



A topological analysis of a geographical model for internet

André Franceschi de Angelis¹, Gonzalo Travieso²,
Carlos Antônio Ruggiero² and Luciano da Fontoura Costa²

Manuscript received on November 10, 2008 / accepted on January 10, 2009

ABSTRACT

This paper presents the analysis of a complex networks model of Internet connectivity named “Geometric Growing Model with n Redundant Edges” (GGM-RE-n). This approach models the growth of Internet while taking into account the physical position of the nodes. The node degree distribution and average radius are investigated for different configurations of the networks. The results provide insights about the progressive evolution of the networks as new edges are added, as well as about the effect of the locality factor over the degree distribution and average radius.

Keywords: Internet, complex networks.

Correspondence to: André Franceschi de Angelis

¹Centro Superior de Educação Tecnológica, Universidade Estadual de Campinas, Rua Paschoal Marmo, 1888, 13484-370 Limeira, SP, Brazil.

E-mail: andre@ceset.unicamp.br

²Instituto de Física de São Carlos, Universidade de São Paulo, Av. do Trabalhador São-Carlense, 400, Caixa Postal 369, 13560-970 São Carlos, SP, Brazil.

E-mails: gonzalo@ifsc.usp.br / toto@ifsc.usp.br / luciano@ifsc.usp.br

1 INTRODUCTION

The “Growing Geometric Model with n Redundant Edges” (GGM-RE- n) [1] was created with the objective of correctly represent the Internet physical infrastructure evolution in developing regions. The model takes care of connections locality and physical constraints imposed by landscape geography on the network evolution. Once the geometry of the underlying space is defined, no additional parameters are required in order to define the GGM evolution. Instead of assuming a pre-specified spatial distribution of all nodes (as in previous geometrical complex networks models), the GGM model involves the progressive incorporation of new nodes, which are connected to the closest existing nodes. Such a dynamics is expected to emulate, to some accuracy, the historical evolution of the Internet in developing regions where, starting from an initial point, the Internet is progressively extended through the addition of new points connected to the closest existing outlet. More specifically, the GGM model involves the following steps:

- (i) the definition of an underlying spatial region (a two dimensional unit square is adopted in the present work);
- (ii) the choice of one of its points as the initial node;
- (iii) the random choice of a new point (uniformly, in the present work) within the underlying region and its connection to the closest existing node; and
- (iv) the previous step is repeated until the desired number of nodes is reached.

Figure 1 shows a GGM network with 1000 nodes and tree topology. There are no redundant edges in this network. There are no cycles. It is possible to see some long range links and several short distance connections between nodes.

The Internet contains a number of cycles. The GGM-RE model can emulate this behavior through the addition of redundant edges. We express the fact that a model has an $n\%$ increase in the number of edges with respect to the GGM model by the abbreviation GGM-RE- n , such that GGM-RE-0 corresponds to the initial model. Figure 2 presents a GGM-RE-20 network with 1000 nodes and 200 redundant edges. In this sample, geographical constraints were not taken into account for the redundant connections. Due chosen parameters, there are more long distance links than observed in real samples of Internet.

We improved our model with the introduction of the “locality factor”, a parameter which takes into account the importance of geographical distance in the establishment of connections.

Figure 3 shows a GGM-RE-20 network with 1000 nodes, 200 redundant edges and locality factor 5. There is a smaller number of long distance links than in the previous example. Most redundant edges are short. Neighbor nodes have higher connection probability than distant ones.

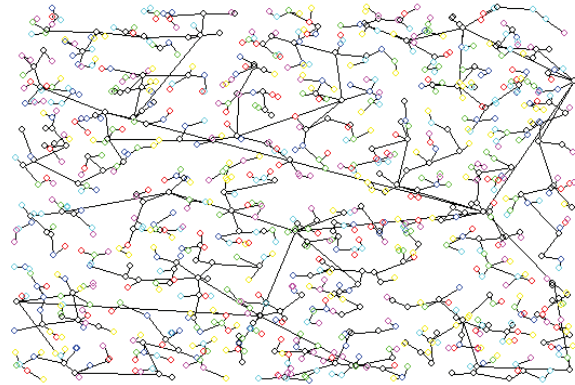


Figure 1 – A GGM-RE-0 network with 1000 nodes.

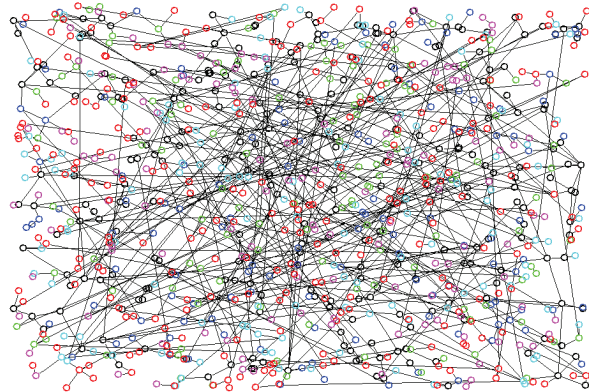


Figure 2 – A GGM-RE-20 network with 1000 nodes, 200 redundant edges and locality factor 0.

