FORGING AND CHANGING SOCIAL NETWORKS AT THE CSSS SUMMER SCHOOL IN BEIJING

Richard CALLAHAN
rjcal@u.washington.edu
University of Washington, USA

Maria Teresa PULIDO
tecigurl@gmail.com
University of the Philippines - Diliman, Philippines

REN Yike
ykren_2004@yahoo.com.cn
Xi'an Jiaotong University, China

SHI Yongren
yongren_shi@163.com
Huazhong University of Science and Technology, China

ABSTRACT

This project investigates how homophily, balance, reciprocity, and preferential attachment influence familiarity among participants in a month-long international summer school in Beijing. 49 participants were asked four times during the second and third weeks of the program to list the participants they knew well and to also list who they would ask for advice. We then use a continuous-time stochastic model within an exponential random graph modeling framework (ERGM) to understand how homophily, reciprocity, preferential attachment and balance influence both the shape of the network at the time of the first survey and subsequent changes in the network. We find that early tie formations and subsequent changes tend to take place as a result of reciprocation. In addition, homophily increases the likelihood of relationships initially but not over time, while balance shapes relationships over time but not initially. Preferential attachment only explains tie formation at the outset, and then only for familiarity but not advice.

INTRODUCTION

This project investigates how participants in the 2007 Santa Fe Institute's Complex Systems Summer School in Beijing (CSSS) come to know each other well. While the research question itself is of limited scope and perhaps most useful to event planners and education policymakers, the methodologies we utilize could be used in a variety of research areas, such as epidemiology, management, and trade. Primarily we seek to understand how the social networks change over time, and we make use of individual characteristics and network characteristics to test how homophily, balance, reciprocity, and preferential attachment shape the formation and development of these familiarity networks.

Early network research suffered from the fact that social networks were measured at static moments in time and little could be said about their dynamic properties. Moreover, estimates of the statistical properties of networks often assumed that edges are formed independently, a condition that almost never holds. Recent advances in exponential random graph models, however, have enabled the statistical estimation of how parameters for nodal attributes and network properties predict both the structure of a network at a static moment and how the network changes. Assuming that edges are conditionally independent given a set of nodal attributes and network processes such as reciprocity and transitivity results in a much more realistic model. Within this framework, researchers can then model network change as a continuous-time stochastic process. This project is an exercise in implementing these techniques, with a few policy implications for the Santa Fe Institute’s future summer programs in China.

Exponential random (or *p**) graphs have been used for about a decade to estimate the effects of individual and pair characteristics on network evolution [1, 2]. Rather than assume that all ties in the network are independent, we make a much more realistic assumption that the probability of observing a tie between any two actors is conditionally independent of the probability of observing a tie between any other pair of actors, given a set of attributes of the individual actors and information about the overall graph structure. Using our conditional independence assumption, we then model the network as a continuous-time stochastic process [3, 4]. We looked into the effect of each candidate characteristic as well as their possible collective effects, how these characteristics affect connections over time, and how they help predict the behavior of the network and/or of the individuals.

Familiarity among the CSSS student relationships is a noteworthy example of a social network to investigate. It was primarily chosen for the large ease and simplicity of gathering respondent data, which naturally leads to an unusually high expected response rate. In addition, the school setting also ensured that the population was controlled in terms of size and that we could obtain the informed consent both of all the survey respondents and the people who the respondents would name in the survey (since participants only named other participants).

Each survey respondent who consented to participate listed up to eight other consenting participants that he/she knew well, and up to eight consenting participants that he/she would go to for personal advice. The survey was disseminated four times during the second and third weeks of the school. “Facebooks" containing the names and pictures of the students, all of whom consented to be part of the study, were also provided to enable respondents to remember names. The survey was double blind in that the identities of the respondents and the people they named were unknown even to the authors. Personal characteristics that are likely to affect student social ties include: academic background, age, assigned hotel room, gender, nationality, and research interests. As such, the authors collected basic demographic information from the respondents as well. The first two surveys netted response rates of greater than 80%, compared to about 60% for the third survey and 40% for the fourth survey. Our analysis makes use of information from the first three surveys.

