SOFTWARE FOR STATISTICAL ANALYSIS OF SOCIAL NETWORKS

Mark Huisman* and Marijtje A. J. van Duijn†‡
University of Groningen

This paper gives a state-of-the-art overview of available software for the statistical analysis of social networks as of Summer 2004. It reviews and compares software packages for social network analysis with respect to their statistical procedures, illustrating some procedures with example data. The choice of routines that were inspected is restricted to procedures for statistical modeling based on probability distributions (e.g., exponential random graph models, QAP correlation, statistical analysis of longitudinal network data). This definition of analysis routines excludes the extensive review of procedure-based routines based on more complex (iterative) algorithms like cluster analysis or eigendecompositions. The paper concludes with some recommendations.

Key words: exponential random graph model, longitudinal network data, statistical modelling, software packages, permutation tests.

1 INTRODUCTION

This paper reviews software for the statistical analysis of social networks. Both commercial and freely available packages are considered. Based on the software page on the INSNA website (see the Appendix for the URL) sixteen software packages that contain procedures for the statistical analysis of social networks were selected. The programs are listed in Table 1.

The choice of software packages reviewed in this paper is based on the availability of statistical procedures in the packages. The programs were selected from the set of programs investigated by Huisman and Van Duijn (2004), who present a general review of software for social network analysis and discuss procedures for all main themes in network analysis: structural and locational properties, roles and positions, dyadic and triadic methods, and statistical models (Wasserman and Faust, 1994). They also include descriptive methods and visualization procedures in their review, although software merely aimed at visualization of

*Department of Psychology/DPMG, Grote Kruisstraat 2/1, 9712 TS, Groningen, the Netherlands, email: m.huisman@ppsw.rug.nl.

†Department of Sociology/Statistics and Measurement Theory, Grote Rozenstraat 31, 9712 TG, Groningen, the Netherlands, email: m.a.j.van.duijn@ppsw.rug.nl.

‡We thank Vladimir Batagelj, Ghi-Hoon Ghim, Andrej Mrvar, Bill Richards, and Andrew Seary for making their software available.

networks were not admitted to the review (see Freeman, 2004, for network visualization). In this paper we only treat dyadic and triadic methods and statistical modelling.

A distinction is made between stand-alone programs and software utility toolkits. These latter packages generally consists of procedure libraries to be used within a specific development environment (e.g., Java or R) or statistical program (e.g., Excel or Gauss). Twelve stand-alone programs were included in the review, and five utility toolkits. The age of the software was not a criterion for selection, although the release dates of the last versions of the majority of the reviewed software were within the last two or three years.

Table 1 describes the main objective or characteristic of each program, the version included in the review, and, for toolkits, the (development) environment. Furthermore, data format, functionality, and support are described. The data format distinguishes three aspects: 1) the type of data the program can handle, 2) the input format, and 3) whether there is an option to indicate missing value codes for network relations.

The description of the functionality includes network visualization options, and the kind of nonstatistical and statistical analyses the software can perform. Three different types of nonstatistical procedures are distinguished: network descriptives, structure and location, and role and position (Wasserman and Faust, 1994, Parts 3-6; see also Huisman and Van Duijn, 2004). The theoretical background of almost all of the obtainable output can be found in Wasserman and Faust (1994) as well. For the statistical procedures five different types of analysis were distinguished. These analysis types are based on different kinds of research topics that are interesting from a statistical point of view: statistical network descriptives, relations within and between networks, relations between networks and actor attributes, comparison and categorization of actors, and network dynamics. In Section 2 these topics and corresponding statistical procedures will be discussed in more detail, together with their availability in software packages.

The amount of support is the final characteristic mentioned in Table 1, distinguishing availability of the program (free or commercial, not listing prices), presence and availability of a manual and/or online help during execution.

The remaining part of Section 2 gives a review the programs UCINET, NetMiner II, MultiNet, and StOCNET with respect to the statistical procedures these programs contain. These programs are considered to be well-known, and either to be general and comprehensive (UCINET and NetMiner) or containing specifically designed statistical procedures (MultiNet and StOCNET). According to Table 1, these programs are most comprehensive with respect to statistical procedures, containing procedures in at least three of the five specified categories. The remaining software packages listed in Table 1 are briefly described in the Appendix. They are considered to be more specialized and their objectives and statistical properties are discussed to a limited extent. In the Appendix a table containing the programs, version numbers, and URL is presented.

A selection of statistical procedures contained in the various packages is illustrated in Section 3 by applying them to the data from Freeman's EIES network (Freeman and Freeman, 1979). The paper concludes with a section comparing the programs and gives some general recommendations. This section is by no means conclusive, if only because by definition a paper like this becomes outdated with publication.

Table 1. Overview Of Selected Programs For Statistical Analysis Of Social Networks, With Version Number, Objectives, Environment (For Toolkits), Data Format (Type, Input, Missing Values), Functionality (Visualization, Nonstatistical Methods, Statistical Methods), And Support (Availability, Help).

				Data			Functionality			Support	
Software	Ver.	Objective	Tool	Type[1]	Input[2]	Miss.	Vis.	Nonstat.[3]	Stat.[4]	Avail.[5]	Help[6]
FATCAT	4.2[7]	contextual analysis		c	ln	yes	yes	d	c	free[7]	o
GRADAP	2.0[7]	graph analysis		c	ln	yes	no	d, sl	d, nn	com[7]	m
JUNG	1.3.0	modeling graphs	JAVA	c	ln	–	yes	d, sl, rp	d	free	m
KliqFinder	0.05	cohesive subgroups		c	m, ln	no	no[10]		c	free	m
MatMan	1.1	structural analysis	Excel	c, a	m	no	no	d, sl	d, nn	com	m, o
MultiNet	4.38	contextual analysis		c, l	ln	yes	yes	d, rp	d, na, c	free	m, o
NetMiner II	2.4.0	visual analysis		c, e, a	m, ln	no	yes	d, sl, rp	d, nn, c	com[8]	m, o
Pajek	0.98	large data exploration		c, a, l	m, ln	yes[9]	yes	d, sl, rp	d	free	m
PermNet	0.94	permutation tests		c	m	yes	no		nn	free	o
PREPSTAR	1.0[7]	data preparation	SPSS	c	m	yes	–		na	free[7]	m
SNA	0.44	general	R	c	m	no	yes	d, sl, rp	d, nn, na	free	m
SNAP	2.5	general	Gauss	c	ln	no	no	d, sl, rp	d, nn	com	m
Snowball	–[7]	hidden populations		e	ln	–	no		d	free[7]	m
StOCNET	1.5	statistical analysis		c	m	yes	no	d	d, nn, na, c, dy	free	m, o
STRUCTURE	4.2[7]	structural analysis		c, a	m	yes[9]	no	sl, rp	na, c	free[7]	m
UCINET	6.29	comprehensive		c, e, a	m, ln	yes	yes	d, sl, rp	d, nn, c	com[8]	m, o

[1] a=affiliation, c=complete, e=ego-centered, l=large networks.
[2] ln=link/node, m=matrix
[3] d=descriptive, sl=structure and location, rp=roles and positions
[4] d=descriptive, nn=relation network-network, na=relation network-actor, c=comparing/categorizing actors, dy=dynamics
[5] com=commercial product, free=freeware/shareware.
[6] m=manual, o=online help.
[7] DOS-program which is no longer updated.
[8] An evaluation/demonstration version is available and/or the program is freely accessible on the internet.
[9] Only missing value codes for actor attributes.
[10] Generates output to create graphics in SAS.

