IJCSS International Journal of Complex Systems in Science

Int. J. Complex Systems in Science
vol. 1 (2011), pp. 37–41

Structural Navigability on Complex Networks

Pau Erola*, Sergio Gómez* and Alex Arenas*

* *Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili*

Abstract. The famous Milgram's small-world experiment revealed that there is something special in the structure of natural and man-made complex systems: without a global view of the network, a message can be routed efficiently between any pair of nodes. Our initial hypothesis is that the community structure, that provides meaningful insights on the structure and function of complex networks, is an important actor in these routing properties. To exploit the modular structure of networks we need to analyze the contribution of each node to the modules. Unfortunately, this analysis involves a huge amount of data. To reduce this problem we propose to build a map using the linear projection theory as a basis of a guided routing. First we project the matrix of contributions of each node of a given network to its modules in a plane using the Truncated Singular Value Decomposition. This two-dimensional plane reveals the structure of modules and their boundaries and we will use it as the map for navigating through the network. Considering that each node only has knowledge about its neighbors, we define a simple greedy routing algorithm to guide the communication among them. We apply our framework to the Internet Autonomous Systems (ASs) network achieving, in high percentage, close to optimal paths.

Keywords: complex networks, singular value decomposition, routing, Internet
MSC 2000: 68M10, 05C75, 15A18

1. Introduction

The famous Milgram's small-world experiment revealed that there is something special in the structure of complex systems [6]: without a global view of the network, a message can be routed efficiently between any pair of nodes in many strongly clustered networks [2]. From a computational perspective, if we have a coordinate system that reflects these routing properties, and a decision algorithm, we must be able to calculate a more or less optimal path. Here we present a methodology to obtain a navigable map of a complex network and how to guide the routing process through it.

Our initial hypothesis is that the community structure, that proves meaningful insights on the structure and function of complex networks, is an important actor in these routing properties. To exploit this, we propose to analyze the contribution of each node to the modules using the linear projection technique presented by Arenas *et al.* [3]. With it, we construct a map to guide the communication among nodes using a simple greedy routing algorithm.

Previous work on understanding the routing information phenomenon was addressed by Boguñá *et al.* [4]. They propose to fit the network structure on an underlaying hidden metric space where the geodesics between to nodes correspond to their shortest path. Here we propose an alternative to the hyperbolic embedding presented by the author using the mesoscopic structure of the network.

This paper is organized as follows. In Section 2 we review the methodology used to construct our map. Section 3 presents our greedy routing approach to surf a network through its structure. Finally, in section 4 we apply our framework to the ASs network.

2. Mapping the modular structure

The analysis of the interrelations between nodes and modules involves a huge amount of data. To get a tractable data set we need to reduce the dimensionality of the problem with a desirable minimum information loss. The methodology proposed by Arenas *et al.* [3] achieves this objective: using the Truncated Singular Value Decomposition they construct a tractable two-dimensional map that reflects the structure of the modules and their interrelations.

In order to project the structure of the network on a plane, first we need to compute the contribution matrix $C_{i\alpha}$ between of each node to its modules. When making the community detection, we must be confident that there is a balance between the number of communities and their size. According to Adamic *et al.* [1], if we consider the intramodular navigation as a search using high degree nodes, the number of steps required to achieve a successful path is $N^{2-\frac{4}{\tau}}$ for a community of N nodes. For high values of N, the performance of our routing will be lower due to the difficulty to find the destination.

The TSVD of the matrix $C_{i\alpha}$ projects the structure of individual modules in a plane $\mathcal{U}_2$. For each of these projections we calculate the polar coordinates (R_n, θ_n) where R_n is the length of the contribution projection vector $\tilde{v}_n$, and θ_n is the angle between $\tilde{v}_n$ and the hortizontal axis. To interpret correctly this outcome we need to know also the intramodular projection $\tilde{e}_\alpha$ of each module, the distinguished direction line of the projection of its internal nodes. With these parameters, we can compute a new pair (R_n, ϕ_n), where $\phi_n = |\theta_n - \theta_{\tilde{e}_\alpha}|$, and the new values $R_{\text{int}} = R \sin\phi$ and $R_{\text{ext}} = R\cos\phi$. R_{int} informs about the internal contribution of nodes to their corresponding modules, and R_{ext} reflects the boundary structure of modules. Both values, R_{int} and R_{ext}, and the angle θ_n are the basis of our routing algorithm described below.

3. Routing algorithm

Finding the shortest paths between nodes using only local information is a big deal. We propose the use of a greedy algorithm to choose the locally optimal step at each stage using the constructed map.

Let us assume we want to go from node i to node j, and let N_i refer to the neighbors of node i. For each node $k \in N_i$ we compute a metaheuristic cost function and select the candidate that minimizes it. This process is repeated until the destination is reached, the current node i does not find a feasible successor or a time constraint is violated. We do not allow loops.

The cost function we use is the following:

$$\text{cost}_k = \begin{cases} \beta\left(\dfrac{\lambda + |\Delta\theta_{k\to j}|}{R_{\text{int}_k}}\right) & \text{if } k \in \alpha_j, \\ \dfrac{|\Delta\theta_{k\to j}|}{R_{\text{ext}_k}} & \text{if } k \notin \alpha_j \ \text{ and } \ R_{\text{ext}_k} > 0, \\ \dfrac{\lambda + |\Delta\theta_{k\to j}|}{R_{\text{int}_k}} & \text{otherwise,} \end{cases} \tag{1}$$

where $|\Delta\theta_{k\to j}|$ is the angular distance between the neighbor node k and the destination node j; R_{int_k} and R_{ext_k} are the intra- and intermodular contribution projections of each node in N_i respectively; α_j is the destination community; and the constants λ and β allow us to control the weight of each term of the equation.

