Library & Information Science Research 38 (2016) 259–264

Contents lists available at ScienceDirect

Library & Information Science Research

State of the art review

An evaluation of exponential random graph modeling and its use in library and information science studies

Ana Dubnjakovic

College of Communications, Department of Library and Information Sciences, University of South Carolina, 800 Sumter Street, Columbia, SC 29208, USA

ARTICLE INFO

Article history:
Received 25 November 2014
Received in revised form 8 January 2016
Accepted 29 August 2016
Available online 6 September 2016

ABSTRACT

Social network analytical tools and theories have long been an accepted part of the research landscape in many social and physical sciences including: sociology, political science, psychology, communications, business, geography, biology, physics, and chemistry as well as library and information science (LIS). Given the level of activity in the social network analysis (SNA) area concerning LIS, it is important to review the latest trends in the SNA stochastic modeling, namely exponential random graph (ERG) models. Unlike previous SNA methods, ERG models offer insight into generative network properties through simultaneous inclusion of structural parameters and attributes in the analysis while accounting for the interdependent nature of network data. Additionally, when Monte Carlo Markov Chain Maximum Likelihood Estimator is used, ERG modeling results in parameter estimates superior to other methods (e.g., MRQAP). The current study will demonstrate the utility of ERG models in LIS through a brief overview of major concepts and techniques in SNA, followed by a detailed description of ERG modeling technique, a review of currently available software used in analysis and a brief examination of its current use in LIS studies.

1. Introduction

Social network analytical tools and theories have long been an accepted part of the research landscape in many social and physical sciences including: sociology, political science, psychology, communications, business, geography, biology, physics, chemistry as well as library and information sciences (LIS). At least two studies in major information sciences journals synthesize the social network analysis impact on the field of library and information sciences. The first examines, in some detail, techniques and theories of social network analysis (SNA) in the context of information exchange studies, including its unique qualities distinguishing it from other research approaches (Haythornthwaite, 1996). Schultz-Jones (2009) provides an update of this study in terms of recent research questions and a detailed literature review of seven disciplines, including LIS, using the SNA approaches to examine information behavior. This study also distinguishes between SNA theory and analytical tools and provides an overview of history and important developments in both. Additionally, recent LIS studies in the areas of bibliometrics, webometrics, knowledge management and user information behavior indicate interest the SNA stochastic modeling. Given the level of activity in the SNA area concerning LIS as demonstrated in the two studies as well as various content areas, it is important to review the latest trends in the SNA stochastic modeling, namely the exponential random graph (ERG) models.

Since ERG modeling requires substantial background in statistics and SNA, the aim of this study is to demystify the technique and promote its usage in LIS studies by demonstrating its unique value and advantages over other SNA descriptive and stochastic analytical tools. In keeping with this goal, the current study will demonstrate the utility of ERG models in LIS through a brief overview of major concepts and techniques in SNA, followed by a detailed description of ERG modeling technique, a review of currently available software used in analysis and a brief examination of its current use in LIS studies.

2. Literature review

2.1. SNA analytical procedures

SNA relies on relational data consisting of nodes, sometimes also referred to as actors, and connecting ties, also known as edges, which can be directed or non-directed (Wasserman & Faust, 1994). Resulting networks can be viewed from a single actor's perspective, termed egocentric, or a whole network perspective focusing on ties as reported by the entire set of actors. In information exchange studies, egocentric networks can provide information about who the actor goes to for information and where they receive it, while the whole network provides insight into information behavior of groups of actors (Haythornthwaite, 1996). Networks can also be of a one-

E-mail address: ana@mailbox.sc.edu.

http://dx.doi.org/10.1016/j.lisr.2016.08.005
0740-8188/

mode variety consisting of a single set of similar nodes (e.g., a set of authors) or two-mode variety consisting of two sets of nodes with ties connecting the two sets (e.g., a set of authors and a set of articles where the author is tied to an article if they are listed as an author). Data are commonly presented in graphs or adjacency matrices suitable for further analysis. An adjacency matrix is a square matrix with as many rows or columns as nodes in the network where the intersections between rows and columns indicate the presence or absence of ties (i.e., number one indicates a tie is present and a zero tie is absent). In undirected graphs (i.e., where all ties are reciprocated), the adjacency matrix is symmetrical, with zeros on the diagonal to avoid loops (i.e., ties to self). Consequently, directed graphs or digraphs are asymmetrical since ties are not necessarily reciprocated.