2 STATISTICAL SOFTWARE - A CLOSER LOOK

2.1 Social Network Data

Social network data consist of two elements: ties forming the network, and actors (or nodes) that are connected by the ties. Assuming that the network consists of n actors, it can be represented by an $n \times n$ *adjacency matrix* X of which the elements X_{ij} are the tie variables from actor i to actor j. Often the ties are dichotomous, indicating the presence ($X_{ij} = 1$) or absence ($X_{ij} = 0$) of a tie between actor i and j. The diagonal elements X_{ii} are ignored, because self-relations are usually undefined. The remaining $n(n-1)$ tie variables are considered the complete network data[1].

Statistical analysis of social network data usually is not easy. The main reason for this is the dependence structure of the network. The ties from and to the same actor are not independent, and therefore statistical analysis cannot be based on broad independence assumptions as is usually done. Also, the ties are often dichotomous, making methods or models based on normal distributions inappropriate.

A *software network* can be created for the programs in Table 1 by defining a relationship between the individual packages. An example is given in Figure 1. In this software network the actors are the seventeen software packages for the statistical analysis of social networks. The ties are defined by the data input functionality of the programs. There are two possibilities to have a tie from one package to another: 1) the programs have the same data system files or there are (direct) import/export functions to exchange data from one program to another, and 2) the programs use the same data format and are able to read general types of data files (e.g., comma separate files, DL files, etc.).

Several characteristics of the programs are also included in the graph. The color of the nodes represents the data format (dark blue is matrix, red is link/node, light green is both), the size represents the number of different types of statistical analyses available (see Table 1), and the shape represents availability of network visualization properties (circle is yes, triangle is no). All characteristics can also be found in Table 1. The graph is created with Pajek by using the spring-embedding algorithm of Kamada-Kawai and including an attribute file with node characteristics in the drawing process.

In the graph the packages are grouped according to data input and format (color). The programs with the largest number of techniques are located in the center, as well as the programs with visualization properties. These are generally the programs that are continually upgraded and under development. The oldest programs (usually DOS-based and no longer updated) are located in the periphery, as are (generally) the toolkits and the highly specialized programs.

2.2 A Categorization of Statistical Procedures

In order to choose the software to analyze network data, the following question should be answered: How do we determine which type of statistical analysis is appropriate? To answer this question, statistical analyses have to be linked to research questions. Five different kinds of research questions emerge, which are interesting from a statistical point of view:

[1] We restrict this review to complete networks ignoring ego-centered network data.

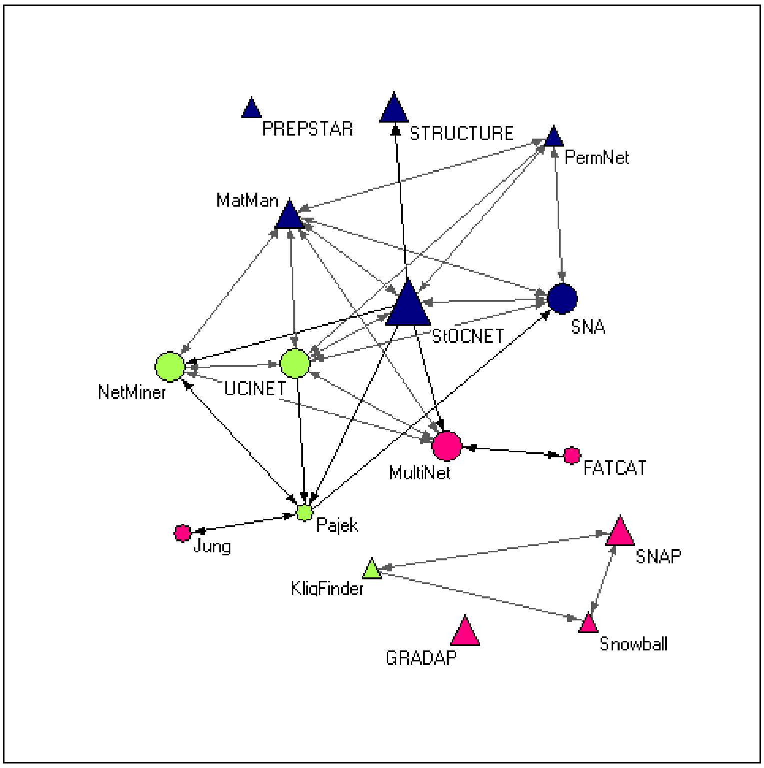

Figure 1. A Network Of Software For Statistical Analysis Of Social Networks (Visualization With Pajek).

1. How can properties of networks and actors be described and summarized?

2. How can the association between ties within one network and between networks be described and modeled?

3. How can the association between network ties and (endogenous or exogenous) actor characteristics be described and modeled?

4. How can actors and their (endogenous or exogenous) characteristics be compared or categorized?

5. How do networks develop over time and how do they influence each other?

Based on these types of research questions, statistical analysis methods were categorized into five different types of analysis (references of methods are given in the discussion of the programs and the analysis of the EIES data):

1. *Statistical network descriptives*, that is, (general) summary descriptives like density, degrees, degree variances, dyad and triad census, combined with distributional properties based on specific distributions (e.g., conditional uniform distributions, or the $\mathcal{U} \mid M, A, N$ or $\mathcal{U} \mid X_{i+}, X_{+j}$ distributions) or permutation tests. Although no statistical procedures, it should be noted that network visualization techniques play an important role in describing networks.

2. *Relations within and between networks*, that is, the prediction of ties within a network or the prediction of one network using other networks. Several types of procedures are available: correlation (QAP) and regression methods (QAP), and exponential random graph models (p_1, p_2, p^*, without actor attributes).

3. *Relations between networks and actor attributes*, which can be exogenous or endogenous. Available methods are crosstabulation of links and nodes (panigrams, see Section 2.5), regression methods, and exponential random graph models with actor attributes (p_2, p^*).

4. *Comparison and categorization of actors.* Available procedures are ANOVAs, some specific procedures like contagion analysis in STRUCTURE, and the statistical counterparts of the nonstatistical analysis of network positions and roles (e.g., cohesive subgroups based on logistic regression or stochastic blockmodeling).

5. *Network dynamics.* Simulation-based procedures are available that require longitudinal network data (e.g., the stochastic actor-oriented models).

Using this categorization of procedures, the availability of routines was investigated in the set of software packages reviewed by Huisman and Van Duijn (2004). The number of different types of analysis that are available for each program can be found in Table 1 (and is visualized in Figure 1). In the remainder of Section 2 the programs (i.e., not utility toolkits) that offer the broadest range of statistical procedures (i.e., procedures in three or more categories) are discussed in more detail. For a more general discussion of these programs (and others) see Huisman and Van Duijn (2004). The order in which the packages are presented is based on generality as well as on age: UCINET, NetMiner, MultiNet, and StOCNET.

