The Multilevel $p_2$ Model

A Random Effects Model for the Analysis of Multiple Social Networks

Bonne J. H. Zijlstra, Marijtje A. J. van Duijn, and Tom A. B. Snijders

University of Groningen, the Netherlands

Abstract. The $p_2$ model is a random effects model with covariates for the analysis of binary directed social network data coming from a single observation of a social network. Here, a multilevel variant of the $p_2$ model is proposed for the case of multiple observations of social networks, for example, in a sample of schools. The multilevel $p_2$ model defines an identical $p_2$ model for each independent observation of the social network, where parameters are allowed to vary across the multiple networks. The multilevel $p_2$ model is estimated with a Bayesian Markov Chain Monte Carlo (MCMC) algorithm that was implemented in free software for the statistical analysis of complete social network data, called StOCNET. The new model is illustrated with a study on the received practical support by Dutch high school pupils of different ethnic backgrounds.

Keywords: social network analysis, random effects, multilevel modeling, $p_2$ model

Introduction

An important goal of social network analysis is to find structure in relations within groups. Usually, one social network at a time is analyzed. This article deals with the analysis of multilevel network data where multiple observations of the same binary relation in different groups are available. In an educational setting, for example, networks may be observed in multiple schools or school classes, forming the higher-level units. Analyzing multiple networks simultaneously provides greater generalizability of research results compared to analyses of single-network data. For multilevel network data, an interesting question is whether all social networks show a common structure. If this turns out not to be the case, subsequent questions are on which aspects the networks differ, and whether there are network attributes to which these differences can be ascribed. The multilevel $p_2$ model has been designed to answer these types of questions. So far, a statistical analysis of multiple networks with the $p_2$ model involved a two-stage estimating procedure, where a meta-analysis was performed on the parameters obtained in separate $p_2$ analyses of each network (Baerveldt, Van Duijn, Vermeij, & Van Hemert, 2004). The multilevel $p_2$ model estimates the parameters more efficiently and, moreover, quantifies the differences between networks by modeling the variability of parameters over networks. For continuous social network data, the social relations model for multiple groups can be used that can be estimated with standard multilevel software (Snijders & Kenny, 1999).

The data under consideration consist of $K$ networks. Each single network is defined by a set of social actors and a relation (e.g., friendship, collaboration) defined on this set. Each relation is expressed by a collection of tie variables: $Y_{ij}$ equals 1 if there is a tie from actor $i$ to actor $j$, and 0 otherwise (where the notation temporarily omits to indicate which of the networks is being referred to). It is assumed that the sets of actors for the $K$ networks are disjoint (e.g., different school classes), and that the content, or meaning, of the relation is the same in each network (e.g., friendship). It should be noted that the actor sets are not required to have the same size for every network.

In the $p_2$ model (Van Duijn, Snijders, & Zijlstra, 2004), the tie variables are regressed on explanatory variables, while the dependence of ties from and to the same actor is modeled using random effects. The multilevel $p_2$ model defines an identically specified $p_2$ model with varying parameters for multiple independent social networks. One of the Bayesian Markov Chain Monte Carlo (MCMC) algorithms developed for the single-network $p_2$ model by Zijlstra, Van Duijn, and Snijders (2005a) can be expanded for the multilevel $p_2$ model. This hybrid Metropolis-Hastings algorithm will be briefly described in the third section, after defining the multilevel $p_2$ model in the next section. Software for the multilevel $p_2$ model is available in the $p_2$ module of StOCNET (Boer, Huisman, Snijders, & Zeggelink, 2003), an open software system for the statistical analysis of social networks.

The multilevel $p_2$ model is applied to network data collected by Chris Baerveldt in 20 Dutch high schools (Baerveldt, 2000; Snijders & Baerveldt, 2003). The central question dealt with in this article is whether these pupils tend to report more practical support from other pupils with the same ethnic background.
The Multilevel $p_2$ Model

The multilevel $p_2$ model is an extension of the $p_2$ model (Van Duijn et al., 2004) that originates from the $p_1$ model proposed by Holland and Leinhardt (1981). These models are defined by specifying the probability of observing one of the four possible outcomes of the pair of two directed ties between each pair of actors, which is called a dyad. Let a dependent network with $n$ actors be denoted by the tie indicator variables $Y_{ij}$ and let the actors $i$ and $j$ have numbers $1, \ldots, n$. Then the $p_1$ model for the probabilities of the two observed ties between actors $i$ and $j$ is defined by

$$P(Y_{ij} = y_{ij}, Y_{ji} = y_{ji}) = \exp(y_{ij} \mu_{ij} + y_{ji} \rho_{ij})/c,$$

$$c = \sum_{y_{ij}, y_{ji} = 0, 1} \exp(y_{ij} \mu_{ij} + y_{ji} \rho_{ij}),$$

where $\alpha_i$ is a sender parameter for actor $i$, $\beta_j$ a receiver parameter for actor $j$, $\mu$ the density parameter, and $\rho$ the reciprocity parameter.

