i,t) =$$
$$= d^{t-t_i} + 2 + \frac{d^{t-t_i} - d}{d-1} = 1 + \frac{dd^{t-t_i} - 1}{d-1}. \qquad (6)$$

Therefore the degree spectrum of the graph is discrete and to relate the exponent of this discrete degree distribution to the power law exponent of a continuous degree distribution for random scale free networks, we use the technique described by Newman in [15] to find the cumulative degree distribution $P_{\rm cum}(k)$. If we denote by $V(t,k)$ the set of vertices that have degree k at step t,

$$P_{\rm cum}(k) = \frac{\sum_{k' \geq k} |V(t,k')|}{|V(t)|} = \frac{2+\sum_{t'_i=1}^{t_i}(2d)^{t'_i}}{\frac{(2d)^{t+1}+2d-2}{2d-1}} =$$
$$= \frac{(2d)^{t_i+1}+2d-2}{(2d)^{t+1}+2d-2} =$$
$$= \frac{(2d)^{t-\frac{\ln(k+\frac{2-k}{d}-1)}{\ln(d)}+1}+2d-2}{(2d)^{t+1}+2d-2}.$$

Fot t large, we obtain,

$$P_{\rm cum}(k) \approx (2d)^{-\frac{\ln(k+\frac{2-k}{d}-1)}{\ln(d)}} = (k+\frac{2-k}{d}-1)^{-\frac{\ln(2d)}{\ln(d)}}$$
$$= k^{-\frac{\ln(2d)}{\ln(d)}}(1-\frac{1}{d}+\frac{2-d}{kd})^{-\frac{\ln(2d)}{\ln(d)}}.$$

For $k >> 1$ this expression gives

$$P_{\rm cum}(k) \approx k^{-\frac{\ln(2d)}{\ln(d)}}(1-\frac{1}{d})^{-\frac{\ln(2d)}{\ln(d)}} \quad (7)$$

The degree distribution follows a power-law $P_{\rm cum}(k) \sim k^{-\gamma}$ with exponent $\gamma = \frac{\ln(2d)}{\ln(d)}$, see Fig. 2.

Therefore the degree distribution is scale-free. Research on networks associated to electronic circuits show that many of them are almost planar, modular and have a small clustering coefficient and in most cases their degree distributions follow a power-law [14,15] with exponent values in the same range than those of $M_d(t)$.

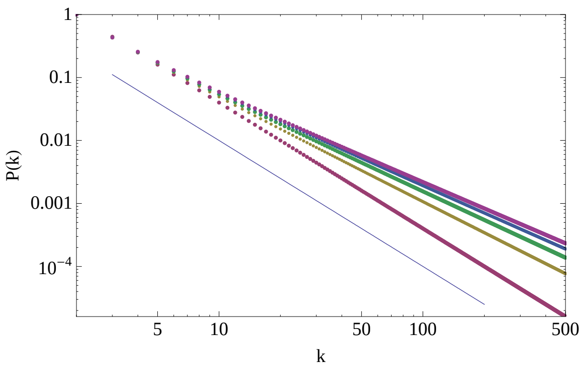

Fig. 2. Log-log representation of the cumulative degree distribution for $M_d(t)$, $d = 2,3,4,5$. The reference line has slope -2.

Correlation coefficient.– We obtain here the Pearson correlation coefficient, $r(d,t)$, for the degrees of the endvertices of the edges of $M_d(t)$ [17].

$$r(d,t) = \frac{|E(t)|\sum_i j_i k_i - [\sum_i \frac{1}{2}(j_i+k_i)]^2}{|E(t)|\sum_i \frac{1}{2}(j_i^2+k_i^2) - [\sum_i \frac{1}{2}(j_i+k_i)]^2} \quad (8)$$

where j_i, k_i are the degrees of the endvertices of the ith edge, with $i = 1, \cdots, |E(t)|$.

To find $r(d,t)$ we look at the degree distribution of the endvertices of the edges in $\tilde{E}(t_i)$ at a given step t_i. We denote by $\langle j,k \rangle$ an edge connecting vertices of degrees j and k.