The physical structure of the social network at a given time, i.e., the presence of connections between respondents was visually inspected and its properties established. The survey results for different time frames were also compared to understand the time evolution of the network. We then examined to what extent some basic theories for network formation and change played a role in shaping this network.

Theoretical Background

This research explicitly tests four theories related to social network formation and change: reciprocity

[5], balance [6], homophily [7] and preferential attachment [8]. Each theory may interpret different aspects of human social behavior. Detailed descriptions of each theory follow.

1. Reciprocity

Reciprocity is a theory that individuals who are liked will return the goodwill. Reciprocity can be positive or negative: kindness is met with kindness and antipathy or indifference is similarly met [9]. Moreover, as Falk and Fischbacher [10] point out, “people evaluate the kindness of an action not only by its consequences but also by the intention underlying this action. A large body of evidence indicates that reciprocity is a powerful determinant of human behavior.” Reciprocal behavior, both positive and negative has been observed in such a variety of social experiments [9] that the authors of the SIENA package control for it by default.

In fact, regardless of the research setting, reciprocity involves preferential choice, i.e. a game. Through reciprocity, a group of actors form a structure of social relations that is usually well-represented by social networks.

2. Balance theory[1]

Initially proposed by Heider [11], balance theory can be applied to either groups of two actors (dyads) or three actors (triads). Here we discuss balance theory as it applies to triads. In brief, “balanced” triads are likely to contain either zero or two negative (e.g., “do not know well”) ties (Figure 1). The product of the ties in a balanced triad is positive. Balanced triads are likely to remain so, while members of unbalanced triads will attempt to resolve their situation such that the triad becomes balanced.

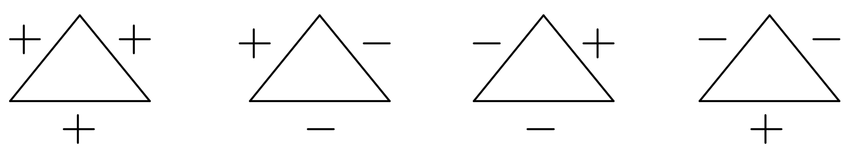

Figure 1. A Set of Balanced Triads.

In extending Heider’s theory with his “A-B-X” model, Newcomb [12] used role theory to focus on communication as a way for an unbalanced triad to resolve itself In Newcomb’s model, one actor A shares information with another actor B about some person or concept X.[2] Also a “nonbalanced” state, actors are indifferent (and so communication may not occur.) In this study, X represents a third person.

This research uses a notion of balance most closely related to Newcomb’s, except slightly extended using graph theory to encompass systems of any number of actors as done by Cartwright and Harary [13]. Given that we assume that familiarity changes through communication, balance is a sensible theory to test.

[1] The concepts presented here and their format are drawn from an overview of homophily theory developed by the Center for Interactive Advertising at the University of Texas at Austin (http://www.ciadvertising.org/studies/student/97_fall/theory/cognitive_coc/balance.htm).

[2] Balance as described by Newcomb and applied to directed graphs can be seen as a theory of outgoing ties; A and B adjust their attitudes towards each other based on what they think about C, not what C thinks about them. Accordingly, the software package we use calculates balance based on outgoing ties only. To compensate for the fact that an actor A’s attitudes towards C may nonetheless change based on C’s attitudes towards A, we always control for the effect of reciprocity when estimating the effect of balance in our models.

3. Homophily[3]

Homophily, which literally means "like attracts like," is the extent to which pairs of actors with shared characteristics are more likely to interact than those who lack these commonalities [14]. Such commonalities include demographic variables such as age and gender, technical know-how, language, social status, and values.

Rogers and Bhowmik [15] note that homophily takes place because actors with much in common are able to communicate more easily. Conversely, communicating with those who are different requires more time and effort, and may also cause cognitive dissonance when an individual receives information inconsistent with his or her knowledge or beliefs.