In the $p_1$ model, all ties coming from or directed toward the same actor are mutually related through—or conditionally independent given—the 2$n$ parameters $\alpha$ and $\beta$. The number of parameters in the $p_1$ model increases linearly with the number of actors, an undesirable property for any model. The $p_1$ model extends the $p_1$ model by including covariate effects for $\alpha$, $\beta$, $\mu$, and $\rho$, while the total number of parameters in this model is reduced by assuming random (instead of fixed) sender and receiver effects $A_i$ and $B_j$. The latter are assumed to be independent, identically bivariate normally distributed with zero means, and covariance matrix $\Sigma$ with diagonal elements $\sigma^2_{\alpha}$ (sender variance), and $\sigma^2_{\beta}$ (receiver variance) and off-diagonal elements $\sigma_{\alpha\beta}$ (sender-receiver covariance). In the $p_2$ model, the sender and receiver parameters are regressed on (actor) covariates $X_1$ and $X_2$ with fixed regression parameters $\gamma_1$ and $\gamma_2$.

$$\alpha_i = X_i \gamma_1 + A_i,$$

$$\beta_j = X_j \gamma_2 + B_j.$$

The density and reciprocity parameters $\mu$ and $\rho$ are regressed on $Z_1$ and $Z_2$ with fixed regression parameters $\delta_1$ and $\delta_2$.

$$\mu_i = \mu + Z_i \delta_1,$$

$$\rho_i = \rho + Z_i \delta_2.$$

Note the added subscripts $i$ and $j$ for the density and reciprocity parameters, which are now assumed to be dyad-specific. The variables $Z_1$ and $Z_2$ specify dyadic covariates depending on the ordered pair of actors $(i, j)$, where $Z_1$ is a symmetric matrix, expressing the mutuality-by-definition of reciprocity, $\rho = \mu$. The multilevel $p_2$ model specifies an identical $p_2$ model for $K$ independent observations of a network relation. To account for differences between the $K$ networks, the multilevel $p_2$ model includes random coefficients for the fixed $p_2$ regression parameters at the network level. The fixed regression parameters become random coefficients by adding random effects $G_i$ and $G_j$ to the regression parameters $\gamma_i$ and $\gamma_j$ for the sender and receiver effects. The density and reciprocity parameters $\mu$ and $\rho$ obtain random effects $M$ and $R$, their regression parameters random effects $D_1$ and $D_2$.

The parameters in the $p_1$ model (Equation 1) for the $k$th network are thus substituted by

$$\alpha_k = X_k \gamma_1 + G_{ik} + A_k,$$

$$\beta_k = X_k \gamma_2 + G_{jk} + B_k,$$

$$\mu_k = \mu + M_k + Z_k \delta_1 + D_{ik},$$

$$\rho_k = \rho + R_k + Z_k \delta_2 + D_{jk}.$$

The vector with random effects for each network $k$ is denoted by $T_k = (M_k, R_k, G_{ik}, G_{jk}, D_{ik}, G_{jk})$, and is assumed to be normally distributed with zero means and covariance matrix $\Omega$. A further assumption, common in multilevel modeling, is that the random effects at the network level $(k)$ are independent from the random effects at the actor level $(i)$.

The multilevel $p_2$ model can be regarded as a three-level random effects model where Level 1 is formed by the tie observations, cross-nested in the actors (Level 2), who are nested in the networks (Level 3). Just like the $p_2$ model allows for covariates at the actor level, the multilevel model allows for explanatory variables at the network level, adding a regression model to the network-specific parts of Equation 2. For instance, the term $\mu + M_k$ can be replaced by $\mu + M_k + W_k \eta_k$, where $W_k$ denotes a vector of network-level covariates such as network size and aggregated actor characteristics.