2.3 UCINET

UCINET 6.0 (Version 6.29; Borgatti et al., 2003) is a comprehensive program for the analysis of social networks and other proximity data. It is probably the best known and most frequently used software package for the analysis of social network data and contains a large number of network analytic routines. The program does not contain procedures to visualize networks, but it is distributed together with the programs Mage, NetDraw, and Pajek, to be used for network visualization. It contains a large number of network descriptives (e.g., centrality, cohesive subgroups, regions, structural holes) and procedure-based methods (e.g., cluster analysis, two-mode scaling, structural equivalence), and has a built-in spreadsheet editor.

Statistical analysis. Various statistical routines are available in UCINET in the categories *statistical descriptives* and *relations within and between networks.* In the latter category there are autocorrelation methods, QAP correlation, and QAP regression procedures (Krackhardt, 1987). It should be noted, however, that QAP tests for individual regression parameter estimates are biased due to multicollinearity and thus the results of QAP regression procedures should be interpreted with care (see Dekker et al., 2003). Also the p_1 model (Holland and Leinhardt, 1981) can be fitted to the data with UCINET and univariate vector methods are available, which are combined with permutation tests.

UCINET also supports procedures in the category *comparison and categorization of actors*. An example of the latter group of methods is ANOVA with attribute vectors and/or rows or columns of the adjacency matrix, representing a sending or receiving actor, as variables. This is different from procedures where all incoming and outgoing links in an adjacency matrix are used as input for an ANOVA (e.g., MultiNet).

2.4 NetMiner II

NetMiner II (Version 2.4.0; Cyram, 2004) is a software tool that combines social network analysis and visual exploration techniques. It allows users to explore network data visually and interactively, and helps to detect underlying patterns and structures of the network. The program is especially designed for the integration of exploratory network analysis and visualization. In order to facilitate this integration the main window of the program contains a map frame in which the results of the analysis are graphically presented and (graphically) inspected.

NetMiner integrates analysis and visualization and therefore has advanced graphical properties. Moreover, almost all results are presented both textually and graphically, contrary to both other programs, except results of statistical analyses. In NetMiner graphical and textual results are directly obtained via the *Explore* function of the main menu. The *Visualize* function produces graphical visualizations of networks with various options. Like UCINET the program contains ample descriptive methods and methods for procedure-based analysis.

Statistical analysis. NetMiner supports a number of standard statistical routines in three of the five defined categories of routines: *descriptive statistics*, *relations within and between networks*, and *comparison and categorization of actors*. In the first category dyadic and triadic methods are available. The second category contains the most procedures: QAP correlations and regression (Krackhardt, 1987, Dekker et al., 2003), Markov chain Monte Carlo simulation tests for several network measures based on the on the $\mathcal{U} \mid X_{i+}, X_{+j}$ and $\mathcal{U} \mid X_{i+}, X_{+j}, M$ distributions (cf. the module ZO in StOCNET), and the p_1 exponential random graph model (Holland and Leinhardt, 1981). For categorization of actors ANOVA's can be performed for which the statistics are given with conventional significance tests (based on probably unwarranted assumptions of independence and normality) and random permutation tests. The ANOVAs, correlations and regressions can also be computed for attribute vectors.

2.5 MultiNet

MultiNet (Version 4.38 for Windows; Richards and Seary, 2003) is a program suitable for the analysis of large data sets and sparse network data. Some of the network analysis methods and procedures in MultiNet were originally contained in separate programs. FATCAT (Version 4.2, Richards, 1993), for instance, performs the same type of categorical social network analysis as MultiNet. Although incorporated in MultiNet, FATCAT is still freely available as a stand alone DOS program that runs under Windows. Another program integrated in MultiNet is PSPAR (Seary, 1999), which estimates the p^* model for sparse matrices (Wasserman and Pattison, 1996, Seary and Richards, 2000).

Like NetMiner, MultiNet contains procedures to give graphical representations of almost all output generated by the analysis routines. It provides not only graphical representations of networks (with eigendecompositions), but also histograms (for distributions), line diagrams (for ANOVAs), and so-called panigrams (for crosstables). Networks are visualized using eigendecompositions, and also adjacency matrices can be presented graphically. It has a few descriptive methods and eigenprocedures, which are used to visualize the networks.

Statistical analysis. MultiNet contains statistical routines to analyze attribute data (e.g., crosstables, ANOVA, correlations). For the analysis of networks, MultiNet provides statistical techniques in the categories *statistical descriptives*, *relations within and between networks*, and *comparison and categorization of actors*. There are procedures for crosstables and χ^2-tests, ANOVA, and correlations. In these procedures, however, all ties within a network are used, thereby assuming independence between all relations (contrary to the procedures in UCINET and NetMiner which assume independence between actors). This assumption will generally not be valid, and the user should therefore be very cautious interpreting the results. Also the p^* exponential random graph model (Wasserman and Pattison, 1996, Seary and Richards, 2000) can be fitted to the network data, using standard logistic regression based on independence assumptions that lead to biased estimates (Snijders, 2002).

2.6 StOCNET

StOCNET (Version 1.5; Boer et al., 2004) is an open software system, in a Windows environment, for advanced statistical analysis of social networks. It provides a platform to make available a number of statistical methods, presented in separate modules, and allows new routines to be easily implemented (Huisman and Van Duijn, 2003). Analyses take place within sessions. A session consists of (a cyclical process of) five steps: 1) data definition, 2) transformation, 3) selection, 4) model specification and analysis, and 5) inspection of results. StOCNET contains a few descriptive methods, but has neither methods for network visualization, nor procedure-based routines.

Statistical analysis. StOCNET contains statistical procedures in all five analysis categories. In the category *statistical descriptives* degree variances, and dyadic and triadic methods are available. Also the module ZO can be used to determine probability distributions of statistics of random graphs based on the $\mathcal{U} \mid X_{i+}, X_{+j}$ and $\mathcal{U} \mid X_{i+}, X_{+j}, M$ distributions (Snijders, 1991; Molloy and Reed, 1995). For the analysis of relations of ties within and between networks, and between networks and actor attributes, StOCNET contains procedures to fit several exponential random graph models: p_1 (Holland and Leinhardt, 1981), p_2 (Van Duijn et al., 2004; the module P2), and p^* (Wasserman and Pattison, 1996; Snijders et al., 2004; the module SIENA).

In the category *comparison and categorization of actors* three modules are available in StOCNET: BLOCKS for stochastical blockmodeling (Nowicki and Snijders, 2001), PACNET for constructing partial algebras using statistical criteria based on conditional uniform random graph distributions (Pattison and Wasserman, 1995, Pattison et al., 2000), and ULTRAS, for estimating latent transitive structures using ultrametrics (Schweinberger and Snijders,