Before recent developments in ERG models were introduced, the majority of SNA techniques mostly focused on descriptions of network properties such as density, in-degree, out-degree, size, centralization/centrality and distance (Wasserman & Faust, 1994). Each of these provides valuable insight into how networks operate. Density provides an account of the level of connectivity of actors in a network, in-degree and out-degree detail numbers of connections attracted and radiating from an actor respectively, while distance provides a measure of actors' reachability. Centrality measures an individual actor's influence in the network while centralization measures the extent to which actors in a network are organized around a central node. Degree centrality posits higher centrality for the actors with the most ties. Closeness centrality assigns it to the actors who are most easily reached, while betweenness centrality considers those with the highest probability of occurring on the shortest path between two randomly chosen nodes to be the most central (Borgatti & Everett, 2006).

Additional advanced SNA analytical approaches concerning whole networks are grouped around notions of cohesion (i.e., the process of grouping actors according to common characteristics), brokerage (i.e., the information diffusion in connection with centralization and bridges), ranking as assessed through measures of prestige, dyadic and triadic analysis and positions as revealed through the blockmodeling procedure (i.e., grouping of structurally equivalent actors into clusters). With few exceptions (e.g., stochastic blockmodels and the quadratic assignment procedure), the advanced SNA analytical tools are best suited to describing properties of networks and fall short when it comes to describing their generative properties. Indeed, ample evidence suggests that most network studies conducted prior to 2003 focus on consequences rather than generative properties of networks (Schultz-Jones, 2009). Given the complexity of human information behaviors, it is reasonable to expect that behaviors giving rise to complex processes such as, for instance, information exchange networks are stochastic and that statistical analyses are needed.

Prior to recent advances in ERG modeling, the few statistical studies conducted mostly focused on inclusion of some of the aforementioned network concepts into the statistical models as independent variables along with other continuous attribute variables relying on techniques such as regression (e.g., Gest, Graham-Bermann, & Hartup, 2001; Oliver & Montgomery, 1996; Reich, 2007). Even from this brief discussion of social network data, it should be immediately apparent that, due to their relational nature, network data violate independence assumptions associated with such analytical tools thereby making the inferences questionable. The problem lies in standard error computation which relies on error variance. When actors are chosen in groups rather than as individuals, the possibility of correlated disturbances increases, making the coefficients unreliable (Allison, 1999). Permutation based regression based on Krackhardt's quadratic assignment procedure (QAP) (Krackhardt, 1987), where the rows and columns in the adjacency matrix are permuted simultaneously in such a way that the network structure is left intact (Snijders, 2011), emerges as the most popular solution. The sampling distribution is generated from the possible combinations of the sample space and the observed statistic is compared to a simulated distribution (Schaeffer, 2012).

Before QAP was introduced, very few options were available to researchers who wished to include attributes in their analysis along with network structural properties. Consequently, the temptation to include structural network properties as independent variables along with other independent continuous variables is understandable. However, even QAP as the earliest form of true stochastic network analysis suffers from notable faults. For instance, while multiple regression QAP (MRQAP) extends the QAP to include examination of more than two relations, both require data manipulation and make no attempt to model network dependencies (Snijders, 2011). In contrast, ERG models provide a statistical framework capable of directly manipulating network data and attributes associated with actors resulting in less statistical noise.

