

A Note on the Self-similarity of Some Orthogonal Drawings^{*}

Maurizio Patrignani

Università di Roma Tre, Italy
patrigna@dia.uniroma3.it

Abstract. Large graphs are difficult to browse and to visually explore. This note adds up evidence that some graph drawing techniques, which produce readable layouts when applied to medium-size graphs, yield self-similar patterns when launched on huge graphs. To prove this, we consider the problem of assessing the self-similarity of graph drawings, and measure the box-counting dimension of the output of three algorithms, each using a different approach for producing orthogonal grid drawings with a reduced number of bends.

1 Introduction

The picture of Fig. 1 shows a huge planar connected random graph drawn within the orthogonal grid layout standard, where nodes are points with integer coordinates and edges are sequences of horizontal and vertical segments of integer length. The drawing was obtained by using the well-known *topology-shape-metrics* approach [10, 3]. It can be noticed that the drawing can be roughly split into rectangular blocks, some very dense, and some others much more sparse. Also, the same pattern seems to occur at different zooming levels inside the denser rectangular blocks.

The property of showing the same structure at different scales is referred to as self-similarity and is often associated with the concepts of fractional dimensionality and fractal. These concepts will be more formally defined in Section 2. In the rest of this section we will use the intuitive notion of self-similarity to introduce the main point of this note and the implications of it.

Self-similar structures, characterized by the property of showing similar patterns at all length scales, abound in nature. Lots of examples can be found in disciplines as different as geology (river basins, coastlines, mountains landscapes, etc.), human and animal physiology (blood vessels, nerves, bronchial tubes, etc.), and biology (trail networks for harvesting ants, etc.)

^{*} Work partially supported by European Commission – Fet Open project COSIN – COevolution and Self-organisation In dynamical Networks – IST-2001-33555, by European Commission – Fet Open project DELIS – Dynamically Evolving Large Scale Information Systems – Contract no 001907, by “Progetto ALINWEB: Algoritmica per Internet e per il Web”, MIUR Programmi di Ricerca Scientifica di Rilevante Interesse Nazionale, and by “The Multichannel Adaptive Information Systems (MAIS) Project”, MIUR Fondo per gli Investimenti della Ricerca di Base.

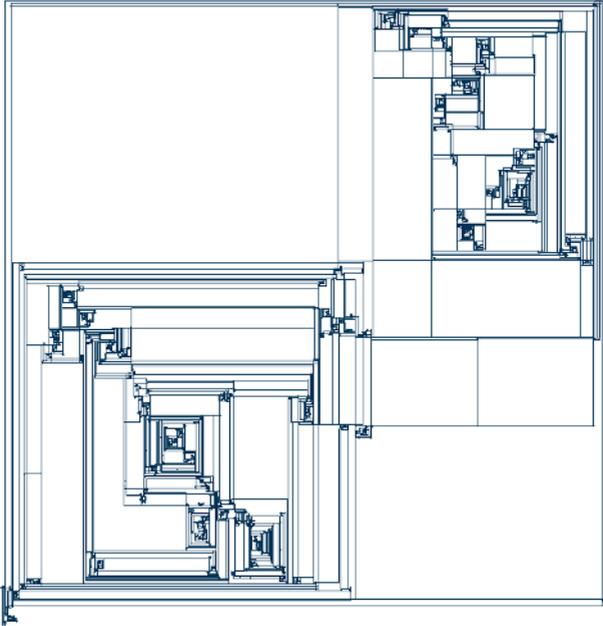


Fig. 1. A drawing with the minimum number of bends of a planar connected graph with 4,700 nodes and 6,155 edges. The number of bends is 2,463.

Although fractals are usually put in relation with chaos, they are sometimes credited with some desirable property caused by their self-similar pattern. However, it is very hard to attribute a fractal structure to a specific optimization problem and such an explicit statement can be searched in vain in the literature. Rigorously defined combinatorial processes, like graph drawing techniques, provide promising examples to study the relationship between fractals and optimization criteria.

This note offers an argument in favor of the hypothesis that self-similarity may be induced by a minimization process. In fact, the line followed by our argument goes the other way around: we start from a well known aesthetic criterion, namely the reduction of the number of bends in orthogonal drawings, and we assess the self-similarity of the produced layouts, a property elusive enough to remain unnoticed for several years.

