Response Neighborhoods in Online Learning Networks: A Quantitative Analysis

Reuven Aviv, Ph.D.
Learning International Network Consortium (LINC), Massachusetts Institute of Technology, Cambridge, U.S. A, and
Department of Computer Science, Open University of Israel
108 Ravutski Street, Raanana 43107, Israel. Tel: +972 9 7781252,
Email: aviv@openu.ac.il

Zippy Erlich, Ph.D.
Department of Computer Science, Open University of Israel
108 Ravutski Street, Raanana 43107, Israel. Tel: +972 9 7781253
Email: zippy@openu.ac.il

Gilad Ravid
Center for Information Technology in Distance Education
Open University of Israel, 108 Ravutski Street, Raanana 43107, Israel
Email: gilad@openu.ac.il

Abstract

Response mechanisms and neighborhoods in networks of online learners are revealed by Statistical Analysis of p* Markov Models for the Networks. Our analysis shows that the minimal-effort hunt-for-social-capital mechanism controls a major behavior of the networks: negative tendency to respond. Differences in the goals, interdependence and pre-assigned roles features of the designs lead to the development of different mechanisms: cognition balance and peer pressure in a team like network, exchange mechanism in a Q&A type forum. These differences lead to the formation of transitive and star-like response triads in the team network and mutual dyads in the Q&A forum. These micro structures lead to differences in the macro-structures of the networks and the buildup of collaborative knowledge. The techniques presented in this work can be extended to other types of mechanisms and networks.

Keywords: Online Learning-Networks, Response-Neighborhoods, p* analysis, Social Network Analysis

Introduction

Building networks is recognized as an essential strategy for online learning. An online network consists of actors who develop certain relations among themselves. For example, some actors only read what others write; some respond to queries posted by others and some influence others to do something (for example to access a web page), etc. More generally, a network is a set of actors – members of groups, web-pages, countries, genes, etc. – with certain possible relations between pairs of actors. The relations may or may not be hierarchical, symmetrical, binary, or other. Network abstraction is thus extremely flexible.

Social Network Analysis (SNA) is a useful tool for studying relations in a network (Wasserman & Faust 1994) . It is a collection of graph analysis methods to calculate specific network structures such as *cohesiveness* and *transitivity*: cohesiveness measures the tendency to form groups of strongly interconnected actors; transitivity measures the tendency to form transitive triad relations (if i relates to j and j relates to k, then i necessarily also relates to k). SNA has been utilized to analyze networks in various areas with actors that include politicians (Faust, Willet, Rowlee & Skvoretz 2002), the military (Dekker 2002), adolescents (Ellen et al. 2001), multi-national corporations (Athanassiou 1999), families (Widmer & La Farga 1999), and terrorist networks (van Meter 2002). SNA methods were introduced into online networks research in Garton, Haythornthwaite et al. (1997). Since then, several scholars have demonstrated the applicability of SNA to specific collaborative learning situations (Haythornthwaite 1998; Lipponen, Rahikainen, Lallimo & Hakkarainen 2001; de Laat 2002; Reffay & Chanier 2002; Aviv, Erlich, Ravid & Geva 2003).

Macro-level SNA identifies network macro-structures such as *cohesiveness*. Micro-level SNA reveals significant underlying microstructures, or neighborhoods, such as transitive triads (Pattison & Robbins 2000; Pattison & Robbins 2002). The neighborhoods identified are the basis for deducing theories that explain their emergence (Contractor, Wasserman & Faust 1999). For example, the theory of cognitive balance explains the emergence of transitive triads, which underlies the macroscopic phenomenon of cohesiveness. The precise definition of a neighborhood is given in section 2.

We examine online networks of learners according to the constructivist perspective (Jonassen et al. 1995). Rafaeli (1988) emphasized that constructive communication is determined by its responsiveness. Accordingly, we analyze the network structures of the responsiveness relation between actors in the online networks. Previous work (Aviv, Erlich & Ravid 2003) demonstrated that certain macrostructures (cohesion, centrality and role groups) are correlated with the design of the networks and with the quality of the constructed shared knowledge. In this study, we extract the micro-level neighborhoods of the same networks. Our goal is to reveal the underlying theoretical mechanisms that control the dynamics of the networks and to correlate them with the design parameters and with the quality of the knowledge constructed by the networks.

Response Neighborhoods

Every ordered pair of actors in an online network has a potential *response tie relation*. The response tie between actor i and actor j is *realized* if i responded to at least one message sent by j to the network; otherwise the response tie is not realized. In addition, a (non-directed) *viewing relation* is realized between a pair of actors if they read the same messages. In a broadcast network, a realized response tie relation is also a realized viewing tie. The reverse is not necessarily true.

A *response neighborhood* (RN) is a sub-set of actors, endowed with a set of prescribed possible response ties between them, all of which are pair-wise statistically dependent. We identified the significant RNs of a network by fitting a p* stochastic Markov model (Wasserman & Pattison 1996) to the response tie data. In this model, every pair of response ties in a RN has a common actor, which is why they are interdependent. Same topology RNs are aggregated into a class of RNs. In the model, every possible class is associated with a *strength parameter* that measures the tendency of the network to realize RNs of that class. The basic ideas and the formulas of the p* Markov model are elaborated in the Appendix. Examples of Markov RNs are presented graphically in Figure 1.