The edges introduced at step t_i are:

1. Edges $\langle 2,2 \rangle$, connecting two vertices introduced at step $t_i > 0$. There are $(2d)^{t_i}/2$ edges (a half of the vertices introduced at step t_i). Notice that there is one edge $\langle 1,1 \rangle$ introduced at $t_i = 0$.
2. Edges $\langle 2, k(i', t_i) \rangle$ connecting vertices of degree two, introduced at step t_i, with all the vertices i' introduced at step $t_{i'}$ with $0 \leq t_{i'} \leq t_i - 1$. For each vertex i' there are $k_g(i', t_i)$ edges:

From the two vertices introduced at $t_{i'} = 0$, see (5), there are $2d^{t_{i'}}$ edges $\langle 2, \frac{d^{t_i+1}-1}{d-1}\rangle$.

From the $(2d)^{t_{i'}}$ vertices introduced at $t_{i'} > 0$, see (6), there are $(2d)^{t_{i'}} d^{t_i - t_{i'}}$ edges $\langle 2, 1 + \frac{d\, d^{t_i - t_{i'}} - 1}{d-1}\rangle$.

Table 1 displays a summary of these results.

Using the results of Table 1, we can find the sums:

$$\begin{aligned}\sum_i j_i k_i = {} & (4 + 16\,d - 51\,d^2 + 41\,d^3 - 8\,d^4 - \\ & - 3\,d^5 + d^6 + d^{t+1}(40 + 8t - 80 \cdot 2^t) + \\ & + d^{t+2}(-184 - 40t + 282 \cdot 2^t) + \\ & + d^{t+3}(306 + 74t - 373 \cdot 2^t) + \\ & + d^{t+4}(-236 - 64t + 227 \cdot 2^t) + \\ & + d^{t+5}(86 + 26t - 63 \cdot 2^t) + \\ & + d^{t+6}(-12 - 4t + 7 \cdot 2^t) + \\ & + d^{2t+2}(10 + 4t) + d^{2t+3}(-43 - 18t) + \\ & + d^{2t+4}(62 + 28t) + d^{2t+5}(-35 - 18t) + \\ & + d^{2t+6}(6 + 4t))/((d-1)^3 \\ & (2d^2 - 5d + 2)(d - 2)),\end{aligned}$$

$$\begin{aligned}\sum_i (j_i + k_i) = {} & \frac{-2}{(2d-1)(d-2)(d-1)^2}(-2 - 3d + \\ & + 10d^2 - 6d^3 + d^4 + d^{t+1}(-4 + 16 \cdot 2^t) + \\ & + d^{t+2}(14 - 37 \cdot 2^t) + d^{t+3}(14 + 26 \cdot 2^t) + \\ & + d^{t+4}(4 - 5 \cdot 2^t) - d^{2t+2} + 3d^{2t+3} - \\ & - 2d^{2t+4}),\end{aligned}$$

$$\begin{aligned}\sum_i (j_i^2 + k_i^2) = {} & -2(-8 - 32d + 186d^2 - 282d^3 + \\ & + 145d^4 + 49d^5 - 96d^6 + 48d^7 - 11d^8 + \\ & + d^9 + d^{t+1}(-72 + 160 \cdot 2^t) + d^{t+2}(384 - \\ & - 728 \cdot 2^t) + d^{t+3}(-750 + 1252 \cdot 2^t) + \\ & + d^{t+4}(606 - 882 \cdot 2^t) + d^{t+5}(-33 - 39 \cdot 2^t) + \\ & + d^{t+6}(-285 + 456 \cdot 2^t) + d^{t+7}(201 - 286 \cdot 2^t) + \\ & + d^{t+8}(-57 + 74 \cdot 2^t) + d^{t+9}(6 - 7 \cdot 2^t) + \\ & + 4d^{3t+2} - 16d^{3t+4} + 17d^{3t+5} + 5d^{3t+6} - \\ & - 19d^{3t+7} + 11d^{3t+8} - 2d^{3t+9})/((d^2 - 2d + 1) \\ & (2d - 1)(d - 2)(d^3 - 2d^2 - 2d + 4)(d - 1)^2).\end{aligned}$$

Replacing these sums into equation (8) we obtain directly, and for any d, a (long) exact analytical expression for the Pearson correlation coefficient of $M_d(t)$. For $d = 2$ the equation is displayed in Eq.(9):

Table 2 shows numerical values of the correlation for different instances of the graphs.

