Chapter 2
The Multiple Flavours of Multilevel Issues for Networks

Tom A.B. Snijders

Away from Atomistic Approaches

It is strange that the assumption that data obtained from human respondents represent independent replications has been so pervasive in statistical models used in sociological research. Sociology, after all, is about the interdependence among individuals, and about the ways in which individuals make up larger wholes such as families, tribes, organizations, and societies. Of course we know some of the reasons for this: statistical models founded on independence assumptions are convenient and have properties that can be mathematically ascertained; surveys are a major means of getting social information and ideally are obtained from probability samples containing a lot of independent operations in obtaining respondents; and, indeed, independence assumptions may yield good first-order approximations for statistical modeling. However, as early as 1959 Coleman (1959, p. 36) made an eloquent plea for taking social structure into account in methods of data collection and analysis. Coleman writes: “Survey methods have often led to the neglect of social structure and of the relations among individuals. (...) But (...) one fact remained, a very disturbing one to the student of social organization. The *individual* remained the unit of analysis. (...) Now, very recently, this focus on the individual has shown signs of changing, with a shift to groups as the units of analysis, or to networks of relations among individuals”. He goes on to discuss methods for survey data collection and for data analysis that reflect this change in perspective, away from the focus on atomistic individuals. The analysis methods he discusses

T.A.B. Snijders (✉)
Department of Sociology, University of Groningen, Grote Rozenstraat 31, 9712 TG Groningen, The Netherlands

Nuffield College, University of Oxford, Oxford, UK
e-mail: tom.snijders@nuffield.ox.ac.uk

E. Lazega, T.A.B. Snijders (eds.), *Multilevel Network Analysis for the Social Sciences*, Methodos Series 12, DOI 10.1007/978-3-319-24520-1_2

include contextual analysis, the precursor of present-day multilevel analysis, and the study of subgroups and cliques, still now of crucial importance in social network analysis. He concludes by saying that these methods "will probably represent only the initial halting steps in the development of a kind of structural research which will represent a truly sociological methodology", and mentions the promise of electronic computers.

In the past half century, since Coleman wrote these words, great advances have been made in methodologies for analyzing groups, or collectives, along with individuals; or, more generally, for simultaneously analyzing variables defined on different domains. The name 'multilevel analysis' has replaced[1] 'contextual analysis'. Great strides also have been taken in the study of relations among individuals, known now as social network analysis. Network analysis likewise treats variables defined in various different domains, such as sets of nodes and sets of node pairs, and it is concerned with groups, but by and large multilevel analysis and social network analysis have developed separately, meeting each other only incidentally. Recently, however, developments in social network analysis have led to combinations of these two strands of methodology. We are still in an early phase of the junction of multilevel analysis and social network analysis, and we may echo Coleman in saying that this book presents some 'initial halting steps' of this junction. This chapter gives an overview of some concepts and techniques that now can be seen as playing important roles in the combination of multilevel and network analysis.

Multilevel Analysis

To be able to discuss multilevel network analysis, we need to present a sketch about 'regular' multilevel analysis.

Origins

Multilevel analysis, as a collection of methods, was born from the confluence of two streams. On the one hand, sociological methodologists had been developing quite some conceptual precision for inference relating individuals to collectives, for which variables need to be combined that are defined in several different domains. On the other hand, statisticians had already extended analysis of variance and regression analysis, the general linear model, to linear models combining fixed with randomly varying coefficients.

Let me first sketch some highlights on the sociological methodology side. Lazarsfeld and Menzel (1961), in their paper *On the relation between individual*

[1] Albeit with a shift of meaning.

and collective properties—written in 1956, reprinted as Lazarsfeld and Menzel (1993)—distinguish variables according to the set of units to which scientific propositions are meant to apply. For propositions about individual and collective properties, they state that there need to be sets of units both at the individual and at the collective level. Here 'individual' may refer to individual humans, but also, e.g., individual organizations or other groupings; 'collective' refers to sets of 'individuals'. Lazarsfeld and Menzel go on to define three types of properties defined for collectives. Analytical properties are obtained by a mathematical operation performed on each member, for example the mean of an individual variable, or the correlation between two variables. Structural properties are obtained by a mathematical operation performed on the relations of each member to some or all of the other members, for example the 'cliquishness' of a network. Global properties, finally, are properties of collectives that cannot be directly deduced from properties of individual members, e.g., the type of government of a city.

As for properties of individuals, Lazarsfeld and Menzel discuss that the correlation between individual variables may be considered as a correlation between the individuals but also between the collectives, pointing to the *ecological fallacy* presented in Robinson (1950): the mistake of regarding associations between variables at one level of aggregation as evidence for associations at a different aggregation level; an extensive review was given by Alker (1969). Researchers became aware of the importance of the different levels, or sets of units, in which variables are defined, and as suggested here the focus was on nested levels, representing individuals and collectives.

During the 1970s, methods for contextual analysis were developed taking into account these levels of analysis, and trying to avoid ecological fallacies. This was called 'contextual analysis' mainly by sociologists (Blalock 1984), and 'multilevel analysis' by educational researchers (Burstein 1980).

Statisticians had a few decades earlier developed models that waited to be discovered by these social scientists. In the analysis of variance, precursor and paradigmatic example of the general linear model, models had been developed where coefficients could themselves be random variables, allowing for the investigation of multiple sources of random variation in, e.g., agricultural and industrial production. Models with only fixed, fixed as well as random, or only random coefficients were called fixed, mixed, and random models, respectively (Wilk and Kempthorne 1955; Scheffé 1959).

In the early 1980s contextual analysis and linear mixed (or generalized linear mixed) models were brought together by several statisticians and methodologists: Mason et al. (1983), Goldstein (1986), Aitkin and Longford (1986), and Raudenbush and Bryk (1986). These researchers also developed estimation algorithms and implemented them in multilevel software packages, making use of the nested structure of the random coefficients to achieve efficiency in the numerical algorithms. The scientific gains from the combination of contextual analysis and random coefficient models are also discussed by Courgeau (2003). A more extensive history of these developments is given in Kreft and de Leeuw (1998).

Hierarchical Linear Model

The prototypical statistical model used in multilevel analysis is the *Hierarchical Linear Model*, which is a mixed effects linear model for nested designs (Raudenbush and Bryk 2002; Goldstein 2011; Snijders and Bosker 2012). This generalizes the well-known linear regression model. It is meant for data structures that are hierarchically nested, such as individuals in collectives, where each individual belongs to exactly one collective. The most detailed level (individuals) is called the lowest level, or level one. The Hierarchical Linear Model is for the analysis of dependent variables at the lowest level. The basic idea is that studying the simultaneous effects of variables defined at the individual level, as well as of other variables defined at the level of collectives, on an individual-level dependent variable requires the use of regression-type models that include error terms for each of those levels separately; the Hierarchical Linear Model is a linear mixed model that has this property.

In the two-level situation—let us say, individuals in groups—it can be expressed as follows. Highlighting the distinction with regular regression models, the terminology speaks of *units* rather than cases, and there are specific types of unit at each level. We denote the level-1 units, individuals, by i and the level-2 units, groups, by j. Level-1 units are nested in level-2 units (each individual is a member of exactly one group) and the data structure is allowed to be unbalanced, such that j runs from 1 to N while i runs, for a given j, from 1 to n_j. The basic two-level hierarchical linear model can be expressed as

$$Y_{ij} \;=\; \beta_0 \;+\; \sum_{h=1}^{r} \beta_h \, x_{hij} \;+\; U_{0j} \;+\; \sum_{h=1}^{p} U_{hj} \, z_{hij} \;+\; R_{ij} \; . \tag{2.1}$$

Here Y_{ij} is the dependent variable, defined for level-1 unit i within level-2 unit j; the variables x_{hij} and z_{hij} are the explanatory variables. Some or all of them may be defined at the group level, rendering superfluous the index i for such variables. Variables R_{ij} are residual terms, or error terms, at level 1, while U_{hj} for $h = 0, \ldots, p$ are residual terms, or error terms, at level 2. In the case $p = 0$ this is called a *random intercept model*, for $p \geq 1$ it is called a *random slope model*. The usual assumption is that all R_{ij} and all vectors $U_j = (U_{0j}, \ldots, U_{pj})$ are independent, R_{ij} having a normal $\mathscr{N}(0, \sigma^2)$ and U_j having a multivariate normal $\mathscr{N}_{p+1}(\mathbf{0}, \mathbf{T})$ distribution. Parameters β_h are regression coefficients (fixed effects), while the U_{hj} are random effects. The presence of both of these makes (2.1) into a linear mixed model. Similar models can be defined for nesting structures with more than two levels, e.g., employees in departments in firms.

In most practical cases, the variables with random effects are a subset of the variables with fixed effects ($x_{hij} = z_{hij}$ for $h \leq p$; $p \leq r$). The Hierarchical Linear Model can then be expressed in the appealing form

$$Y_{ij} \;=\; (\beta_0 + U_{0j}) \;+\; \sum_{h=1}^{p} (\beta_h + U_{hj}) \, x_{hij} \;+\; \sum_{h=p+1}^{r} \beta_h \, x_{hij} \;+\; R_{ij} \; , \tag{2.2a}$$

which shows that it can be regarded as a regression model defined for the groups separately, with group-specific intercept

$$(\beta_0 + U_{0j}) \tag{2.2b}$$

and group-specific regression coefficients

$$(\beta_h + U_{hj}) \tag{2.2c}$$

for $h = 1, \ldots, p$; variables X_h for $p + 1 \le h \le r$ have regression coefficients that are constant across groups. This pictures the Hierarchical Linear Model as a linear regression model defined by the same model for all groups, but with regression coefficients that differ randomly between groups.

Going back to the teachings of Lazarsfeld and Menzel, it can be concluded that multilevel analysis elaborates the inference about individual and collective properties as a system of nested samples drawn from nested populations: a population of individuals nested in a population of groups (or collectives). The fact that, in practice, groups will be finite, whereas the populations are mathematically considered as if they were infinite, is usually glossed over in research aiming to generalize to social mechanisms or processes (as distinct from descriptive survey research about concrete groups, without the aim of generalization to other groups) (see Cox 1990; Sterba 2009).