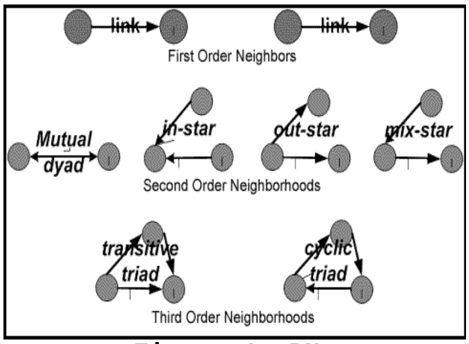

Figure 1. RNs

Tendencies to form RNs of a certain class are the result of underlying mechanisms. Several candidate mechanisms, postulated by certain network emergence theories are briefly described below. See (Monge and Contractor (2003) for an extensive survey.

In this research we consider the set of Markov classes of RNs listed in Table 1.

RN Class	Participating Actors & Prescribed Response Ties
link	All pairs: (i→j) or (j→i)
resp$_i$	All pairs: (i→j) **fixed i**
trigg$_i$	All pairs: (j→i) **fixed i**
mutuality	All pairs: (i→j) and (j→i)
out-stars	All triplets: (i→j) and (i→k)
in-stars	All triplets: (i→j) and (k→j)
mixed-stars	All triplets: (i→j) and (j→k)
transitivity	All triplets: (i→j) and (j→k)and (i→k)
cyclicity	All triplets: (i→j) and (j→k) and (k→i)

Table 1. Classes of RNs

The theory of social capital (Burt 1992) postulates efficient connectivity in the hunt for a social capital mechanism. In an online *broadcast* network, efficiency means forming zero response ties because a response tie is a redundant viewing tie, so actors prefer to remain passive. This mechanism predicts a tendency for not creating RNs of any class. Thus, other mechanisms are responsible for creating responsiveness.

Exchange and resource dependency theories (Homans 1958; Willer 1999) postulate an information exchange mechanism in which actors prefer to forge ties with potentially "resource-promising" peers. This mechanism creates tendency for RNs of class *mutuality*.

The theory of generalized exchange (Bearman 1997) postulates an information exchange mechanism via mediators. This theory then predicts tendencies for n-link cycles, in particular RNs from the *cyclicity* class.

Theories of collective action (Marwell & Oliver 1993) postulate a social pressure mechanism that induces actors to contribute to the goal of the network if threshold values of "pressing" peers, existing ties, and central actors are met (Granovetter 1983; Valente 1996). In that case, actors will respond to several others, forging *out-stars* RNs.

Contagion theories (Burt 1987; Contractor & Eisenberg 1990) postulate that the exposure of actors leads to a contagion mechanism that uses social influence and imitation to create groups of equivalent actors with similar behaviors (Carley & Kaufer 1993). Contagion predicts a tendency for RNs of the various *star* classes.

Theories	Predicted Tendencies	Hypotheses
Social capital	Few single tie links	H1: *link* < 0
Collective action	If thresholds met then respond to several others	H2: if thresholds met then *out-stars* > 0
Exchange	Tendency to reciprocate	H3: *mutuality* > 0
Generalized exchange	Tendency to respond cyclically	H4: *cyclicity* > 0
Contagion	Respond to same as others	H5: *out-stars*> 0; *in-stars* > 0; *mixed-stars* > 0
Cognitive balance	Respond via several paths	H6: *transitivity* > 0
Uncertainty reduction	Attract many responses	H7: *in-stars* > 0
Exogenous factors: Students	No tendencies to respond/trigger	H8: {*resp*$_i$ = 0 | i Є students} H9: {*trigg*$_i$ = 0 | i Є students}
Exogenous factors: Tutors	Personal tendencies to respond/trigger	H10: {*resp*$_i$ > 0 | i = tutor} H11: {*trigg*$_i$ > 0 | i = tutor}

Table 2: Research Hypotheses

Theories of cognitive balance (Cartwright & Harary 1956) postulate a cognition balance mechanism with a drive to overcome dissonance and achieve cognition consistency among actors. This drive is implemented by *transitivity* RNs.

The uncertainty reduction theory (Berger 1987) postulates drives in actors to forge links with many others to reduce the gap of the unknown between themselves and their environment; this theory predicts a tendency to create *in-stars* (responses to triggering actors) RNs.

Finally, responsibilities of actors influence their residual personal tendencies toward response ties. In this study, students did not have pre-assigned responsibilities, predicting that the students' RNs $resp_i$ and $trigg_i$ will be insignificant. The tutors' residual tendencies will be significant, due to their roles.

The theories, and predicted tendencies stated as Research Hypotheses, are presented in Table 2.

The Analysis

We analyzed recorded transcripts of two online networks of students at the Open University of Israel. These networks were established for 17 weeks during the Fall 2000 semester (19 participants) and the Spring 2002 semester (18 participants) as part of an academic course in Business Ethics. Each network included one tutor. The designs of the activities of the two networks were different. The Fall 2000 network was designed as a goal-directed collaborative team, whereas the Spring 2002 network was a Q&A forum. Hence we have labeled the networks "team" and "forum," respectively.

The *team* network engaged in a formal debate. Participants registered and committed to active participation, with associated rewards in place. Students took the role of an "advisory committee" that had to advise a company on how to handle the business/ethical problem of cellular phone emissions. The debate was scheduled as a 5-step process of moral decision-making, with predefined goals (Geva 2000). A unique feature of the team network was that the goals of the debate were to reach consensus up to the point of writing a joint proposal to an external agency. The *forum* network was open to all students in the course. Participants were asked to raise questions on issues relating to the course. We followed the social interdependence theory of cooperative learning (Johnson & Johnson 1999) to characterize the networks according to four groups of parameters: interdependence, promotive interaction, pre-assigned roles, and reflection. The two networks differ in most of the design parameters. Table 3 summarizes the differences between the designs of the two networks.