	$t = 1$	$t = 2$	$t = 3$	$t = 10$
$d = 2$	-0.1667	-0.0886	-0.0460	-0.0003
$d = 10$	-0.4091	-0.2338	-0.1174	-0.0009
$d = 100$	-0.4901	-0.2057	-0.0934	-0.0007

Table 2. Correlation coefficient at steps $t = 1, 2, 3, 10$ for several values of d.

From the values of the correlation coefficient we see that this family of graphs has the degrees of the endvertices negatively correlated, large degree vertices tend to be connected with low degree vertices, and the graphs are disassortative.

Step t_i	Edges at step t_i	Number	Edges at step $t > t_i$
$t_i = 0$	$\langle 1, 1 \rangle$	1	$\langle \frac{d^{t+1}-1}{d-1}, \frac{d^{t+1}-1}{d-1} \rangle$
$1 \leq t_i \leq t$	$\langle 2, 2 \rangle$	$\frac{(2d)^{t_i}}{2}$	$\langle 1 + \frac{dd^{t-t_i}-1}{d-1}, 1 + \frac{dd^{t-t_i}-1}{d-1} \rangle$
	$\langle 2, \frac{d^{t_i+1}-1}{d-1} \rangle$	$2d^{t_i}$	$\langle 1 + \frac{dd^{t-t_i}-1}{d-1}, \frac{d^{t+1}-1}{d-1} \rangle$
$2 \leq t_i \leq t$ $1 \leq t_{i'} \leq t_i - 1$	$\langle 2, 1 + \frac{dd^{t_i - t_{i'}}-1}{d-1} \rangle$	$(2d)^{t_{i'}} \cdot d^{t_i - t_{i'}}$	$\langle 1 + \frac{dd^{t-t_i}-1}{d-1}, 1 + \frac{dd^{t-t_{i'}}-1}{d-1} \rangle$

Table 1. Number of edges in $M_d(t)$ according to the degrees of their endvertices.

$$r(2,t) = \frac{4^t t^2 - 2^{t+1}t + 3 \cdot 2^{2t+1}t - 2^{3t+2}t + 13 \cdot 4^t - 3 \cdot 2^{t+1} + 4^{2t+1} - 3 \cdot 2^{3t+2} + 1}{2^{4t+1}t^2 + 2^{2t+1}t - 2^{3t+3}t + 2^{4t+3}t - 2^t + 5 \cdot 4^t - 2^{3t+4} + 3 \cdot 2^{4t+3} - 3 \cdot 2^{5t+2}} \quad (9)$$

We notice that most technological and biological networks have this property, see [15].

For $d >> 1$, we obtain

$$r(d,t) \approx \frac{1}{1 - 3 \cdot 2^{t-1}},$$

which for t large gives $r(d,t) \sim 0$.

Diameter.– At each iteration step we introduce, for every generating edge, $2d$ new vertices. These vertices are among them at distance at most 3. As each vertex joins the graph of the former step through one new edge, the diameter will increase by exactly 2 units. Therefore $D(t) = D(t-1)+2$, $t \geq 2$. As $D(1) = 3$, we have that the diameter of $M_d(t)$ is $D(t) = 3 + 2 \cdot (t-1)$, $t \geq 1$. Therefore, from Eq. 1, and as for t large, $t \sim \ln |V(t)|$ we have in this limit that $D(t) \sim \ln |V(t)|$.

Average distance.– The average distance of $M_d(t)$ is defined as:

$$\bar{D}(t) = \frac{1}{|V(t)|(|V(t)|-1)/2} \sum_{i,j \in V(t)} d_{i,j}\,, \quad (10)$$

where $d_{i,j}$ is the distance between vertices i and j. In what follows, $S(t)$ will denote the sum $\sum_{i,j\in V(t)} d_{i,j}$.