Non-nested Data Structures

It soon transpired that the relevant data structures are not always nested, because social structures often are not. A basic example in studies of school effectiveness is that neighborhoods may also be an important factor for student achievement, but schools will have students coming from diverse neighborhoods while neighborhoods will have students attending different schools. This leads to a data set where students are nested in schools and also nested in neighborhoods, but schools and neighborhoods are not nested in each other; the term used for non-nested category systems is 'crossed', so that this would be called a cross-nested data structure. To present an extension of model (2.1) for such a cross-nested data structure, consider again a data structure with individuals i nested in groups j but now also nested in aggregates k of a different kind (in the example of the previous sentence, neighbourhoods). Denote by $k(i,j)$ the aggregate k to which individual i in group j belongs. In the simplest extension there is only a random intercept V_k associated with k, leading to the equation

$$Y_{ij} = \beta_0 + \sum_{h=1}^{r} \beta_h x_{hij} + U_{0j} + \sum_{h=1}^{p} U_{hj} z_{hij} + V_{k(i,j)} + R_{ij} \,. \tag{2.3}$$

The default assumption for the V_k is that again they are independent and normally distributed with mean 0 and constant variance, and independent of the U and R variables. A further extension is to mixed-membership models (Browne et al. 2001), in which individuals may be partial members of more than one group.

Frequentist and Bayesian Estimation

Multilevel models such as (2.2), in which parameters vary randomly between groups, provide a natural bridge between the frequentist paradigm in statistics, which treats parameters as fixed quantities which are unknown, ‘out there’, and the Bayesian paradigm, which treats parameters as random variables; in both paradigms, of course, the observations are the material that helps us get a grip on the values of the parameters. In the multilevel case, the random variation of parameters can be linked to a frequency distribution of parameters in the population of groups, which may be estimated from empirical data. Accordingly, this bridging ground is often called *empirical Bayes* (see, e.g., Raudenbush and Bryk 2002, and Chapter 5 of Gelman et al. 2014). Bayesian estimators[2] for the parameters such as (2.2a) and (2.2b), using the sample of groups to get information about the corresponding population, are called *empirical Bayes estimators*. For the parameters β, σ^2, and $\mathbf{T}$ in (2.1), frequentist as well as Bayesian estimators have been developed.

Especially for non-nested data structures, Bayesian estimators may have algorithmic advantages, and Bayesian Markov chain Monte Carlo (‘MCMC’) algorithms are often employed (Draper 2008; Rasbash and Browne 2008) for such more complex models. These are algorithms which use computer simulations, very flexible but also much more time-consuming than traditional algorithms. Today, Bayesian methods for multilevel analysis are often proposed and used without much attention paid to the distinct philosophical underpinnings. This lack of attention does not, however, take away the differences. The Bayesian approach can be a useful way to account for prior knowledge; this is discussed for the special case of multilevel analysis by Greenland (2000), and elaborated more practically in Chapter 5 of Gelman et al. (2014). Using this approach requires, however, that one pays attention to the sensitivity of the results to the choice of the prior distribution. In addition there are interpretational differences, but these may be less important because of the convergence between frequentist and Bayesian approaches discussed in Gelman et al. (2014, Chapter 4).

[2] In frequentist terminology these are not called estimators but predictors, because they refer to statistics that have the purpose to approximate random variables.

What Is a Level?

The various extensions of the basic multilevel model have made even more pressing the question '*What is a level?*' which has harrowed quite a few researchers even in the case of the more basic nested models. The mathematical answer is that, for applications of linear mixed or generalized linear mixed models, a level is a system of categories for which it is reasonable to assume random effects. More elaborately, this means that we assume that the categories j on which the variables U_j are defined (which are latent variables in model (2.1)) may be regarded has having been sampled randomly from some universe or population $\mathscr{G}$, making the U_j into independent and identically distributed random variables, and our aim is to say something about the properties of the population $\mathscr{G}$ rather than about the individual values U_j of the units in our sample. In the case that the U_j are one-dimensional quantities, the property of interest concerning population $\mathscr{G}$ could be, e.g., the variance of U_j. In practical statistical modeling, the assumption that the units in the data were randomly sampled from the population is usually taken with a grain of salt (again cf. Cox 1990; Sterba 2009). The essential assumption is *residual exchangeability*, which can be described as follows. The random effects, R_{ij} and U_j in (2.1) and also V_k in (2.3), are residuals given that the explanatory variables x_{hij} are accounted for; these residuals are assumed to be exchangeable across i and j (or k) in the sense that they are random and as far as we know we have no *a priori* information to distinguish them for different units in the data. Any R_{ij} could be high or low just as well as any $R_{i'j}$ in the same group j or any $R_{i'j'}$ in a different group j'; any U_{0j} could be high or low just as well as any other $U_{0j'}$; etc.

In this sense, multilevel analysis is a methodology for research questions and data structures that involve several sources of unexplained variation, contrasting with regression analysis which considers only one source of unexplained variation. Employing the Hierarchical Linear Model, as in (2.1) or its variants with additional levels, gives the possibility of studying contextual effects on the individual units. But also in more complex structures where nesting is incomplete, random effects will reflect multiple sources of unexplained variation. In social science applications this can be fruitfully applied to research questions in which different types of *actor* and *context* are involved; e.g., patients, doctors, hospitals, and insurance companies in health-related research; or students, teachers, schools, and neighborhoods in educational research. The word 'level' then is used for a type of unit, or a category system, for which a random effect is assumed. The basic phenomenon we are studying will be at the most detailed level (patients or students, respectively), and the other levels may contribute to the variation in this phenomenon, e.g., as contexts or other actors.

Lazarsfeld and Menzel (1961, first page) mentioned that, to be specific about the intended meaning of variables, we should 'examine (them) in the context of the propositions in which they are used'. This focus on propositions also sheds light on the question about what can be meaningfully considered as a '*level*' in multilevel analysis. We have to distinguish between the individual level, which is the level of the phenomena we wish to explain, the population of units for which the dependent variable is defined; and higher, collective levels, which do not need to be mutually nested, but in which the individuals are nested. To be a level requires, in the first place, that the category system is a population—a meaningfully delimited set of units with a basic similarity and for which several properties may be considered, such as a well-defined set of schools, of companies, of meetings. A category system then is a meaningful higher level if it is a population that we wish to use to explain[3] some of the variability in our phenomenon and also, potentially or actually, we may be interested in finding out which properties of the categories/units explain the variability associated with this category system.

To illustrate this, suppose we are interested in the phenomenon of juvenile delinquency as our dependent variable, and we consider neighborhoods as collectives. The individual level is, e.g., a set of adolescents living in a certain area at a certain time point; the dependent variable is their delinquency as measured by some instrument. We may observe that neighborhoods differ in average juvenile delinquency, and we then may wonder about the properties of neighborhoods—perhaps neighborhood disorder, of which a measurement may be available—that are relevant in this respect. This step, entertaining the possibility that there might be specific properties of neighborhoods associated with their influence on juvenile delinquency, and analyzing this statistically, is the step that makes the neighborhood a meaningful 'level' in the sense of multilevel analysis. In the paradigm of multilevel analysis we will then further assume that in addition to the effect of disorder there may be other neighborhood effects, but conditional on the extent of disorder and perhaps other neighborhood properties that we take into account, the neighborhoods are exchangeable (as far as we know) in their further, residual, effects.

The fact that we are interested in statistically analyzing the effect of the categories on the dependent variable also implies that for a level to be meaningful in a practical investigation, the total number of its units in the data set should be sufficiently large: a statistical analysis based on a sample of, say, less than 10 units usually makes no sense.

Dependent Variables at Any Level

The Hierarchical Linear Model is considered a model for dependent variables at the lowest level of the nesting hierarchy. However, it is so amazingly flexible that it can just as well be used for complex configurations of multiple dependent variables defined for several different levels. This was proposed, quite casually, already by

[3] 'Explaining' is meant here in the simple statistical sense, without considering deeper questions of causality.

Goldstein (1989a,b). It is also explained in Goldstein (2011, Section 5.3). The basic idea can be made clear by showing, for a two-level structure, the model for interdependent dependent variables $Y^{(1)}$ at level 1 and $Y^{(2)}$ at level 2. Denoting by x_h and z_h any explanatory variables and by w_h explanatory variables at level 2, the model reads

$$Y_{ij}^{(1)} = \beta_0 + \sum_{h=1}^{r} \beta_h \, x_{hij} + U_{0j} + \sum_{h=1}^{p} U_{hj} \, z_{hij} + R_{ij} \tag{2.4a}$$

$$Y_{j}^{(2)} = \gamma_0 + \sum_{h=1}^{q} \delta_h \, w_{hj} + V_j \, , \tag{2.4b}$$

where $(U_{0j}, \ldots, U_{pj}, V_j)$ is a $(p+2)$-dimensional random residual at level 2, with a multivariate normal distribution. By using products with dummy variables this can be written as a single Hierarchical Linear Model, see Goldstein (2011, p. 150). Not all multilevel modeling software will allow for this complexity, but Goldstein's program MLwiN (Rasbash et al. 2014) handles such models straightforwardly.

This model for a two-level nested hierarchy allows studying a dependent variable $Y^{(2)}$ at the higher level, and the idea can be extended to other multilevel structures, not necessarily nested.

An equivalent model was proposed independently by Croon and van Veldhoven (2007) and further elaborated by Lüdtke et al. (2008). These authors proposed models where the regression of level-1 variables is on latent level-2 variables, thus allowing analysis methods that correct for unreliability of measurement of level-2 variables. They developed and investigated estimators using structural equation modeling. Recently, similar models were elaborated for latent classes, i.e., discrete rather than normally distributed latent variables (Bennink et al. 2013).

Models for Social Networks

This section gives an overview of some statistical models for explaining social networks, as represented by directed graphs; we will focus on models and issues that are related to the treatment of multilevel networks in the next section. A wider overview of statistical models for networks is given in Snijders (2011).

The nodes $1, \ldots, n$ of the digraph refer to social actors, and ties are represented by tie variables Y_{ij} with the value 1 if a tie $i \to j$ exists, and 0 otherwise. The digraph then can be represented by its adjacency matrix $\big(Y_{ij}\big)_{[1 \le i,j \le n]}$. Y denotes the random digraph and y one outcome, or realization of it; henceforth we shall usually denote outcomes, or deterministic variables, by small letters and random variables by capitals.