Parameter	Team	Forum
Registration & commitment	Yes	No
Interdependence: deliverables	Yes	No
Interdependence: tasks & schedule	Yes	No
Interdependence: resources	Yes	No
Reward mechanism	Yes	No
Interdependence: reward	No	No
Promotive interaction: support & help	Yes	No
Promotive interaction: feedback	Yes	No
Promotive interaction: advocating achievements	No	No
Promotive interaction: monitoring	Yes	No
Pre-assigned roles: tutor	No	Yes
Pre-assigned roles: students	No	No
Reflection procedures	No	No
Individual accountability	Yes	No
Social skills	Yes	Yes

Table 3: Design of Networks

The p* model of the *team* network has 43 classes of RNs, each with its explanatory and parameter: 18 $resp_i$, 18 $trigg_i$, *link*, *mutuality*, *transitivity*, *cyclicity*, and the three *stars*. Similarly, the model of the forum network includes 45 classes of RNs: 19 $resp_i$, 19 $trigg_i$, *link*, *mutuality*, *transitivity*, *cyclicity*, and the three *stars*. The explanatories count the number of RNs that were completely realized in the networks. The strength parameters represent the tendency to create (or not) neighborhoods from the classes.

The analysis (see the appendix) revealed three significant classes of RNs for the *team* network, and four significant classes of RNs for the *forum* network. The strength parameters are presented in Table 4.

Class	θ_K	*SE*	Wald		$\exp(\theta_K)$
Team					
link	-3.13	.32	97.5	.000	.043
out-star	.18	.06	9.6	.002	1.199
transitivity	.31	.06	23.9	.000	1.366
Forum					
link	-2.6	.8	10.29	.001	.076
$resp_{18}$	6.1	.12	26.78	.000	456.28
mutuality	6.2	1.38	20.61	.002	519.92
in-stars	-3.2	.91	12.39	.000	.041

Table 4: Revealed RNs

In Table 4, θ_K is the MPLE (maximal pseudo-likelihood estimator) for the strength parameter of class K of RNs; *SE* is an estimate of its associated standard error, $\exp(\theta_K)$ measures the increase (or decrease, if θ_K negative) in the conditional odds of creating a response tie between any pair of participants if that response tie completes a new RN of class K.

We tested the hypotheses that $\theta_K = 0$ by the Wald parameter $(\theta_K/SE)^2$ which is assumed to have chi square distribution. Table 4 shows that all these null hypotheses were rejected with extremely small *p* values.

The statistical distributions of the MPLEs and the Wald parameters are unknown (Robins & Pattison 2002), so inferences are not precise in the pure statistical sense.

Results

Few classes of RNs are significant: 3 in the *team*, 4 in the *forum*. In particular, the personal classes of RNs of students, $resp_i$ and $trigg_i$, are *not* significant. This corroborates hypotheses H8 and H9. The relative importance of the classes of RNs is depicted by their contributions to the goodness of fit of the Markov models. These are presented in Figure 2.

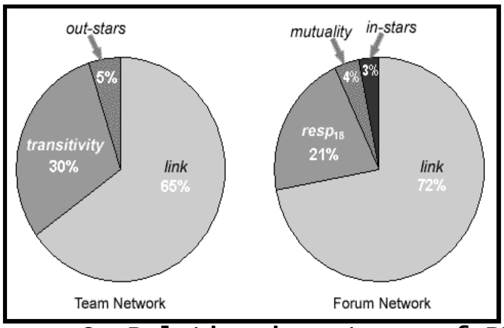

Figure 2. Relative importance of RNs

Figure 2 shows that the global class *link* of the single response tie RNs is the most significant in both networks. Table 4 shows that in both networks the strength parameter θ of the *link* class is negative. This means that the major observed phenomenon in both networks is a significant tendency for not responding. As elaborated above, this can be explained by basic self-interest – minimizing the effort required to forge a response tie vs. the possible social capital reward, given that every response tie is a redundant viewing tie. This supports hypothesis H1. This is a feature of every broadcast network, irrespective of the design of the network.

Actual responsiveness is formed by neighborhoods of other classes. These neighborhoods are quite different in the two networks. The significant RNs in the *team* network are from the global classes *transitivity* and *out-stars*. The significant RNs in the *forum* network are from the personal class $resp_{18}$, and from the global classes *mutuality* and *in-stars*. We will consider each of these RNs below.

The *team* network has a positive tendency to create transitive RNs. Specifically, the likelihood of setting up a response tie from any actor i to any other actor j is enhanced (by 1.37) if that tie completes a transitive triangle RN. No such tendency exists in the *forum* network.

These tendencies can be explained by the cognitive balance theory. It seems that the design of the *team* network leads to the cognition balance mechanism, by which dissonance between actors and between their perceptions of objects is resolved by balanced paths of communication. This can be attributed to the interdependence built into the design of the network and to the particular goal which forced the participants to reach consensus during the online debate (in order to submit joint proposals). The *forum* network, on the other hand, was a series of typical short, limited scope Q&A sessions, usually related to an assignment. There was no drive to settle conceptual inconsistencies regarding past issues, or dissonance in perceptions regarding others. Thus, no cognitive balance mechanism was needed and none was established. This explains why H6 was accepted for the *team* network but not for the *forum*.