2.2. Exponential random graph (p^*) models

The basic assumption underlying the logic of ERG models is that the observed network is a result of some unknown stochastic process. The proposed model aims to explain this stochastic process by testing a set of hypotheses derived from theory or prior research and represented by the structural properties of the observed network (Robins, Pattison, Kalish, & Lusher, 2007). Specifically, ERG models test whether a generative process in a network occurs more frequently than expected by chance. Broadly speaking, these generative processes can be explained through network self-organization processes characterized by activity/popularity, reciprocity, closure and brokerage, actor attributes characterized by effects of the sender, effects of the receiver and their interaction, and exogenous contextual factors such as other networks or special factors (Lusher, Koskinen, & Robins, 2013). For instance, researchers could test if preferential attachment (i.e., new actors link to actors with high indegrees) can be modeled by including appropriate parameters into the model.

When exogenous factors and attributes are not included, structural configurations shape the form of the model. The simplest structural configuration in a directed network is an arc. Higher order parameters include star and triangle configurations. Stars generally reflect the degree distribution in a network and range from 2 stars to k stars. In directed networks, they can appear in the form of in-stars (i.e., all nodes are connected to the central node but not to each other), out-stars (i.e., the central node is connected to other nodes but they are not connected to each other) and mixed stars (i.e., some combination of in-stars and out-stars). Degree distribution reflects popularity and activity effects. For example, in the context of information exchange, degree distribution could provide valuable information about the existence of hubs (i.e., nodes receiving multiple ties) that play an important role in the way information is transferred through the network. Triangles reflect the process of closure, appearing in directed networks as transitive triangles where a node's connection to two other nodes increases the likelihood those nodes will be connected (e.g., a friend of my friend is a friend), and cyclic triangles where ties are unidirectional. In transitive triads, one node receives 2 ties, one node sends 2 ties and one receives and sends one tie. When modeling information exchange networks, the prevalence of such structures could indicate that a node receiving 2 ties has the most valuable information. In undirected networks, triadic relationships appear in 4 possible configurations (i.e., no ties, one tie, two ties, or all three ties). Dominance of any of these configurations indicates to what degree the nodes in that network are isolated, appear in couples, structural holes (i.e., when a node is connected to 2 other nodes but those nodes are not connected to each other) or clusters.

Attributes, in network parlance, represent individual characteristics of actors and can be dichotomous, categorical, and/or continuous. In ERG models, attributes are considered exogenous and network structural properties endogenous to the model. If the actors' attributes affect their involvement in the network in such a way that they become more active (e.g., similar actors might share more information) those effects are known as sender effects. If, in turn, they become more popular (e.g., actors receive more ties because they have valuable information) those are known as receiver effects (Lusher et al., 2013). Obviously, these processes can also be the result of purely structural properties as outlined above. The value of ERG models is precisely in their ability to distinguish between the two effects.

Exogenous contextual factors usually come in the form of dyadic covariates although they can also include an entire network. For example, a study could include a matrix of geographic distances in a model to assess the effect such distances might have on information sharing in a certain region. If a dyad is comprised of two actors and ties between them, dyadic covariates can be seen as deterministic variables dependent on a pair of actors (Snijders, 2009). An ERG model can then be used to investigate if the covariate influences the appearance of a corresponding tie in the network of interest (Lusher et al., 2013).

The ERG model estimation process is described as follows (Robins, Pattison, Kalish, & Lusher, 2007).

1. Each tie is a random variable and for each i and j who are distinct members of the set of n actors, random variable Y_{ij} is defined as $Y_{ij} = 1$ if there is a network tie from actor i to actor j and $Y_{ij} = 0$ if the tie is absent. Y, which may be directed, undirected or valued, is the matrix of all variables and y is the matrix of the observed network. It should be noted, however, that the ERG modeling is, at present, limited to binary ties.
2. A dependence hypothesis which leads to a dependence graph is specified. This includes hypotheses about the local social processes assumed to generate network ties. Ties may be independent or a reciprocity process may be at work implying some form of dyadic dependence although social circuit dependence (i.e., tie variables within a social circuit or four-cycle are considered conditionally dependent) is most frequently used in practice. Node level attributes may also affect tie formation (e.g., actor's education level might affect information sharing in a network).
3. A specific model is generated from the specified dependence graph. Each configuration or parameter in the model is related to the presence of specific configurations. These are configurations of interest and the model is built from localized patterns represented by those configurations.
4. Homogeneity and/or other constraints on the parameter space are introduced. In order to make models identifiable, certain parameters are considered equal for all ties or constrained to other parameters.
5. Model parameters are estimated and interpreted. Parameter dependence structure dictates the complexity of this step. Generally, most structures based on realistic models will be complex. Markov chain Monte Carlo (MCMC) estimation algorithm, based on computer simulation, represents a significant improvement in this regard.

