COSTNET Training Event, München, February 13, 2019

Statistical methods for longitudinal network analysis

Tom A.B. Snijders
University of Groningen
University of Oxford

Networks are relational structures between nodes, represented mathematically by graphs and directed graphs, possibly with values on the nodes and/or the edges. Mathematical models for networks have been constructed in many different ways. Physicists, computer scientists, economists, social scientists, and biologists tend to have, to some extent, differently structured network data and different research questions. In addition, purposes of analysis may be different, aiming either at mathematical aggregate properties or at inference about mechanisms generating network data – or at still other goals. This course day will present models for networks for the purpose of statistical inference and inspired by social science applications, focusing on models for longitudinal data consisting of directed graphs.

Models for network are considered where the node set is fixed and the set of arcs is random. A characterizing feature of network data is the complicated dependence structure between arcs: e.g., reciprocation, transitivity, centralization, or other differences between nodes. Next to such dependencies, statistical models must be able to represent effects of covariates. The most frequently occurring type of longitudinal network data in the social sciences is repeated measures, also called panel data, where a relation on a given node set is observed at two or more time points. Continuous-time stochastic processes, observed at discrete time points, here can be used fruitfully. In this approach it is assumed that an unobserved sequence of changes takes place between the observation moments, and a Markov chain model for these unobserved changes allows representing feedback and thereby the dependencies that are typical for network data. The so-called Stochastic Actor-oriented Model is a class of Markov chain models where the transition rates are modeled in a way that can be represented as resulting from choices made by the social actors who are represented by the nodes. The state space can consists of one digraph (directed graph, representing a social network), but also several interdependent digraphs, or one or more digraphs together with a space of nodal attributes.

The lectures will start with a presentation of the type of questions posed by social scientists for this type of data, and an outline of the actor-oriented model, for various specifications of the state space. Several estimation procedures will be presented: method of moments, maximum likelihood, Bayesian. The implementation of these methods depends on various tricks to improve computational efficiency. These are of great practical importance, and some of the mathematical issues will be discussed. For practical applications, model specification and goodness of fit assessment are important. An R package, RSiena, is available for carrying out these procedures. Some attention will be paid to the interaction between statisticians and users. Finally, some open problems will be mentioned.

Tom A.B. Snijders. (2017). Stochastic Actor-Oriented Models for Network Dynamics. *Annual Review of Statistics and Its Application*, 4, 343-363.

http://www.stats.ox.ac.uk/~snijders/siena/