Introducing the personal class $resp_{18}$ to the model of the *forum* network increases its goodness of fit by 21%. The tendency of N18 – the Tutor - to respond is significant. Specifically, in the *forum* network the odds of setting up a response tie ($i \rightarrow j$) increases (by 1,280) if actor i is the Tutor. In contrast, the personal class of the tutor's responses in the *team* network, $resp_1$, is statistically insignificant. This simply means that the tutor of the *team* network, P1, showed no tendency to respond.

This difference is attributed to role-assignment designs of the two networks. The tutor of the *forum* network was assigned the job of responder. The tutor of the *team* network was – deliberately – not assigned that role. This results in a difference their tendency to create the personal class of RNs. A similar observation, mentioned above, is that none of the students in either network showed a significant personal residual tendency to respond, which supports hypothesis H8. This again is attributed to the fact that students were not assigned any particular role. Similarly, in both networks every actor could trigger others by posting a question. No student was pre-assigned the role of trigger. This is reflected in the insignificance of the $trigg_i$ class of neighborhoods (consisting of a single response tie towards actor i), in agreement with hypothesis H9.

We see that the tutors in both networks had no significant tendency to trigger others, contrary to assumption H11. This is because the tutors' behavior was not controlled by roles but by other factors. In the *forum* network, the tutor served only as a helper or responder; no initiation of discussion was designed; accordingly, no triggering role was assigned to the tutor. In the *team* network, discussion was initiated by the tutor, but the design of the collaborative work dictated that the tutor should step aside. The tutor was therefore not responsible for triggering others.

Incorporating the *out-stars* class increases the goodness of fit of the Markov model for the *team* network by 5% but has no significance for the *forum* network. This means that in the *team* network the likelihood of forging a response tie from any actor i to an actor j is enhanced (by 1.2) if the tie completes an *out-star*. No such tendency is observed in the *forum* network.

The tendency to create *out-stars*, that is, to forge more than one response tie can be explained by the contagion theory (hypothesis H5) and the theory of collective action (hypothesis H2). Contagion theory predicts tendencies toward both *in-stars* and *mixed-stars*, but these predictions were not supported by the data for either network. Thus, hypothesis H5 was rejected for both networks. In general, contagion by exposure, as found in friendship relations, is a time-consuming process which, presumably, could not be developed during the short lifetime of the networks (one semester).

H2 was accepted for the *team* network but rejected for the *forum* network. This theory assumes the development of peer pressure, provided that network density and centrality are above threshold values. This condition is apparently fulfilled for the *team* network, but not for the *forum* network. The process of developing peer pressure has to overcome the basic tendency for passiveness. In the *team* network, appropriate initial conditions – commitments, interdependence, and in particular promotive interactions – were set up, and peer pressure was maintained by the tight schedule of common sub-goals imposed on the network. None of these features were designed into the *forum* network, hence no peer pressure was developed, and no drive for collective action arose.

The *mutuality* class of RNs accounts for 4% of the goodness of fit of the Markov model for the *forum* network. It has no significance for the *team* network. This means that in the *forum* network the likelihood of setting up a response tie from any actor i to any actor j is enhanced (by 5,000) if that tie closes a mutual tie. (As stated elsewhere in this paper, the actual number is not precise). No such tendency for *mutuality* RNs exists in the *team* network.

Mutuality RNs are constructed on the basis of the exchange mechanism postulated by the theories of exchange and resource dependency. Actors select their partners for response according to their particular resource-promising state. In the *forum* network the actors prefer to forge response ties (if at all) with partner(s) who usually respond to them – which in this network is the tutor. The tutor is an a priori resource-promising actor as result of her pre-assigned role. This kind of exchange calculus is not developed in the *team* network because actors in that network cannot identify a priori resource-promising actors. Hence H3 is accepted for the *forum* network but rejected for the *team* network.

Predicted Hypotheses and Tendencies	**Results and explanation**
H1: *link* < 0 Few single tie links	Supported for both networks
H2: If large density, centrality, and size, then out-stars > 0 Respond to several others	Supported only in *team*; lack of promotive interactions in *forum*
H3: *mutuality* > 0 Tendency to reciprocate to resource promising partners	Supported only in *forum*; non-existence of a priori resource-promising actors in *team*.
H4: *cyclicity* > 0 Tendency to respond cyclically to resource-promising partner	Rejected for both networks; no need for information exchange via mediators
H5: *out-stars* > 0; *in-stars* > 0; *mixed-stars* > 0; *transitivity* > 0 Respond to same as other equivalent actors	Rejected for both networks; contagion process could not develop in the short lifetime
H6: *transitivity* > 0 Respond via several paths	Supported only in *team*; difference in consensus reaching requirements and interdependence
H7: *in-stars* > 0 Attract responses from several others	Rejected for both networks; uncertainties were clarified by the design (in *team*) and by the tutor (in *forum*)
H8: {respi = 0 | i Є students} H9: {triggi = 0 | i Є students} H10: {respi > 0 | i = tutor} H11: {triggi > 0 | i = tutor} Residual personal tendencies to respond or trigger only to actors with pre-assigned roles	H8, H9: Supported for both networks; no pre-assigned role of responders to students H10: Supported in *forum*, but not in *team*; differences due to differences in pre-assigned roles of the tutor H11: rejected for both; no pre-assigned role of triggers to students

Table 5: Summary of Results

The *in-stars* class of neighborhoods accounts for 3% of the goodness of fit of the Markov model to the *forum* network but has no significance in the *team* network. In that network the likelihood of setting up a response tie from i to j decreases if this tie complements an *in-star* neighborhood, that is, if some other actor already has a response tie with j. Contagion theory and the theory of uncertainty reduction both predict a positive tendency for *in-stars* RNs. This prediction is not fulfilled. Hypotheses H5 and H7 are rejected for both networks. As mentioned above, the fact that a contagion process did not develop can probably be attributed to the short lifetime of the networks (one semester). In addition, it seems that there was no need in either network to reduce uncertainties by attracting responses from several sources: in the *forum* network, the tutor was assigned this role; in the *team* network, the rules of the game were clearly explained in the document detailing the design of the forum.