The general form of ERGMs describing a probability distribution of graphs on n nodes can be written as:

$$\Pr(Y = y) = (1/\kappa)\exp\left\{\sum \eta_A A g_A(y)\right\} \tag{1}$$

where η_A is the parameter corresponding to the configuration A; $g_A(y) = \Pi\, y_{ij} \in A\, y_{ij}$ is the network statistic corresponding to configuration A; k is the normalizing constant ensuring proper probability distribution and the summation is over all configurations A. The probability of observing the graph is dependent on the presence of various structural configurations included in the model.

Recently, higher order models including resulting parameters involving more than three nodes have been proposed (Snijders, Robins, & Handcock, 2006). These parameters offer more realistic representations of the degree distributions (e.g., alternating k star) and the transitivity structures (e.g., alternating k triangle) in the network. The alternating k star parameter takes into account all star configurations in the graph and the positive alternating k star parameter suggests higher order stars (i.e., hubs) are likely present in the network. Conversely, a positive alternating k triangle would suggest a denser area of clustering triangles in the network. Both parameters can be included in a model simultaneously and partial conditional dependence between any two disjoint pairs of nodes applies if ties are observed between the two nodes in each pair (Pattison & Robins, 2002). The new, higher order parameters introduced by Snijders are also treated in a follow up study along with the Monte Carlo maximum likelihood estimation (Robins, Pattison, Kalish, & Lusher, 2007). Additionally, *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications* (Lusher et al., 2013), a recent seminal text on ERG modeling, is an excellent reference and an invaluable tool for the LIS researchers and beyond.

2.3. Software

ERG modeling software ranges from relatively simple programs incorporating user graphic interface such as PNet[1] to programs making extensive use of command language such as the R package in Statnet.[2] PNet uses MCMC maximum likelihood estimation (MCMCMLE) procedure and, in its current form, includes packages capable of estimating one and two-mode networks (BPnet), multilevel networks (MPNet), snowball sampled networks (SnowPNet), social influence models (IPNet), and longitudinal models (LPNet). As compared to Statnet, PNet has more extensive specifications for directed and multivariate networks and the information about convergence is indicated more clearly.

Statnet, part of the R package, can make both MCMCMLE and approximate MCMCMLE estimates. The chief advantage of this program over others is that the advanced users can write their own packages that can be included into the R environment and take advantage of connections to other statistical software within the R environment.

3. ERG modeling using PNet

3.1. Model specification

Model specification is a crucial step in ERG modeling. Model parameters represent subsets of actors and ties that connect them. A list of selected commonly used parameters in social sciences available in PNet can be seen in Table 1.[3] A correctly specified model will, ideally, only include network generating parameters that fit the hypotheses tested and describe data in an adequate way. If a parameter is positive, it appears in the model more frequently than can be expected by chance. Most models, for both directed and undirected networks, will at least include a density (edge or an arc), degree distribution (stars) and closure (triangle) parameters (Lusher et al., 2013).