The cost function sets two scenarios, when our neighbor k belongs to the same community α_j than the destination node j, and when not. If $k \in \alpha_j$ the cost function prioritizes the nodes closest in community connectivity and those who have more internal links. Since $|\Delta\theta_{k\to j}|$ is small for nodes within the same community, we add a shift value λ to reduce its fast annealing. We also use the weight constant $\beta < 1$ in order to not leave the community α_j due to the attraction of external hubs not in α_j, which otherwise would be given larger priority.

A different case is when $k \notin \alpha_j$. Then, the function seeks the boundary node that is highly connected externally and is closer to the communities to whom j is connected to. Since nodes N_i may be only internally connected to their community α_i ($R_{\text{ext}_k} = 0$) we consider these cases using only their intramodular contribution projections.

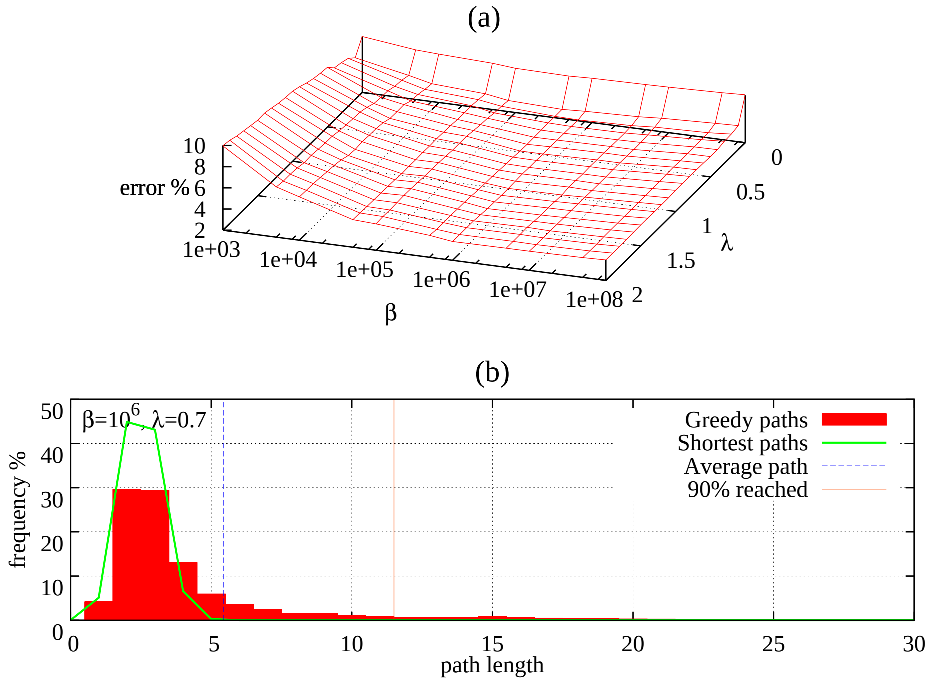

Figure 1: The top plot (a) shows the routing unsuccessful paths as a function of the λ and β parameters. The bottom figure (b) shows the distribution of the path lengths of 10^5 routes compared with the shortest path distribution.

4. Application to a real network

Here we apply the described routing technique to the Internet ASs network. Internet is a collection of more than 23,000 computer networks each known as an autonomous system (AS). In the last few years, Internet is experiencing an explosive growth that is compromising its scalability [7]. With our local navigability, we present an alternative to the Border Gateway Protocol that requires an unnecessary huge amount of data exchange to maintain an updated view of the network topology.

To apply the proposed mapping, first we need to detect the communities of the network. To conduct this, we use mainly the Extremal Optimization [5] algorithm. We apply it recursively to obtain communities of the desired size, resulting in 241 communities. Then we construct the $C_{i\alpha}$ matrix and project it on a plane using the TSVD. With these projections, we compute the values R_{int_n}, R_{ext_n}, and θ_n for each node.

Our test set consists of a 10^5 node pairs randomly selected. The constants λ and β were adjusted to the values that minimize the percentage of unsuccessful paths, as is showed in Figure 1a. Our greedy routing algorithm achieves 96.5% success with an average path of 5.4 steps. The Figure 1b presents the path length distribution of the performed paths. Note that the average path achieved is far from the median value due to the long tail of the distribution. Almost 70% of the paths are below the average optimal path length.

In summary, the experimental results reinforce our hypothesis that the community structure is pervasive in the routing properties of the complex systems. Our greedy algorithm establishes the mechanism to navigate by the structure of the network explaining and simplifying the routing process.

Acknowledgements

This work has been supported by the Spanish MICINN FIS2009-13730-C02-02, and the Generalitat de Catalunya 2009-SGR-838. PE acknowledges a URV PhD grant.

References

[1] A.L. Adamic, R.M Lukose and B.A. Huberman, *arXiv:cond-mat/0204181v2* (2002).

[2] R. Albert and A.L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).

[3] A. Arenas, J. Borge-Holthoefer, S. Gomez and G. Zamora-Lopez, *New Journal of Physics* **12**, 053009 (2009).

[4] M. Boguñá, D. Krioukov and kc claffy, *Nature Phys.* **5**, 74 (2008).

[5] J. Duch and A. Arenas, *Physical Review E* **72**, 027104 (2005).

[6] J. M. Kleinberg, *Nature* **406**, 845 (2000).

[7] D. Meyer, L. Zhang, K. Falls and A. Brady, *RFC4984*, The Internet Architecture Board (2000).