The Basic Multilevel Nature of Social Network Analysis

Social network analysis (Wasserman and Faust 1994; Carrington et al. 2005) is fundamentally a multilevel affair with a focus on relations rather than attributes, thereby combining the actor level and the dyadic level. A basic issue for social network analysis is the study of how relations—the dyadic level—and individual characteristics—the monadic level—impinge on one another. This has led to models studying how a given, fixed network influences individual actor attributes, with a variety of network autocorrelation models (e.g. Doreian 1980; Leenders 2002) and models for social influence (Friedkin 1998). Network autocorrelation models use correlation structures to represent dependencies between the values of linked actors. In this volume, they are used in the contributions by Agneessens and Koskinen (2015) and Bellotti et al. (2015). Another way to model this was proposed by Tranmer et al. (2014), who used the multiple membership models of Browne et al. (2001) to represent network effects on individual outcomes. This has the limitation that the network effects are represented only by additive random effects of the affiliations of the individual, and the advantages of flexibility in choosing these affiliations (which can include, e.g., clique or other subgroup memberships) and the possibility to combine this with other random effects, representing other types of context. This method is used in this volume in Tranmer and Lazega (2015).

In the literature about social support and social capital, multilevel models have been used for studying characteristics of ties in egocentric networks, taking into account the hierarchical structure of ties nested in egocentric networks (van Duijn et al. 1999). In this field, Wellman and Frank (2001) specifically paid attention to the importance of including in the model not only attributes calculated for the actor and the dyadic level, but also for the network level more generally.

This chapter focuses, however, on models for networks where the collection of ties itself is the dependent variable. While in traditional models for social networks the focus was on the relations, and individual attributes were considered quite circumspectly or as an afterthought, modern statistical methods representing network data are in the realm of generalized linear models and incorporate dyadic as well as actor attributes in a very straightforward way; we see this, e.g., in MRQAP modeling (Dekker et al. 2007), the p_2 model (van Duijn et al. 2004), latent space models (Hoff et al. 2002), exponential random graph models (Lusher et al. 2013), and stochastic actor-oriented models (Snijders 2001). The presence of variables defined at different levels does not by itself bring these models close to the Hierarchical Linear Model, however—the exception being the p_2 model.

As discussed in Snijders (2011), there are several quite different approaches for representing network dependencies in probability models that can be used as a basis for statistical inference. Leaving aside conditionally uniform models (which cannot incorporate general attributes) and MRQAP (which controls for network structure but does not represent it), we can distinguish latent variable models, of which the p_2 model, latent space models, and stochastic block models (Nowicki and Snijders 2001) are major representatives; exponential random graph models; and stochastic actor-oriented models as the main approaches.

p_2 Model

Let us begin with the p_2 model. For a network represented by a digraph on n nodes, it postulates the existence of random sender effects $U = (U_1, \ldots, U_n)$ and random receiver effects $V = (V_1, \ldots, V_n)$. As proposed in van Duijn et al. (2004), conditionally on (U, V) and given dyadic covariates $x_h = \big(x_{hij}\big)_{[1 \leq i,j \leq n]}$ (some or all if which may depend only on i or only on j, making them actor covariates), in the p_2 model the probability distribution for each dyad (Y_{ij}, Y_{ji}) is given by

$$P\{(Y_{ij}, Y_{ji}) = (a, b) \mid U, V\} = c_{ij} \exp\Big(a \Big(\sum_h \beta_h x_{hij} + U_i + V_j\Big) + b \Big(\sum_h \beta_h x_{hji} + U_j + V_i\Big) + a\, b\, \rho\Big) \tag{2.5}$$

where $a, b \in \{0, 1\}$ and c_{ij} is a norming constant independent of (a, b). One of the covariates will be constant, representing the intercept. ρ is a reciprocity parameter. Variables U_i and V_i are, respectively, the latent sender and receiver effects at the actor level, and can be correlated for the same actor i, but are independent across different i. Conditional on (U, V), the dyads (Y_{ij}, Y_{ji}) are assumed to be independent but there is dependence between Y_{ij} and Y_{ji} with a strength depending on parameter ρ. In this way, random effects are used to represent those dependencies between network ties that follow from actor differences, while the model also represents tendencies toward reciprocity. In the bestiary of statistical models, this qualifies as a generalized linear mixed model, and therefore is akin to the Hierarchical Linear Model.

It should be noted that the p_2 model is a close relative of the so-called Social Relations Model (Kenny and La Voie 1985; Kenny et al. 2006), a random effects model with a similar structure for continuous relational variables Y_{ij} assumed to have normal distributions. The relation between the Social Relations Model and the Hierarchical Linear Model was discussed in Snijders and Kenny (1999).

Latent Space Models

Another latent variable model for networks is the latent metric space model, proposed by Hoff et al. (2002). Here the nodes in the network are assumed to have locations in a metric space, and the probability of a tie depends on the distance between the nodes. Denoting the location of node i by α_i, and the distance between α_i and α_j by $d(\alpha_i, \alpha_j)$, the probability of a tie in this model is given by

$$\text{logit}\big(P\{Y_{ij} = 1 \mid \alpha\}\big) = -d(\alpha_i, \alpha_j) + \sum_h \beta_h x_{hij} \tag{2.6}$$

where again x_{hij} are values of covariates with logistic regression coefficient β_h. This expresses that actors who are closer to each other, controlling for covariates, have a larger probability of being tied. Although the model was formulated for arbitrary metric spaces, it is being applied mainly for 2- or 3-dimensional Euclidean spaces.

This model was extended by Handcock et al. (2007) to a random effects model for the locations according to a mixture model, with the purpose to represent clusters of actors. Krivitsky et al. (2009) further extended this to a model where also the actors have main effects for activity U_i and popularity V_j,

$$\text{logit}\big(P\{Y_{ij} = 1 \mid \alpha, U, V\}\big) = -d(\alpha_i, \alpha_j) + \sum_h \beta_h x_{hij} + U_i + V_j \tag{2.7}$$

where the U_i and V_i are (unfortunately!) assumed to be independent.

One of the attractive features of the latent Euclidean space models is their visual interpretation: an estimated 2-dimensional model corresponds directly to a graphical layout of the network, where ties will correspond to relatively short distances.

Exponential Random Graph Models

The Exponential Random Graph Model, fondly abbreviated to ERGM, is a generalized linear model for graphs and digraphs, representing the dependence between the ties in a direct way. It was proposed by Frank (1991) and Wasserman and Pattison (1996), and is treated in the extensive recent textbook by Lusher et al. (2013).

This model is defined by the probability function

$$P_\theta\{Y = y\} = \exp\Big(\sum_h \theta_h u_h(y) - \psi(\theta)\Big), \tag{2.8}$$

where y is the digraph, the $u_h(y)$ $(h = 1, \ldots, p)$ are statistics of the graph, and θ is a p-dimensional parameter. The function $\psi(\theta)$ takes care of the normalization requirement that the probabilities sum to 1. There may be covariates defined on the nodes, and on the dyads, on which the $u_h(y)$ may depend. This is still an extremely general model, and Snijders et al. (2006) discussed how to specify it in practically feasible and fruitful ways, avoiding the so-called 'near-degeneracy'. Lusher et al. (2013, Chapter 6) contains an extensive presentation of statistics $u_h(y)$ that may be included in the specification of an ERGM.

The dependence on actor and dyadic covariates can be implemented by defining some of the $u_h(y)$ as sums of ties weighted by covariates, such as

$$u_h(y) = \sum_{i,j} v_i y_{ij}$$

for the sender effect of an actor covariate V, or

$$u_h(y) = \sum_{i,j} v_{ij} y_{ij}$$

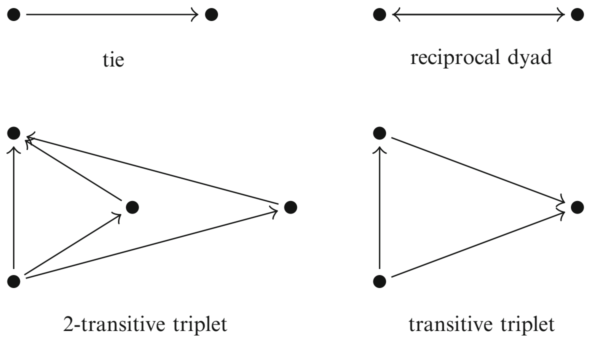

Fig. 2.1 Examples of subgraphs, counts of which are used in ERG models

for a dyadic covariate V. Dependence between tie variables, such as reciprocity and transitivity, is expressed by defining some of the $u_h(y)$ to be counts of subgraphs like those in Fig. 2.1. The literature mentioned explains this more fully, e.g., Lusher et al. (2013, Chapter 7).

Stochastic Actor-Oriented Models

Longitudinal network data potentially give much more information about the antecedents as well as consequences of network configurations than cross-sectional data. They also require more effort to collect, but there already are a large number of longitudinal network data sets, and their number is growing.

The Stochastic Actor-oriented Model ('SAOM'; Snijders 2001) is a statistical model for network dynamics that has been developed for the interdependent dynamics of networks and(monadic) actor variables (Steglich et al. 2010) and various other network structures. We sketch it here for the case of interdependent networks and actor variables, calling the latter 'behavior' just as a general term, and denoting the 'behavior' of actor i by Z_i. The network is Y, the vector of behaviors for all actors is $Z = (Z_1, \ldots, Z_n)$. The method assumes that data are available for a number of discrete observation moments, the panel waves, and that the process of change in network and behavior runs on in between the observation moments. The probabilities of changes in network ties depend on the network configurations in which the actor is involved who sends the ties; this can be formulated in a model where the changes in network and in behavior result from choices by the actors. The interpretation is that actors control their outgoing ties and their behavior, subject to constraints determined by network context, attributes, and path dependence (inertia).

In the basic model, the network is a directed graph and the behavior is a discrete variable with a finite number (say, 2–10) of ordered categories, integer coded $(1, 2,$ etc). The time parameter is continuous, meaning that at any moment between the observations, a change in tie or behavior is possible. The model is a Markov chain, which means that the probabilities of change at any moment depend only on the current state (y, z) of the network and behavior, together with the available covariates. The dynamic process is defined as follows. At random moments, the frequency of which is determined by 'rate functions', a randomly selected actor i gets the opportunity to change either one outgoing network tie Y_{ij} or the behavior Z_i. The behavior can change only by unit steps, $+1$ or -1. The actor can also let the network and behavior stay as it is. The network tie to be changed, or the change in the behavior, is determined probabilistically by the so-called *evaluation functions* and the current state of the network and behavior (y, z). There are separate evaluation functions for the network and the behavior, and the probability of a particular change is greater when it would lead to a higher change in the evaluation function.