2 Background

Central to the definition of fractal is that of dimension [4]. The dimension which we generally are familiar with is the so-called topological dimension. The *topological dimension* D_T of a set is always an integer and is zero if the set is totally disconnected, while it is n if the intersection of the set with the boundary of an arbitrarily small neighborhood of one of its points has topological dimension

$n - 1$. Thus, the topological dimension of a collection of points is zero, that of a collection of curves is one, that of a collection of surfaces is two, and so on.

While the topological dimension is based on the concept of intersection, an alternative and more tortuous definition of dimension is based on the concept of cover, and is obtained from the definition of Hausdorff measure. The Hausdorff measure $\mathcal{H}^p(S)$ of a set S is defined parametrically with p , where p is a non negative real. Thus, there are as many Hausdorff measures as many values for p (we will provide a rigorous definition of Hausdorff measure in the next paragraphs.) If there is a value D_H in $[0, \infty]$ such that $\mathcal{H}^p(S) = \infty$ for p in $[0, D_H)$ and $\mathcal{H}^p(S) = 0$ for p in $(D_H, \infty]$, then D_H is the *Hausdorff dimension* of S .

In order to define the Hausdorff measure $\mathcal{H}^p(S)$, used to find the Hausdorff dimension, we need some further definitions. The *diameter* $|U|$ of a set U is the greater distance apart of any pair of its points. A δ -*cover* of a set S is a countable collection of sets $\{U_i\}$ of diameter at most δ whose union contains S . $\mathcal{H}_\delta^p(S)$ is the infimum of the sum of the p -powers of the diameters of a δ -cover of S , i.e., $\mathcal{H}_\delta^p(S) = \inf(\sum_{i=0}^\infty |U_i|^p : \{U_i\} \delta\text{-cover of } S)$. Finally, $\mathcal{H}^p(S) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^p(S)$.

Fractals are defined by Mandelbrot [6] as geometric sets whose Hausdorff dimension D_H strictly exceeds their topological dimension D_T .

As it could be difficult to directly compute the dimension D_H of a set S , often a variant of the above definition is considered imposing the additional constraint that all the sets in $\{U_i\}$ have equal diameter. With this hypothesis, given an arbitrary value p different from D_H , by studying how $\mathcal{H}_\delta^p(S)$ goes to ∞ (when p is in $[0, D_H)$) or to 0 (when p is in $(D_H, \infty]$) for $\delta \rightarrow 0$, we can infer D_H .

In fact, $\mathcal{H}_\delta^p(S) = \inf(\sum_{i=0}^\infty |U_i|^p) = \inf(\delta^{p-D_H} \sum_{i=0}^\infty |U_i|^{D_H}) = \delta^{p-D_H} \mathcal{H}_\delta^{D_H}$. Assume that the diameter of S is finite and let N be the cardinality of the δ -cover that gives the infimum of $\sum_{i=0}^\infty |U_i|^p$. We have that $\mathcal{H}_\delta^p(S) = N\delta^p = \mathcal{H}_\delta^{D_H} \delta^{p-D_H}$. Thus, for $\delta \rightarrow \infty$, $N = C\delta^{-D_H}$, where $C = \mathcal{H}^{D_H}(S)$. It follows that measuring the exponent of the power law that relates the number N of covering sets to their diameter δ gives indirectly the value of the dimension we are searching for. Based on these considerations, a simple method, called *box-counting*, can be devised to compute a dimension D_B , which is usually near to D_H . It consists of partitioning the plane into equal-sized squares of side δ , and of plotting on a log-log scale the number N of squares intersecting S with respect to δ . The slope of this curve gives D_B .

3 Measuring the Self-similarity of Orthogonal Drawings

In order to compute the box-counting dimension of a graph drawing, we use the tool by Leejay Wu and by Christos Faloutsos [11], which in addition to being well known and widely used, is fast and rigorously designed. Since the input of such a tool is a geometric object consisting of a collection of points, we replaced each node with a point and each edge with a dotted line (see Figs 2.a and 2.b).

