

**Dr Johan Koskinen,  
Cathie Marsh Centre for Census and Survey Research (CSSR),  
University of Manchester**

**Title:** Missing data in social networks: Problems and prospects for model-based inference

**Abstract:** The problems of missingness in network data are considerably more complex than in other areas of social science and there is a need to move beyond ad hoc approaches. This paper discusses general issues relating to missing data in social network studies and argues for model-based statistical inference using exponential random graph models (ERGM). We discuss the problem of fitting ERGMs when data are partially observed, and we consider the use of ERGMs as an efficient and principled way of imputing missing data. We define a standard set of assumptions – the standard model - under which Bayesian inference is admissible. The standard model allows us to formulate interesting and important deviations that need future research, particularly in the cases of incomplete, overlapping informant reports, missing attribute data, network “doppelgangers” (i.e. actors with multiple aliases who appear as multiple nodes in the data), and “covert” actors (nodes suspected to exist but for whom there is no data). In regard to covert actors, we show by example using the Sageman (2004) terrorist data that snowball conditional estimation with increasing numbers of “ghost” nodes can yield valid inference and adequate imputations. In regard to fuzzy actors (missing attributes) using a well-known data set we show that not only may a modified ERGM still be used but also that it serves as a reasonable explicit imputation model.