Specifically, the model has two components, a waiting model for timing of changes and a choice model for outcome of changes. The timing component is relatively simple. It must satisfy the consequence of the Markov assumption that waiting times between changes have an exponential distribution; to this are added considerations of interdependence between actors, and interdependence between networks and behavior. The assumption is that each actor has a rate function $\lambda_i^Y(y, z)$ for the network and a rate function $\lambda_i^Z(y, z)$ for the behavior. The waiting time for the next opportunity for a change in an outgoing tie of actor i is exponentially distributed with parameter $\lambda_i^Y(y, z)$, and for the next opportunity for a change in behavior of actor i it is exponentially distributed with parameter $\lambda_i^Z(y, z)$. At any given moment, the briefest of these waiting times across all actors is selected, the choice model is activated, which usually will lead to a change in state, and then the model starts again with the new state.

To define the choice model, suppose that the current state of the network and behavior combination is $(y^{(0)}, z)$, and actor i gets the opportunity for a network change. Then the set $\mathscr{C}$ of possible networks that could result from this change opportunity is composed of all networks y' for which in comparison with network $y^{(0)}$ exactly one outgoing tie $i \to j$, for some $j \neq i$, is either added or dropped; and, in addition, the network $y^{(0)}$ itself, representing no change. Denote the evaluation[4] function for the network for actor i by $f_i^Y(y, z)$, defined for all possible network-behavior configurations (y, z). The probability that the resulting network is y' is given by

$$P\{\text{next } Y = y'\} = \frac{\exp\big(f_i^Y(y', z)\big)}{\sum_{y \in \mathscr{C}} \exp\big(f_i^Y(y, z)\big)} \qquad (y' \in \mathscr{C}) \,. \tag{2.9}$$

For behavior changes the set of possible changes has only 3 elements: up, stay, down; and the evaluation function for behavior is used. For the rest, all is analogous. The dynamic process then consists of a repetition of these steps, where the result of the previous step is always he starting point of the next.

[4]We restrict the discussion to specifications with only an evaluation function; see Ripley et al. (2015) for more general models.

The heart of the model is the specification of the evaluation functions. These are defined as linear combinations of theoretically argued and/or empirically necessary characteristics of the network and the behavior,

$$f_i^Y(y,z) \;=\; \sum_h \beta_h^Y s_{ik}^Y(y,z) \text{ and } f_i^Z(y,z) \;=\; \sum_h \beta_h^Z s_{ik}^Z(y,z) \,. \qquad (2.10)$$

These characteristics $s_{ik}^Y(y,z)$ and $s_{ik}^Z(y,z)$ are called 'effects.' On the network side, these can be dependent on the network position of actor i. For example, tendencies toward reciprocity and transitivity, respectively, can be represented by positive parameters for the reciprocity and transitive triplets effects,

$$s_{ik}^Y(y) \;=\; \sum_j y_{ij}\, y_{ji} \,, \qquad\qquad s_{ik}^Y(y) \;=\; \sum_{j,h} y_{ij}\, y_{jh}\, y_{ih}$$

as in Fig. 2.1; but, contrasting with ERG modeling, the role of actor i is now special, as it is used to denote the focal actor of whom the evaluation function is being considered.

The network and behavior dynamics become interdependent when some of the effects for network change of actor i, $s_{ik}^Y(y,z)$, depend on behavior z, not only on the behavior of the actor i but also of the other actors. E.g., the cross-product 'ego × alter behavior' interaction term

$$s_{ik}^Y(y,z) \;=\; \sum_j y_{ij}\, z_i\, z_j$$

will reflect (if it has a positive coefficient) that actors who have themselves a higher value of z_i will have a larger probability to create and maintain ties with other actors j accordingly as these in their turn have a higher z_j value. On the other side, some of the effects for behavior change of actor i, $s_{ik}^Z(y,z)$, can depend on the network y. An example is the 'average behavior alter' effect

$$s_{ik}^Z(y,z) \;=\; \frac{z_i \sum_j y_{ij}\, z_j}{\sum_j y_{ij}} \,,$$

defining $0/0 = 0$. If its coefficient is positive, this effect will imply that actors whose connections have on average a higher z_j value, will themselves tend to increase more, or decrease less, in their own z_i value. In models including such effects, the changes in the network lead to changes in the change probabilities for behavior and vice versa: the actors are each others' changing environment.

These dynamic models can be studied by computer simulation which is also how parameters are estimated: see the mentioned literature. Further information is at http://www.stats.ox.ac.uk/~snijders/siena/.

Choice of Model

The range of statistical network models is starting to be bewildering and it may be helpful to point out some differences in their properties. All these models can incorporate fixed effects of quite arbitrary covariates, so the difference is only in how they represent structural network features.

The p_2 model represents only three aspects of networks: differences between actors in popularity (indegrees) and activity (outdegrees), and reciprocity. Further structural features such as transitivity are not modeled.

The latent Euclidean space models represent networks by embedding the actors, as nodes, in a 2- or 3-dimensional Euclidean space. (More dimensions are possible but unusual.) This is visually very attractive. Network dependencies such as reciprocity, transitivity, and higher-order dependencies are represented only as consequences of this embedding. On the one hand the model is inflexible in the representation of network dependencies, as there are no free parameters for this purpose: the tendencies towards reciprocity and transitivity follow jointly from the spatial arrangement of the nodes, and cannot further be tuned. On the other hand the model is very flexible in choosing the locations of the nodes. This has a downside: the likelihood surface for the location of the nodes is often quite multimodal, a problem that is not really resolved by giving the locations a probability distribution as in a random effects model. I think it is doubtful that the intricacies of social space can be well represented by Euclidean space.

The Exponential Random Graph Model represents network dependence directly by using subgraph count statistics as statistics $u_h(y)$ in (2.8), as discussed in Lusher et al. (2013, Chapters 6, 7). A large number of triadic and higher-order structures can be considered, and are indeed used in practical network research, as is illustrated by the same book. The Stochastic Actor-oriented Model represents network dependencies, somewhat similarly, by the effects $s_k(y)$ in (2.10), and here also a large array of structural effects can be considered (Ripley et al. 2015).

An illuminating difference between ERGM and SAOM models on the one hand, and latent variable models (spatial or otherwise) on the other hand, is the consequence of restriction to a smaller set of nodes and the importance of network delineation. The former models do not allow restriction to a random subset of nodes; for the ERGM this was elaborated in Snijders (2010). The reason is that ERGMs and SAOMs represent dependencies, and cutting off arbitrary nodes would be an amputation. For the latent variable models, on the other hand, it is conceptually unproblematic to consider only a subset of nodes: if a random subset of nodes with their incoming and outgoing ties is dropped, the information available in the data is reduced but the model formulation of the rest remains intact. In practice, however, it appears that working with a somewhat restricted node set in ERGMs and SAOMs usually does not strongly change results except for the fact that the data set is less informative, so this difference may be more important theoretically than practically.

This issue may be regarded as a practical advantage of latent variable models, but it also highlights that these represent networks in a descriptive way but not in their essential dependence structure.

In some research the focus is on the structural dependencies directly, and then the ERGM and SAOM will be preferable. In other research the estimates of the random effect variances (sources of variability) and the posterior predictions of the random effects and spatial locations may be important, leading to preference for a latent variable model.

Whether the latent variable approach or the directly structural approach of the ERGM and the SAOM yield a better representation of empirical social networks is still an open question. In a sense this question is ill-posed because both models have flexible opportunities for model specification, so a poor fit may always be remedied by a more appropriate specification. Other open questions include: how important a good fit for such models is in practice; and how robust conclusions can be for a model that fits poorly on characteristics that has a poor fit on characteristics that are secondary to the main research questions.

Multilevel Network Analysis

The combination of the terms 'multilevel' and 'social networks' leads to a multiplicity of directions. Above it was mentioned that social networks combine different types of units—social actors and social ties—and variables can and will be defined on both of these sets. Varieties of the ERGM (Daraganova and Robins 2013) and of the SAOM (Steglich et al. 2010) combine dependent network variables with dependent actor variables. But this volume is about other combinations. The current section is about *multilevel network analysis*: the combined network analysis for several independent groups. Section "Analysis of Multilevel Networks" is about *analysis of multilevel networks*: the analysis of structures with nodes of several types, connected by ties of several types.

Why Combine Several 'Parallel' Networks?

Multilevel network analysis, where the term 'multilevel' is used in the sense of hierarchical nesting, is a combined network analysis for several groups, applying the same model to each group. We then have several networks, with different actor sets and assumed to be mutually independent, that may be combined in a single analysis with a common model but allowing parameter values to be different. Why should we do this?

In general, multilevel analysis may have several main purposes. I formulate them for the case where individuals are the lower-level units and groups the higher-level units. These purposes are entwined, and the salience of each of them will differ depending on the application considered.

⇒ Obtain results from the combination of data sets about multiple groups, taking into account the ‘random’ variability between individuals within groups as well as the ‘random’ variation between groups, with standard errors (or other measures of uncertainty of the results) that account for these two sources of variation.
⇒ Increase the amount of information (sample size) compared to analyzing a single group.
⇒ Generalize to the population of groups.
⇒ Test effects of group-level variables.
⇒ Analyze the groups jointly in a way that allows more detail and precision than would be possible when analyzing the groups separately. This sometimes is formulated by saying that the analysis of each group ‘borrows strength’ (Morris 1983) from the other groups, which is possible because of the assumption that this group is a member of the same population as the other groups. This is related to the idea of ‘empirical Bayes’ estimates mentioned in section “Frequentist and Bayesian Estimation”.

All except the last purpose are also, potentially, goals of meta-analysis (e.g., Hedges and Olkin 1985). The main difference between multilevel analysis and meta-analysis is that, usually, meta-analysis is a two-step procedure, using finished analyses of the single groups and combining these in overall conclusions, whereas multilevel analysis usually unites these two parts of the analysis. Meta-analysis also can be more liberal with respect to the model assumptions concerning the group level. The correspondence between meta-analysis and multilevel analysis is discussed in Raudenbush and Bryk (2002), Chapter 7, and Snijders and Bosker (2012), Section 3.7. A two-step approach can also be used in multilevel analysis provided that the groups individually are large enough, cf. Achen (2005).