Students at the CSSS School were required to engage in interdisciplinary collaboration in teams that included both Chinese students and students from other countries. Many students mentioned that one of the most challenging aspects of the group collaboration was communication. Accordingly, we would anticipate an effect of homophily in our models but would want to control for participation in the same working group.

4. Preferential attachment

Briefly, preferential attachment is a theory of network change that states that the probability that an actor will link to a new actor in the future is proportional to the number of ties that the actor currently has. The concept has existed since at least 1923 [16], but Albert and Barabasi [8] first used the term "preferential attachment." Here we test the hypothesis that students who are popular (have a high in-degree, or are "known well" by many people) will become even more so over time.

Methods

Recent theoretical advances and developments in multilevel modeling allow researchers to construct causal models of dynamic social behavior. Two approaches for static (one time point) and dynamic social network models are, respectively, the Exponential Random Graph Model or ERGM [17] and Snijders's Actor-Oriented Dynamic Model, which builds upon the ERGM [18]. In this section, we will introduce both of these approaches and formalize them into statistical models for the summer school familiarity networks. We also explain how the SIENA software can be used for estimating these social network models.

1. Exponential Random Graph Model

The Exponential Random Graph Models (ERGM), also called the p* class of models, is derived from the p1 model class which restricts the relation of actors' behaviors in the dyadic independence assumption. A major goal of developing statistical models for social networks is to identify the particular pattern of an observed network from a large set of possible patterns. Which structural characteristics are most important in shaping the form of the network? The probability of occurrence of all possible networks is represented by a probability distribution and those graphs with substantial

---

[3] The concepts presented here and their format are drawn from an overview of homophily theory developed by the Center for Interactive Advertising at the University of Texas at Austin (http://www.ciadvertising.org/SA/summer_02/chjin/Net_ad/Homophily%20Theory.html).

levels of structural characteristics that reflect the generative processes actually taking place in reality are likely to have higher probability of occurrence than graphs with few such characteristics [17]. The goal of the statistical model as mentioned above is to identify model parameters that yield a "good fit" to the data in that they generate graphs with structural characteristics close to the observed graph. Robins et al. outline five steps to build an ERGM: (1) Regard network ties as random variables, (2) develop a hypothesis for how these variables are dependent, (3) ascertain the structural characteristics implied by the assumed dependence model, (4) reduce the number of parameters to be estimated through constraints, and (5) estimate and interpret the model parameters. The general form of exponential random graph model can be built in the following form:

$$\Pr(\mathbf{Y} = \mathbf{y}) = \left(\frac{1}{\kappa}\right) \exp\left\{\sum_{A} \eta_A g_A(\mathbf{y})\right\} \tag{1}$$

First the model sums all the possible structural characteristics over all configurations A; η_A is the parameter to be estimated which is corresponding to the configuration A and the $g_A(\boldsymbol{y})$ is the specific network structural characteristics A (in our summer school friends network, it could be interpreted as the density, reciprocity and transitive triangles). $g_A(\boldsymbol{y}) = 1$ if the particular structural characteristic is observed in the network **y**, otherwise it will equal to 0. Finally, k is a normalizing constant which ensures that Pr(Y=**y**) is a proper probability density.

Because calculating the set of parameters that optimizes Pr(Y=$\boldsymbol{y}$) exactly quickly becomes intractable, network researchers frequently make use of Markov Chain Monte Carlo procedures to estimate these parameters. Snijders [19] calculates the parameter set maximizing Pr(Y = $\boldsymbol{y}$) through the Conditional Method of Moments, and has made this procedure available through the SIENA software package. He has also developed a procedure for maximum likelihood estimation and has implemented this in SIENA as well, although at the time of this writing the article describing this procedure has not yet gone to press. The two estimation procedures yielded approximately the same parameter estimates for our data set. Bayesian methods are also available, although the algorithm runs slowly. The approach starts simulation from a set of parameters and then compares each generated graph with the observed graph to refine this parameter value. The process repeats until the parameters converge to a stable point that maximizes the likelihood of generating the observed network. At that point the parameters can be fixed at their final values, and additional simulations can generate a covariance matrix and information matrix for these parameters, and thus standard errors for the parameter estimates [20].