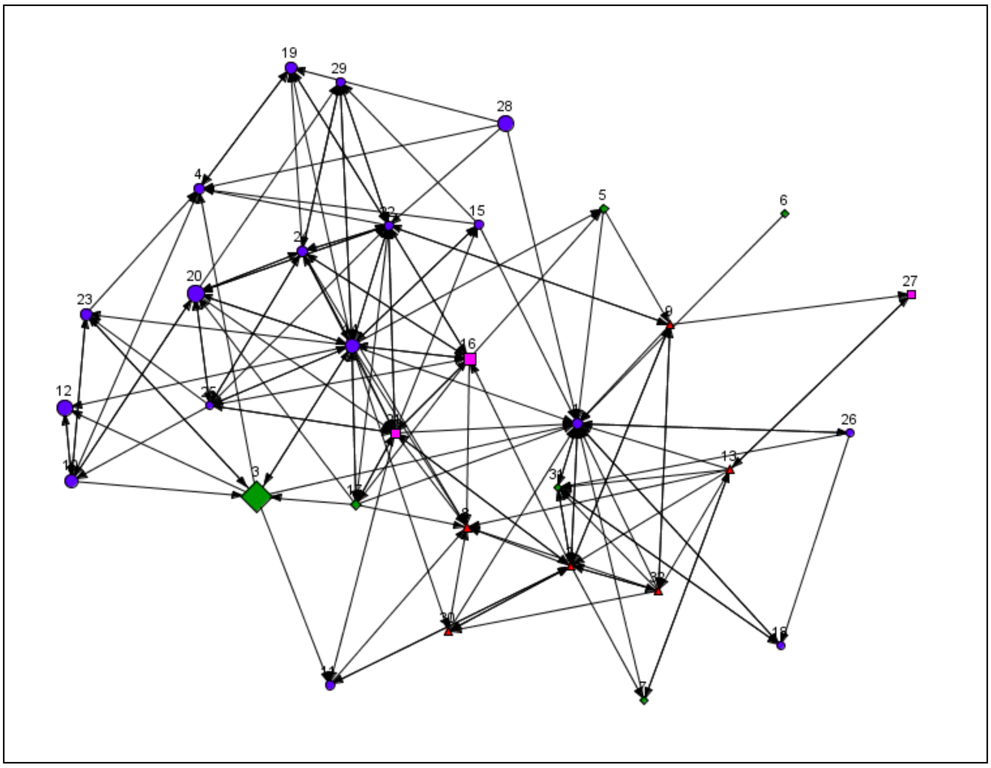

Figure 2. The EIES Acquaintanceship Network, First Time Point (Visualization With NetMiner).

2003). Longitudinal network data can be analyzed with the module SIENA to investigate the development of networks over time (Snijders, 2001, 2004).

3 ANALYZING FREEMAN'S EIES NETWORK

In this section a selection of procedures mentioned above is illustrated with the data from Freeman's EIES network (Freeman and Freeman, 1979). The data consist of three one-mode networks with two relations on a set of actors ($n = 32$). The data were collected as part of a study of the impact of the Electronic Information Exchange System (EIES). Two relations were recorded: the number of messages sent and acquaintanceship. The acquaintanceship relation is longitudinal, measured at two time points, ranging from 0 (did not know the other) to 4 (close personal friend). Because for most analyses the data need to be binary, the ties in the acquaintanceship network are dichotomized: 1 for values larger than 2 (friend, close friend), 0 for other values (not knowing, not having met, having met). The data set contains two actor attribute variables: primary disciplinary affiliation (sociology, anthropology, statistics and mathematics, psychology), and the number of citations (social science citation index).

The graph of the first acquaintanceship network is presented in Figure 2, obtained with NetMiner. The actors' discipline is depicted by color and shape (resp. blue diamond, red circle, magenta circle, and green box), citations by the size of the nodes. The complete data

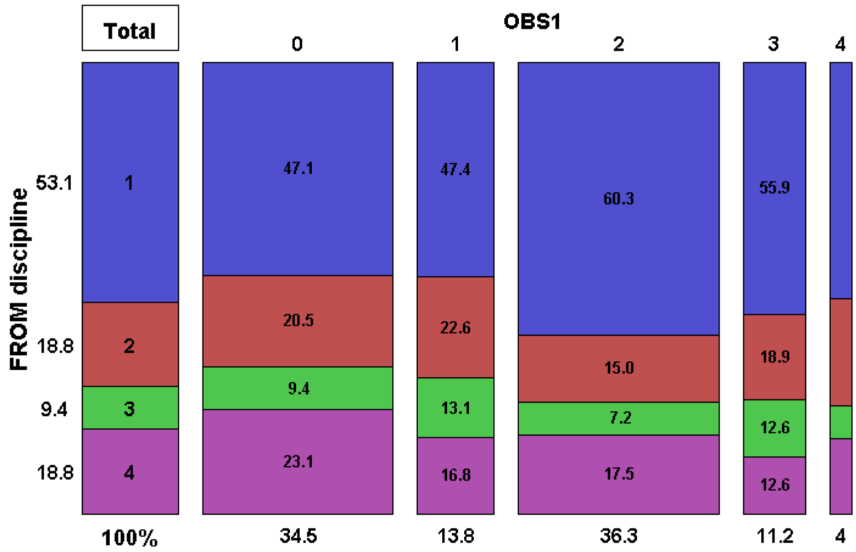

Figure 3. Panigram Of Discipline And The Incoming Links Of The EIES Acquaintanceship Data (First Observation).

set can be found in Wasserman and Faust (1994, p. 745–748) and is one of the standard data sets distributed with UCINET and StOCNET.

Crosstables (NetMiner, MultiNet)
NetMiner offers the option to create crosstabulations of vectors (actor attributes) or matrices (networks). Significance tests can be based on the classic independence assumption or on permutation tests. MultiNet also offers both kinds of crosstables (for attributes or networks), but also the option to use an attribute as row variable and a network as column variable (or vice versa). The resulting crosstables are visualized using panigrams, which can be used to explore the association within networks (out- and in-degrees, i.e., sender and receiver effects) or the association between networks and an attribute.

In Figure 3 a panigram of discipline and incoming links (receiver effects) of the first observation of the acquaintanceship network is presented. The links can take the values 0 to 4 ('have not met' to 'close friends'), discipline the values 1 to 4. For example, 20.5% of the links with value 0 ('have not met') come from actors in discipline 2 (anthropogy). The χ^2 statistic equals 23.4 (df = 12, $P < 0.05$), which indicates a significant association between the variables (with sociologists receiving more friendship choices). The association between discipline and outgoing links (sender effects) of the first acquaintanceship network is also significant (results not reported here). These tests should be interpreted with care, because they are based on the assumption that all ties in the network are independent, which generally does not hold.

QAP correlation (UCINET, NetMiner)
Computation of QAP correlations between the three EIES matrices gives the correlations as presented in Table 2, with p-values indicating the percentage of random correlations that are

Table 2. QAP Correlations In The EIES Network Data (p-values In Parentheses, 2500 Permutations).

	Acquaintanceship			
	time 1		time 2	
Acquaintanceship time 2	0.809	(0.00)	—	—
Messages sent	0.240	(0.00)	0.347	(0.00)

as large as the observed correlation in 2500 permutations (see Krackhardt, 1987). Besides Spearman correlations, UCINET also calculates the simple matching coefficient, the Jaccard coefficient and Goodman-Kruskal's gamma.

The QAP-correlation found by NetMiner between the two time points of the acquaintanceship data is 0.818 ($p < 0.001$, 2500 simulations). The difference between the NetMiner and UCINET correlations is due to the fact that NetMiner treats the diagonal values as valid, while UCINET does not take them into account (the default option; this option can be changed in UCINET, not in NetMiner).