While we assume that the same model applies to all groups, they will have different parameter values. In addition, groups will usually have different sizes and different distributions of explanatory variables; in consequence, the standard errors resulting from analyses per single group will also differ across groups. To be used in a valid way, meta-analytic and two-step approaches should take these differences into account—which is automatic in multilevel analysis via the Hierarchical Linear Model.

For multilevel network analysis, any or all of these purposes may apply. One major purpose is to generalize to a population of networks. It was noted by Snijders and Baerveldt (2003) and Entwisle et al. (2007) that traditional social network analysis focused on the analysis of single networks, while nevertheless usually implying that the mechanisms and processes uncovered have a larger validity than only for the particular group under study. But these authors also noted that more and more studies are being done where data is collected for multiple networks considered to be similar. On the level of networks, traditional social network research mostly was based on $N = 1$ studies. To have a statistical basis for generalizing to a wider population, however, one needs to analyze data for several networks that may be regarded, in some sense, as replications of each other. The target population then will be a population of networks, and almost always will be somewhat vaguely described and perhaps have a somewhat hypothetical nature. This is often the case for the populations at higher levels in multilevel analysis. Above, Cox (1990) and Sterba (2009) were already mentioned as references about this topic; some further philosophical considerations about the use of probability models for multilevel and network data are presented in Sections 1.1.1 and 14.1.1 of Snijders and Bosker (2012) and on pages 135–137 of Snijders (2011). The practical question is whether a particular collection of networks is homogeneous enough with respect to the social processes taking place to justify pursuing a common conclusion by using all of them together; as well as to justify applying a common statistical model, with parameters that are allowed to vary from group to group according to a joint probability distribution in the population of groups.

The 'replications' may be network studies in several similar schools, several similar companies, etc. The *Adolescent Society* study of Coleman (1961) was based on detailed investigations of friendship networks in 10 schools, juxtaposed as 10 interconnected case studies. More recent examples such as the *PROSPER* study (Moody et al. 2011), the *ASSIST* study (Campbell et al. 2008; Steglich et al. 2012), and the *School Social Environments* study (Light et al. 2013) have provided network data to be analyzed by multilevel or meta-analytic means.

Two-Step Meta-for-Multilevel Network Analysis

In the following model for two-step meta-analysis, the population at the higher level is made explicit. It is assumed that independent groups—in the meta-analysis case these may be individual studies or publications—are combined, being regarded as a sample from a population of groups. The focus often is on one parameter at a time, so that the parameter is one-dimensional and denoted by θ. The dependent variable at the group level is the parameter estimate from group k, denoted by $\hat{\theta}_k$. The assumption of the random effects model for meta-analysis (cf. p. 210 in Raudenbush and Bryk 2002; Snijders and Bosker 2012, p. 37) is

$$\hat{\theta}_k = \theta_k + R_k = \mu_\theta + E_k + R_k \, . \tag{2.11}$$

Here θ_k is the true parameter in group k; R_k is the estimation error within this study; μ_θ is the mean of parameter θ in the population of groups; and E_k is the deviation of this group from the population mean. R_k reflects within-group variability and E_k reflects between-group variability. From the point of view of estimating θ_k, R_k is regarded as error variation and E_k as true variation.

These are independent residuals both with expected value 0. The secret of this analysis method is that the within-group analysis provides us with an estimate of the standard error $\sigma_k = \text{s.e.}(\hat{\theta}_k)$ which is the standard deviation of R_k, and we act (almost always) as if we know this standard error exactly. Armed with this extra information we can estimate not only μ_θ but also $\text{var}(E_k) = \text{var}(\theta_k)$ without the 'hat' on top of θ, the 'true between-group variance' of θ_k; as opposed to

$$\text{var}(\hat{\theta}_k) = \text{var}(R_k) + \text{var}(E_k) ,$$

which is the 'observed between-group variance'.

If the number of groups is large enough, such a study also permits the assessment of effects of variables X_h at the group level, by entering them in the model as predictor variables:

$$\hat{\theta}_k = \mu_\theta + \sum_h \beta_h x_{hj} + E_k + R_k , \tag{2.12}$$

where x_{hj} is the value of X_h for group k. In most practical cases the number of networks in a data set for a multilevel network analysis will be not very large, so the number of variables X_h of which the effect can be studied will be low.

For model (2.11) an explicit estimator in a network context was suggested by Snijders and Baerveldt (2003), using a method derived by Cochran (1954). The maximum likelihood (ML) or restricted maximum likelihood (REML) estimators under the assumption that R_k and E_k have normal distributions will usually be more efficient. This can be calculated by multilevel software such as HLM (Raudenbush et al. 2011) and MLwiN (Rasbash et al. 2014), and by R packages such as metafor (Viechtbauer 2010). This two-step approach was used for multilevel network analysis, e.g., by Lubbers (2003) and Schaefer et al. (2011) who combined ERGM analyses for several groups; and by Mercken et al. (2012) and Huitsing et al. (2014) who combined Stochastic Actor-oriented Models for several groups.

Integrated Multilevel Network Analysis

The other possibility is to integrate the within-network and between-network models in one joint model and analyze this in one simultaneous analysis. The generic way to do this is by postulating a between-network probability model, where the parameters of the within-network model are supposed to be drawn independently from a common across-network distribution: in other words, a random effects model. This is more complicated than the two-step approach, and for every type of within-network model a multilevel model has to be specifically elaborated. The integrated approach is sketched in Sweet et al. (2013, Section 2), who call this the *Hierarchical Network Model*.

The great potential advantage to this is the possibility of 'borrowing strength' as was mentioned above. In many settings where network data are collected, the groups are rather small—e.g., school classes with sizes between 20 and 40—and for each group separately an analysis might be possible only with a quite meagre model specification. The consequence then is that various effects of focal interest may have to be left out because the data for each individual group does not support parameter estimation for a truly interesting model, or the possibilities of controling for additional or competing mechanisms are reduced. In such a case, a random effects multilevel model can be very helpful; sometimes the analysis may even be impossible without it. In addition, an integrated random effects multilevel model will often be more efficient, and an integrated analysis may be in itself more attractive than a two-step analysis.

The first multilevel network analysis model of this kind was presented by Zijlstra et al. (2006), a multilevel version of the p_2 model. To define this extension, indicate the groups by k and the tie variable from actor i to actor j in group k by Y_{kij}. The simplest multilevel version of the p_2 model (2.5), containing random intercepts W_k for the groups, then is given by

$$\begin{aligned} P\{(Y_{kij}, Y_{kji}) = (a, b) \mid U, V, W\} \\ = c_{kij} \exp\Big(a\Big(\sum_h \beta_h x_{hij} + W_k + U_i + V_j\Big) \\ + b\Big(\sum_h \beta_h x_{hji} + W_k + U_j + V_i\Big) + a\, b\, \rho\Big), \end{aligned} \tag{2.13}$$

again for $a, b \in \{0, 1\}$, where c_{kij} does not depend on a or b. This means that (on the logistic scale) there is a random main effect for the groups, but further they are similar. More elaborate models can be obtained by adding random slopes for some of the X_h, and the reciprocity coefficient ρ may also get a random effect.

Several applications of this model were published, e.g., by Vermeij et al. (2009) and Rivellini et al. (2012).

There is a lot of recent and current activity in extending other network models to multilevel versions. Sweet et al. (2013) elaborated their 'Hierarchical Network Model' for the case of the latent Euclidean space model, and presented an application with a random intercept and an (unfortunately, uncorrelated) random slope. In another publication (Sweet et al. 2014) these authors elaborated a multilevel version of the hierarchical mixed membership latent block model of Airoldi et al. (2008). Koskinen and Snijders (2016) are working on a multilevel extension of the Stochastic Actor-oriented Model, and a brief documentation of this is given in Ripley et al. (2015).

Hierarchical Structures

Much like the situation of multilevel analysis with the Hierarchical Linear Model and its variants, multilevel network analysis is also a hierarchical type of model for a hierarchical data structure. Estimation for this hierarchical data structure again may be regarded as empirical Bayes estimation, where the group-level parameters θ_k have a frequency distribution about which we get information thanks to the observed sample of groups. The analysis of each group borrows strength from the data of the other groups. Therefore, multilevel network analysis is particularly appropriate for combining the data of many small networks, each of which would be too small to permit analysis by a suitably specified ERGM or SAOM.

For single-level as well as multilevel network analysis, frequentist as well as Bayesian estimation methods have been proposed. Bayesian methods are potentially more compatible with the hierarchical nature of multilevel network analysis, and may be helpful for incorporating prior knowledge in cases where the number of groups is rather small. More research is needed to make meaningful comparisons between estimation methods, be they Bayesian or frequentist, for these complicated models.

Analysis of Multilevel Networks

Brass et al. (2004) proposed that for network studies in organizational research, it is important to consider both intra-organization and inter-organization networks. Lazega et al. (2008) pioneered a study with a linked intra- and inter-organizational design. Models and methods for the complex network structures that are necessary for the analysis of such designs are now in an early stage of development, and this volume aims to contribute to this domain.

A multilevel network (Wang et al. 2013) can be defined as a network with nodes of several types, where a distinction is made between types of ties according to the types of nodes that they connect. Thus, if types of nodes are A, B, C, etc., there is a distinction between $A - A$, $B - B$, $C - C$ ties, etc., and also between $A - B$, $A - C$, etc., ties. The first are intra-type, the second inter-type ties. Some of the networks may be the networks of interest, others may be fixed constraints, still others may be non-existent or otherwise outside of consideration. The intra- and inter-organization network of Brass et al. (2004) and Lazega et al. (2008) is composed of organizations (type A) and their members (type B), where $A - A$ ties can be organizational cooperation, competition, etc., while $B - B$ ties can be interpersonal collaboration, acquaintance, etc. The primary two-mode $A \times B$ network then will be the membership or affiliation network, where the simplest situation is one of complete nesting, and each individual is a member of exactly one organization; the $B \times A$ network may be superfluous, and then could be defined formally as an empty network. The design will be especially interesting if $B - B$ ties between members of different organizations are also recorded, so that interpersonal ties within as well as between organizations can be included in the analysis. Another example is the co-evolution of a one-mode and a two-mode network as studied by Snijders et al. (2013), where A is a set of individual students, B a set of companies, the $A \times A$ network represents friendship or advice ties, while the two-mode $A \times B$ network represents that the student is potentially interested to work for this company; $B \times B$ and $B \times A$ networks were not used.