2. Actor-Oriented Dynamic Model

Social science suggests that the social dynamics of networks could be driven by local actor who is represented as the node. One class of statistical model to investigate the longitudinal data of social dynamical networks is actor-oriented dynamic model proposed by Snijders [3, 18]. This statistical model for network dynamics treats the network change "as an endogenous dynamic process that evolves in the continuous time". In other words, network evolution is treated as a continuous-time Markov process, implying that the state of the graph at time $t+1$ depends only on the state of the graph at time t. This strong assumption becomes more realistic once we incorporate both network characteristics and attribute data into describing the state of the graph at time t. Consistent with the approach for static networks, the simulation also tries to determine a parameter set that maximizes the likelihood of observing the actual network at time $t+1$.

The basic elements of this method are abstracted by Snijders in the following four steps:
1. The actors control their outgoing ties independently.

2. Decisions are never made simultaneously: at any single moment in time, only one variable $X_{ij}(t)$ may change.
3. Changes are made by the actors to optimize their situation according to an objective function.
4. The objective function also contains a random element expressing aspects of the environment not modelled explicitly.

The specification of the model based on the four steps above has up to three ingredients: the rate function, the objective function and a gratification function.

Rate function

The rate function denotes the frequency with which actors are presented with the opportunity to change their ties, and network characteristics or attribute data may affect this rate function. For simplicity, here we estimate the rate function as a constant but allow the rate to change from one time point t_m to the next time point t_{m+1}.

Objective function

In objective function, it is assumed that each actor locally control its outgoing relation from time t_m to t_{m+1}. It is assumed that actors are trying to optimize their configuration from one time to another, and that they have all information required to calculate their own objective function. The probability for the new situation of social network is then given by the multinomial logit expression:

$$f(\mathrm{x}) = \sum_k \beta_k s_k(\mathrm{x}) \qquad (2)$$

β_k represents the vector of k parameters to be estimated, and $s_k(x)$ denotes the attribute data or network characteristics that determine the evolution of the networks. Snijders [3] and Snijders et al. [21] contain lists of possible effects.

Gratification function

The gratification or "endowment" function indicates the instant gratification experienced by an actor I when changing the relation with an actor j. The idea of including this function is that in the short run an actor may regret breaking a tie, even if the actor would benefit in the long run, and the painful effect of breaking a tie may be such a burden that the actor never chooses to break the tie [3]. For the sake of simplicity, we include no endowment effects in our model.

Model Estimation

Pairing the rate and objective functions together, we have that

$$q_{ij}(\mathrm{x}) = \rho \frac{\exp\left(f(\mathrm{x}(i,j,1-x_{ij}))\right)}{\exp\left(f(\mathrm{x}(i,j,0))\right) + \exp\left(f(\mathrm{x}(i,j,1))\right)} \qquad (3)$$

where $q_{ij}(\boldsymbol{x})$ represents the probability that an actor i will change his tie with actor j between two time intervals given $\boldsymbol{x}$, the state of the graph at the first time interval. ρ denotes the rate at which actors are presented with new choices (held constant for all actors), and the fraction represents the odds that actor i will change the tie given that he or she is presented with the choice to do so. The evaluation of the configuration is thus defined as a function of the actor's position in the network and depends on parameters that can be estimated from the data by a Markov chain Monte Carlo procedure.

The Problem of Degeneracy

In estimating model parameters, an additional problem is raised by the fact that the effect of specific network configurations may be exaggerated under certain configurations. Consider the effect of the number of transitive triangles, for example, a characteristic similar to balance for discrete moments in time. As graph density increases, adding each additional edge to a network will increase the number of observed triangles in the graph, far more than in a graph of low density. While this phenomenon does not affect of the parameter values that are most likely to generate the actual network, it may make them more difficult to estimate. In the case of the triangles in a dense graph, once the estimation algorithm wanders into a section of the parameter space in which the effect of transitivity appears extremely strong, wandering back out again may take place with extremely low probability, such that in finite time the algorithm overestimates the importance of transitivity. Accordingly, we make use of some new methods for defining network configurations that reduce the likelihood of the estimation algorithm getting caught in a corner of the parameter space [21].