The p_1 model (UCINET, NetMiner, StOCNET)
The p_1 probability distribution of an adjacency matrix X is expressed in terms of probabilities of mutual, asymmetric, and null dyads in the network. These probabilities are modeled as functions of three sets of parameters: expansivess and popularity of each actor, and reciprocity. Fitting the p_1 model to the first observation of the dichotomized EIES acquaintanceship data with UCINET gives estimates of the network density and reciprocity (-3.45 and 4.39), and for each actor their expansiveness and popularity (not presented). Expected values and residuals to inspect the fit of the model are given as well. The reciprocity parameter shows that there is a tendency for mutual acquaintanceships between the researchers.

NetMiner and StOCNET can also be used to obtain parameter estimates, expected values, and fit statistics. These two programs give the same parameter estimates, which differ slightly from those obtained with UCINET. This difference is probably due to different implemented estimation algorithms. All three programs report the goodness-of-fit statistic G^2, but they report each different degrees of freedom (see Holland and Leinhardt, 1981, or Wasserman and Faust, 1994).

The p_2 model (StOCNET)
The p_2 model is a random effects model with the dyadic ties as the dependent variable (Van Duijn et al., 2004). The sender and receiver parameters, fixed in the p_1 model, are regressed on available (categorical or continuous) actor attributes. If no attributes are available, the regression model reduces to random sender and receiver effects. Likewise, the density and reciprocity parameters, can be linked to other available networks (*dyadic covariates*), without a random component. Dyadic covariates can also be computed from the actor attributes, for instance by taking their difference or absolute difference, which are both standard options in the P2 module. Thus, dissimilarity matrices are created. If the actor attribute is categorical dichotomous (dis)similarity matrices can be constructed, comparable to the block-parameters in the estimation of the p^* model in MultiNet (see below). Unlike the p^* model, the p_2 model

Table 3. p_2 **Estimates For The EIES Acquaintanceship Data (First Observation, Significant Effects) Obtained With StOCNET.**

		Model 1		Model 2	
Effect	Parameter	Est.	S.E.	Est.	S.E.
Density	μ	-2.45	0.22	-2.79	0.29
	Dissim. citation (abs. diff.)			-0.017	0.005
	Dissim. citation (diff.)			-0.013	0.003
	Similarity discipline			0.64	0.18
Reciprocity	ρ	2.71	0.32	2.36	0.32
Sender	Variance σ_A^2	1.04	0.24	1.01	0.24
	Citations			0.028	0.0082
Receiver	Variance σ_B^2	1.06	0.25	0.98	0.23
Sender-receiver	Covariance σ_{AB}	-0.57	0.20	-0.40	0.18

does not contain network effects other than reciprocity.

Table 3 contains parameter estimates for the fixed and random effects of two models. The first model does not include any actor attributes, the second does. In the second model, dissimilarity with respect to citation has a significant negative effect on density, in two ways: expressed as the absolute difference of the actors' number of citations, and expressed as the simple difference of the actors' number of citations. The first effect implies that the probability of an acquaintance relation decreases the more actors differ with respect to their citations; the second indicates a directional effect that actors whose citations are high tend to choose less often actors whose citations are low. The second effect can be viewed as a refinement of the positive sender effect for citation which indicates that the probability of an outgoing acquaintanceship relation (irrespective of the receiver attributes) increases with the number of citations. The positive effect of similarity with respect to discipline indicates that actors tend to choose more within their own discipline group. There is a general reciprocity effect, but this is not differentiated according to dyadic attributes.

The p^* model (MultiNet, StOCNET)

Contrary to the p_1 model, which assumes dyad independence, and the p_2 model, which models dependence between dyads through random actor effects, the p^* model takes network dependence into account by choosing suitable network statistics. The model can be formulated as a logit model and can be estimated by maximum pseudolikelihood using logistic regression estimation techniques, (Wasserman and Pattison, 1996). The package PREPSTAR (see the Appendix) can be used to preprocess network data into a format suitable for fitting the model in SPSS or SAS using logistic regression.

MultiNet offers a procedure to fit p^* models to large networks by pseudolikelihood based on sparse methods. The method fits the model parameters to triad statistics selected by the user. Blockparameters can be obtained by fitting models of which the blockstructure is defined by one or more (categorical) actor attributes. MultiNet was used to estimate a p^* model for the first observation of the acquaintanceship data. The model included the effects

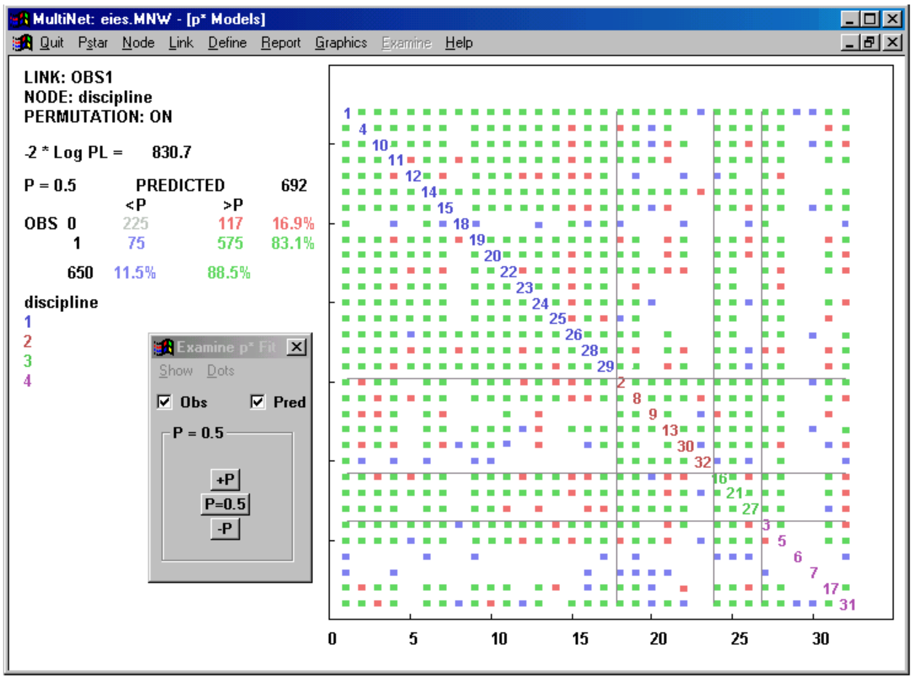

Figure 4. p^* **Graphic Display Window Of MultiNet Showing Results For The EIES Acquaintanceship Data (First Observation).**

density, reciprocity, transitivity, and the blockparameter 'choice within blocks' were included. The blocks were defined by the attribute discipline. All estimates were significant and are reported in Table 4, together with the estimates obtained with StOCNET.

Figure 4 shows the p^* graphic display window provided by MultiNet. It shows the adjacency matrix with correctly predicted links (green), the false negatives (blue), and false positives (red). It also gives a classification table (a standard option in logistic regression) to evaluate the model.

In the SIENA module of StOCNET, MCMC estimation with the Robbins-Monro algorithm of p^* model is implemented. As Snijders (2002) notes, both the pseudolikelihood estimation and MCMC estimation using the Geyer and Thompson (1992) method are unsatisfactory. The pseudolikelihood estimate is not a function of the complete statistic and has unknown properties. This leads in any case to underestimation of the standard errors of the estimates. MCMC estimation is not satisfactory either, because the simulation of random graph distributions turns out to be a complicated matter due to bimodality and poor mixing properties of the Metropolis-Hastings and Gibbs algorithms, which leads to convergence problems. See Wasserman and Robins (2004) and Snijders et al. (2004) for an extended discussion of estimation of p^* models. More developments in this area are to be expected.