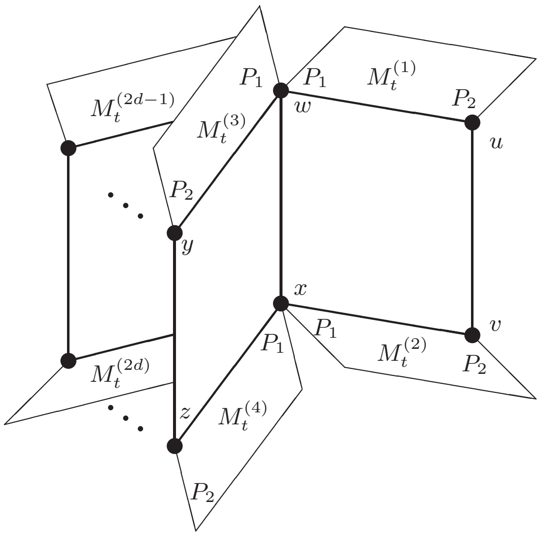

Fig. 3. $M_d(t+1)$ is obtained from the juxtaposition of $2d$ copies of $M_d(t)$.

The modular recursive construction of $M_d(t)$ allows us to calculate the exact value of $\bar{D}(t)$. At step t, $M_d(t+1)$ is obtained from the juxtaposition of $2d$ copies of $M_d(t)$, which we label $M_{d,t}^{(\eta)}$, $\eta = 1, 2, \cdots, 2d$, see Figures 1 and 3. Whenever possible, we drop the subscript d and represent $M_{d,t}^{(\eta)}$ as $M_t^{(\eta)}$ to keep the notation uncluttered. The copies are connected one to another at $2d+2$ vertices which we call *connecting vertices* such as u, v, w, x, y, and z in Fig. 3. Thus, the sum of distances S_{t+1} satisfies the following recursion:

$$S_{t+1} = 2d\, S_t + \Delta_t. \tag{11}$$

where Δ_t is the sum over all shortest path length whose endpoints are not in the same $M_t^{(\eta)}$ branch.

To compute Δ_t, we classify the vertices of $M_d(t+1)$ into two categories: the two vertices with the largest degree (i.e., w and x in Fig. 3) are called *hubs*, while any othe vertex is named a non-hub vertex. Thus Δ_t can be obtained by summing the following path lengths that are not included in the distance between vertex pairs of $M_t^{(\eta)}$: length of the shortest paths between non-hub vertices, length of the shortest paths between a hub and non-hub vertices, and length of the shortest paths between hubs (for example, d_{uv}, d_{ux}, and d_{uz}).

Let us denote $\Delta_t^{\alpha,\beta}$ as the sum of all shortest paths between non-hub vertices, whose endpoints are in $M_t^{(\alpha)}$ and $M_t^{(\beta)}$, respectively. Thus, $\Delta_t^{\alpha,\beta}$ rules out the paths with endpoints at the connecting vertices belonging to $M_t^{(\alpha)}$ or $M_t^{(\beta)}$. For example, each path contributing to $\Delta_t^{1,2}$ does not end at vertex u, v, w or x, and each path contributing to $\Delta_t^{1,4}$ does not end at vertex u, w, x or z. According to its value, $\Delta_t^{\alpha,\beta}$ can be split into three classes, where the three representatives are $\Delta_t^{1,2}$, $\Delta_t^{1,3}$, and $\Delta_t^{1,4}$, and the cardinality of the three classes are d, $d(d-1)$, and $d(d-1)$, respectively. Analogously, the length of the shortest paths between a hub and all non-hub vertices can be classified into two classes, while the shortest paths between hubs can be partitioned into three classes with path lengths equal to 1, 2, or 3.