The negative tendency toward *in-stars* RNs means that participants in the *forum* network deliberately avoid responding again to the same actor. This phenomenon is explained by the theory of social capital: responding again to an actor is a waste of energy; it decreases the structural autonomy of the responder.

Neither network shows a tendency for *mixed-stars* or *cyclicity* classes of RNs. *mixed-stars* is predicted by contagion theory, hypothesis H5; the tendency for *cyclicity* is predicted by the theory of generalized exchange, hypothesis H4. Both hypotheses were rejected for both networks. As mentioned above, it is plausible that the contagion mechanism could not develop during the short lifetime of the networks. The theory of generalized exchange relies on knowledge transfer through intermediaries, who seem to be unnecessary in online broadcast networks.

Our findings, according to hypotheses, are summarized in Table 5.

Conclusions

Our analysis shows that the minimal-effort hunt-for-social-capital mechanism, predicted by the theory of social capital & transaction costs controls a large part of the behavior of both networks: a negative tendency to respond. This is a feature of every broadcast network, independent of design.

Differences in the goals, interdependence, and the promotive interaction features of the designs of the two networks lead to the development of different mechanisms: cognitive balance, predicted by the balance theory, and peer pressure, predicted by the collective action theory developed in the *team* network, but not in the *forum* network. An exchange mechanism developed in the *forum* network, but not in the *team* network. In addition, the unique pre-assigned role of the tutor in the *forum* network gave rise to the responsibility mechanism in that network, but not in the *team* network. The differences in the mechanisms led to the formation of different sets of RNs, transitive triads and out-stars in the *team* network, mutual dyads in the *forum* network. These RNs show up macroscopically as differences in cohesion and in distribution of response power and in knowledge construction (Aviv et al. 2003).

It should be noted that the important contagion mechanism did not develop in either network. This mechanism, if developed, would have led to social influence and imitation in attitudes, knowledge, and behavior, which would have developed all kinds of *star* RNs. The required design parameters – promotive interaction – were in place in the *team* network, but it seems that the lifetime of the network was too short for the development of this mechanism. This idea should be explored in longer-lived networks.

There are obvious limitations to the conclusions drawn here. First, we have considered only two networks. In order to capture the commonality, as well as the differences in design, neighborhoods, and mechanisms of online networks, one needs to consider a larger set of networks of different sizes, topics, and, in particular, with different designs. Furthermore, one should consider a set of relations embedded in these networks. One possibly relevant relation between actors is common interest, which can be captured by common keywords in transcripts and/or common sets of visited web-pages.

Another limitation lies in restricting ourselves to Markov neighborhoods. Pattison and Robbins (2002) emphasized the possible importance of non-Markovian neighborhoods and brought initial evidence of the empirical value of models that incorporate such neighborhoods. Thus, the dependence structures can, and perhaps should, be treated as a hierarchy of increasingly complex dependence structures.

It seems that SNA, and in particular p*, can be a useful research tool for revealing network architectures and mechanisms of online networks. There are numerous directions for future research. One direction is "network-covariate interaction." Several studies, such as Lipponen, Rahikainen et al. (2001), revealed that certain participants take on the roles of influencers (who trigger responses) or of celebrities (who attract responses). Others are isolated – no-one responds to them or is triggered by them. The question is whether this behavior depends on individual attributes or whether this is universal and found across networks. Another direction is "network dynamics," an inquiry into the time development of network structures. When do cliques develop? Are they stable? What network structures determine their development? Yet another direction is "large group information overload." It is well known that the dynamics of large groups leads to boundary effects that occur when the group and/or the thread size increase (Jones, Ravid & Rafaeli 2002). How are these manifested in online networks?

One practical implication of the methodology used here is the possibility for online monitoring and evaluation of online networks, by embedding SNA tools into network support environments. This can provide the instructor an intuitive understanding of the student's interactions within the network (Saltz, Hiltz & Turoff 2004).

Acknowledgement

The authors of this paper thank Gila Haimovic for her careful and productive review of the manuscript.

Appendix: Key Ideas of the p* Markov Model and the Estimation Procedure

Any ordered pair of actors in a network has a potential response tie relation. The response tie between actor i and actor j is realized if i responded to at least one message sent by j to the network. Otherwise it is not realized. The state of the network of g actors is then defined by the gXg response matrix **r**: $r_{ij} = 1$ if a response tie between i and j is realized, otherwise $\mathbf{r}_{ij} = 0$. The states of the response ties are assumed to be the result of stochastic mechanisms. The probability that the response matrix will actually be in a state **r**, Pr(**r**), is an exponential function of a linear combination of p state-dependent explanatory variables or explanatories, $\{z_1(\mathbf{r}), z_2(\mathbf{r}), \ldots, z_p(\mathbf{r})\}$. Each explanatory $z_i(\mathbf{r})$ has an associated unknown strength parameter θ_i.

