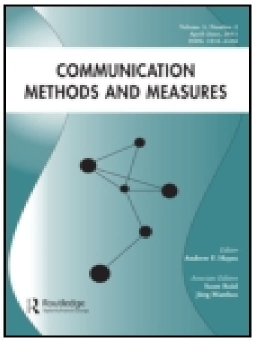

Communication Methods and Measures

ISSN: 1931-2458 (Print) 1931-2466 (Online) Journal homepage: https://www.tandfonline.com/loi/hcms20

Dynamic Models of Communication in an Online Friendship Network

Brooke Foucault Welles, Anthony Vashevko, Nick Bennett & Noshir Contractor

To cite this article: Brooke Foucault Welles, Anthony Vashevko, Nick Bennett & Noshir Contractor (2014) Dynamic Models of Communication in an Online Friendship Network, Communication Methods and Measures, 8:4, 223-243, DOI: 10.1080/19312458.2014.967843

To link to this article: https://doi.org/10.1080/19312458.2014.967843

Published online: 09 Dec 2014.

Communication Methods and Measures, 8:223–243, 2014

ISSN: 1931-2458 print/1931-2466 online
DOI: 10.1080/19312458.2014.967843

Dynamic Models of Communication in an Online Friendship Network

Brooke Foucault Welles
Northeastern University

Anthony Vashevko
Stanford University

Nick Bennett and Noshir Contractor
Northwestern University

In this article, we argue for the usefulness of relational event network analysis to study online communication networks. Unlike other network analytic techniques that require online communication data to be summarized prior to analysis, relational event network analysis uses un-summarized time-stamped data to track the dynamic evolution of communication networks. To illustrate, we use relational event network analysis to analyze the evolution of a communication network within the virtual world *Second Life*. Results suggest that there are different patterns of communication among nonfriends and friends within the network. Nonfriends tend to communicate with those they have communicated with in the past, reciprocate communication, and close communication triads. Friends tend not to communicate with those they have communicated with in the past, instead preferring to reciprocate communication and close triads. We discuss implications for the study of online communication and identify directions for future research using relational event network analysis.

Online communication data and advanced network analysis techniques have enabled scholars to examine communication patterns and practices with increasing precision and detail. In recent years, the prevalence of automated online data collection, including social network data collected passively from social media, online games and virtual worlds, have only expanded these possibilities. Although studies of social networks in Computer Mediated Communication (CMC) date back 20 years or more (Garton, Haythornthwaite, & Wellman, 1997), prior research has focused almost exclusively on static networks, or "snapshots," of communication patterns aggregated or taken at a single point in time. These snapshots mask the dynamic processes that produced the networks of study, and as a result, much of what we know about online communication is predicated on assumptions of homogeneity over time.

Correspondence should be addressed to Brooke Foucault Welles, Northeastern University Department of Communication Studies, 360 Huntington Avenue, 204 Lake Hall, Boston, MA 02115. E-mail: b.welles@neu.edu

In this article, we argue that online communication is better understood as a series of *relational events* and introduce a new method for analyzing *relational event networks* that allows communication researchers to directly examine how the dynamic, heterogeneous communication processes influence the creation, maintenance, and dissolution of communication ties over time. This technique was developed by Butts (2008) to study communication among emergency first-responders and later extended by Brandes, Lerner, and Snijders (2009) to examine to predict military hostility and cooperation among nation states. We argue that the technique is especially well suited for contemporary CMC data and will enable more accurate theorizing about how communication processes produce patterns of communication observed in CMC networks.

The remainder of this article is organized in five parts. First, we give a brief overview of how social network analysis methods have been applied to the study of CMC. Second, we discuss the limitations of prior work, highlighting how analytic techniques can obscure available detail in CMC data. Third, we introduce a new technique called *relational event network analysis* (Brandes et al., 2009; Butts, 2008) and describe how the technique can overcome some of the limitations of prior work. Fourth, we demonstrate the strengths of relational event network analysis and provide an example of its implementation using communication network data drawn from the virtual world *Second Life* as an example. Finally, we discuss the results of the example analysis and describe the benefits of relational event network analysis for the study of CMC more broadly.

SOCIAL NETWORK ANALYSIS AND CMC

In their seminal paper, Garton et al. (1997) argue for the use of social network analysis to understand CMC. Social networks are composed of a set of two or more individuals (nodes) connected by one or more relationships (ties). There are a number of different ways to conceive of nodes and ties in CMC (Rosen, Barnett, & Kim, 2011). For example, nodes in a CMC network may be individual people connected by acts of communication (e.g., sending an email, posting on a Facebook wall), websites connected by hyperlinks, or user accounts connected by encoded relationships (e.g., Twitter followers, Facebook friends). Social network analysis examines the social structure represented by nodes and ties, describing and/or statistically modeling the particular pattern of nodes and ties in an observed network.