In Table 4 the results are given of fitting the p^* model to the first observation of the EIES data. Maximum pseudolikelihood estimates were obtained with MultiNet, and MCMC estimates with the Robbins-Monro algorithm were obtained with the SIENA module in StOCNET. It was not possible to estimate the p^* model unconditionally. As soon as the transitivity effect was added to the model, no convergence was obtained. It was possible to obtain es-

Table 4. Pseudolikelihood Estimates (MultiNet) And Markov Chain Monte Carlo Robbins Monro p^* Estimates (StOCNET) For The EIES Acquaintanceship Data (First Observation).

	Pseudo-likelihood		MCMC Robbins Monro conditional on ties			
Effect	Est.	S.E.	Est.	S.E.	Est.	S.E.
Density	-3.61	0.22				
Reciprocity	1.94	0.23	2.15	0.31	2.20	0.30
Transitivity	0.32	0.036	0.17	0.01	0.17	0.01
Dissimilarity discipline	0.55	0.22	0.25	3.32		

timates of the p^* model conditional on the number of ties, which means that no density effect is estimated. The convergence of the model with the dissimilarity (or block) effect of discipline was unsatisfactory as well, which shows in the large standard error for this effect. The convergence of the conditional model (fixing the number of relations and thus limiting the outcome space) with only reciprocity and transitivity was acceptable. The estimates for reciprocity and even their standard errors are similar for pseudolikelihood and MCMC. The estimates for transitivity and the similarity (block) effect of discipline are quite different.

Network dynamics
Both time points of the dichotomized acquaintanceship networks are analyzed with the dynamic actor-oriented model of Snijders (2001, 2004), implemented in the StOCNET module SIENA. The estimated effects of the SIENA model are presented in Table 5 (see Snijders, 2004, for a discussion on the interpretation of the parameters). The rate parameter shows that on average the actors made about 2.5 relationship changes in the period between the observations. In the evolution of the acquaintanceship network, a clear reciprocity effect and a transitivity-type effect are present, the latter being specified as a tendency away from indirect relations. There is also a tendency for popular others (i.e., others who receive many choices). No significant attribute effects were found[2].

4 RECOMMENDATIONS

Examination of the software network in Figure 1 shows that three layers of packages can be distinguished. The first layer contains the programs located in the periphery, which have few ties to and from other programs. These programs contain few statistical procedures and are either outdated (DOS-based) and no longer updated (PREPSTAR, STRUCTURE, FATCAT, GRADAP, Snowball, SNAP), or rather new and therefore still under development and/or highly specialized (KliqFinder, JUNG). It should be noted that some DOS-based programs (e.g., STRUCTURE, Snowball) contain unique procedures that are not available in other packages.

[2] Snijders and Van Duijn (1997) analyzed another dichotomization of the EIES data: not knowing/having met vs. having met/being friends. They found different effects (especially effects of the attribute citation) influencing the evolution of this 'meeting' network.

Table 5. Estimated (Significant) Effects For The Evolution Of The EIES Acquaintanceship Data Obtained with StOCNET.

Effect	Est.	S.E.
Constant change rate	2.47	
Density (out-degree)	-1.80	0.52
Reciprocity	2.06	0.39
Indirect relations	-0.27	0.13
Popularity	6.40	1.05

The second layer consists of four packages that have many ties with other programs, but either are procedure libraries, which require a programming environment (SNA for R, Matman for Excel), or contain relatively few statistical procedures (Pajek, PermNet). Figure 1 shows that Pajek contains the least statistical procedures, but it has many direct ties from other programs (i.e., the other programs have export functions to Pajek) due to its advanced visualization options.

The third layer contains the (stand-alone) programs that contain the largest amount of statistical procedures and the largest variation in types of analysis procedures. These are the programs presented in Section 2, of which the procedures were used to analyze the EIES data in Section 3: UCINET, NetMiner, MultiNet, and StOCNET. The former two programs are comprehensive and contain many non-statistical procedures as well, the latter two are more specialized, of which StOCNET is especially designed to perform statistical analyses.

We conclude this paper with a summary of the four packages mentioned above and presented in Section 2. We scored the software at the three categories defined in Table 1: data manipulation (data entry was found not to be a problem for any program), functionality (network visualization, non-statistical, and statistical methods), and support (availability of a manual and a online help). Furthermore, we scores the user-friendliness of the programs. The scores are given in Table 6. A + is used to indicate that it is good (or at least sufficient), ++ that it is very good or strong, a − that it has shortcomings, a 0 that it is lacking, and a +− that it is undecided (having both good and bad parts). We will explain the scores, especially the negative ones, further below.

Obviously, we try to present an objective, substantiated view, but we admit that we cannot give a completely unbiased opinion. We also stress that it is impossible to make a fair comparison between the packages, because their objectives are different, which leads to different functionalities. For instance, the aim of StOCNET is not to compete with but to be an addition to existing software, and therefore it contains few non-statistical and no visualization procedures.

In two programs, MultiNet and StOCNET, data manipulation obtained the score +− because they contain relatively few options. The visualization aspect of UCINET is meager, but this

Table 6. Scores for the packages presented in Section 2.

	Data Manipulation	Functionality			Support		User-friendliness
		Visual.	Non-stat.	Stat.	Manual	Help	
MultiNet	+−	++	+	+−	+−	++	+
NetMiner	++	++	++	+−	+	+	++
StOCNET	+−	0	0	++	+	+	+
UCINET	++	+[1]	++	+−	+	+	+

[1] The program NetDraw for network visualization is distributed with UCINET

is compensated by export possibilities to specialized network visualization software and the option to call NetDraw within UCINET. StOCNET does not have any visualization options, but this is compensated via export possibilities to NetMiner and Pajek, which score very well with respect to visualization.

The scores for the non-statistical and statistical methods are indicative of the number of different features. Apart from a few descriptive methods supporting the statistical analysis, StOCNET lacks non-statistical procedures. Descriptive methods are rather sparse in MultiNet, but it contains unique methods based on eigendecompositions. NetMiner and UCINET contain a large number of descriptive and procedure-based methods.

As StOCNET is designed for statistical network analyses is contains the largest amount of (unique) statistical procedures. The other three programs do contain a number of – sometimes exclusive – statistical methods, but they are presented uncritically whereas some warning would definitely be warranted for the ANOVA procedures, estimation of the p^* model, and QAP regression.

All four programs offer good support and are regularly updated and new versions are released. Although MultiNet's manual is, at the time of writing, incomplete, the program has good, interactive, online help. The manuals and help functions, however, are not always updated at the same rate, making it difficult to use new features.

With respect to user-friendliness, NetMiner stands out, because of its interface where visualization, data, and procedures are integrated. It remains, however, hard to compare the different packages, as we already pointed out at the beginning of this section. We leave it to the reader of this chapter to decide which software to use for the statistical social network analysis s/he wishes to do.

APPENDIX

Table 7 contains the URLs of all programs presented in Table 1 and visualized in Figure 1. Also the URL of the INSNA website (International Network for Social Network Analysis) is added.

GRADAP (Version 2.0; Sprenger & Stokman, 1989)
The software package GRADAP (GRAph Definition and Analysis Package), an environment for analyzing graphs and networks, is an organized set of programs explicitly developed to analyze network data represented as graphs. It includes statistical procedures to calculate

Table 7. URLs of all reviewed programs and software toolkits.