Let Ω_t^{α} be the set of non-hub vertices in $M_t^{(\alpha)}$, then the total sum Δ_t is given by

$$\begin{aligned}\Delta_t = d\Delta_t^{1,2} + d(d-1)\left(\Delta_t^{1,3} + \Delta_t^{1,4}\right) + 2d(d+1)\\ \sum_{j\in\Omega_t^2} d_{wj} + 2d(d-1)\sum_{j\in\Omega_t^4} d_{uj} +\\ + (d+1)d_{uv} + d(d+1)\, d_{uy} + d(d-1)\, d_{uz},\end{aligned} \tag{12}$$

where $d_{uv} = 1$, $d_{uy} = 2$, and $d_{uz} = 3$ are easily seen.

Having Δ_t in terms of the quantities of $\Delta_t^{1,2}$, $\Delta_t^{1,3}$, $\Delta_t^{1,4}$, $\sum_{j\in\Omega_t^2} d_{wj} \sum_{j\in\Omega_t^2} d_{uj}$, and $\sum_{j\in\Omega_t^2} d_{uj}$, the next step is to explicitly determine these quantities. To this end, we classify non-hub vertices in $M_d(t+1)$ into two different parts according to their shortest path lengths to either of the two hubs (i.e. w and x). Notice that the vertices w and x themselves are not partitioned into either of the two parts represented as P_1 and P_2, respectively. The classification of vertices is shown in Fig. 3). For any non-hub vertex φ, we denote the shortest path length from φ to w, x as a, and b, respectively. By construction, a and b can differ at most by 1 since vertices w and x are adjacent. Then the classification function $class(\varphi)$ of vertex φ is defined to be

$$class(\varphi) = \begin{cases} P_1 & \text{for } a < b, \\ P_2 & \text{for } a > b. \end{cases} \tag{13}$$

It should be mentioned that the definition of the vertex classification is recursive. For instance, class P_1 and P_2 in $M_t^{(1)}$ belong to class P_1 in $M_d(t+1)$, class P_1 and P_2 in $M_t^{(2)}$ belong to class P_2 in $M_d(t+1)$, and so on. Since the two hubs w and x are symmetrical, in the graph we have the following equivalent relations from the viewpoint of class cardinality: classes P_1 and P_2 are equivalent one to each other. We denote the number of vertices in network $M_d(t)$ that belong to class P_1 as N_{t,P_1}, and the number of vertices in class P_2 as N_{t,P_2}. By symmetry, we have $N_{t,P_1} = N_{t,P_2}$, which will be abbreviated as N_t hereafter. It is easy to see that

$$N_t = \frac{|V(t)|}{2} - 1 = \frac{d(2d)^t - d}{2d-1}. \tag{14}$$

For a vertex φ in $M_d(t+1)$, we are also interested in the smallest value of the shortest path length from φ to either of the two hubs w and x. We denote the shortest distance as this value by f_φ, and it can be defined as

$$f_\varphi = min(a,b). \tag{15}$$

Let δ_{t,P_1} (δ_{t,P_2}) denote the sum of f_φ for all vertices belonging to class P_1 (P_2) in $M_d(t)$. Again by symmetry, we have $\delta_{t,P_1} = \delta_{t,P_2}$ that will be written as δ_t for short. Taking into account the recursive method of constructing $M_d(t)$, we notice that the vertex classification follows also a recursion. Therefore we can write the following recursive formula for δ_{t+1}:

$$\delta_{t+1} = 2d\,\delta_t + d\,N_t + d. \tag{16}$$

Substituting equation (14) into equation (16), and considering the initial condition $\delta_0 = 0$, equation (16) is solved inductively

$$\delta_t = \frac{2d - 2d^2 - (2d)^{1+t} + d(2d)^{1+t} - dt(2d)^t + dt(2d)^{1+t}}{2(2d-1)^2}. \tag{17}$$