Fig. 2.c shows the box-counting log-log plot, produced by the package [11]. Between points B and C the -1.64 slope of the curve gives the fractal dimension induced by the self-similarity of the “holes” in the drawing. Note that before point A , as each non empty box covers a single point of the drawing, the dimen-

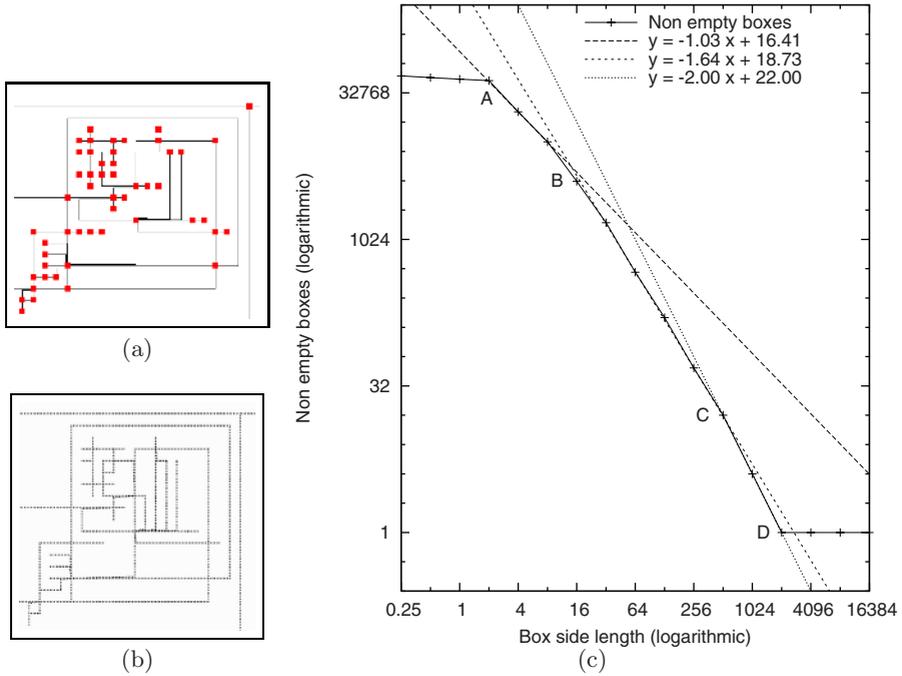


Fig. 2. (a) A portion of the drawing of a graph and (b) the same portion where segments have been replaced with dots (five dots for each grid segment). (c) The box counting log-log plot for a drawing of a graph with 485 vertices.

sion appears to be zero. Between points *A* and *B*, since a non empty box rarely intersects more than a single segment, the line with slope -1.03 is measuring the dimension of a collection of segments. Between points *C* and *D*, the line with slope -2.00 is measuring the dimension of a continuous surface, since boxes are largest than the largest “holes” of the drawing. Finally, after point *D*, the whole drawing fits into a single box, and the slope of the line is zero again.

Using Pigale generator [2], we created three test suites of planar connected, biconnected and triconnected graphs ranging from 500 to 3,000 edges, increasing each time by 500 edges, 10 graphs for each type. After the generation multiple edges and self-loops were removed. Fig. 3.a shows the number of edges before and after simplification. For all the three test suites the number of nodes is roughly half of the number of edges before simplification.

Orthogonal layouts were produced by applying three different algorithms corresponding to the three approaches described hereunder.

OFV: The *orthogonal-from-visibility* approach consists of constructing a planar orthogonal grid drawing of a biconnected planar graph starting from a visibility representation of the same graph [3]. Our implementation handles graphs of degree greater than four inserting at most two bends per edge.

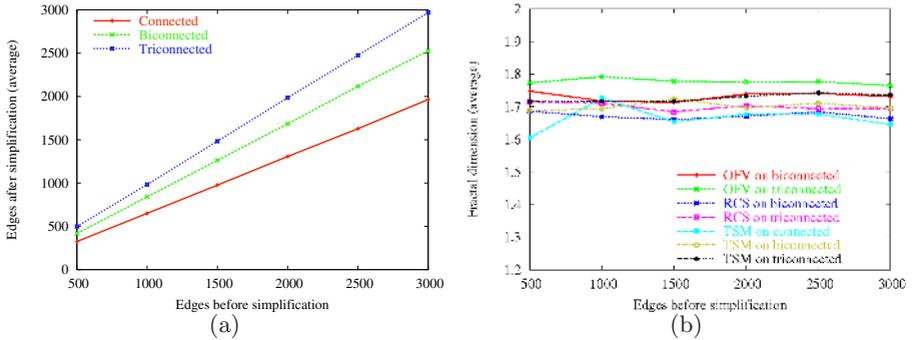


Fig. 3. (a) Average number of edges before and after simplification for the three test suites. (b) The D_B fractal dimensions for orthogonal drawings computed with various approaches.