Fig. 2.2 Adjacency matrix for combined node set

$$\begin{array}{c c} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} \text{one-mode } A\times A & \text{two-mode } A\times B \\ \text{two-mode } B\times \text{A} & \text{one-mode } B\times B \end{pmatrix} \end{array}$$

This kind of multilevel network can potentially be studied by extensions of the models mentioned above. This is sketched in the following sections for the Exponential Random Graph Model and the Stochastic Actor-oriented Model.

A representation that is quite generally useful for handling multilevel relational structures was proposed by Wasserman and Iacobucci (1991). This defines a combined node set as the union or disjoint union of the A, B, etc., node sets. The combined node set allows treating the various one-mode and two-mode networks as subgraphs of an overall graph, with its associated adjacency matrix as in Fig. 2.2.

If some of the within-type or between-type networks are undefined, meaningless, or not studied for other reasons, the corresponding sub-matrices can be defined as structurally null blocks, i.e., having all entries equal to 0.

Exponential Random Graph Models for Multilevel Networks

Mathematically, model (2.8) can be used straightforwardly for multilevel networks, because it defines a general exponential family of graphs (directed or non-directed), and the node set can be taken as the union or disjoint union of the A, B, etc., node sets, as mentioned above. The outcome space of graphs can be restricted so that certain blocks in the adjacency matrix are fixed; e.g., a two-mode network of affiliations of individuals to organizations might be considered an exogenously fixed datum of the analysis.

Of course, turning the general ERG model into a model for multilevel networks in this way is not as easy as it might seem from the previous sentences. The model must be specified in a way that corresponds to the differences between the node sets; and the existing algorithms must be tuned for the estimation of parameters in the model. This was accomplished by Wang et al. (2013). The following is a very brief sketch.

To express the ERGM for a multilevel network with two node sets A and B, let us refer to the one-mode $A\times A$ and $B\times B$ networks by A and B (a manageable misuse of notation) and to the two-mode $A \times B$ cross-level network by X. Then the multilevel network can be denoted by (y_A, y_B, y_X), and the vector of statistics $s(y)$ in (2.8) can be split into parts depending on each of y_A, y_B, and y_X separately, and each of their combinations; leading to the formulation of the multilevel ERGM as

$$\begin{aligned} P_\theta\{(Y_A, Y_B, Y_X) = (y_A, y_B, y_X)\} = \exp\big(&\theta_A\, s_A(y_A) \;+\; \theta_B\, s_B(y_B) \\ &+\; \theta_X\, s_X(y_X) \;+\; \theta_{AX}\, s_{AX}(y_A, y_X) \\ &+\; \theta_{BX}\, s_{BX}(y_B, y_X) \;+\; \theta_{ABX}\, s_{ABX}(y_A, y_B, y_X) - \psi(\theta)\big)\,, \end{aligned} \tag{2.14}$$

where $\theta = (\theta_A, \theta_B, \theta_X, \theta_{AB}, \theta_{AX}, \theta_{BX}, \theta_{ABX})$. The θ and s symbols all denote vectors. This decomposes the model in parts with the following statistics:

$s_A(y_A)$ internal dependence of the one-mode network A, specified as in Lusher et al. (2013, Chapter 6).

$s_B(y_B)$ internal dependence of the one-mode network B, analogous.

$s_X(y_X)$ internal dependence of the two-mode network X, specified as in Lusher et al. (2013, Section 10.2).

$s_{AX}(y_A, y_X)$ bivariate interdependence between the A and X networks; interdependence between a one-mode and a two-mode network is not treated specifically in the ERGM literature (as far as I know), but since two-mode networks have less structural features than one-mode networks, the directions for specifying bivariate networks given in Lusher et al. (2013, Section 10.1) can be followed.

$s_{BX}(y_B, y_X)$ bivariate interdependence between the B and X networks, analogous.

$s_{ABX}(y_A, y_B, y_X)$ three-way interdependence between the A, B, and X networks, to which Wang et al. (2013) is specifically devoted. For example, a basic three-way effect expressing the multilevel structure is the effect that ties between individuals will tend to go together with ties between the organizations they are members of. This is the $C4AXB$ effect discussed in their Section 6.5.

In practice, all cross-level dependencies will be crucial in giving a meaningful representation of the multilevel network, and the three-way interdependence represented by $s_{ABX}(y_A, y_B, y_X)$ will often be the main point of scientific interest. The other parameters are also interesting in their own right. Wang et al. (2013) find that including three-way and other between-level dependencies may simplify the intra-network models compared to modeling the A, B, and X networks independently, which reflects the theoretical notion that internal structure will be shaped depending on external or contextual demands, pressures, and possibilities, and 'controlling for' the between-level dependencies gives a purified view of the intra-network mechanisms.

As is mentioned in the discussion of Wang et al. (2013), the determination of the levels in a multilevel network can be done in several ways, depending on the aims of the research. One possibility is to define node sets based on their different nature and way of connecting to other nodes, such as individuals and organizations. Another possibility is to distinguish nodes of the same basic kind by attributes, thus permitting a model with arbitrary differences between the ways in which the nodes relate to other nodes, depending on these attributes. The discussion above focuses on the first method, but the multilevel ERG model can be applied also to the other way of determining node sets.

In this volume, this model is applied in several varieties. Two chapters in this volume provide examples of the nested case. Both are about managers in companies. The study by Brennecke and Rank (2015) is concerned with the interdependence of the knowledge sharing network between managers (B) and the R&D collaboration network between the companies (A). Zappa and Lomi (2015) study advice and communication relations between managers (B) in subsidiaries

(A) of an international multi-unit industrial group. The cross-level relation (X) is membership affiliation, the within-A relation is the hierarchical reporting relation between the subsidiaries.

Hollway and Koskinen (2015) apply the multilevel ERGM to a study about multilateral fisheries treaties, where the node sets are the countries (A) and the multilateral treaties (B). This is a crossed rather than nested design because countries can be members of several treaties. The chapter by Brailly et al. (2015) considers one node set of buyers and another of sellers, where moreover the buyers as well as the sellers are nested in their respective organizations. This is analyzed as two separate bipartite buyer $\times$ seller networks, one for the organizations and one for the individuals, where some of the variables of the other level of aggregation (individuals and organizations, respectively) are obtained by projection (aggregation or disaggregation).

The second way of determining the levels is represented by Wang et al. (2015), who present an application of the multilevel ERG model where the two node sets are entrepreneurial and non-entrepreneurial farmers, who differ so strongly in their network structures that a multilevel ERGM is able to give a much better representation than a regular one-mode network analysis. An exploratory method for derivation and specification of hypotheses in multilevel ERG models is proposed by Zhu et al. (2015).

Stochastic Actor-Oriented Models for Multilevel Networks

For the Stochastic Actor-oriented Model likewise, the basic mathematical model explained in section "Stochastic Actor-Oriented Models" can be used,[5] if it is specified in accordance with the multilevel structure. The actor-oriented nature of this model requires specifying something about agency: which sets of actors will be specified as those making the choices? In the standard actor-oriented model for two-mode networks (Koskinen and Edling 2012; Snijders et al. 2013) with node sets A and B, there is agency in only one node set, so ties are regarded as being directed from A to B and determined by the actors of type A.

Again, we consider a multilevel network with two node sets, A and B. In this discussion we leave out the dependent behavioral variable, but it could be added in a rather direct way. In the general situation there could possibly be ties from A to B as well as ties from B to A; for the current exposition the second kind of tie will be ignored, so that again we consider two one-mode networks internal, respectively, to the actor sets A and B; and one two-mode network X supposed to be directed from A to B, with agency in the A nodes.

The specification of the model for the RSiena package (Ripley et al. 2015) is possible by employing the representation with a combined node set $A \cup B$ as above

[5] I thank James Hollway for pointing out this possibility.

$$\begin{array}{cc} & \begin{array}{cc} A & \quad B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} \text{internal } A & 0 \\ 0 & \text{internal } B \end{pmatrix} \\ & \text{networks } A,B \end{array} \qquad \begin{array}{c} \begin{array}{cc} A & \qquad B \end{array} \\ \begin{pmatrix} 0 & \text{two-mode } A \times B \\ [\text{two-mode } B \times A] & 0 \end{pmatrix} \\ \text{network } X \end{array}$$

Fig. 2.3 Two dependent networks for combined node set

but now with two dependent networks, as displayed in the block structure for the adjacency matrices shown in Fig. 2.3. The reason why the data must be separated and treated as two dependent networks instead of one as in the ERGM (Fig. 2.2) will be explained further below.

To avoid confusion, in the rest of this section we shall refer to the original networks as the one-mode and two-mode networks, and to the two constructed networks used for the analysis in RSiena as the multi-networks. Both multi-networks have node set $A \cup B$. The multilevel network is specified as a multivariate network of two multi-networks, consisting of

(1) a one-mode multi-network containing the two one-mode networks as diagonal blocks, and off-diagonal blocks that are structurally 0;
(2) another one-mode multi-network containing the $A \times B$ network as an off-diagonal block; all the rest are structurally zero blocks. If the data structure would also include a $B \times A$ two-mode network with agency in the B nodes, this could be included as the $B \times A$ off-diagonal block in the second multi-network. If the $A \times B$ network would be a fixed context and not a dependent variable (e.g., if it denotes an externally given membership structure), then the second multi-network would be replaced by a dyadic covariate.