Program	Ver.	URL
FATCAT	4.2	http://www.sfu.ca/~richards/Pages/fatcat.htm
GRADAP	2.0	http://www.assess.com/Software/GRADAP.htm
JUNG	1.0	http://jung.sourceforge.net/index.html
KliqFinder	0.05	http://www.msu.edu/~kenfrank/software.htm
MatMan	1.1	http://www.noldus.com/products/index.html?matman/index
MultiNet	4.38	http://www.sfu.ca/~richards/Multinet/Pages/multinet.htm
NetMiner II	2.4.0	http://www.netminer.com/NetMiner/
Pajek	0.98	http://vlado.fmf.uni-lj.si/pub/networks/pajek/default.htm
PermNet	0.94	http://www.meijigakuin.ac.jp/~rtsuji/en/software.html
PREPSTAR	1.0	http://kentucky.psych.uiuc.edu/pstar/index.html
SNA	0.44	http://erzuli.ss.uci.edu/R.stuff/
statnet		http://www.stat.washington.edu/ handcock/567/statnet.html
SNAP	2.5	http://www.soc.ucsb.edu/faculty/friedkin/Software/Software.htm
Snowball		http://stat.gamma.rug.nl/snijders/socnet.htm
StOCNET	1.5	http://stat.gamma.rug.nl/stocnet/
STRUCTURE	4.2	http://gsbwww.uchicago.edu/fac/ronald.burt/teaching/STRUC.EXE
UCINET	6.29	http://www.analytictech.com/ucinet.htm
INSNA		http://www.sfu.ca/~insna/

degree variances, correlations, and models for the distribution of in- and outdegrees. It is only available as a DOS application and will not be updated to a Windows environment.

JUNG (Version 1.4; White et al., 2004)
The Java Universal Network/Graph (JUNG) framework is a software library with JAVA routines that provides a common and extendible language for the modeling, analysis, and visualization of data that can be represented as a graph or network. JUNG supports a variety of representations of graphs (e.g., directed, undirected) and the current version includes some algorithms for statistical analyses (degree distributions and statistical moments). It also provides a visualization framework to construct tools for data exploration.

KliqFinder (Version 0.05; Frank, 2003)
KliqFinder is the Windows version of the Fortran and SAS-based program KliqueFinder. It is aimed at identifying cohesive subgroups and produces a so-called crystallized subgroup representing the subgroups and their relations within and between the clusters. The subgroups are identified in an iterative algorithm maximizing the log odds of a tie within the group (Frank, 1995, 1996). For the graphical representation of the subgroups, the program SAS is called from within KliqFinder,

MatMan (Version 1.0; Noldus, 2001)
An add-in for Microsoft Excel, MatMan is aimed at performing specific matrix manipulations, common in ethological research, for sociomatrices, behavioral profile data, and transition matrices. Furthermore, social dominance and correlation analyses can be performed.

Pajek (Version 0.98; Batagelj & Mrvar, 2004)

Pajek is a network analysis and visualization program, specifically designed to handle large data sets. The main goals in the design of Pajek are 1) to facilitate the reduction of a large network into several smaller networks that can be treated further using more sophisticated methods, 2) to provide the user with powerful visualization tools, and 3) to implement a selection of efficient network algorithms (Batagelj & Mrvar, 1998). It can handle multiple networks simultaneously, as well as 2-mode networks, and time event networks. The structure of the program is based on several data structures (networks, partitions, permutations, clusters, hierarchies, and vectors) and on transitions among these structures. The graphical properties are advanced, supporting 2D and 3D visualizations of networks, which can be saved in several graphical formats. The graph of the software network in Figure 1 is drawn with Pajek.

The program contains very few basic statistical procedures. Attributes of nodes (including structural properties that can be expressed as attributes), which are available as partitions and vectors, can be included in statistical analyses: computation of correlations, linear regression, and cross-tabulation (including some measures of association). For networks some dyadic and triadic methods (dyad and triad census) are available However, the statistical package R can be called with Pajek data structures (networks and vectors) and the statistical procedures available in R can be used, for instance those of the SNA library (see below).

PermNet (Version 0.94; Tsuji, 1997)
The program PermNet (PERMutation NETworks) contains a set of permutation tests for social network data. It provides symmetry tests, transitivity tests for real-valued data, and a triad census test for binary data (cf. NetMiner and the module ZO of the StOCNET software).

PREPSTAR (Version 1.0; Crouch & Wasserman, 1998)
Software necessary to preprocess network data into a format suitable to perform p^* analyses in SPSS or SAS. The procedures for logistic regression in the latter two programs are used to obtain parameter estimates.

R routines (Butts, 2004, Handcock et al., 2003)
There are several collection of routines to be used in R: SNA (Version 0.44, 'Carter's archive', Butts, 2004), and statnet and graphtesting (Handcock et al., 2003). SNA contains many well-documented procedures for performing various kinds of social network analyses ranging from general analyses such as mutuality, betweenness or centrality to specific analyses such as QAP and p^* analyses, or blockmodeling. It also contains visualization routines.

Statnet and graphtesting (Handcock et al., 2003) are two other libraries for visualization and statistical analysis of network data in R. The former package performs maximum likelihood estimation of exponential random graph models using MCMC (see also Snijders et al., 2004), and also allows graphs to be plotted. The packages are available at the website 'Statistical Software for Social Networks' (the URL is given in Table 5). The R routines can be called from the program Pajek.

SNAP (Version 2.5; Friedkin, 2001)
SNAP is a collection of GAUSS routines for network analysis. It includes procedures for

calculating many graph theoretical properties of graphs and nodes, the triad census, QAP correlations and regressions, and fitting social influence models and the expected value model of social power.

SNOWBALL (Snijders, 1994)
SNOWBALL is a DOS program for the estimation of the size of a hidden population from a one-wave snowball sample, implementing the estimates proposed by Frank & Snijders (1994). Snowball sampling is a term used for sampling procedures that allow the sampled units to provide information not only about themselves but also about other units. This is advantageous when rare properties are of interest.

STRUCTURE (Version 4.2; Burt, 1991)
STRUCTURE supports network models within five types of network analysis. These are autonomy (analysis of structural holes), cohesion (detection of cliques), contagion, equivalence (analysis of structural or role equivalence and blockmodeling), and power (analysis of network prominence and equilibrium). It is a command-driven DOS program that needs an input file containing commands for data management and network analysis. After opening the input file, the program executes the required routines without the possibility of user interaction.

STRUCTURE contains two routines for statistical modeling of the network data. The first one is contagion analysis (*comparison and categorization of actors*). It is based on the principle that an attribute (behavior) of actors that is affected by contagion results in network correlation (see Scott, 1991). This is modeled via simple linear regression with the attribute value of one actor as the dependent variable and the weighted average of the values of the same attribute of the other actors as the independent variable, where the weights reflect the structure of the network.

The second statistical procedure is analysis of network equilibrium (*relations between networks and attributes*). Network equilibrium is analyzed by predicting how relations in a network will change if powerful actors could initiate any relation they want. This prediction is based on a linear regression model that predicts the value of equilibrium relations from observed relations. Power distributions are obtained with eigenprocedures (see Scott, 1991)

REFERENCES

Batagelj, V. & Mrvar, A. (1998). Pajek: A program for large network analysis. *Connections*, *21*, 47–57.