For testing the statistical model of summer school social network, we make use of the software package SIENA, which stands for Simulation Investigation for Empirical Network Analysis. SIENA is a software package that carries out statistical estimation for the Exponential Random Graph Model and the Actor-Oriented Dynamic Model as stated above. It can be downloaded on the website http://stat.gamma.rug.nl/siena.html. This software is best used for networks of small-to-medium size (150 actors or less, depending on the number of parameters to estimate). For large networks, we recommend using the **statnet** package in the **R** programming environment.

Model Specification and Results

1. Network at Time of First Survey

Our first set of models captures the state of the "knowing well" and "personal advice" networks as of the time the first questionnaire was administered, and tests the four theories outlined above both separately and within the same model. As controls, we include being assigned to the same room and working within the same project group. Also, as a network-related control we use network density (defined as the number of pairs of directed edges that have the same sender), but in calculating density we include a penalty that keeps the estimated importance of density from exploding in quite dense graphs, hence the name "alternating out-2-stars" [21].

To test for homophily we include three demographic characteristics: age, gender, and whether pairs of respondents were either either of the same cultural background or whether one was Chinese and one was from a different country. Reciprocity is estimated with one parameter, and preferential attachment is measured as a count of the number of pairs of edges with the same receiver (popularity), again with a penalty included to reduce the likelihood of degeneracy and labeled "alternating in-2-stars."

Finally, we estimate the effect of transitivity, which is similar to balance but which applies to a graph at a single moment in time. In a word, we observe whether triads are balanced (my friend's friend is also my friend), but this is not quite Newcomb's theory in that we are not testing to see if the graphs are becoming more balanced over time. To do so, we include two parameters: alternating transitive *k*-triangles and alternating independent 2-paths. The latter, paths of length 2, represents a precondition for triangle formation. A significant effect of transitive *k*-triangles in the presence of the parameter for 2-paths thus implies that triangles are more likely to be formed than would be explained by randomly adding edges to a graph containing many paths of length 2. Again, we include penalty functions to reduce the chance of observing a degenerate solution [21].

Table 1 shows the results of our model for the first time point. Most of the parameters for network characteristics such as reciprocity, transitivity, and popularity are not easy to interpret directly into statements about the probability of an edge forming or breaking, as the exact effect of the values of the parameters in $\boldsymbol{\beta}$ on $\mathbf{q}$ in Equation 3 depends on the state of the network. Similarly, the magnitudes of these parameters cannot be compared to each other in a straightforward way. That said, whether the parameters are positive or negative and significantly different from zero does give useful information about whether these configurations affect the graph.

Table 1. Parameters Estimated Using an ERGM for Questionnaire 1.

Parameter	**Knowing Well**	**Advice**
Roommates	3.509* (.459)	1.731* (.339)
Same Project Group	1.08* (.13)	.756* (.15)
Same Gender	-.005 (.12)	.174 (.14)
Both Chinese or Both Foreigners	.704* (.12)	.58* (.15)
Same Age Group	.239* (.13)	.280* (.13)
Reciprocity	2.32* (.26)	2.24* (.32)
Alternating transitive *k*-triangles (transitivity)	.070 (.07)	.150 (.10)
Alternating out-2-stars (density)	1.88* (.24)	1.48* (.19)
Alternating in-2-stars (popularity)	.147* (.51)	-.223 (.35)
Alternating independent 2-paths	-.182* (.03)	-.093* (.04)

* Indicates that the parameters are significantly different from zero with 95% confidence.

In short, we find that both reciprocity affect the odds of knowing someone well and asking them for advice, while transitivity does not. Popularity matters for familiarity but not personal advice; in other words, if a number of people know some actor *i* well, another actor *j* is more likely to also state that they know *i* well, whereas the number of people who would ask *i* for advice has no bearing on whether *j* will choose to ask *i* for advice. The control for density is unsurprisingly significant: ties are more likely to be observed in dense graphs.