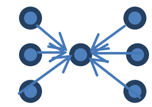
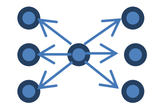
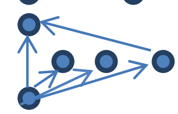
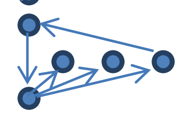
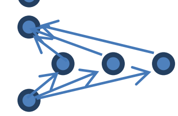
For example, researchers interested in modeling student information exchange network would include the density parameter to account for frequency of exchange. A positive reciprocity parameter would indicate the students are more likely to share information when others share information with them. A negative A-in-S parameter would indicate there is no overly popular source of information in the network, while a negative A-out-S parameter would suggest students tend to choose uniform number of others to share the information with. Positive transitivity parameter would indicate students tend to form clusters of information sharing. In other words, if student A, who student B goes to for information, goes to seek information from student C, then student B is likely to seek information from student C as well. A negative cyclic closure parameter would indicate there is no tendency for generalized information exchange. In conjunction with strong evidence for transitive closure (i.e. a positive significant AT-T parameter), a negative cyclic closure parameter would also suggest the information

[1] http://www.swinburne.edu.au/fbl/research/transformative-innovation/our-research/MelNet-social-network-group/PNet-software/index.html.

[2] http://statnet.csde.washington.edu/.

[3] For a full list see PNet User Manual available at http://sna.unimelb.edu.au/__data/assets/pdf_file/0006/662865/PNetManual.pdf.

Table 1
Commonly used structural parameters in PNet.

Parameter	PNet designation	Graphic representation
Arc (i.e., density or tendency to form ties)	Arc	
Reciprocity (i.e., tendency to form mutual ties)	Reciprocity	
Popularity (i.e., differential tendency to receive ties)	A-in-S	
Activity (i.e., differential tendency to send ties)	A-out-S	
Transitive closure (i.e., a friend of a friend is a friend)	AT - T	
Cyclic closure (i.e., tendency for generalized exchange)	AT - C	
Alternating two path (i.e., multiple connectivity)	A2P - T	

sharing process in this network is hierarchical. However, it should be noted that a recent meta-analysis study of 29 friendship networks using stochastic actor oriented models found evidence for significantly lower friendship reciprocation within transitive groups, concluding that the negative tendency toward forming three cycles was the result of neglecting to control for this tendency (Block, 2015). Block suggests modeling reciprocation in transitive triplets and possibly even omission of tree-cycles from analysis.

In addition to purely structural parameters, PNet also handles a variety of actor-relation effects or attributes, exogenous to the network. Attributes included can be categorical and/or continuous. If the attributes are categorical, dummy variables should be used just as in linear regression analysis. It should be noted that while attributes are treated as in regression analysis monadic and dyadic attributes are assigned different roles. Sender effects, receiver effects and homophily are some of the most frequently used actor attributes included in PNet. As previously noted, sender effects explain to what extent sending ties is contingent on a particular actor attribute. Similarly, receiver effects measure the extent to which actor attributes account for receiving ties, while homophily measures the effect of similarity in actor attributes on tie formation (e.g., like attracts like). For example, academic status (e.g., freshmen, sophomore, junior, etc.) could be an attribute used to measure sender effect in the information exchange study we used to illustrate commonly used structural parameters' role in generating networks. A positive sender effect parameter, in this case, would indicate students' seniority affects their tendency to share information (e.g., freshmen are more likely to share information than seniors). Similarly, a negative receiver effect would indicate there was no evidence students' academic status affects their tendency to receive information (e.g., freshmen are just as likely to receive information as juniors). A positive homophily parameter would indicate students of similar academic status are more likely to share information (e.g., seniors are more likely to share information with other seniors than freshmen or juniors). Although a single attribute (academic status) was used for the purposes of this discussion, PNet is capable of handling multiple attributes in a single model. For example, in addition to academic status the study could also have included other student attributes such as gender, GPA and major. Actor-relation effects work in conjunction with the structural parameters to explain network tie formation. Failing to include all relevant parameters in the study would result in an incomplete model at best and prevent the model from converging at worst.