Social network analysis offers a number of advantages over traditional statistical approaches for the study of CMC. Because network analysis involves modeling individuals embedded in social contexts, it is well suited for the study of phenomena that originate across analytic levels (individual, dyad, whole-network) (Monge & Contractor, 2003). Social network analysis is also uniquely able to model the interdependence of online communication. Unlike other data collected from individuals, CMC data is relational, involving two or more people and the message(s) they exchange. The decision to communicate online depends, in part, on motivations of the message sender, the receiver, and the social context they are both embedded in. Although more traditional statistical techniques assume independence among data, social network analysis approaches model these interdependencies directly. Finally, although it is possible to study the sparse, unbounded networks characteristic of CMC using other approaches, social network analysis *focuses* on these configurations, with well-developed statistical techniques for handling large amounts of missing data (Garton et al., 1997; Haythornthwaite, 1996).

Since Garton and colleagues' call, many researchers have used network approaches to study CMC (see Lazer et al., 2009, for review). Prior work typically uses one of three approaches. The first approach focuses on networks from the perspective of individuals, or *ego-networks.* Ego-networks are a particular type of social network representation that consists of one individual (the *ego)* and his/her direct ties to other individuals (the *alters)*. For example, a Facebook user account can be represented as the ego-network of an individual Facebook user and the ties among his/her Facebook friends. Researchers focusing on CMC ego-networks have made a number of useful observations. Prior work has determined that online ego-networks have approximately the same structure of nodes and ties as offline ego-networks, and that the number of active ties in online ego-networks, on average, does not differ significantly from the number of active ties in offline ego-networks (Arnaboldi, Conti, Passarella, & Pezzoni, 2012). Researchers have also discovered that individual characteristics including age, gender and personality are all predictive of the number of ties in online ego-networks (Amichai-Hamburger & Vinitzky, 2010; Thelwall, 2008). Researchers have also determined that diverse, highly interconnected networks are associated with a number of benefits, including social support and access to social capital (Ellison, Steinfield, & Lampe, 2007; Stefanone, Kwon, & Lackaff, 2011).

A second network-based approach to studying CMC involves describing the characteristics of large CMC networks as a whole, where the unit of analysis is an entire network, rather than individuals or dyads within the network. Mislove, Marcon, Gummadi, Druschel, and Bhattacharjee (2007), for example, discovered that many online networks are characterized by the hub-and-spoke pattern of relationships observed in offline networks, where a small number of popular individuals connect a larger number of less popular individuals in a single network. This finding has been confirmed elsewhere (Ahn, Han, Kwak, Moon, & Jeong, 2007), and also extended to include signed networks consisting of positive and negative relationships (Leskovec, Huttenlocher, & Kleinberg, 2010) and to describe distinct types of CMC users, based on the structure of the network surrounding them (Kumar, Novak, & Tomkins, 2010).

More recently, researchers have begun to use inferential network statistics to test hypotheses about how the particular arrangement of nodes and ties in CMC networks came to be. These analyses can be performed at any level of analysis (individual, dyad, whole-network) or across many levels at once (Monge & Contractor, 2003). For example, some researchers have used hierarchical regression modeling to examine the factors involved in creating network ties and/or the effects of network structure on the actors involved (Ratan, Chung, Shen, Poole, & Williams, 2010; Shen, Monge, & Williams, 2012). Others have used a specialized class of inferential statistics for networks called exponential random graph modeling (ERGM) to perform statistical tests of the likelihood of certain patterns of communication and collaboration within online networks (Hunter, Handcock, Butts, Goodreau, & Morris, 2008; Keegan, Gergle, & Contractor, 2012; Shumate & Palazzolo, 2010). These techniques allow researchers to make statistical inferences about dynamic network processes that emerge over time, but they rely on static network data. Neither hierarchical regression modeling nor ERGM allow for the direct observation of dynamic network processes.

Less frequently, researchers have attempted to map online communication network dynamics using panel data, drawing inferences about changes observed in network configurations among the same set of individuals observed at two or more points in time (Lewis, 2011; Lewis, Kaufman, Gonzalez, Wimmer, & Christakis, 2008). One inferential statistical technique, stochastic actor-oriented network analysis (SIENA), is used to examine how network structure influences the behavior of the individuals within a network over time (Snijders, 2001). Researchers using stochastic actor-oriented network analysis sample data from the same network at multiple points in time (T1, T2, . . .) and use statistics to infer what caused the network to change from a particular configuration of nodes and ties at T1 to a new configuration of nodes and ties at T2.