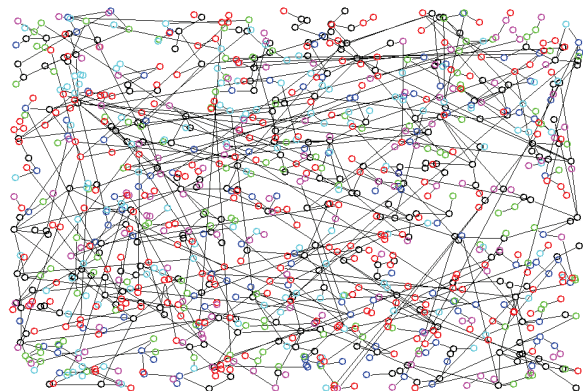


Figure 3 – A GGM-RE-20 network with 1000 nodes, 200 redundant edges and locality factor 5.

2 PREVIOUS RESEARCH

In a previous research [1], the GGM-RE-n model was used to study the effect of geographical constraints over the execution of tasks in complex networks [2], taking into account computational grids built over the Internet. The investigation focused on the distribution dynamics and task processing on the grid. The simulations partitioned the initial problem into independent tasks of the same duration, executed by the nodes, under coordination of a master node. The communication costs were proportional to the number of links traversed between the master and each slave node. The time spent by the communication through a link was used as the simulation time unit. The grid general performance was directly affected by the network topology [3]. It was demonstrated that the topology influences small tasks, in contrast to its relatively small importance for long duration tasks. The improvement of grid efficiency by the inclusion of redundant edges was determined based on the Barabási-Albert model [4] used as an additional comparison reference. The data showed that higher locality of connections tends to reduce the processing performance and that the spatial constraints represent a limiting factor on network efficiency. Figure 4 shows a summary of these results.

3 RESEARCH

This work is concerned with topological properties of GGM-RE-n complex networks. We intend to compare several measures of

these networks to Brazilian Internet samples to evaluate the model fitness. In this paper, we investigate node degree distribution and average radius of the GGM-RE-n networks. The independent variables we used were percentage of redundant edges addition and locality factor. We used Scilab 4.1.2 [5] to generate networks and to perform a set of calculations over them. Analytical solutions for these measures were not determined. All networks were built with 1000 nodes and we considered average values for 50 samples in each configuration.

We performed calculations with 20, 40, 60, 80 and 100% of redundant edges. It means 200 to 1000 new edges added to each 1000 nodes network. The original GGM tree becomes a much more connected graph as we proceed with this process. It is clear that higher values of redundant edges are not practical due to associated infra-structure costs.

Locality factor is an important parameter for the edges addition algorithm. Low values of locality generate a larger set of long range edges into network. Higher values produce short distance connections and highlight the geographical constraints. We used 0, 1, 10, 50 and 100 for locality factor.

4 RESULTS

Considering node degree distribution, no hubs have been found in the networks. The maximum node degree observed during experiments was 27 which occurred once. No zero degree nodes have

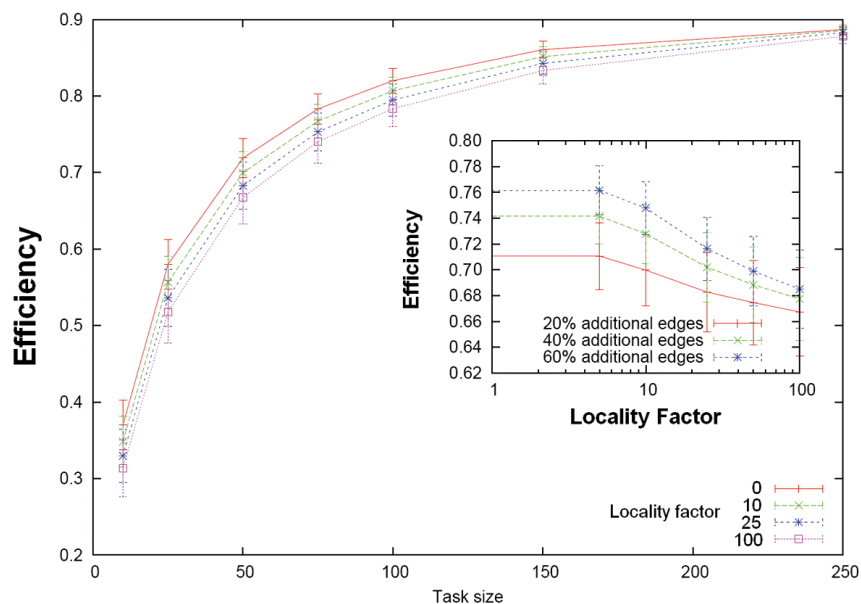


Figure 4 – Parallel average efficiency $\langle E \rangle$ as a function of size L of tasks for the GGM-RE model with 20% additional edges and locality factor from 0 to 100. The inset shows the effect of the locality factor on efficiency, for different connectivities and $L = 50$.

been found; this is to be expected considering the construction algorithm of GGM networks. We noted a small influence of the locality factor on node degrees distribution (Fig. 5). The average degree of the network depends on the redundant edge percentage and it is invariant due to the locality. No power-laws were identified for the node degrees of GGM-RE-n networks.