3.2. Estimation and goodness of fit

Estimation in PNet is based on MCMCMLE. As previously noted, MCMCMLE, relies on computer simulation and is the preferred method of estimation due to its ability to produce reliable standard errors. Through this procedure, random networks are compared with the specified, observed network until the parameters converge (Snijders, 2002). One of the main reasons for model degeneracy is the inclusion of effects implying the assumption of Markov dependence. Models including the alternating effects are almost universally non-degenerate (Robins, Snijders, Wang, Handcock, & Pattison, 2007). PNet for Dummies contains helpful hints on how to identify parameters contributing to model degeneracy in PNet (Harrigan, 2007).

Goodness of fit in PNet is indicated by the convergence t statistic ratio. If the model is a good fit, the difference between observed and estimated parameters will be negligible. Acceptable values for the convergence t-ratios for estimated parameters are <0.10, and <2.00 for unestimated parameters (i.e. those not explicitly included in the model). A good final model would only include parameters within those values. For example, if all parameter t statistics, except for the transitive triangle (AT-T), in the aforementioned information exchange model fell below 0.10, the model would still need to be respecified (e.g., by adding more parameters) or more simulations would need to be run until the AT-T parameter fell within the acceptable range.

4. ERG modeling in LIS studies

ERG modeling is a relatively recent method in the information behavior studies. A single recent study provides insight into how consultants' information seeking from human and digital knowledge sources is influenced by their relationships with both types of knowledge sources and the characteristics of the knowledge domain in which information seeking takes place (Su & Contractor, 2011). Researchers took full advantage of ERGM ability to handle multiple networks and tested models at three different levels: the knowledge domain level (i.e., how members of a specific consulting team seek information in a specific knowledge domain), the team level (i.e., aggregates all team members' information seeking across all knowledge domains identified by the team), and the whole network level (i.e., aggregates team members' information seeking across all knowledge domains and teams) but the results were mostly reported at the overall network level.

Knowledge management (KM) studies lend themselves well to network approach, and this is currently an area in LIS where ERG models are used the most. Jiang, Gao, Chen, and Roco (2014) used ERG models to study the effects of network structures and public funding on knowledge diffusion networks of patent inventors, simultaneously modeling knowledge-sharing and knowledge-mentoring processes. Skerlavaj, Dimovski, and Desouza (2010) studied patterns and structures of intra-organizational learning networks within a knowledge-intensive organization, simultaneously modeling multiple cognitive and social processes theories. Sosa, Gargiulo, and Rowles (2014) examined the effects of a common third party on an interdependent communication team network. Their final model revealed that when the common third party fell in the middle of a communication chain between the potential source and the potential recipient of technical communication, their presence increased the likelihood of transitive structures and improved communication between the two teams. However, when the cyclic exchange was already present, the presence of the third party hindered the communication between the interdependent teams.

Lungeanu, Huang, and Contractor (2014) use ERG modeling in their bibliometric study of 1103 grant proposals submitted to the National Science Foundation. Their results indicated overall grant collaborations are more likely among individuals with longer tenure, from lower institutional tier, lower H-index, and with higher levels of prior co-authorship and citation relationship. In addition, relationship for successful proposals also indicated collaborations between females, and lower levels of citation relationship. Fanelli and Glanzel (2013) looked for the evidence supporting the Hierarchy in Sciences 19th century hypothesis, which states that "moving from simple and general phenomena to complex and particular, researchers lose ability to reach theoretical and methodological consensus." They used ERG modeling on a network of 29,000 papers from 12 disciplines and their results confirmed the original hypothesis provided the best rational framework to understand disciplines' diversity.

ERG modeling has also made inroads into webometric studies. The more traditional webometric study approach was to regress the count of inbound hyperlinks on the characteristic websites and the actors associated with them in order to identify reasons for those hyperlinks. Lusher and Ackland (2011) used ERG modeling to control for structural network parameters while studying the hyperlinking behaviors of Australian asylum advocacy groups, finding that they exhibit many of the characteristics of a social network. Yang and Yu (2014) use ERGs to model the Chinese diabetes Sina Weibo micro-blog network growth over time. Jung, Park, Wu, and Park (2014) examined the nature of online citizen participation in the field of policy analysis and management and identified patterns of citizens' e-participation and relationships between citizens, governments, and various organizations involved in policymaking processes through social media. Gonzalez-Bailon (2009) uses the economic resources of the producers of the websites as a proxy to their wider pool of resources and their presence in traditional news media as a proxy to their status to add additional relevant sociological dimensions to webometric studies regarding website visibility.