For example, Hurme, Veermans, Palonen, and Järvelä (2008) used stochastic actor-oriented network modeling to investigate whether individuals' participation in online class discussion forums was influenced their demographic attributes, their position in the discussion network, or both. By examining changes in the communication network on a weekly basis over the course of five weeks, they discovered that network structure, and not individual attributes, accounted for most of the variance in participation patterns. Specifically, they found betweenness centrality (measured as the average number of links between any given node and all other nodes in the network) and reciprocity were better predictors of participation than individual demographic attributes including gender and class rank (Hurme et al., 2008). These conclusions speak to processes that happen over time, although notably, the exact nature of these processes is inferred, not observed.

Recent advances in exponential random graph modeling techniques (including the software packages LPNet and TERGM) have also enabled the addition of time-based information into network models (Hanneke, Fu, & Xing, 2010; Krivitsky & Handcock, 2014; Wang, Robins, & Pattison, 2006). Similar to stochastic actor-oriented models, these techniques use panel data or indicators of the duration of network ties (e.g., how long a tie has existed) to model longitudinal changes in network structure. The chief distinction is that LPNet and TERGM are tie-based network models, used to estimate the formulation and dissolution of network ties over time, while stochastic actor-oriented modeling is actor-based, used to estimate the behavior of actors within the network over time.

Stochastic actor-oriented network modeling, LPNet and TERGM all rely on modeling static network *snapshots* to make inferences about the dynamic processes that caused the network to change over time. When these approaches are applied to the study CMC, they focus on stable, long-term network structures and ignore the dynamic, short-term patterns of communication that generate those structures. Consider, for example, two partners in a stable and frequent communication relationship: period after period, they stay in touch, communicating often and replying quickly to one another. Snapshot-based network models can predict transitions into and out of such stable relationships, but they have difficulty capturing the importance of factors such as reciprocal communication for ongoing relationship maintenance.

This limits theorizing about CMC to aggregate patterns of communication, and eliminates the possibility of explaining any variance that might occur during, or resulting from the processes that produce those patterns. It is also an awkward fit for many contemporary CMC datasets, which directly capture communication events (such as individual tweets or IM messages), but only loosely capture the status of relationships between users, if they capture relationship status at all (Gilbert & Karahalios, 2009). Instead, to examine networked CMC processes and improve theoretical precision, researchers can use of an emerging class of network analytics that uses *relational event data* to track the dynamic evolution of online communication networks (Brandes et al., 2009; Butts, 2008).

A NEW APPROACH: RELATIONAL EVENT NETWORK MODELING

Relational events are interactions between people that happen at specific points in time. For example, when one individual communicates with another, that act of communication with a sender, a receiver, and a moment in time is a relational event. The aggregation of these relational events across a group of people communicating with one another can be represented as a *relational event network*, and a new class of network statistics allows for *relational event network modeling* over time.

Advantages of Relational Event Network Modeling for CMC

There are at least two reasons why relational event network modeling can improve the study of CMC. The chief advantage to relational event network analysis over other techniques is that it can explicitly identify the time dependence of network communication patterns. For instance, one of the most basic time-dependent patterns in CMC is inertia—the tendency for people to continue to communicate with those they have communicated with in the past. Traditional network statistics have difficulty capturing this tendency, and therefore the effect of inertia is usually overlooked in models of CMC behavior. Event-based network analysis allows researchers to accurately model the effect of inertia, as well as higher-order time-based communication patterns such as reciprocity and triadic closure, which have been modeled inferentially based on observations of static networks in the past.

Although limitations in data collection may have prevented the effective use of event-based techniques in the past, online data present an ideal opportunity to examine communication events, refining our understanding of how communication networks emerge and evolve. To that end, relational event network analyses enable researchers to leverage the uniquely detailed data available for many online communication networks, retaining detail about each act of communication as it occurs. In the past, network data, both online and offline, tended to be composed of relational *states* such as whether or not a given pair of individuals reported being friends. Increasingly, as communication moves online, network data include records relational *events* in addition to (or sometimes, instead of) relational states. In many CMC datasets, in addition to records of relationships among individuals (e.g., Facebook friends, Twitter followers), we also have records of specific acts of communication among individuals (e.g., wall posts, re-tweets). Moreover, unlike in the offline world, where records of specific acts of communication are difficult and expensive to collect, records of online communication are often collected and cataloged automatically and unobtrusively. As a result, researchers now have access to pristine records of online communication where each instance of communication is logged, time-stamped, and stored in a database awaiting analysis.

