STATISTICAL METHODS FOR SOCIAL NETWORK ANALYSIS WITH APPLICATIONS IN ECONOMICS

Carlo Drago

Available at: http://works.bepress.com/carlo_drago/126/
Statistical Methods for Social Network Analysis with Applications in Economics

Carlo Drago

September 2012

Working Paper 1/1
Statistical Methods for Social Network Analysis with Applications in Economics

Carlo Drago
Department of Mathematics and Statistics,
University of Napoli “Federico II”,
80126, Italy.
Email: carlo.drago@unina.it

5 September 2012

JEL classification: C0,C2,C30,C38,C60,C63,C88
Statistical Methods for Social Network Analysis with Applications in Economics

Abstract
In this research project we show the aims and the goal for the research in the academic year 2012/2013

Keywords: Social Network Analysis
1 Aims of the Project: Analysing Network Structure and Performances

There was in last years a growing interest in analysing and modelling social relational data (DiEsposito & Zaccarin 2011, Goldenberg et al. 2010). Important results in this field in empirical applications was obtained as well by recent advances in statistical modeling of random networks which allow to analyse more in depth contemporary networks (Snijders et al. 2006, Handcock et al. 2008, Handcock & Gile 2010, Snijders et al. 2010, Morris et al. n.d.). In this way the analysis of relational data was becoming increasingly important in various predictive contexts (Otte & Rousseau 2002, Bell et al. 2002, Asur & Huberman 2010).

In particular in the Social Sciences and in Finance there was important works which consider the growth (or the declining) of the sectors to the social network structure. So the relevance of the social network analysis cannot be undervalued. A first clear example can be seen in economic networks in which innovation and knowledge can flows over the networks and can produce social capital (Potts et al. 2008, Kim 2012). The social networks act in this sense like spillovers in producing innovations and technological advances (Sorenson & Singh 2007, Owen-Smith & Powell 2004) In these context the structure of the network and the internal structures of coordination could be decisive in producing social capital (Lin 1999) and creating an higher performance in terms of outcomes like economic growth.

In other cases the network structure can produce some disruptive effects by considering the internal existence of multiple connections between some central actors. In this sense a particular network configuration can spread some phenomena like epidemics or financial crises. In financial context, for example, there is a recent work of Bargigli & Gallegati (2012) which detect and analyse the structure of network communities in the credit sector as a direct source of systemic risk. In other case the different network models can be cause of financial instability Nier et al. (2007). Another relevant application related to the banking sector was recently presented by (Battiston et al. 2012, Gai & Kapadia 2010, Cassar et al. 2001) and detect the relevancy of the network structure in the dynamics of the the financial crises.

In other cases the network structure can be considered related with the performances of the economic actors in the network which can be considered playing a strategical game in which networks can spread information and strategies. Drago et al. (2011), Santella et al. (2007, 2008) analyzes the impact of the structure of the interlocking directorship network in Italy and in Europe by considering the effects of the gender attributes on the economic performances of the
firms (Drago et al. 2011, 2012).

In Social Sciences various recent works has detected specific relationships between the network evolution and structure Borgatti & Everett (1992) and the economic growth (Giles 2012, Butler & Hansen 1991). By considering the framework presented in Jackson (2008) (see also Del Monte (1992) there are very relevant open questions in literature which could be considered. In particular it is possible to consider the effects of the network structure and the effects of the cooperation of network actors (evolutive communities see Li et al. (2008), Gulati (1995)) on the economic growth. Is the cooperation between the different actors important? Is possible to detect and to measure this type of phenomenon? Is possible to measure the impact of public policies on this type of cooperation?

At the same time is possible to consider real problems in economic systems which can be solved by using social network analysis. Another open problem in literature is the analysis can be the impact of the public policies on the network structure and on this way the effects on the economic growth. At the same time can be considered the problem to measure of the performances in network optimal communities (with applications for example in finding the best firm aggregations or industrial districts). Finding statistical methods to detect the optimal sub-networks (and measuring their performances) can be useful in designing incentives in a public policy scheme. At the same time another relevant point can be to analyse the network structures and his dynamics to discriminate the “real” sub-networks (or the network communities) in order to detect the dynamics of the cooperation. As well another important application can be considering the analysis of the impact of public policy measures on the sub-networks. By considered these motivations it is necessary to propose new statistical methodologies to extract some latent information in the networks which can be useful to understand the network performances. Various statistical research proposals can be considered to reach these goals. The application on real cases is an important part of the research process. These research problems are not exhaustive of existant problems but reflects some relevant points open in literature and the general point of view it is considered in the research.

2 Development of the Project

2.1 Statistical Methods

The development of the research is based on two lines: working on statistical new methods and working toward real data applications. So the principal objective is to innovating the literature
by proposing statistical new methods and new algorithms by showing as well their properties. First of all there can be considered methods to detecting the economic aggregations in the considered networks (in literature also defined like as communities). Secondly can be considered methods for the analysis of the dynamics of the networks. Thirdly can be considered the methods for analyzing the impact of the public policies. A fourth objective is to analyze large networks and considering some statistical methods to handle these types of networks. For example methods which can allow to show the “relevant” part of the network like pruning. A fourth objective is network modelling considering some relevant explanatory variables.

The innovative statistical methods are compared with the methods proposed in the recent literature considering both a theoretical level and analysing their properties and characteristics. As well the methods will be tested and compared by intensively experimentating them by simulation experiments. Simulated data can be referred to existant networks and their known empirical characteristics. The application on real data is another way to show the usefulness of the considered methods and the proposals.

All the innovative methods can be developed by programming as R functions and packages (or using another suitable language) and allow to be replicated on different data and problems.

2.2 Applications

Applications are defined on real cases on real data. In this sense real problems are considered. Possible applications can be related but are conditioned to the data availability. Possible examples not excluding other possibilities of research are:

1. Industrial districts
2. Geographical networks and road networks
3. Interlocking directorship networks
4. Ownership networks
5. Migration networks
6. Social media networks
7. Documents networks based on textual data
At the same time the possibility to use new methodologies and new algorithms developed in a language as R, allow the collaboration between different researchers (for example REPOS researchers working on specific topics) on common problems.

### 2.3 Software

The methods will be proposed both theoretically and using the R language ([R Development Core Team 2011](https://www.r-project.org)). The original software can be transformed in an original R package and the methods can be usefully applied to different networks and datasets over time and space.
References


Del Monte, A. (1992), Recent developments in the theory of industrial organization, Univ of Michigan Pr.


**URL:** [http://www.R-project.org/](http://www.R-project.org/)