Outline of the MCMC Estimation Algorithm

For random effects models, maximum likelihood estimation of the fixed effects and of the variances of the random effects requires integration over the random parameters. In the case of the multilevel $p_2$ model, these are practically intractable integrals, which is why an MCMC algorithm is applied.

In MCMC estimation methods, prior distributions need to be specified for the model parameters, which, together with the data, determine the posterior distributions for the parameters, as follows from Bayes’s theorem. In these algorithms, a Markov chain generates a sample of all parameters. If convergence of the chain can be postulated after an initial burn-in period, then this is a sample from the posterior distribution and Monte Carlo estimates can be calculated from this sample. This section provides a succinct overview of how the posterior distributions of the multilevel $p_2$ model parameters are obtained and how these are sampled from.

The multivariate distribution of the data and the fixed and random parameters has a probability density that is proportional to the joint density

$$P(Y, C, \Psi, T, \Theta, \Omega),$$

where $Y=(Y_1', \ldots, Y_K')$ contains the observed data for all networks, with $Y_k'$ the vector of all dyads ($Y_{ik}$, $Y_{jk}$) in network $k$. $C$ contains all pairs of random actor effects ($A_{ik}$, $B_{jk}$), and $\Psi$ denotes their 2 $\times$ 2 covariance matrix. Vector $\theta = (\mu, \rho, \gamma_1, \gamma_2, \delta_1, \delta_2)$ is the vector of fixed parameters, and $T=(T_1', \ldots, T_K')$ is the vector of random effects of the parameters in $\theta$ with covariance matrix $\Omega$.

A convenient way to obtain a sample from the multivar-
iate distribution of all variables in Equation 3 is by means of the Gibbs sampler, which involves drawing subsequent random variables from each distribution formed by a separate set of parameters, conditional on all other parameters and \( Y \) (see, e.g., Chib and Greenberg, 1995).

For some of the parameters in Equation 3, the posterior distributions are easy to sample from due to conditional independence and convenient conjugate priors. Assuming that \( \Sigma, \theta \) and \( \Omega \) are mutually independent, the multivariate density can be factorized as

\[
P(Y, C, \Sigma, \theta, T, \Omega) = P_1(C|\theta, T)P_2(C, \Sigma)p_3(\theta)p_4(T, \Omega)p_5(\Omega).
\]

Here, \( P_1(Y|C, \theta, T) \) is the conditional likelihood of the multilevel \( p_2 \) model given the fixed parameters and random actor and network parameters—that is, the probability of network \( k \) according to Equation 1, with the substitutions defined in Equation 2, multiplied over all networks. As can be seen from Equation 4, the density of the random actor effects, \( C \), depends only on \( \Sigma \) and the density of the random effects of the fixed parameters, \( T \), depends only on \( \Omega \).

Prior probability densities \( P_2(\Sigma) \) and \( P_5(\Omega) \) are assumed to have inverse Wishart(\( \Sigma \), \( v_\Sigma \)) and inverse Wishart(\( \Sigma \), \( v_\Sigma \)) distributions, respectively, the natural conjugate priors for the covariance matrices of normally distributed random variables (see, e.g., Press, 1989, p. 141). The covariance matrices of the Wishart prior distributions are chosen as \( \Sigma = I \) and \( \Sigma = I \), where \( I \) is the identity matrix. The degrees of freedom, \( v_\Sigma \), are chosen as the number of dimensions plus one, representing little prior information. Consequently, the posterior inverse Wishart distribution of \( \Sigma^{-1} \) has \( 3 + v_\Sigma \) degrees of freedom and covariance matrix \((C'C + I)^{-1}\). The number of degrees of freedom for \( \Omega^{-1} \) is \( v_\Omega \) with covariance matrix \((TT + I)^{-1}\) (see, e.g., Box & Tiao, 1973, p. 427).

A priori, the parameters in \( \theta \) are assumed to follow independent normal distributions with zero means and variances for \( \mu \) and \( \rho \) equal to 100. The variances of the regression parameters \( \gamma_1, \gamma_2, \delta_1, \) and \( \delta_2 \) are set to 100 divided by the observed variance of the corresponding covariate. Thus, the variance of a parameter for a "standardized" covariate is equal to 100 as well. Since parameters in \( \theta \) are on a logistic scale, a standard deviation of 10 implies that 33% of the observations are larger than the absolute value of 10. This prior for \( \theta \) may thus be regarded as very lightly informative and reflects the fact that almost any statistician will be surprised when seeing log odds ratios larger than 10.