The center of a graph is defined as the node for which the largest of the shortest paths to all the other nodes is minimal. Radius is defined as the maximum path length from center to another node. Redundant edges introduce cycles into network;

as a consequence the radius should decrease. Small values of locality lead to lower average radius due to the existence of longer links. Figure 6 shows the average radius for the examined networks.

5 CONCLUSIONS

We have generated a set of GGM-RE-n networks and have performed calculations over it. Our goal was to investigate the topological properties of these networks. We found that the addition of redundant edges: a) reduces the number of terminal nodes;

Nodes Degrees

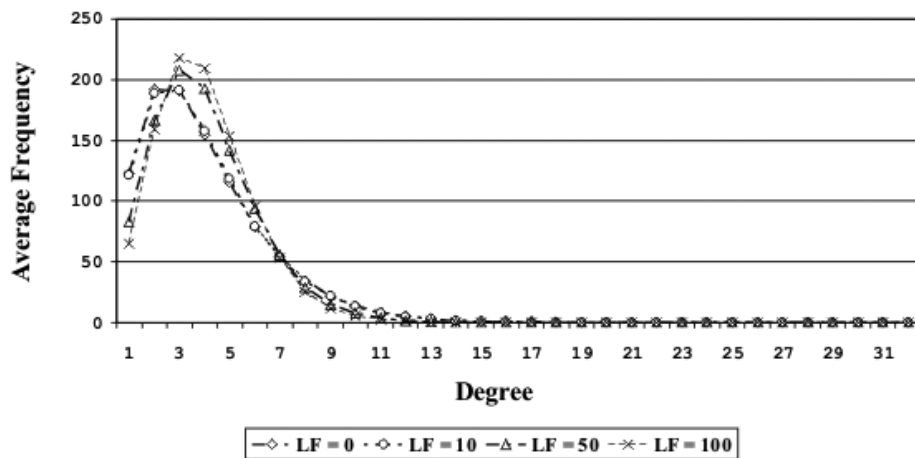


Figure 5 – Average frequency of degree values for 50 GGM-RE-n networks with 1000 nodes; 100% of additional redundant edges and Locality Factor (LF) 0, 10, 50 and 100.

Average Radius

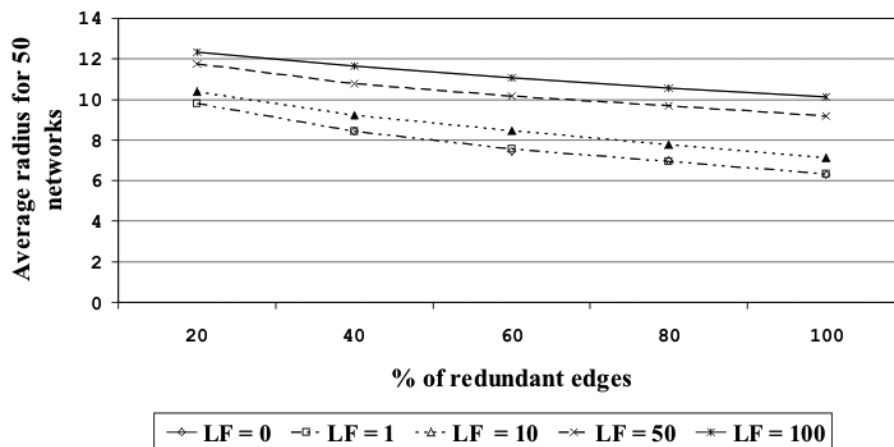


Figure 6 – Average radius for 50 GGM-RE-n networks with 1000 nodes; 20, 40, 60, 80 and 100% of additional redundant edges and Locality Factor (LF) 0, 1, 10, 50 and 100.

b) does not produce hubs; c) improves the frequency of nodes with degrees between 2 and 9; d) shifts the frequency curves slightly to the right.

Locality factor improvement: a) increases the number of nodes with degrees between 3 and 6; b) causes a small and non-linear reduction of maximum network degree; c) increases the average radius; d) does not significantly alter the node degree curves.

Future research will involve the analytical description of the observed behavior and the determination of the model fitness to represent the Internet in regions of interest.

ACKNOWLEDGMENTS

Carlos Antônio Ruggiero and André F. de Angelis acknowledge FAPESP (2003/08269-7) for financial support. Luciano da F.

Costa acknowledges FAPESP (05/00587-5) and CNPq (301303/06-1) for financial support.

REFERENCES

- [1] DE ANGELIS AF, COSTA LF, TRAVIESO G & RUGGIERO CA. 2008. *International Journal of Modern Physics C*, 19(6): 1–7.
- [2] NEWMAN MEJ. 2003. *SIAM Review*, 45: 167.
- [3] COSTA LF, TRAVIESO G & RUGGIERO CA. 2005. *Eur. Phys. Journal B*, pp. 119–128.
- [4] BARABÁSI A-L & ALBERT R. 1997. *Science*, 286: 509.
- [5] SCILAB. The open source platform for numerical computation. <<http://www.scilab.org/>>.