Social Networks 54 (2018) 26–40

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

Using modern methods for missing data analysis with the social relations model: A bridge to social network analysis

Terrence D. Jorgensen[a,*], K. Jean Forney[b], Jeffrey A. Hall[c], Steven M. Giles[d]

[a] Department of Child Education and Development, University of Amsterdam, Postbus 15776, 1001NG Amsterdam, the Netherlands
[b] Florida State University, 1107 W Call St, Tallahassee, FL 32306, USA
[c] University of Kansas, Bailey Hall, 1440 Jayhawk Blvd. Rm 102, Lawrence, KS 66045, USA
[d] Wake Forest University, 1834 Wake Forest Road, 125 Carswell Hall, Winston-Salem, NC 27109, USA

ARTICLE INFO

Article history:
Received 8 January 2017
Received in revised form 13 October 2017
Accepted 10 November 2017

Keywords:
Body image
Dyadic data
Imputing round-robin data
Missing data
Peer influences
Social relations model

ABSTRACT

Social network analysis identifies social ties, and perceptual measures identify peer norms. The social relations model (SRM) can decompose interval-level perceptual measures among all dyads in a network into multiple person- and dyad-level components. This study demonstrates how to accommodate missing round-robin data using Bayesian data augmentation, including how to incorporate partially observed covariates as auxiliary correlates or as substantive predictors. We discuss how data augmentation opens the possibility to fit SRM to network ties (potentially without boundaries) rather than round-robin data. An illustrative application explores the relationship between sorority members' self-reported body comparisons and perceptions of friends' body talk.

Introduction

The focus of our paper is how the social relations model (SRM; Kenny, 1994; Kenny et al., 2006) can be utilized to model a social relational system in a bounded network. The SRM is traditionally applied to data gathered from a so-called "round-robin design," in which all possible reciprocal perceptions of members in a closed network are recorded. Social network analysis (SNA) typically models the structure of a network comprised of ties between nodes. In this paper, we propose a methodological bridge between SNA and SRM, such that the criterion for recording dyad-level perceptions is whether a directed (or reciprocated) tie between the pair exists. This bridge is built on modern advances in missing-data analysis.

Traditionally, SRM parameters are estimated using random-effects ANOVA to partition a single outcome (Warner et al., 1979) into components associated with the ego,[1] alter, and dyadic relationship. Extensions of the SRM allow ego and alter effects to correlate with other ego or alter characteristics (e.g., Brunson et al., 2016; Kenny, 1994; Kwan et al., 2004). We utilize the multilevel modeling (MLM) framework to fitting the SRM (Snijders and Kenny, 1999), using a flexible Bayesian approach (Hoff, 2005; Lüdtke et al., 2013). One advantage of using Bayesian estimation methods is that missing data can be treated as unknown parameters to be estimated along with the model's fixed and random effects. Although Lüdtke et al. (2013) and Hoff (2005) hinted at this advantage of fitting the SRM in a Bayesian paradigm, the method of fitting the SRM to partially observed data has yet to be developed. We contribute to the SRM literature by (a) elaborating on missing-data mechanisms in the context of the SRM and (b) demonstrating how ignorable missing data can be accommodated using a Bayesian approach. We contribute to SNA literature by demonstrating how (a) perceptions of alters and (b) self-reported characteristics of egos can be modeled simultaneously to answer questions about within-network perceptions. Given the ability to fit the SRM to incomplete round-robin data, we propose that SRM parameters can be interpreted with regard to ties in a social network rather than to round-robin data.

After introducing an extended SRM for partially observed data, we apply it to self-reported body attitudes and body comparison to illuminate the nature of peer-perceptions about body talk. Our investigation explores the association among perceptions of one's peers, peers' perceptions of oneself, and peers' self-report in order to assess whether perceptions of peers are related more to char-

* Corresponding author.
E-mail addresses: T.D.Jorgensen@uva.nl (T.D. Jorgensen), forney@psy.fsu.edu (K.J. Forney), hallj@ku.edu (J.A. Hall), gilessm@wfu.edu (S.M. Giles).

[1] In previous SRM literature, egos (self) and alters (peer) have been referred to as actors and partners, respectively, in the context of behavioral observations, or as perceivers and targets in the context of interpersonal perception.

https://doi.org/10.1016/j.socnet.2017.11.002
0378-8733/

acteristics of the perceiver (ego) or to characteristics of the peer being perceived (alter). To more fully account for the influence of the ego on perceptions of alters, we also consider the relationship of an ego's perceptions with the ego's weight-related attitudes (i.e., drive for thinness). To date, no publication has explored whether the relationship between an ego's own behavior and her perception of her alters' corresponding characteristics is moderated by her own drive for thinness.