$$\Pr(\mathbf{r}) = \exp\{\theta_1 z_1(\mathbf{r}) + \theta_2 z_2(\mathbf{r}) + \ldots + \theta_p z_p(\mathbf{r})\} / K(\theta_1, \theta_2, \ldots, \theta_p) \quad \text{(A1)}$$

The Hamersley-Clifford theorem (Besag, 1974, 1975) that the pair $z_N(\mathbf{r})$, θ_N are associated with one Response Neighborhood, **K**. An RN is set of actors and prescribed possible response ties between them, all of which are pair-wise statistically dependent. Actors in a neighborhood may be physically far apart, but due to certain inherent mechanisms, their possible response ties are all statistically interdependent. An RN may be completely or partially realized, or not realized at all. The explanatory $z_K(\mathbf{r})$ measures whether the RN K is completely realized, in which case it is 1. Otherwise it is zero. The strength parameter θ_K quantifies the probabilistic tendency to realize the RN.

In a Markov neighborhood (Frank & Strauss 1986), every two prescribed response ties have one actor in common. Such dependency is natural in online network: Forging response ties is an effort, so one actor's response ties are conceivably interdependent. Examples of Markov RNs are graphically presented in Figure 1.

The isomorphism invariance approximation aggregates same-topology RNs into isomorphism classes, each having one common strength parameter and one explanatory. The explanatory counts the number of RNs of the particular class that are realized in the network. The strength parameter quantifies the probabilistic tendency of the network for realizing RNs of the class. In this research we consider the set of Markov isomorphism classes listed in Table A1. The three left-hand columns in the table define the membership of actors in each class and the prescribed possible response ties, the name of the associated strength parameter (which also serves as the name of the class itself), and the formula for deriving the explanatory variables (counters) from the response matrix **r**.

Participating Actors & Prescribed Response ties	Strength Parameter θ	Explanatory $z_K(\mathbf{r})$ (counter)	Effects: If θ > 0 is significant → enhanced tendency to create
All pairs {i, j}| (i→j) or (j→i)	*link*	$L(\mathbf{r})=\Sigma_i\Sigma_j r_{ij}$	links (either direction)
All pairs {i, j}| (i→j) fixed i	*resp*$_i$	$R_i(\mathbf{r}) = \Sigma_j r_{ij}$	responses
All pairs {j, i}| (j→i) fixed i	*trigg*$_i$	$T_i(\mathbf{r}) = \Sigma_j r_{ji}$	triggers
All pairs {i, j}| (i→j) AND (j→i)	*mutuality*	$M(\mathbf{r}) = \Sigma_i\Sigma_j r_{ij} r_{ji}$	mutual responses
All triplets {i, j, k}| (i→j) AND (i→k)	*out-stars*	$OS_2(\mathbf{r}) = \Sigma_i\Sigma_j\Sigma_k r_{ij} r_{ik}$	star-responses
All triplets {i, j, k}| (i→j) AND (k→j)	*in-stars*	$IS_2(\mathbf{r}) = \Sigma_i\Sigma_j\Sigma_k r_{ij} r_{kj}$	star-triggers
All triplets {i, j, k}| (i→j) AND (j→k)	*mixed-stars*	$MS_2(\mathbf{r}) = \Sigma_i\Sigma_j\Sigma_k r_{ij} r_{jk}$	mixed trigger-responses
All triplets {i, j, k}| (i→j) AND (j→k) AND (i→k)	*transitivity*	$TRT(\mathbf{r}) = \Sigma_i\Sigma_j\Sigma_k r_{ij} r_{jk} r_{ik}$	transitive triads

All triplets {i, j, k}| (i→j) AND (j→k) AND (k→i)	*cyclicity*	CYT(**r**) $=\Sigma_i\Sigma_j\Sigma_k r_{ij} r_{jk} r_{ki}$	cyclic triads

Table A1: Isomorphism Classes of RNs and Explanatories used in Study

The probability function then takes the following form:

$$\Pr(\mathbf{r}) = \exp\{\underline{\theta}' \bullet \underline{z}(\mathbf{r})\} / k(\underline{\theta}) \tag{A2}$$

Where the vector of explanatories consists of the counters listed in Table A1:

$$\underline{z}(\mathbf{r}) = \{L(\mathbf{r}), Ri(\mathbf{r}), Ti(\mathbf{r}), M(\mathbf{r}), OS_2(\mathbf{r}), IS_2(\mathbf{r}), MS_2(\mathbf{r}), TRT(\mathbf{r}), CYT(\mathbf{r})\} \tag{A3}$$

and the strength parameters measure the tendencies for realizing the RNs of the corresponding Markov classes:

$$\underline{\theta} = \{\mathit{link}, \mathit{resp}_i, \mathit{trigg}_i, \mathit{mutuality}, \mathit{out\text{-}stars}, \mathit{in\text{-}stars}, \mathit{mixed\text{-}stars}, \mathit{transitivity}, \mathit{cyclicity}\} \tag{A4}$$

L(**r**) counts the number of RNs of class *link* that were actually realized in the network whose response matrix is **r**. *link* measures the common tendency to form single response ties; that is, to respond or to trigger. If *link* is negative, it measures the tendency not to form response ties. R_i(**r**) counts the number of RNs of the *resp*$_i$ class that were realized in the network, and *resp*$_i$ measures the residual tendency (or non-tendency) of actor i to respond, above and beyond the common tendency measured by *link*. Similarly, T_i(**r**) counts the number of neighborhoods of class *trigg*$_i$ that were actually realized, and *trigg*$_i$ measures the residual capability of actor i to attract responses to his/her previous messages; that is, to trigger others. M(**r**) counts the number of realized mutual dyads, OS_2(**r**), IS_2(**r**), and MS_2(**r**) count the number of realized *star* RNs, TRT(**r**) and CYT(**r**) count the number of realized transitive and cyclical RNs. The corresponding strength parameters measure the tendency to realize RNs of these classes. Note that the explanatories count only completely realized neighborhoods.