Figure 2 translates the attribute-related controls and homophily parameters into first differences. Circles denote point estimates, while the red and blue lines denote confidence intervals. Being roommates has an enormous effect on familiarity, for example: respondents who are roommates are, all else being equal, 33 times more likely to state that they know each other well compared to respondents who are not roommates.

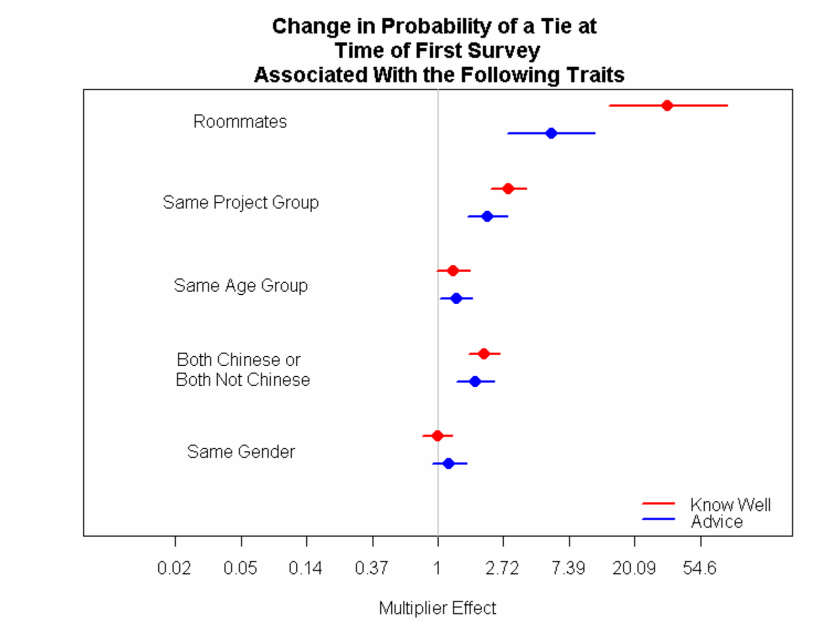

Figure 2. Effect of Shared Attributes on the Log Odds of Tie Observation for Questionnaire 1.

2. Evolution of the Networks Over Time

Our second analysis models how the "knowing well" and "personal advice" networks change between the first and second and second and third surveys. We did not make use of the fourth survey due to a low response rate. We then test the four theories outlined above within the same model, using the same controls and parameter estimates. One limitation to our analysis is that the new specifications for Exponential Random Graph Models in May 2007 have not been extended in SIENA to the Stochastic Actor-Oriented Model, so we are somewhat concerned that we are running into degenerate solutions. That said, the picture we see is of a network that has reached an equilibrium.[4] Figure 3 shows that actor attributes such as age, gender, and nationality have no effect on network change. Even roommates or members of the same project group who had not already said they were familiar with each other at the time of the survey only slightly tend to name each other as alters in later surveys. Reciprocity is still significant, indicating that new ties tend to form as a form of reciprocation. The negative effect of density appears to be an artifact of the survey design: as respondents could only name up to eight people on the survey, at a certain point density becomes a constraint on selecting new people as alters.

---

[4] As further evidence that the network was reaching an equilibrium, we also observed that the rate parameter $\boldsymbol{\rho}$ for the period between the first and second surveys was greater than that for the interval between the second and third surveys; individuals were presented with about five choices to change a tie during the first interval compared to about three during the second.

Table 2. Parameters Estimated Using an ERGM for Questionnaire 1.

Parameter	Knowing Well	Advice
Reciprocity	1.0963* (0.1840)	1.0951* (0.1830)
Density	-1.8201* (0.3014)	-1.7056* (0.3062)
Roommates	0.6081* (0.2837)	0.5710 (0.2945)
Groups	0.3215* (0.1418)	0.3049* (0.1438)
Balance	4.7378* (1.2070)	4.9407* (1.4567)
Transitive Triplets	0.0973* (0.0383)	0.0935* (0.0394)
Alter's Popularity	1.1375 (1.2167)	0.8461 (1.3543)
Same gender	0.0827 (0.1476)	0.0665 (0.1350)
Same age	0.2017 (0.1468)	0.1856 (0.1441)
Both Chinese or Both not Chinese	0.2123 (0.1468)	0.1899 (0.1451)

* Indicates that the parameters are significantly different from zero with 95% confidence.