Batagelj, V. & Mrvar, A. (2004). *Pajek: Package for large networks. Version 0.98.* Ljubljana: University of Ljubljana.

Boer, P., De Negro, R., Huisman, M., Snijders, T.A.B., Steglich, C.E.G., & Zeggelink, E.P.H. (2004). *StOCNET: An open software system for the advanced statistical analysis of social networks. Version 1.5.* Groningen: ICS/Science Plus Group, University of Groningen.

Borgatti, S.P., Everett, M.G., & Freeman, L.C. (2002). *UCINET 6 for Windows: Software for social network analysis.* Harvard: Analytic Technologies.

Burt, R.S. (1991). *STRUCTURE. Version 4.2.* New York: Columbia University.

Butts, C.T. (2004). *Package SNA: Tools for social network analysis. Version 0.44.* Irvine: University of California.

Crouch, B. & Wasserman, S. (1998). A practical guide to fitting p^* social network models. *Connections*, *21*, 87–101.

Cyram (2004). *Cyram NetMiner II. Version 2.4.0.* Seoul: Cyram Co., Ltd.

Dekker, D., Krackhardt, D., & Snijders, T.A.B. (2003). Multicollinearity robust QAP for multiple-regression. CASOS paper presented at the NAACSOS Conference June 22-25, 2003, Omni William Penn, Pittsburgh, PA.

Frank, K.A. (1995). Identifying cohesive subgroups. *Social Networks*, *17*, 27–56.

Frank, K.A. (1996). Mapping interactions within and between cohesive subgroups. *Social Networks*, *18*, 93–119.

Frank, K.A. (2003). *KliqFinder for Windows. Version 0.05.* East Lansing, MI: Michigan State University.

Frank, O. & Snijders., T.A.B. (1994). Estimating the size of hidden populations using snowball sampling. *Journal of Official Statistics*, *10*, 53–67.

Freeman, L.C. (2004). Graphical techniques for exploring social network data. In: P.J. Carrington, J. Scott, & S. Wasserman (eds.), *Models and methods in social network analysis.* Cambridge: Cambridge University Press.

Freeman, S.C., & Freeman L.C. (1979). The networks network: A study of the impact of a new communications medium on sociometric structure. Social Science Research Reports No. 46. Irvine, CA: University of California.

Friedkin, N.E. (2001). *SNAP: Social network analysis procedures for GAUSS. Version 2.5.* Maple Valley, WA: Aptech Systems, Inc.

Geyer, C.J. & Thompson, E.A. (1992). Constrained Monte Carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society B*, *54*, 657–699.

Handcock, M.S., Hunter, D.R., Butts, C.T., & Morris, M. (2003). *The 'statnet' and 'graphtesting' packages.* Washington: University of Washington.

Holland, P.W., & Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*, *76*, 33–50.

Huisman, M., & Van Duijn, M.A.J. (2003). StOCNET: Software for the statistical analysis of social networks. *Connections*, *25*, 7–26.

Huisman, M., & Van Duijn, M.A.J. (2004). Software for social network analysis. In: P.J. Carrington, J. Scott, & S. Wasserman (eds.), *Models and methods in social network analysis.* Cambridge: Cambridge University Press.

Hyatt, A., Contractor, N., Ferrone, T., Han, K.K., Hsu, F., Kochhar, S., Palazzo, E., Chunke, S., & Willard, B. (2003). *Blanche. Version 4.6.5.* Urbana-Champaign: University of Illinois.

Krackhardt, D. (1987). QAP partialling as a test of spuriousness. *Social Networks*, *9*, 171–186.

Molloy, M. & Reed, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, *6*, 161–179.

Noldus Information Technology (2001). *MatMan. Version 1.0 for Windows.* Wageningen: Noldus Information Technology bv.

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

Pattison, P. & Wasserman, S. (1995). Constructing algebraic models for local social networks using statistical methods. *Journal of Mathematical Psychology*, *39*, 57–72.

Pattison, P., Wasserman, S., Robbins, G., & Kanfer, A. (2000). Statistical evaluation of algebraic constraints for social networks. *Journal of Mathematical Psychology*, *44*, 536–568.

Richards, W.D. (1993). *FATCAT. Version 4.2.* Burnaby: Simon Fraser University.

Richards, W.D., & Seary, A.J. (2003). *MultiNet. Version 4.38 for Windows.* Burnaby: Simon Fraser University.

Schweinberger, M. & Snijders, T.A.B. (2003). Settings in social networks: a measurement model. In: R.M. Stolzenberg (ed.), *Sociological Methodology.* London: Basil Blackwell.

Scott, T. (1991). Network models. In: Burt, R.S. (ed.) *STRUCTURE. Version 4.2 reference manual.* New York: Columbia University.

Seary, A.J. (1999). *PSPAR: Sparse matrix version of PSTAR.* Burnaby: Simon Fraser University.

Seary, A.J. & Richards, W.D. (2000). Fitting to p^* models in MultiNet. *Connections*, *23*, 84–101.

Snijders, T.A.B. (1991). Enumeration and simulation models for 0-1 matrices with given marginals. *Psychometrika*, *56*, 397–417.

Snijders, T.A.B. (1994). *SNOWBALL.* Groningen: University of Groningen.

Snijders, T.A.B. (2001). The statistical evaluation of social network dynamics. In: M.E. Sobel, & M.P. Becker (eds.), *Sociological Methodology.* London: Basil Blackwell.

Snijders, T.A.B. (2002). Markov chain Monte Carlo estimation of exponential random graph models. *Journal of Social Structure*, *3*. `http://zeeb.library.cmu.edu:7850/JoSS/snijders/Mcpstar.pdf`

Snijders, T.A.B. (2004). Models for longitudinal network data. In: P.J. Carrington, J. Scott, & S. Wasserman (eds.), *Models and methods in social network analysis.* Cambridge: Cambridge University Press.

Snijders, T.A.B., Pattison, P.E., Robins, G.L., & Handcock, M.S. (2004). New specifications for exponential random graph models. Submitted for publication. `http://stat.gamma.rug.nl/sprh_d.pdf`

Snijders, T.A.B. & Van Duijn, M.A.J. (1997). Simulation for statistical inference in dynamic network models. In: R. Conte, R. Hegselmann, & P. Terna (eds.), *Simulating social phenomena.* Berlin: Springer.

Sprenger, C.J.A. & Stokman, F.N. (1989). *GRADAP: Graph definition and analysis package. Version 2.0.* Groningen: iec. *Pro*GAMMA.

Tsuji, R. (1997). *PermNet. Version 0.94.* Sapporo: Hokkaido University.

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

Wasserman, S. & Faust, K. (1994). *Social Network Analysis. Methods and applications.* Cambridge: Cambridge University Press.

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.

Wasserman, S. & Robins, G. (2004). An introduction to random graphs, dependence graphs, and p^*. In: P.J. Carrington, J. Scott, & S. Wasserman (eds.), *Models and methods in social network analysis.* Cambridge: Cambridge University Press.

White, S., O'Madadhain, J., Fisher, D., & Boey, Y-B. (2004). *JUNG. Version 1.4.* Irvine, California: University of California.