Title: Temporal evolution of the structure of social networks

Director: Dr. Àlex Arenas Moreno

Co-director: Dr. Sergio Gómez Jiménez

Research group: Algorithms Embedded in Physical Systems

Masters: MIA; MEIS - Sistemes Intel·ligents

Description:

The study of complex networks has become a research field of great interest due to its presence in all kind of situations (e.g. see review [1]). For example, there are complex networks in fields as diverse as biology (metabolic networks, neural networks), ecology (food webs), sociology (collaboration networks), finance (trade networks), infrastructures (transportation networks, distribution networks), technology (P2P, Internet, WWW) or psychology (word association networks), to name a few. A complex network is a graph with certain non-trivial topological properties, which do not appear in other simple network models. Some of these properties are: scale-free degree distributions, high clustering coefficients, assortativity (or dissortativity), hierarchical structures, and community structures. On the other hand, simple networks such as random networks or grids have very homogeneous nodes, which invalidate them as models of reality.

In our research group we have been studying the community structure of complex networks for a long time. A community could be defined as a set of nodes highly linked between them, but poorly linked to the rest of the network. There are many ways of defining communities, each one having their own pros and cons. The most successful was introduced by Newman and Girvan [2], based on the definition of a magnitude called modularity. Given a partition of the network in communities, modularity measures the difference between the number of internal links and the number of links which would correspond to a random network (preserving the original degrees of all nodes). The larger the modularity, the better the classification of nodes in communities.

In most cases, the study of community structure has been applied to static networks, ignoring the temporal evolution of the topology of the network. This is an excessive simplification, since many networks, in particular social networks, are always changing. An exception could be [3], in which the temporal evolution of a collaboration network is analyzed.

The objective of this research project is to introduce our group in the study of the temporal evolution of the structure of social networks. First, we will have to start with the most basic ingredient: the data. Taking advantage of the existence of different data sources in Internet, we will have to be able to extract the data, save them in a database, and prepare the mechanisms to build time-evolving networks. The system will have to be flexible enough to cope with different kinds of data. For instance, we want to obtain scientific collaboration networks from arXiv [4] or scientific publications of a certain field, or networks of actors participating in the same films from the Internet Movie Database [5].

Once we have the time-evolving networks, the next step will be the study of the evolution of their basic structural parameters. Finally, we will start the analysis of the temporal evolution of the community structure. The degree of development of this last part will depend on the success in the previous stages.

The working plan is the following:

1. Introduction to the problem: study of the basic bibliography on complex networks, structural parameters, and community structure.
2. Design and implementation of a database to store the data, from which it will be possible to extract complex networks at different time windows.
3. Analysis of the webs which will be the data sources (arXiv, IMDb, etc.), and design, implementation and test of the crawlers or methods to extract the information and save them in the previous database.
4. Design, implementation and test of the tools to query the database and generate the time-evolving complex networks.
5. Analysis of the evolution of the main structural parameters of the generated complex networks.
6. Adaptation of our community detection algorithms to these time-evolving networks.
7. Analysis of the evolution of the communities' structure.

References:

[1] M. E. J. Newman, The structure and function of complex networks, SIAM Review 45 (2003) 167-256; arXiv:cond-mat/0303516
[2] M. E. J. Newman and M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2004) 026113
[3] M. Ausloos and R. Lambiotte, Eur. Phys. J. B 57 (2007) 89-94
[4] http://arxiv.org
[5] http://www.imdb.com/