With all distributions in Equation 3 defined, the conditional posterior distributions for each of the model parameters can be derived. By repeatedly sampling from these distributions, eventually a sample from the multivariate posterior distribution of all parameters is obtained. From the conditional distributions of the covariance matrices \( \Sigma \) and \( \Omega \) can be obtained directly from their inverse Wishart distributions, whereas sampling of the random effects \( C \) and \( T \) and the fixed parameters \( \theta \) requires using a Metropolis-Hastings algorithm (Metropolis, Rosenbluth, Rosenbluth, Teller & Teller, 1953). For this purpose, we used a similar type of random walk algorithm as in Zijlstra et al. (2005a), in which the variance of the proposals is based on its esti-
Table 1. Descriptive statistics for the 20 schools in the Dutch Social Behavior Study data.

<table>
<thead>
<tr>
<th>School</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of pupils</td>
<td>50</td>
<td>43</td>
<td>55</td>
<td>38</td>
<td>73</td>
<td>96</td>
<td>62</td>
<td>39</td>
<td>91</td>
<td>42</td>
</tr>
<tr>
<td>Mean degree</td>
<td>2.98</td>
<td>2.30</td>
<td>1.47</td>
<td>1.95</td>
<td>3.22</td>
<td>2.16</td>
<td>2.71</td>
<td>3.33</td>
<td>2.44</td>
<td>1.52</td>
</tr>
<tr>
<td>Mean reciprocal degree</td>
<td>1.56</td>
<td>1.12</td>
<td>0.62</td>
<td>0.89</td>
<td>1.70</td>
<td>0.96</td>
<td>0.90</td>
<td>1.54</td>
<td>1.10</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Table 2. Parameter estimates of a multilevel $p_2$ analysis of the Dutch Social Behavior Study data.

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>Covariate</th>
<th>Parameter</th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Sender</td>
<td>Boy</td>
<td>$\gamma_{s1}$</td>
<td>-0.11</td>
</tr>
<tr>
<td>Receiver</td>
<td>Boy</td>
<td>$\gamma_{s1}$</td>
<td>-0.12</td>
</tr>
<tr>
<td>Density</td>
<td>$\mu$</td>
<td></td>
<td>-3.58</td>
</tr>
<tr>
<td>Gender</td>
<td>$\delta_{s1}$</td>
<td></td>
<td>1.26</td>
</tr>
<tr>
<td>Dutch</td>
<td>$\delta_{s2}$</td>
<td></td>
<td>0.50</td>
</tr>
<tr>
<td>Moroccan</td>
<td>$\delta_{s3}$</td>
<td></td>
<td>0.40</td>
</tr>
<tr>
<td>Turkish</td>
<td>$\delta_{s4}$</td>
<td></td>
<td>0.53</td>
</tr>
<tr>
<td>Surinamese</td>
<td>$\delta_{s5}$</td>
<td></td>
<td>0.30</td>
</tr>
<tr>
<td>Girl</td>
<td>$\delta_{s6}$</td>
<td></td>
<td>-0.14</td>
</tr>
<tr>
<td>Reciprocity</td>
<td>$\rho$</td>
<td></td>
<td>4.19</td>
</tr>
<tr>
<td>Gender</td>
<td>$\delta_{r1}$</td>
<td></td>
<td>-0.45</td>
</tr>
<tr>
<td>Dutch</td>
<td>$\delta_{r2}$</td>
<td></td>
<td>-0.33</td>
</tr>
<tr>
<td>Moroccan</td>
<td>$\delta_{r3}$</td>
<td></td>
<td>-0.32</td>
</tr>
<tr>
<td>Turkish</td>
<td>$\delta_{r4}$</td>
<td></td>
<td>-0.03</td>
</tr>
<tr>
<td>Surinamese</td>
<td>$\delta_{r5}$</td>
<td></td>
<td>0.02</td>
</tr>
<tr>
<td>Girl</td>
<td>$\delta_{r6}$</td>
<td></td>
<td>0.16</td>
</tr>
</tbody>
</table>

Actor-level random effects

| Sender variance | $\sigma^2_s$ | 1.13 | 0.09 | 0.92 | 0.97 | 1.31 | 1.37 |
| Receiver variance | $\sigma^2_r$ | 0.44 | 0.05 | 0.33 | 0.36 | 0.54 | 0.57 |
| Sender receiver covariance | $\sigma_{sr}$ | -0.57 | 0.05 | -0.71 | -0.67 | -0.47 | -0.44 |