We now return to compute equation (12). For convenience, we use $\Gamma_t^{\eta,i}$ to denote the set of non-hub vertices belonging to class P_i in $M_t^{(\eta)}$. Then $\Delta_t^{1,2}$ can be written as

$$\begin{aligned} \Delta_t^{1,2} &= \sum_{\substack{r\in\Gamma_t^{1,1}\bigcup\Gamma_t^{1,2} \\ s\in\Gamma_t^{2,1}\bigcup\Gamma_t^{2,2}}} d_{rs} \\ &= \sum_{\substack{r\in\Gamma_t^{1,1} \\ s\in\Gamma_t^{2,1}}} (d_{rw} + d_{wx} + d_{xs}) + \sum_{\substack{r\in\Gamma_t^{1,1} \\ s\in\Gamma_t^{2,2}}} (d_{rw} + d_{wv} + d_{vs}) \\ &\quad + \sum_{\substack{r\in\Gamma_t^{1,2} \\ s\in\Gamma_t^{2,1}}} (d_{ru} + d_{ux} + d_{xs}) + \sum_{\substack{r\in\Gamma_t^{1,2} \\ s\in\Gamma_t^{2,2}}} (d_{ru} + d_{uv} + d_{vs}) \\ &= 8N_t\delta_t + 6(N_t)^2. \end{aligned} \tag{18}$$

Analogously, we find

$$\Delta_t^{1,3} = 8N_t\delta_t + 4(N_t)^2 \tag{19}$$

and

$$\Delta_t^{1,4} = 8N_t\delta_t + 8(N_t)^2. \tag{20}$$

Next we will determine other quantities in equation (12), with $\sum_{j\in\Omega_t^2} d_{wj}$ given by

$$\begin{aligned}\sum_{j\in\Omega_t^2} d_{wj} &= \sum_{j\in\Gamma_t^{2,1}} (d_{wx} + d_{xj}) + \sum_{j\in\Gamma_t^{2,2}} (d_{wv} + d_{vj}) \\ &= 2\,\delta_t + 3\,N_t.\end{aligned} \tag{21}$$

Analogously, we can obtain

$$\sum_{j\in\Omega_t^4} d_{uj} = 2\,\delta_t + 5\,N_t. \tag{22}$$

Substituting equations (18), (19), (20), (21) and (22) into equation (12), we have the final expression for cross distances Δ_t,

$$\begin{aligned}\Delta_t &= 1 + 5d^2 + 4d(4d-1)N_t + 6d(2d-1)(N_t)^2 + \\ &+ 8d\delta_t[d + (2d-1)N_t] = \\ &= \frac{1}{(1-2k)^2}(1 - 4d + 5d^2 - 2d^3 + (1-d)(2d)^{2+t} + \\ &+ (5+2t)4^{1+t}d^{4+2t} - (7+2t)d^2(2d)^{1+2t}).\end{aligned} \tag{23}$$

Inserting equation (23) into equation (11) and using the initial condition $S_0 = 1$, equation (11) is solved inductively,

$$\begin{aligned}S_t &= \frac{1}{(-1+2d)^3}(-1 + 4d - 5d^2 + 2d^3 + 2^{1+t}d^{1+t} - \\ &- 7\cdot 2^{2t}d^{2+2t} + 3\cdot 2^{1+2t}d^{3+2t} - 2^{1+t}d^{1+t}t + \\ &+ 3\cdot 2^{1+t}d^{2+t}t - 2^{2+t}d^{3+t}t - 2^{1+2t}d^{2+2t}t + \\ &+ 2^{2+2t}d^{3+2t}t).\end{aligned} \tag{24}$$

Substituting equation (24) into equation (10) yields the exact analytic expression for the average path length of $M_d(t)$ as

$$\begin{aligned}\bar{D}(t) &= (-1 + 4d - 5d^2 + 2d^3 + 2^{1+t}d^{1+t} - 7\cdot 2^{2t}d^{2+2t} + \\ &+ 3\cdot 2^{1+2t}d^{3+2t} - 2^{1+t}d^{1+t}t + 3\cdot 2^{1+t}d^{2+t}t - \\ &- 2^{2+t}d^{3+t}t - 2^{1+2t}d^{2+2t}t + 2^{2+2t}d^{3+2t}t) \\ &/\ ((-1+2d)(-1+d+2^t d^{1+t})(-1+2^{1+t}d^{1+t})).\end{aligned} \tag{25}$$

Notice that for a large iteration step, $t\to\infty$, $\bar{D}(t) \simeq t \sim \ln|V(t)|$, which shows a logarithmic scaling of the average distance with the order of the graph. As we have a similar behavior for the diameter, the graph is small-world.