RCS: The *relative-coordinates scenario* is based on the incremental construction of the drawings. We applied the ‘simple algorithm’ described in [7] for drawing high degree biconnected graphs. Intersections may be introduced even for planar graphs. Each edge is guaranteed to have a single bend.

TSM: For the *topology-shape-metrics* approach we used the Boyer and Myrvold algorithm [1] to compute a planar embedding. During the subsequent orthogonalization step, the total number of bends was minimized by using the algorithm in [10], modified as described in [5] for handling high degree graphs. Finally, the compaction step was performed with the heuristic based on the rectangularization of the faces [10].

As it can be seen from Fig. 3.b, the D_B fractal dimension computed for the various kinds of graphs is consistently above the D_T of the point sets, which is one. Hence, the drawings are actually fractals, and their fractal dimension is about 1.7 for all types of graphs and for all the three approaches. In all cases the correlation computed by package [11] for the line approximating the curve on the log-log graph is less than -0.999 , suggesting a high reliability for the computed D_B values.

4 Conclusions and Open Problems

Although it is well known that random graphs do not exhibit self-similar properties, we wanted to make sure that the input graphs were not self-similar themselves. The self-similarity of graphs and networks has been widely studied. Several real-life networks, for example, are known to show a self-similar node-degree distribution [9, 12]. We plotted the node-degree distribution for the three test-suites and verified on a log-log scale that the curves are not power laws.

Also, we wanted to check whether the planarization step, exploited by both the **OFV** and the **TSM** approaches, folded the faces one into the other, introducing self-similarity. Again, we found that the node-degree distribution of the dual graphs does not follow a power law.

Finally, we tried alternative compaction heuristics for the **TSM** approach, without noticing significant changes of the D_B fractal dimension computed.

Some other issues deserve a deeper investigation: Is there a way to avoid self-similar patterns in orthogonal grid drawings by introducing a few extra bends? Can alternative measures of fractal dimension, like the *correlation dimension* [8], help deepening our understanding of this phenomenon? Do other graph drawing standards also produce self-similar drawings of large graphs?

Acknowledgments

We would like to thank Paolo Branchesi, Giuseppe Di Battista, Maurizio Pizzonia, and Roberto Tamassia for encouragement and interesting discussions.

References

1. J. Boyer and W. Myrvold. Stop minding your P's and Q's: A simplified $O(n)$ planar embedding algorithm. In *10th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 140–146, 1999.
2. H. de Fraysseix and P. Ossona de Mendez. P.I.G.A.L.E – Public Implementation of a Graph Algorithm Library and Editor. SourceForge project page <http://sourceforge.net/projects/pigale>
3. G. Di Battista, P. Eades, R. Tamassia, and I. G. Tollis. *Graph Drawing*. Prentice Hall, Upper Saddle River, NJ, 1999.
4. K. Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley & Sons Ltd, second edition, 2003.
5. U. Fößmeier and M. Kaufmann. Drawing high degree graphs with low bend numbers. In F. J. Brandenburg, editor, *Graph Drawing (Proc. GD '95)*, volume 1027 of *Lecture Notes Comput. Sci.*, pages 254–266. Springer-Verlag, 1996.
6. B. B. Mandelbrot. *The Fractal Geometry of Nature*. W. H. Freeman and Company, New York, 1977.
7. A. Papakostas and I. Tollis. Efficient orthogonal drawings of high degree graphs. *Algorithmica*, 26:100–125, 2000.
8. M. Schroeder. *Fractals, chaos, power laws: minutes from an infinite paradise*. W. H. Freeman and Company, New York, 1991.
9. G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos. Power laws and the as-level internet topology. In *IEEE/ACM Transactions on Networking (TON)*, volume 11, pages 514–524. ACM Press, 2003.
10. R. Tamassia. On embedding a graph in the grid with the minimum number of bends. *SIAM J. Comput.*, 16(3):421–444, 1987.
11. L. Wu and C. Faloutsos. FracDim, jan 2001. Perl package available at <http://www.andrew.cmu.edu/~lw2j/downloads.html>
12. S.-H. Yook, H. Jeong, and A.-L. Barabasi. Modeling the Internet's large scale topology. In *Proc. PNAS'02*, 2002.