Here we use transitive triangles as a control for balance: just as two-paths were used to identify the real effect of transitivity in the first analysis, here we use transitive triangles to separate the process of actors forging or breaking ties in order to have transitive relationships from the process of changing current configurations to make them more balanced. We find that the effect of balance appears to be huge (although the effect of the parameters on **q** in Equation 3 depends on the state of the graph.) At the very least balance is strongly significant, indicating that either actors indeed are changing their relationships to decrease cognitive dissonance or we have a degenerate solution. This is the sort of situation where some additional qualitative research could inform our results, although we did not follow up by asking individuals if their relationships changed based on their friends' opinions of others.

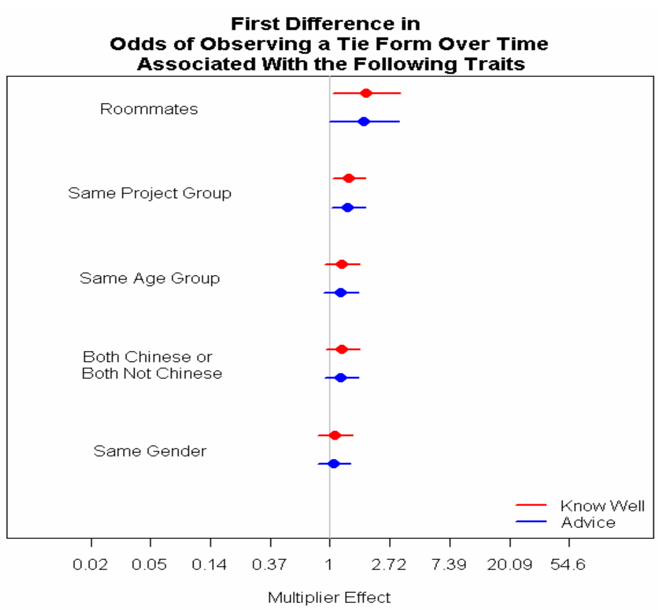

Figure 3. Effect of Shared Attributes on the Log Odds of Tie Observation Over Time.

CONCLUSION

We find that early tie formations and subsequent changes tend to take place as a result of reciprocation. Preferential attachment only explains tie formation at the outset, and then only for familiarity but not advice. In addition, homophily increases the likelihood of relationships initially but not over time, while balance shapes relationships over time but not initially. We are somewhat concerned that our model is overestimating the effect of balance on network dynamics, given the potential for the estimation algorithm to become trapped in degenerate solutions.

Friendships appear to be established early on in the Complex Systems Summer School – Beijing 2007, with 'nationality' having greater significance than either age or gender. If the purpose of arranging a Chinese and a non-Chinese as roommates is to increase mingling across nationalities, then the plan apparently worked and may be justified given the observed homophily effects.

A wide variety of networks can be modeled using the methods presented above; however, each network would depend on a different set of actor characteristics. Because the conditional dependence assumptions imply that only one particular model is “correct” in order to satisfy the Markov assumption, this research approach relies heavily on theory. In the absence of a strong theory informing research, we recommend supplementing surveys with field work for more accurate results.

FUTURE WORK

Much still needs to be done to reduce the danger of obtaining degenerate estimations of model parameters for stochastic models of network dynamics. Model parameters estimated with Bayesian methods are also often more robust than their counterparts, and running such analyses would serve as a good check against reporting degenerate parameter estimates. We also recommend supplementing the surveys with actual field work for more accurate results.

ACKNOWLEDGEMENTS

This work is part of a larger project done in collaboration with the following colleagues: Pei Liquan (University of Massachusetts – Amherst, USA), Wang Chenwei (Beijing University of Posts & Telecommunications, China), Yuan Jing (Tsinghua University, China), Zhang Hantian (Stanford University, USA), and Zhang Jialin (Tsinghua University, China).