Wasserman and Pattison (1996) reformulated the exponential form of Pr(**r**) into a logit form, which provides both an insight into the precise meaning of "tendency" and a useful procedure for estimating the strength parameters. The logit form of the Markov model is presented in equation A5:

$$w_{ij} \equiv \log [\Pr(r_{ij} = 1 | r^c_{ij}) / \Pr(r_{ij} = 0 | r^c_{ij})] = \Sigma_K \theta_K d_K(r^c_{ij}, ij) \tag{A5}$$

The left side is the logit – the log of the conditional odds of a pair of actors (i, j) to realize a response tie (i → j). Here the odds (the ratio between the probability for realizing and not realizing a response tie) is conditioned on all other response tie states, denoted by r^c_{ij}, held fixed. The logit w_{ij} is a linear combination of the "change statistics" $d_K(r^c_{ij}, ij)$.

$$d_K(r^c_{ij}, ij) = z_K(r^c_{ij}, r_{ij} = 1) - z_K(r^c_{ij}, r_{ij} = 0) \tag{A6}$$

The change statistic $d_K(r^c_{ij}, ij)$ counts the increase in the number of RNs of class K when the response tie (i → j) flips from "non-realized" to "realized." It is 1 if (i → j) completes a whole RN; otherwise it is zero.

The logit form (A5) provides a simple interpretation of the strength parameters. Suppose that an explanatory z_K(**r**) with strength parameter θ is significant. If this happens then the conditional odds for the realization of the response tie (i → j) from any actor i to any actor j will be enhanced by e^{θ} if this envisaged response tie will make a new RN of class K realized completely. This will happen if the network already has an almost complete realization of the neighborhood: only (i → j) is missing. Otherwise the conditional odds do not change. Note that if the strength parameter θ is negative, the conditional odds will decrease, so the network has the opposite tendency.

The logit form provides one method for estimating the strength-parameters. Here A5 is considered as a binary logistic regression equation: the response tie variable is the dependent variable. There are g(g-1) cases: each ordered pair of actors (i, j) is one case. The values of r_{ij} (1 or 0) for all cases are the observed response ties. The independent variables are the "change statistics" $d_K(r^c_{ij}, ij)$. The coefficients of the change statistics, θ_N, are the unknown strength parameters.

To solve A5 for θ_N, one constructs the pseudo log likelihood function:

$$PL(\underline{\theta}) \equiv \Sigma_{ij}\log [Pr(r_{ij} = 1| r^c_{ij}) / Pr(r_{ij} = 0| r^c_{ij})] = \Sigma_{ij}\Sigma_K\theta_K d_K(r^c_{ij}, ij) \quad (A7)$$

$PL(\underline{\theta})$ is the log of the product of all the conditional probabilities. It is considered a function of the unknown strength parameters $\underline{\theta} = \{\theta_1, \theta_2, \dots \theta_p\}$, with the response tie states **r** fixed at the observed values. The estimators of the strength parameters are then the values of θ_1, θ_2, … θ_p that maximize $PL(\underline{\theta})$. These are the MPLEs. The problem with this method is that one cannot assume that the estimators have the same statistical (chi squared) distributions as their MLE (maximum likelihood estimator) counterparts. Significance intervals based on this assumption can at best be considered defendable approximations. This study attempts to identify the relative strength of the most important explanatories, with no claim to provide precise numerical values for the actual values of their strength parameters.

In this research the actual values of the change-statistics $d_K(r^c_{ij}, ij)$ were calculated from the observed response **r** matrix using PREPSTAR (Anderson et al., 1999). The MPLEs for the strength parameters were then obtained by solving equation (A5) using the binary logistic procedure of SPSS. See Crouch and Wasserman (1998) for examples and details.

The estimated value of -2* $PL(\underline{\theta})$, is an estimate for the goodness of fit of the model. In the best case, when the product of the conditional probabilities is 1, -2* $PL(\underline{\theta})$ is zero. In general, this is a positive number called Pseudo Log Likelihood Deviance (PLLD) signifying that the model is not perfect. The decrements in the PLLD caused by introducing a class of RNs into the model measures the contribution of that class to the goodness of fit of the model.

Acknowledgements

The authors of this paper thank Gila Haimovic for her critical review of the manuscript.

References

Athanassiou, N. (1999). "International management research and social networks." *Connections* **22**(2): 12-21.

Aviv, R., Z. Erlich & G. Ravid (2003). "Network analysis of knowledge construction in asynchronous learning networks." *Journal of Asynchronous Learning Networks* **7**(3).

Aviv, R., Z. Erlich, G. Ravid & A. Geva (2003). "Network analysis of knowledge construction in asynchronous learning networks." *Journal of Asynchronous Networks* **7**(3): 1-23.

Bearman, P. (1997). "Generalized exchange." *American Journal of Sociology* **102**: 1383-1415.

Berger, C. R. (1987). Communication under uncertainty. *Interpersonal processes: New directions in communication research*. E. Roloff and G. R. Miller. Newbury Park, CA, Sage: 39-62.

Burt, R. S. (1987). "Social contagion and innovation: Cohesion versus structural equivalence." *American Journal of Sociology* **92**: 1287-1335.