The rate functions and evaluation functions have to be differentiated according to the node sets. For the evaluation functions, this differentiation leads to the following structure, where $f_i^A(y_A; y_B, y_X)$ is the evaluation function for actors in A for their ties with other A actors, relevant for Y_A as the dependent variable; $f_i^B(y_B; y_A, y_X)$ the evaluation function for actors in B for their ties with other B actors, for dependent variable Y_B; and $f_i^X(y_X; y_A, y_B)$ the evaluation function for actors in A for their ties to B actors, relevant for dependent variable Y_X:

$$\begin{aligned} f_i^A(y_A; y_B, y_X) &= \theta^A s_A(y_A) + \theta^{AX} s_{AX}(y_A; y_X) + \theta^{ABX} s_{ABX}(y_A; y_B, y_X) \\ f_i^B(y_B; y_A, y_X) &= \theta^B s_B(y_B) + \theta^{BX} s_{BX}(y_B; y_X) + \theta^{BAX} s_{BAX}(y_B; y_A, y_X) \\ f_i^X(y_X; y_A, y_B) &= \theta^X s_X(y_X) + \theta^{XA} s_{XA}(y_X; y_A) \\ &\qquad + \theta^{XB} s_{XB}(y_X; y_B) + \theta^{XAB} s_{XAB}(y_X; y_A, y_B) \,. \end{aligned} \tag{2.15}$$

The functional dependence of these evaluation functions on the other one-or two-mode networks reflects inter-network dependence. The arguments before the semicolon have the role of dependent variable, those after the semicolon are the explanatory variables. Because of the endogeneity according to the Markov model, the state of the Markov process being (y_A, y_B, y_X), the dependent variables also are used as explanations for their own further changes. This model contains more terms compared to the decomposition (2.14) of the ERG model for multilevel networks, because the multivariate associations between two networks are represented in the SAOM—with its 'co-evolution' aspect—as two interdependent one-sided influences.

The separation into two multi-networks (or more, for structures with more than two actor sets) is necessary to separate the choice models. In the SAOM for one network the changes in all the outgoing ties of an actor are considered together, as options in one choice process. Putting the ties of A actors to other A actors in a different network than their ties to B actors means that the ties are chosen in separate, interdependent choice processes; if these ties were put into one multi-network the choices of ties to A would be weighed against ties to B and vice versa, and this would be less natural, given that node sets A and B are of a different nature and $A - A$ ties are conceptually different from $A - B$ ties. The construction of two multi-networks represents that for the A actors there are two distinct but interrelated choice processes, corresponding to the two dependent variables Y_A and Y_X in (2.15), for both of which the agency is with the A actors.

This implies that the multilevel SAOM, contrasting with the multilevel ERGM, is aimed firstly at representing network structures where the several node sets, and especially the ties between several different node sets, are of a different nature. It is less suitable for representing node sets of the same basic kind, differentiated only by an attribute. The different kinds of ties in the multilevel SAOM are distinguished also by having their own timing models, which play no role in the multilevel ERGM. The notion that compensation between different outgoing ties of one actor (e.g., a collaboration tie from i to j_1 may serve the same purpose for i as a collaboration tie from i to j_2) is meaningful for ties of the same kind, but less so between the different sets of ties $A - A$ versus $A - B$, is built into the choice model and also in the model specification for the SAOM—the choice of the effects in (2.15)—, whereas for the ERGM it is only built into the choice of the effects (2.14).

A Forward Look

Multilevel analysis of networks (section "Multilevel Network Analysis") is a natural and important development as more and more data sets are collected that contain similar 'parallel' networks in multiple groups—disconnected groups, or at least, sets of groups for which the inter-group connections are being ignored in the analysis. One of its great advantages is that it allows the study of contextual effects at the network level, i.e., the effects of network-level variables. The analysis of multilevel networks (section "Analysis of Multilevel Networks"), on the other hand, is a different and greater conceptual step. It permits studying in one model the structure of ties between several different node sets, which has some similarity to developments in multilevel analysis that permit studying dependent variables at any level, as discussed in section "Dependent Variables at Any Level". Thereby it enables the representation of social systems with multiple agency and of the structural effects of combined agency patterns. Applications of multilevel ERGMs have started to appear and are contained in this volume; applications of multilevel SAOMs will be coming. These new techniques may well have interesting repercussions on theory development.

The research program heralded by Coleman (1959) has flourished in the past half century with the development of multilevel analysis and social network analysis. Their combination is a young branch on this tree, or rather two branches, one being multilevel analysis of networks and the other the analysis of multilevel networks. This book reflects some of its recent developments and hopefully contributes to further blossoming.

References

Achen, C. H. (2005). Two-step hierarchical estimation: Beyond regression analysis. *Political Analysis, 13*, 447–456.

Agneessens, F., & Koskinen, J. (2015). Modelling individual outcomes using a Multilevel Social Influence (MSI) model: Individual versus team effects of trust on job satisfaction in an organisational context. In E. Lazega & T. A. B. Snijders (Eds.), *Multi-level network analysis for the social sciences: Theory, methods and applications* (pp. 81–106). Cham: Springer.

Airoldi, E. M., Blei, D. M., Fienberg, S. E., & Xing, E. P. (2008). Mixed membership stochastic blockmodels. *The Journal of Machine Learning Research, 9*, 1981–2014.

Aitkin, M., & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society, Series A, 149*, 1–43, with discussion.

Alker, H. R. (1969). A typology of ecological fallacies. In M. Dogan & S. Rokkan (Eds.), *Quantitative ecological analysis in the social sciences* (pp. 69–86). Cambridge, MA: MIT Press.

Bellotti, E., Guadalupi, L. & Conaldi, G. (2015). Comparing fields of sciences: Multilevel networks of research collaborations in Italian academia. In E. Lazega & T. A. B. Snijders (Eds.), *Multi-level network analysis for the social sciences: Theory, methods and applications* (pp. 213–244). Cham: Springer.

Bennink, M., Croon, M. A., & Vermunt, J. K. (2013). 'Micro–Macro multilevel analysis for discrete data: A latent variable approach and an application on personal network data. *Sociological Methods & Research, 42*, 431–457.

Blalock, H. M. (1984). Contextual effects models: Theoretical and methodological issues. *Annual Review of Sociology, 10*, 353–372.

Brailly, J., Favre, G., Chatellet, J., & Lazega, E. (2015). Embeddedness as a multilevel problem: A case study in economic sociology. In E. Lazega & T. A. B. Snijders (Eds.), *Multi-level network analysis for the social sciences: Theory, methods and applications* (pp. 245–272). Cham: Springer.

Brass, D. J., Galaskiewicz, J., Greve, H., & Tsai, W. (2004). Taking stock of networks and organizations: A multilevel perspective. *Academy of Management Journal, 47*, 795–817.

Brennecke, J., & Rank, O. N. (2015). Knowledge networks in high-tech clusters: A multilevel perspective on interpersonal and inter-organizational collaboration. In E. Lazega & T. A. B. Snijders (Eds.), *Multi-level network analysis for the social sciences: Theory, methods and applications* (pp. 273–294). Cham: Springer.

Browne, W. J., Goldstein, H., & Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling, 1*, 103–124.
Burstein, L. (1980). The analysis of multilevel data in educational research and evaluation. *Review of Research in Education, 8*, 158–233.
Campbell, R., Starkey, F., Holliday, J., Audrey, S., Bloor, M., Parry-Langdon, N., Hughes, R., & Moore, L. (2008). An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): A cluster randomised trial. *The Lancet, 371*, 1595–1602.
Carrington, P. J., Scott, J., & Wasserman, S. (2005). *Models and methods in social network analysis*. New York: Cambridge University Press.
Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics, 10*, 101–129.
Coleman, J. S. (1959). Relational analysis: The study of social organizations with survey methods. *Human Organization, 17*, 28–36.
Coleman, J. S. (1961). *The adolescent society*. New York: The Free Press of Glencoe.
Courgeau, D. (2003). From the macro-micro opposition to multilevel analysis in demography. In: D. Courgeau (Ed.), *Methodology and epistemology of multilevel analysis; Approaches from different social sciences* (pp. 43–91). Springer.
Cox, D. R. (1990). Role of models in statistical analysis. *Statistical Science, 5*, 169–174.
Croon, M. A., & van Veldhoven, M. J. P. M. (2007). Predicting group-level outcome variables from variables measured at the individual level: a latent variable multilevel model. *Psychological Methods, 12* 45–57.
Daraganova, G., & Robins, G. (2013). Autologistic attribute models. In D. Lusher, J. Koskinen, & G. Robins (Eds.), *Exponential random graph models* (pp. 102–114). Cambridge: Cambridge University Press.
Dekker, D., Krackhardt, D., & Snijders, T. A. B. (2007). Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. *Psychometrika, 72*, 563–581.
Doreian, P. (1980). Linear models with spatially distributed data: Spatial disturbances or spatial effects? *Sociological Methods & Research, 9*, 29–60.
Draper, D. (2008). Bayesian multilevel analysis and MCMC. In J. de Leeuw & E. Meijer (Eds.), *Handbook of multilevel analysis* (pp. 77–139). New York: Springer.
Entwisle, B., Faust, K., Rindfuss, R. R., & Kaneda, T. (2007). Networks and contexts: Variation in the structure of social ties. *American Journal of Sociology, 112*, 1495–1533.
Frank, O. (1991). Statistical analysis of change in networks. *Statistica Neerlandica, 45*, 283–293.
Friedkin, N. E. (1998). *A structural theory of social influence*. Cambridge: Cambridge University Press.
Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: Chapman & Hall/CRC.
Goldstein, H. (1986). Multilevel mixed linear model analysis using iterative generalized least squares. *Biometrika, 73*, 43–56.
Goldstein, H. (1989a). Efficient prediction models for adult height. In J. M. Tanner (Ed.), *Auxology 1988: Perspectives in the science of growth and development* (pp. 41–48). London: Smith-Gordon.
Goldstein, H. (1989b). Models for multilevel response variables with an application to growth curves. In R. Darrell Bock (Ed.), *Multilevel analysis of educational data* (pp. 107–125). San Diego, CA: Academic.
Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). London: Edward Arnold.
Greenland, S. (2000). Principles of multilevel modelling. *International Journal of Epidemiology, 29*, 158–167.
Handcock, M. S., Raftery, A. E., & Tantrum, J. M. (2007). Model-based clustering for social networks (with discussion). *Journal of the Royal Statistical Society, Series A, 170*, 301–354.
Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic.
Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association, 97*, 1090–1098.

Hollway, J., & Koskinen, J. H. (2015). Multilevel bilateralism and multilateralism; States' bilateral and multilateral fisheries treaties and their secretariats. In E. Lazega & T. A. B. Snijders (Eds.), *Multi-level network analysis for the social sciences: Theory, methods and applications* (pp. 315–332). Cham: Springer.