5. ERGM limitations

While ERG modeling represents a significant improvement in stochastic analysis of relational data, it should not be regarded as a panacea. As previously noted, due to the complexity inherent in network data, models can be difficult to specify and inclusion of inappropriate parameters will result in a degenerate model. Ideally, researchers would be guided by theory and include only hypothesized attributes and structural network parameters that fit the data well. ERGMs are also very sensitive to missing data. Even a small number of missing ties can have a very real and pronounced effect on how structural network parameters are represented in the observed network and consequently estimated. One solution is to treat missing responses as random and simulate them during the course of estimation so that the vector of observed statistics is substituted for the expected statistics conditional on the part that was observed (Lusher et al., 2013). Complex ERG models will also require lots of computing power and time to run. However, with recent advances in technology, this is only true for extremely complex models and is soon likely to be of no consequence. Network size can also be a factor. ERG models for networks larger than a few hundred nodes can result in problems with model specification and fit. Additionally, as currently implemented, ERG modeling is only suitable for binary network analysis. Valued networks need to be dichotomized or an alternative method of stochastic network analysis (e.g., MRQAP) needs to be used.

6. Conclusion

The purpose of this study was to demonstrate the benefits of ERG modeling in LIS. Researchers in information behavior, knowledge management, bibliometrics and webometrics are increasingly recognizing its potential as an essential method for stochastic analysis of relational data. Most other network analytical approaches are only appropriate for descriptive studies or require substantial data manipulation. Using traditional statistical techniques such as regression for network measures is inappropriate as it violates the independence assumptions required for such analyses. Additionally, as demonstrated in this overview, ERG models can also include actor attributes along with structural network parameters in a single model, opening up additional possibilities for researchers in LIS area. Since ERG modeling is a relatively new analytical tool in social sciences, a follow up study of its use in additional LIS areas is recommended to further illustrate its utility in such studies and encourage more widespread use.

Ana Dubnjakovic is the head of the Music Library at the University of South Carolina, where she also serves as the adjunct professor in the School of Music, teaching advanced research courses. Before coming to USC, Ana was the college librarian for Performing Arts and Foreign Languages at Virginia Tech. She has also served as a librarian at the Aspen Music Festival, the University of Louisville, and Herzing College in Louisiana. Ana has a master's of library and information science from Louisiana State University, a master of music in music composition from the University of New Orleans, and is a doctoral candidate in the library and information sciences program at the University of South Carolina. Her research interests focus on information seeking motivation in the university setting and quantitative methodologies in library an information science field. She co-authored a book, *A Practical Guide to Electronic Resources in the Humanities* (Elsevier, 2010), and is co-editor-in-chief of *Music Reference Services Quarterly*.

References

Allison, P. D. (1999). *Multiple regression: A primer*. Thousand Oaks, CA: Pine Forge Press.

Block, P. (2015). Reciprocity, transitivity, and the mysterious three-cycle. *Social Networks*, *40*, 163–173.

Borgatti, S. P., & Everett, M. G. (2006). A graph-theoretic perspective on centrality. *Social Networks*, *28*, 466–484.

Fanelli, D., & Glanzel, W. (2013). Bibliometric evidence for a hierarchy of the sciences. *PloS One*, *8*(6), e66938.

Gest, S. D., Graham-Bermann, S., & Hartup, W. W. (2001). Peer experience: Common and unique features of number of friendships, social network centrality, and sociometric status. *Social Development*, *10*(1), 23–40.