Social network analysis and the social relations model

A traditional goal of SNA is to identify and characterize the structure of a network, typically using graph theory (Wasserman and Faust, 1994). When data are collected from people within a bounded network, individuals (i.e., egos) typically identify connections (i.e., ties) with other network members (i.e., alters), for example, by nominating friends from a list of peers in a classroom. The presence or absence of these ties is meaningful in that these associations describe the type of network (e.g., a friendship network; an advice network), and ties reflect a particular type of relationship between network members. In SNA, ties are often directed, wherein an ego identifying alter(s) as friend(s), which represents the ego's out-degree, and when alter(s) identify ego as a friend it represents the ego's in-degree (Wasserman and Faust, 1994). Each tie can also have a weight (Valente, 2010). This weight could be the strength of the relationship (qualifying the tie as weak or strong) or other information about the link, such as the type of advice or information shared. When ties have weight, it is called valued network data (Valente, 2010). In addition to graphically illustrating the structure of the network (i.e., egos represented as "nodes" and ties represented as links or "edges"), many metrics are used to characterize individuals (e.g., centrality), dyads (e.g., connectivity, reciprocity), groups (e.g., clustering, closure), or the network as a whole (e.g., density, mean vertex degree) (Wasserman and Faust, 1994). Network structure can also be explained or predicted, for instance, using the p_2 model (van Duijn et al., 2004) or exponential random graph models (Robins et al., 2007a; Robins et al., 2007b).

Rather than focusing only on network structure, researchers collecting network data often pose research questions about individual-level outcomes. Valente (2010) makes a strong case for the importance of SNA in understanding health- and disease-related phenomena, particularly regarding the process of behavioral or attitudinal influence. With SNA, each ego's exposure, or "the degree to which a focal individual's alters engage in a particular behavior" (Valente, 2010, p. 65), can be modeled. Exposure is a possible mechanism to explain diffusion of innovations or changes in health behaviors, but standard generalized linear models would be inappropriate to explain individuals' behaviors because observations would not be independent (Kenny et al., 2006). SNA, on the other hand, accounts for interdependency among the observations, and can take into account characteristics of both alters and weights of the ties. Ego-network traditions have sometimes relied upon perceptual data of one's alters as a possible mechanism of measuring exposure, but Valente (2010) cautions that perceptions of one's friends by egos are biased and cannot be taken as an accurate estimate of the actual behavior or attitudes of alters. Importantly, the nature of these perceptional biases (e.g., to be congruent with one's own perceptions) is not something typically modeled in SNA. Valente (2010) even offers examples of prior attempts to reconcile perceptual data with partner self-reports, but does not mention the SRM as a methodological option for doing so. This paper offers SNA researchers a new and statistically appropriate way to model perceptions and biases.

The SRM (Kenny, 1994; Kenny et al., 2006) is fundamentally concerned with how individuals perceive each other (i.e., interpersonal perceptions). In their chapter addressing SNA, Kenny et al. (2006) discuss the similarities between p_1 and SRM, suggesting that the former SNA technique is an extension of SRM for binary data. Kenny et al. (2006) admit that using the SRM for dichotomous SNA is "not entirely appropriate" due to differences in measurement of the ties/perceptions between network members (p. 313). That is, interpersonal perceptions are usually interval level measurements, not dichotomous measures. Interpersonal perceptions can be decomposed into person-level and dyad-level components, allowing investigation of how perceptions relate to each other (i.e., reciprocity) within the network. Data can also be collected on self-perceptions—or relevant attitudes or behaviors—to reveal, for instance, how others' perceptions correlate with self-perceptions (i.e., self–other agreement) or actual behaviors (i.e., accuracy). Many of the conditions of the SRM, including data on perceptions of alters (not just the presence or absence of a tie), interval level measurement, the assumption of primarily reciprocal ties (i.e., bidirectional ties), are atypical in most SNA designs (Valente, 2010). Although researchers are undoubtedly interested in the perceptions of alters in SNA, such as the weight of the ties in relation to alter characteristics or behaviors, it is very rarely done in practice.

The present investigation will extend the SRM to traditional bounded SNA data, which includes perceptions of ties and self-reported behavioral and attitudinal characteristics. This collection of all relevant information about network members—indications of ties as well as characteristics of individuals and of tied dyads—has been referred to as a social relational system (van Duijn et al., 2004; Wasserman and Faust, 1994). The present investigation will be valuable not only for dealing with missing round-robin data, but also for researchers who are interested in exploring the attributes and interpersonal perceptions only among tied network members in existing relationships, in which case ratings from members who are not closely tied would be irrelevant to their research question. For instance, researchers may be interested only in how friends perceive each other, rather than in perceptions among all possible peers. Past research suggests that close relationships are particularly influential and important in understanding health behavior (Valente, 2010). Researchers may not be particularly interested in examining network structure, but instead would use the structure of the network (i.e., presence of directed ties between egos) to define the sampling frame. To apply the SRM to such data, several data management and analysis barriers must be overcome, particularly accommodating the fact that data from a traditional round-robin design would be "missing" when data are gathered from only a subset of all possible dyads.

Missing data mechanisms

Inferences drawn about parameters estimated from partially observed data can be biased to the degree that the missing data are not *ignorable*. Rubin (1976) defined three mechanisms of missingness, some of which can be considered ignorable (Enders, 2010, p. 13; Little et al., 2014), depending on which analytical method is used. If the probability of observation depends on the values of the missing observations themselves, then data are said to be missing not at random (MNAR; Rubin, 1976). Data can also be considered MNAR if missingness depends on variables that are not observed, or are not included in the analysis model. If variables related to missingness are observed and included in the analysis model, then data are said to be missing at random (MAR), given the observed data. That is, whether data are missing is unrelated to the missing data, conditional on the observed data. If missingness is unrelated to missing data even without conditioning on observed data, then data are said to be missing completely at random (MCAR). Only multiple imputation or maximum likelihood methods can return unbiased point and *SE* estimates after adequately incorporating variables