The authors greatly benefited from conversations with Dan Hruschka, Jon Pepper, and Raissa D’Souza.

This work was partially supported by the US National Science Foundation, via award OISE 0623953.

REFERENCES

[1] Robins, G.L., Snijders, T.A.B., Wang, P., Handcock, M., and Pattison, P. “Recent Developments in Exponential Random Graph (p*) Models for Social Networks,” *Social Networks* **29** (2), 192-215 (2007).

[2] Wasserman, S., Pattison, P.E., 1996. Logit Models and Logistic Regressions for Social Networks. I. An Introduction to Markov Graphs and *p**. Psychometrika 61, 401–425.

[3] Snijders, T.A.B. “The Statistical Evaluation of Social Network Dynamics,” *Sociological Methodology*, 361-395 (2001).
[4] Huisman, M. and Snijders, T.A.B. “Statistical Analysis of Longitudinal Network Data with Changing Composition,” *Sociological Methods & Research* **32** (2003), 253-287.
[5] Fehr, E. and Gächter, S. “Fairness and Retaliation: The Economics of Reciprocity,” *Journal of Economic Perspectives* **14** (3), 159-181(2000).
[6] Wasserman, S. and Faust, K. “Social network analysis: Methods and Applications”, Cambridge University Press (1994).
[7] McPherson, M., Smith-Lovin, L., and Cook, J.M. “Birds of a Feather: Homophily in Social Networks,” *Annual Review of Sociology* (2001).
[8] Albert, R. and Barabasi, A.L. “Statistical Mechanics of Complex Networks,” *Reviews of Modern Physics* **74** (1): 47-97 (2002).
[9] Neilson W., and Stowe J. “An 'Unintentional' Theory of Reciprocity, http://faculty.fuqua.duke.edu/~stowe/bio/Professional%20Page/unintentional.pdf (2004).
[10] Falk A., and Fischbacher U. “A Theory of Reciprocity,” Working paper series, ISSN 1424-0459 (2000).
[11] Heider, F. “Attitudes and Cognitive Organization,” *Journal of Psychology* **21**, 107-112 (1946).
[12] Newcomb, T. M. “An Approach to the Study of Communicative Acts,” *Psychological Review* **60**, 393-404 (1953).
[13] Cartwright, D. and Harary, F. “Structural Balance: A Generalization of Heider's Theory,” *Psychological Review* **63**, 277-293 (1956).
[14] Touhey, J.C. “Situated Identities, Attitude Similarity, and Interpersonal Attraction,” *Sociometry* **37**, 363-374 (1974).
[15] Rogers, E.M. and Bhowmik, D.K. Homophily-Heterophily: Relationship Concepts for Communication Research. Human Information Processing and Diffusion. In *Communication*. (1972).
[16] D’Souza, Raissa. “Network Growth Models,” Lecture at the Santa Fe Institute Complex Systems Summer School, July 2007.
[17] Robins, G., Pattison, P., Kalish, Y., and Lusher, D. “An Introduction to Exponential Random Graph (p*) Models for Social Networks,” *Social Networks* **29**, 173–191 (2007).
[18] Snijders, T. A. B.“Stochastic Actor-Oriented Models for Network Change,” *Journal of Mathematical Sociology* **21**, 149-72 (1996).
[19] Snijders, T.A. “Markov Chain Monte Carlo Estimation of Exponential Random Graph Models,” *Journal of Social Structure* **3**:2, (2002).
[20] Snijders, Tom A. B., Christian Steglich, Michael Schweinberger, and Tom Huisman. *Manual for SIENA Version 3*. Retrieved from http://stat.gamma.rug.nl/siena.html (2007).
[21] Snijders, Tom A.B., Pattison Philippa E., Robins, Garry L., and Handcock, Mark S. “New Specifications for Exponential Random Graph Models.” *Sociological Methodology*, 99-153 (2007).