Statistics > Methodology

Confidence sets for network structure

Edoardo M. Airoldi, David S. Choi, Patrick J. Wolfe

(Submitted on 31 May 2011)

Latent variable models are frequently used to identify structure in dichotomous network data, in part because they give rise to a Bernoulli product likelihood that is both well understood and consistent with the notion of exchangeable random graphs. In this article we propose conservative confidence sets that hold with respect to these underlying Bernoulli parameters as a function of any given partition of network nodes, enabling us to assess estimates of 'residual' network structure, that is, structure that cannot be explained by known covariates and thus cannot be easily verified by manual inspection. We demonstrate the proposed methodology by analyzing student friendship networks from the National Longitudinal Survey of Adolescent Health that include race, gender, and school year as covariates. We employ a stochastic expectation-maximization algorithm to fit a logistic regression model that includes these explanatory variables as well as a latent stochastic blockmodel component and additional node-specific effects. Although maximum-likelihood estimates do not appear consistent in this context, we are able to evaluate confidence sets as a function of different blockmodel partitions, which enables us to qualitatively assess the significance of estimated residual network structure relative to a baseline, which models covariates but lacks block structure.

Comments:	17 pages, 3 figures, 3 tables
Subjects:	**Methodology (stat.ME)**; Social and Information Networks (cs.SI); Physics and Society (physics.soc-ph)
Journal reference:	Statistical Analysis and Data Mining, vol. 4, pp. 461-469, 2011
DOI:	10.1002/sam.10136
Cite as:	**arXiv:1105.6245 [stat.ME]**
	(or **arXiv:1105.6245v1 [stat.ME]** for this version)

Submission history

From: Edoardo Airoldi [view email]
[v1] Tue, 31 May 2011 11:49:04 GMT (496kb,D)

Which authors of this paper are endorsers?