Accepted Manuscript

Dealing with reciprocity in dynamic stochastic block models

Francesco Bartolucci, Maria Francesca Marino, Silvia Pandolfi

PII: S0167-9473(18)30011-2
DOI: https://doi.org/10.1016/j.csda.2018.01.010
Reference: COMSTA 6551

To appear in: *Computational Statistics and Data Analysis*

Received date : 19 September 2016
Revised date : 18 January 2018
Accepted date : 19 January 2018

Please cite this article as: Bartolucci F., Marino M.F., Pandolfi S., Dealing with reciprocity in dynamic stochastic block models. *Computational Statistics and Data Analysis* (2018), https://doi.org/10.1016/j.csda.2018.01.010

Dealing with Reciprocity in Dynamic Stochastic Block Models

Francesco Bartolucci[a], Maria Francesca Marino[b], Silvia Pandolfi[a,*]

[a]*Department of Economics, University of Perugia*
[b]*Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence*

Abstract

A stochastic block model for dynamic network data is introduced, where directed relations among a set of nodes are observed at different time occasions and the blocks are represented by a sequence of latent variables following a Markov chain. Dyads are explicitly modeled conditional on the states occupied by both nodes involved in the relation. With respect to the approaches already available in the literature, the main focus is on reciprocity. In this regard, three different parameterizations are proposed in which: (*i*) reciprocity is allowed to depend on the blocks of the nodes in the dyad; (*ii*) reciprocity is assumed to be constant across blocks; and (*iii*) reciprocity is ruled out. The assumption of conditional independence between dyads (referred to different pairs of nodes and time occasions) given the latent blocks is always retained. Given the complexity of the model, inference on its parameters is based on a variational approach, where a lower bound of the log-likelihood function is maximized instead of the intractable full model log-likelihood. An approximate likelihood ratio test statistic is proposed which compares the value at convergence of this lower bound under different model specifications. This allows us to formally test for both the hypothesis of no reciprocity and that of constant reciprocity with respect to the latent blocks. The proposed approach is illustrated via a simulation study based on different scenarios. The application to two benchmark datasets in the social network literature is also proposed to illustrate the effectiveness of the proposal in studying reciprocity and identifying groups of nodes having a similar social behavior.

Keywords: Dyads, EM algorithm, hidden Markov models, likelihood ratio test, variational inference

1. Introduction

A number of social, behavioral, and biological phenomena can be naturally represented in terms of networks. In this literature, the relation between units, that is, "actors" or "nodes",

*Corresponding author. Silvia Pandolfi, Department of Economics, University of Perugia, Via A. Pascoli 20. 06123 Perugia, ITALY. Phone: +39 075 5855278

Email addresses: francesco.bartolucci@unipg.it (Francesco Bartolucci), mariafrancesca.marino@unifi.it (Maria Francesca Marino), silvia.pandolfi@unipg.it (Silvia Pandolfi)