4 Conclusion

The graphs $M_d(t)$ introduced and studied here are planar, modular, have a disassortative degree hierarchy and are small-world and scale-free. Another relevant characteristic of the graphs is their clustering zero. A combination of a low clustering coefficient, modularity, and small-world scale-free properties can be found in some real networks, in particular in technical and biological networks [15,14], and most of them are also disassortative. As examples, the largest benchmark considered in [14] –a network with 24097 nodes, 53248 edges, average degree 4.34 and average distance 11.05– has a degree distribution which follows a power-law with exponent 3.0, and it has a small clustering coefficient $C = 0.01$ and other network properties (diameter, average distance, average degree, etc.) are also in the same range than those of the graph $M_6(4)$, see [15] . Also, the *S. cerevisiae* protein-protein interaction network has 1870 and 2240 edges, see [18], and again most of its network parameters, as described in [15] can be compared directly with those of the graph $M_6(3)$.

Finally, we should emphasize that the planar property and the deterministic character of the family, in contrast with more usual probabilistic approaches, should facilitate the exact determination of other network parameters and the development of new network algorithms that then might be extended to real-life complex systems.

Acknowledgments

F. Comellas and A. Miralles are supported by the Ministerio de Educación y Ciencia, Spain, and the European Regional Development Fund under project TEC2005-03575. L. Chen and Z. Zhang are supported by the National Natural Science Foundation of China under Grant No. 60704044.

References

1. D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature **393** 440–442 (1998).
2. A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science **286** 509–512 (1999).
3. C. Song C, S. Havlin, H.A. Makse, Self-similarity of complex networks, Nature **433** 392–395 (2005).
4. A.-L. Barabási, R. Albert, H. Jeong, Mean-field theory for scale-free random networks, Physica A **272** 173–187 (1999).
5. S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, Pseudofractal scale-free web, Phys. Rev. E **65** 066122 (2002).
6. F. Comellas, G. Fertin, A. Raspaud, Recursive graphs with small-world scale-free properties, Phys. Rev. E **69** 037104 (2004).
7. Andrade J S, Herrmann H J, Andrade R F S and da Silva L R 2005 Apollonian Networks: Simultaneously scale-free, small world, Euclidean, space filling, and with matching graphs. Phys. Rev. Lett. **94** 018702 (2005).
8. Doye J P K and Massen C P Self-similar disk packings as model spatial scale-free networks Phys. Rev. E **71** 01628 (2005).
9. Z.Z. Zhang, F. Comellas, G. Fertin, L.L. Rong, High dimensional Apollonian networks, J Phys A: Math Gen **39** 1811–1818 (2006).
10. Chung F, Lu L, Dewey T G and Galas D J 2003 Duplication models for biological networks *J. Comput. Biol.* **10** 677–687.
11. Z.Z. Zhang, S. Zhou, L. Fang, J. Guan, Y. Zhang, Maximal planar scale-free Sierpinski networks with small-world effect and power law strength-degree correlation, Europhys. Lett. **79** 38007 (2007).
12. F. Comellas, Z. Zhang, L. Chen. Self-similar non-clustered planar graphs as models for complex networks. J. Phys. A: Math. Theor. **42** 045103 (2009).
13. F. Comellas, A. Miralles. Modeling complex networks with self-similar outerplanar unclustered graphs. Physica A doi:10.1016/j.physa.2009.02.004 (2009) .
14. R. Ferrer i Cancho, C. Janssen, R.V. Solé, Topology of technology graphs: Small world patterns in electronic circuits, Phys. Rev. E **64** 046119 (2001).

15. M.E.J. Newman, The structure and function of complex networks, SIAM Review **45** 167–256 (2003).
16. R. Diestel, Graph Theory, 3rd. edition, Springer-Verlag, Heidelberg, Germany, 2005.
17. M.E.J. Newman, Assortative mixing in networks, Phys. Rev. Lett. **89** 208701 (2002).
18. H. Jeong, S. Mason, A.-L. Barabási, Z.N. Oltvai, Lethality and centrality in protein networks, Nature **411** 41–42 (2001).