Gonzalez-Bailon, S. (2009). Opening the black box of link formation: Social factors underlying the structure of the web. *Social Networks*, *31*(4), 271–280.

Harrigan, N. (2007). *PNET for dummies.* Retrieved from http://www.sna.unimelb.edu.au/pnet/pnet.html

Haythornthwaite, C. (1996). Social network analysis: An approach and technique for the study of information exchange. *Library and Information Science Research*, *18*, 323–342.

Jiang, S., Gao, Q., Chen, H., & Roco, M. C. (2014). The roles of sharing, transfer, and public funding in nanotechnology knowledge-diffusion networks. *Journal of the Association* *for Information Science and Technology*, *66*, 1017–1029. http://dx.doi.org/10.1002/asi.23223

Jung, K., Park, S. J., Wu, W., & Park, H. W. (2014). A webometric approach to policy analysis and management using exponential random graph models. *Quality and Quantity*. http://dx.doi.org/10.1007/s11135-014-0010-2

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

Lungeanu, A., Huang, Y., & Contractor, N., S. (2014). Understanding the assembly of interdisciplinary teams and its impact on performance. Journal of Informetrics, 8(1), 59–70.

Lusher, D., & Ackland, R. (2011). A relational analysis of an online social movement. *Journal of Social Sturucture*, *12*(5), 1–35.

Lusher, D., Koskinen, J., & Robins, G. (2013). *Exponential random graph models for social networks: Theory, methods, and applications*. New York, NY: Cambridge University Press.

Oliver, A. L., & Montgomery, K. (1996). A network approach to outpatient service delivery systems: Resources flow and system influence. *Health Services Research*, *30*, 771–789.

Pattison, P., & Robins, G. (2002). Neighborhood-based models for social networks. *Sociological Methodology*, *32*, 301–337.

Reich, S. M. (2007). *Do nice guys finish last? The role of prosocial and aggressive behavior in peer interactions*. (Doctoral dissertation). Retrieved from Dissertation Abstracts International, 67, 4740.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, *29*(2), 179–191.

Robins, G., Snijders, T., Wang, P., Handcock, M., & Pattison, P. (2007). Recent developments in exponential random graph (p*) models for social networks. *Social Networks*, *29*(2), 192–215.

Schaeffer, D. R. (2012). Homophily through nonreciprocity: Results of an experiment. *Social Forces*, *90*(4), 1271–1295. http://dx.doi.org/10.1093/sf/sos065

Schultz-Jones, B. (2009). Examining information behavior through social networks: An interdisciplinary review. *Journal of Documentation*, *65*, 596–631.

Skerlavaj, M., Dimovski, V., & Desouza, K. (2010). Patterns and structures of intra-organizational learning networks within a knowledge-intensive organization. *Journal of Information Technology*, *25*(2), 189–204.

Snijders, T. A. B. (2002). Markov chain Monte Carlo estimation of exponential random graph models. *Journal of Social Structure*, *3*(2), 1–40.

Snijders, T. A. B. (2009). Longitudinal methods of network analysis. *Encyclopedia of complexity and systems science* (pp. 5998–6013). New York, NY: Springer.

Snijders, T. A. B. (2011). Statistical models for social networks. *Annual Review of Sociology*, *37*, 131–153.

Snijders, T. A., Robins, G., & Handcock, M. (2006). New specifications for exponential random graph models. *Sociological Methodology*, *36*, 99–153.

Sosa, M., Gargiulo, M., & Rowles, C. (2014). Inter-team technical communication in complex new product development projects. *INSEAD Working Papers Collection*, *42*, 1–39.

Su, C., & Contractor, N. (2011). A multidimensional network approach to studying team members' information seeking from human and digital knowledge sources in consulting firms. *Journal of the American Society for Information Science and Technology*, *62*, 1257–1275. http://dx.doi.org/10.1002/asi.21526

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

Yang, D., & Yu, G. (2014). Statistical analysis and exponential random graph modeling for micro-blog network. *Journal of Information Science*, *40*(1), 3–14.