Huitsing, G., Snijders, T. A. B., Van Duijn, M. A. J., & Veenstra, R. (2014). Victims, bullies, and their defenders: A longitudinal study of the coevolution of positive and negative networks. *Development and Psychopathology, 26*, 645–659.

Kenny, D. A., Kashy, D.A., & Cook, W.L. (2006). *Dyadic data analysis*. New York: Guilford Press.

Kenny, D. A., & La Voie, L. (1985). Separating individual and group effects. *Journal of Personality and Social Psychology, 48*, 339–348.

Koskinen, J. H., & Edling, C. (2012). Modelling the evolution of a bipartite network – Peer referral in interlocking directorates. *Social Networks, 34*, 309–322.

Koskinen, J. H., & Snijders, T. A. B. (2016). Multilevel longitudinal analysis of social networks. In preparation.

Kreft, I. G. G., & de Leeuw, J. (1998). *Introducing multilevel modeling*. London: SAGE.

Krivitsky, P. N., Handcock, M. S., Raftery, A. E. & Hoff, P. D. (2009). Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models. *Social Networks, 31*, 204–213.

Lazarsfeld, P. F., & Menzel, H. (1961). On the relation between individual and collective properties. In A. Etzioni (Ed.), *Complex organizations: A sociological reader* (pp. 422–440). New York: Holt, Rinehart and Winston.

Lazarsfeld, P. F., & Menzel, H. (1993). On the relation between individual and collective properties. In R. Boudon (Ed.), *Paul Lazarsfeld: On social research and its language* (pp. 172–189). Chicago: The University of Chicago Press.

Lazega, E., Jourda, M.-T., Mounier, L. & Stofer, R. (2008). Catching up with big fish in the big pond? Multi-level network analysis through linked design. *Social Networks, 30*, 159–176.

Leenders, R. Th. A. J. (2002). Modeling social influence through network autocorrelation: Constructing the weight matrix. *Social Networks, 24*, 21–47.

Light, J. M., Greenan, C. C., Rusby, J. C., Nies, K. M., & Snijders, T. A. B. (2013). Onset to First alcohol use in early adolescence: A network diffusion model. *Journal of Research on Adolescence, 23*, 487–499.

Lubbers, M. J. (2003). Group composition and network structure in school classes: A multilevel application of the p^* model. *Social Networks, 25*, 309–332.

Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods, 13*, 203–229.

Lusher, D., Koskinen, J., & Robins, G. (2013). *Exponential random graph models*. Cambridge: Cambridge University Press.

Mason, W. M., Wong, G. Y., & Entwisle, B. (1983). Contextual analysis through the multilevel linear model. *Sociological Methodology, 14*, 72–103.

Mercken, L., Steglich, C. E. G., Sinclair, P., Holliday, J., & Moore, L. (2012). A longitudinal social network analysis of peer influence, peer selection, and smoking behavior among adolescents in British schools. *Health Psychology, 31*, 450–459.

Moody, J., Brynildsen, W. D., Osgood, D. W., Feinberg, M. E., & Gest, S. (2011). Popularity trajectories and substance use in early adolescence. *Social Networks, 33*, 101–112.

Morris, C. N. (1983). Parametric empirical bayes inference: Theory and applications. *Journal of the American Statistical Association, 78*, 47–55.

Nowicki, K., & Snijders, T. A. B. (2001). Estimation and prediction for stochastic blockstructures. *Journal of the American Statistical Association, 96*, 1077–1087.

Rasbash, J., & Browne, W. J. (2008). Non-hierarchical multilevel models. In J. de Leeuw & E. Meijer (Eds.), *Handbook of multilevel analysis* (pp. 301–334). New York: Springer.

Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2014). *A users' guide to MLwiN*. Centre for Multilevel Modelling, University of Bristol, Bristol.

Raudenbush, S. W., & Bryk, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education, 59*, 1–17.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (2nd ed.). Newbury Park, CA: SAGE.
Raudenbush, S. W., Bryk, A. S., Cheong Y. F., Congdon, R. T., & du Toit, M. (2011). *HLM 7: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.
Ripley, R. M., Snijders, T. A. B., Bóda, Z., Vörös, A., & Preciado, P. (2015). Manual for Siena version 4.0. Technical report, Oxford: University of Oxford, Department of Statistics; Nuffield College.
Rivellini, G., Terzera, L., & Amati, V. (2012). Individual, dyadic and network effects in friendship relationships among Italian and foreign schoolmates. *Genus, 67*, 1–27.
Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *Sociological Review, 15*, 351–357.
Schaefer, D. R., Simpkins, S. D., Vest, A. E., & Price, C. D. (2011). The contribution of extracurricular activities to adolescent friendships: New insights through social network analysis. *Developmental Psychology, 47*, 1141–1152.
Scheffé, H. (1959). *The analysis of variance*. New York: Wiley.
Snijders, T. A. B., & Kenny, D. A. (1999). The social relations model for family data: A multilevel approach. *Personal Relationships, 6*, 471–486.
Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. In M. E. Sobel & M. P. Becker (Eds.), *Sociological Methodology – 2001* (Vol. 31, pp. 361–395). Boston/London: Basil Blackwell.
Snijders, T. A. B. (2010). Conditional marginalization for exponential random graph models. *Journal of Mathematical Sociology, 34*, 239–252.
Snijders, T. A. B. (2011). Statistical models for social networks. *Annual Review of Sociology, 37*, 131–153.
Snijders, T. A. B., & Baerveldt, C. (2003). A multilevel network study of the effects of delinquent behavior on friendship evolution. *Journal of Mathematical Sociology, 27*, 123–151.
Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: SAGE.
Snijders, T. A. B., Lomi, A., & Torlò, V. (2013). A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Social Networks, 35*, 265–276.
Snijders, T. A. B., Pattison, P. E., Robins, G. L., & Handcock, M. S. (2006). New Specifications for Exponential Random Graph Models. *Sociological Methodology, 36*, 99–153.
Steglich, C. E. G., Sinclair, P., Holliday, J., & Moore, L. (2012). Actor-based analysis of peer influence in A Stop Smoking In Schools Trial (ASSIST). *Social Networks, 34*, 359–369.
Steglich, C. E. G., Snijders, T. A. B., & Pearson, M. A. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology, 40*, 329–393.
Sterba, S. K. (2009). Alternative model-based and design-based frameworks for inference from samples to populations: From polarization to integration. *Multivariate Behavioral Research, 44*, 711–740.
Sweet, T. M., Thomas, A. C., & Junker, B. W. (2013). Hierarchical network models for education research: Hierarchical latent space models. *Journal of Educational and Behavioral Statistics, 38*, 295–318.
Sweet, T. M., Thomas, A. C. & Junker, B. W. (2014). Hierarchical mixed membership stochastic blockmodels for multiple networks and experimental interventions. In E. M. Airoldi, D. Blei, E. A. Erosheva, & S. E. Fienberg (Eds.), *Handbook of Mixed Membership Models and Their Applications*. Boca Raton: Chapman & Hall/CRC Press.
Tranmer, M., & Lazega, E. (2015). Multilevel models for multilevel network dependencies. In E. Lazega & T. A. B. Snijders (Eds.), *Multi-level network analysis for the social sciences: Theory, methods and applications* (pp. 107–124). Cham: Springer.

Tranmer, M., Steel, D., & Browne, W. J. (2014). Multiple-membership multiple-classification models for social network and group dependences. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 177*, 439–455.

van Duijn, M. A. J., Snijders, T.A. B., & Zijlstra, B. H. (2004). p_2: A random effects model with covariates for directed graphs. *Statistica Neerlandica, 58*, 234–254.

van Duijn, M. A. J., van Busschbach, J. T., & Snijders, T. A. B. (1999). Multilevel analysis of personal networks as dependent variables. *Social Networks, 21*, 187–210.

Vermeij, L., Van Duijn, M. A. J., & Baerveldt, C. (2009). Ethnic segregation in context: Social discrimination among native Dutch pupils and their ethnic minority classmates. *Social Networks, 31*, 230–239.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1–48.

Wang, P., Robins, G., Pattison, P., & Lazega, E. (2013). Exponential random graph models for multilevel networks. *Social Networks, 35*, 96–115.

Wang, P., Robins, G. L., & Matous, P. (2015). Multilevel network analysis using ERGM and its extensions. In E. Lazega & T. A. B. Snijders (Eds.), *Multi-level network analysis for the social sciences: Theory, methods and applications* (pp. 125–144). Cham: Springer.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. New York/Cambridge: Cambridge University Press.

Wasserman, S., & Iacobucci, D. (1991). Statistical modelling of one-mode and two-mode networks: Simultaneous analysis of graphs and bipartite graphs. *British Journal of Mathematical and Statistical Psychology, 44*, 13–43.

Wasserman, S., & Pattison, P. (1996). Logit models and logistic regression for social networks: I. An introduction to Markov graphs and p^*. *Psychometrika, 61*, 401–425.

Wellman, B., & Frank, K. A. (2001). Network capital in a multi-level world: Getting support from personal communities. In N. Lin, K. Cook, & R. S. Burt (Eds.), *Social capital: Theory and research* (pp. 233–273). Chicago: Aldine De Gruyter.

Wilk, M. B., & Kempthorne, O. (1955). Fixed, mixed, and random models. *Journal of the American Statistical Association, 50*, 1144–1167.

Zappa, P., & Lomi, A. (2015). Knowledge sharing in organizations: A multilevel network analysis. In E. Lazega & T. A. B. Snijders (Eds.), *Multi-level network analysis for the social sciences: Theory, methods and applications* (pp. 333–354). Cham: Springer.

Zhu, M., Kuskova, V., Wasserman, S., & Contractor, N. (2015). Correspondence analysis of multirelational multilevel network affiliations: Analysis and examples. In E. Lazega & T. A. B. Snijders (Eds.), *Multi-level network analysis for the social sciences: Theory, methods and applications* (pp. 145–172). Cham: Springer.

Zijlstra, B. J. H., Van Duijn, M. A. J. & Snijders, T. A. B. (2006). The multilevel p_2 model. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 2*, 42–47.

Springer
http://www.springer.com/978-3-319-24518-8

Multilevel Network Analysis for the Social Sciences
Theory, Methods and Applications
Lazega, E.; Snijders, T.A.B. (Eds.)
2016, VIII, 375 p. 65 illus., 41 illus. in color., Hardcover
ISBN: 978-3-319-24518-8