Network-level random effects of fixed parameters

| Density variance | $\Omega_{1,1}$ | 0.21 | 0.08 | 0.08 | 0.10 | 0.41 | 0.52 |
| Reciprocity variance | $\Omega_{2,2}$ | 0.24 | 0.11 | 0.08 | 0.10 | 0.52 | 0.68 |
| Density-reciprocity covariance | $\Omega_{1,2}$ | -0.13 | 0.08 | -0.45 | -0.33 | -0.02 | 0.00 |

Note. The covariate effects for density and reciprocity are similarity effects; burn-in length and sample size of the MCMC estimation algorithm were 8,000 and 40,000 respectively.

estimates of the mean and standard error of the posterior distributions of the parameters are reported. The variances of the density and reciprocity parameters are rather low, especially compared to the sender variance. Next, the network-level covariate “percentage of Dutch pupils” was included in the model, but this turned out to be nonsignificant. Moreover, the quality of the MCMC sample deteriorated, probably because there was little information in the data about the effect of the additional model parameter.

The reported results in Table 2 are based on a burn-in sequence of 8,000 iterations and a sampling sequence of 40,000 iterations. By inspecting the trace plots, one can investigate how well the sampled parameters have converged to a stable distribution. In Figure 1 the traces appear stable, although some sudden jumps are noticeable, which demonstrates the mutual dependence of the parameters discussed above.

The model in Table 2 can be regarded as a so-called random intercept model (cf. Goldstein, 2003) with random effects $M$ and $R$ for the intercepts $\mu$ and $\rho$ on the network level. The parameter estimates show that reported practical support relations are more prevalent between pupils with the same gender. This is pointed out by the strong positive density effect of “similarity gender.” The interpretation of this effect is moderated by the “similarity girl” effect, which shows that the increased practical support within the same
gender is stronger for boys than for girls. The negative similarity gender effect for reciprocity shows that reported practical support relations within the same sex is not a doubled density effect but slightly less, making reciprocal same-gender dyads more likely than asymmetrical dyads. Furthermore, boys and girls do not differ strongly with respect to sending and receiving tendencies.

From fellow pupils with the same ethnic background, more practical support is received, following from the positive density effects for the covariates indicating similarity in ethnic background. For Dutch pupils only a clear effect of reciprocity was found, modifying the double density effect for reciprocal dyads. The posterior distributions for the greater reciprocity effects among pupils with the same ethnic backgrounds are much wider for the other ethnic groups, containing large positive as well as negative values, which may be the result of the fact that for some of these groups our sample contains only a small number of children.

Concluding Remarks

With the multilevel $p_2$ model multiple parallel network observations can be analyzed. Such data are likely to be gathered in, for instance, educational settings. Compared to the two-stage meta-analytic approach used by Baerveldt et al. (2004), the advantage is clear: All data can be analyzed in a single model, resulting in an increase in power to detect possible actor or dyadic covariate effects. Moreover, it is possible to investigate whether these effects differ over networks and whether these differences can be explained by network covariates. In the application, a small amount of between-network variability was found. In general, however, one should be careful not to include too many random effects for the fixed parameters at the network level. Including $t$ random effects implies estimating $(t+1)/2$ parameters in $\Omega$. This means that the number of parameters soon becomes large compared to the number of networks. It also implies that the multilevel $p_2$ model is more easily applied in data sets with a larger number of networks.

The results obtained in the example show that reported practical support is more prevalent among pupils with the same ethnic background. Further research is intended to combine multiple types of relations with a multilevel data structure in a single model. Then the current model can be extended with, for example, the data on emotional support that were analyzed in Baerveldt et al. (2004).

Software for the multilevel $p_2$ model is available in StOCNET (Boer et al., 2003).

Acknowledgments

This research was supported by the Netherlands Organisation for Scientific Research, grant NWO-425-10-0003-P. The authors would like to thank Chris Baerveldt for his permission to use the data.

References

2nd International Network Sampling Workshop, Maastricht, the Netherlands.

Bonne J. H. Zijlstra

University of Groningen
Faculty of Behavioural and Social Sciences
Grote Rozenstraat 31
NL-9721 TG Groningen
The Netherlands
Tel. +31 50 363 6199
Fax +31 50 363 6226
E-mail b.j.h.zijlstra@rug.nl

© 2006 Hogrefe & Huber Publishers