Burt, R. S. (1992). *Structural holes*. Cambridge, MA, Harvard University Press.

Carley, K. M. & D. S. Kaufer (1993). "Semantic connectivity: An approach for analyzing symbols in semantic networks." *Communication Theory* **3**(3): 183-213.

Cartwright, D. & F. Harary (1956). "Structural balance: A generalization of Heider's theory." *Psychological Review* **63**: 277-293.

Contractor, N. S. & E. M. Eisenberg (1990). Communication networks and new media in organizations. *Organizations and communication technology*. J. Fulk and C. W. Steinfeld. Newbury Park, CA, Sage: 143-172.

Contractor, N. S., S. Wasserman & K. Faust (1999). Testing multi-level, multi-theoretical hypotheses about networks in 21 century organizational forms: An analytic framework and empirical example.

Crouch, B. & S. Wasserman (1998). "A practical guide to fitting p* social network models via logistic regression." *Connections* **21**: 87-101.

de Laat, M. (2002). Network and content analysis in an online community discourse. *Proceedings of the computer support for collaborative learning (CSCL 2002) conference*. G. Stahl. Mahwah, NJ, Lawrence Erlbaum: 625-626.

Dekker, A. (2002). "Applying social network analysis concepts to military C4ISR architectures." *Connections* **24**(3): 93-103.

Ellen, J. M., et al. (2001). "Social network centrality and sexual experience among a household sample of urban African-American adolescents." *Connections* **24**(2): 76-81.

Faust, K., K. Willet, D. Rowlee & J. Skvoretz (2002). "Scaling and statistical models for affiliation networks: Patterns of participation among soviet politicians during the Brezhnev era." *Social Networks* **24**: 231-259.

Frank, O. & D. Strauss (1986). "Markov graphs." *Journal of the American Statistical Association* **81**: 832-842.

Garton, L., C. Haythornthwaite & B. Wellman (1997). "Studying online social networks." *Journal of Computer Mediated Communications* **3**(1).

Geva, A. (2000). "The internet and the book: Media and messages in teaching business ethics." *Teaching Business Ethics* **4**(1): 85-106.

Granovetter, M. (1983). "Threshold models of diffusion and collective behavior." *Journal of Mathematical Sociology* **9**: 165-179.

Haythornthwaite, C. (1998). Network structure among computer-supported distance learners: Multiple relations, multiple media, and time.

Homans, G. C. (1958). "Social behavior as exchange." *American Journal of Sociology* **19**: 22-24.

Johnson, D. W. & R. T. Johnson (1999). *Learning together and alone: Cooperative, competitive and individualistic learning*. Needham Heights, MA, Allyn and Bacon.

Jonassen, D., M. Davidson, M. Collins, J. Campbell & B. Haag (1995). "Constructivism and computer mediated communication in distance education." *American Journal of Distance Education* **9**(2): 7-25.

Jones, G. Q., G. Ravid & S. Rafaeli (2002). *An empirical exploration of mass interaction system dynamics: Individual information overload and usenet discourse*. 35rd Annual Hawaii International Conference on System Sciences, Big Island, Hawaii, IEEE.

Lipponen, L., M. Rahikainen, J. Lallimo & K. Hakkarainen (2001). Analyzing patterns of participation and discourse in elementary students' online science discussion. *European perspectives on computer-supported collaborative learning*. Proceedings of the first European conference on CSCL. P. Dillenbourg, A. Eurelings and K. Hakkarainen. McLuhan Institute, University of Maastricht: 421-428.

Marwell, G. & P. E. Oliver (1993). *The critical mass in collective action: A micro social theory*. Cambridge, MA, Cambridge University Press.

Monge, P. R. & N. S. Contractor (2003). *Theories of communication networks*. New York, NY, Oxford University Press.

Pattison, P. E. & G. L. Robbins (2000). *Neighborhood-based models for social networks*. The International Conference on Social Networks, Vancouver, Canada.

Pattison, P. E. & G. L. Robbins (2002). "Neighborhood-based models for social networks." *Sociological Methodology* **22**(1): 301-337.

Rafaeli, S. (1988). From new media to communication. *Sage annual review of communication research: Advancing communication science*. Beverly Hills, CA, Sage: 110-134.

Reffay, C. & T. Chanier (2002). Social network analysis used for modeling collaboration in distance learning groups. *Lecture notes in computer science (LNCS).* S. A. Cerri, G. Guarderes and F. Paraguaco: 31-40.

Robins, G. L. & P. E. Pattison (2002). *Workshop on exponential random graph (p*) models for social networks*. Kennedy School of Government, Harvard University.

Saltz, J., S. R. Hiltz & M. Turoff (2004). *Student social graphs: Visualizing a student's online social network*. Computer Supported Cooperative Work 2004, Chicago, IL.

Valente, T. (1996). "Social network thresholds in the diffusion of innovation." *Social Networks* **18**: 69-89.

van Meter, K. M. (2002). "Terrorists/liberators: Researching and dealing with adversary social networks." *Connections* **24**(3): 66-78.

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

Wasserman, S. & P. E. Pattison (1996). "Logit models and logistic regression for social networks. Part I. An introduction to Markov graphs and p*." *Psychometrika* **60**: 401-426.

Widmer, E. D. & L. A. La Farga (1999). "Boundedness and connectivity of contemporary families: A case study." *Connections* **22**(2): 30-36.

Willer, D. (1999). *Network exchange theory*. Westport, CN, Praeger.