Prior studies that have used panel and/or summary data to examine online communication networks have largely done so artificially; sampling panels or creating summary data by aggregating more detailed relational event records, mainly because methods for longitudinal network analysis could not previously handle any other type of input. For example, although they presumably had access to each individual record of communication across all five observed weeks, Hurme and colleagues (2008) examined changes between five network snapshots, each taken a week apart, rather than examining all of the observed data directly. Event-based network analysis retains the full detail available in records of online communication, treating each instance of communication as a variable in the overall evolution of the communication network. Therefore, at minimum, the technique is better suited to the level of detail available in CMC data, and an event-based network analysis approach may also enable new and/or newly detailed opportunities to understand the micro-processes involved in communication networks that are obscured using other techniques.

Conceptualizing Relational Event Network Modeling

When two individuals communicate in a group context, their communication patterns will depend not only on their own characteristics at that moment, but also on the history of their communication patterns, as well as the communication patterns between everyone else in the group. That is, any given relational event within a relational event network is simultaneously the *dependent variable* predicted by all prior relational events in that network, and the *independent variable* predicting all future relational events in that network. These event-based dependencies historically have been ignored in network analysis. Even stochastic actor-oriented network analysis, which examines the longitudinal evolution of networks, models the persistence and change of networks of static ties and not of transient relations (Snijders, 2001).

Relational event network analysis models relational event dynamics directly, examining how past relational events influence future relational events within a network observed over time. The technique was developed by Butts (2008) to analyze network dependent, time-stamped event data and was later applied to human and organizational interactions by de Nooy (2008) and Brandes and colleagues (2009). In a relational event network model, an event happens when one individual communicates with another individual in the network at a specific time. Each event in the model depends on the history of past events, and, in particular, on the network structure of those past interactions. For example, in Butts's (2008) original application of the model, he examined emergent patterns of communication among emergency first-responders to the 9/11 terrorist attacks on the World Trade Center in New York City. Based on a static snapshot of the network of communications among the emergency first-responders immediately following the attacks, Butts noticed several hubs in the network that did not correspond with organizational role, as might be expected; that is, the people coordinating the communication were not organizational leaders. He proposed two competing explanations for the emergence of these hubs. On one hand, the hubs may arise as a function of heterogeneity in the tendency to communicate. In an emergency situation, some responders may communicate more than others (either because of personality, training or both), and those that communicate most end up becoming hubs by virtue of communicating more frequently than everyone else. On the other hand, it may be that the hubs simply communicated *first* and became more salient targets for communication within the network as a result. That is, the hubs may have naturally had a similar tendency to communicate as others in the network, but because they idiosyncratically communicated first, they ended up becoming the recipients of more communication than others, and emerged as hubs as a result (Butts, 2008). Although traditional static network analyses cannot discriminate between these possibilities, using relational event based network modeling, Butts determined the latter explanation better accounted for the emergence of hubs in the network. Early patters of communication replicated over time, leading to the emergence of hubs within the first-responder network.

In a CMC communication network, an event-based network model could be used to explain a number of emergent communication patterns, including the overall tendency to communicate and the presence (or absence) of theoretically specified patterns of communication within the network, such as reciprocal dyads, closed triads, or hubs. To do so, the model assumes that one person communicates with another at a rate (or *frequency*) that is determined, in part, by the structure of the whole communication network, and the model attempts to estimate which structures in the network best determine the frequency of communication. Conditional on the event history, the possible interactions among all individuals at any given moment are assumed to be independent of one another—two people communicate without considering who else in the network might want to do so at the same time. The relational event network model estimates the frequency at which any given pair of individuals in the network will communicate with one another as a function of the all of the past communication events.

Half-Life

Central to the relational event network model is the idea of a "half-life" on communication events. The half-life is added to the model estimation to account for the diminishing effects of communication events over time. The reasoning is that events that just happened are more likely to influence communication behavior than events that happened a long time ago. To accommodate this idea, each time a specified half-life of time passes, the weight of statistics occurring outside that half-life are halved. This half-life parameter is not estimated by the model but is a parameter chosen by the researcher to maximize fit and the time scale of interest for the study (see *Model Selection* below for details on our procedure).

Model Formalization

This article relies on the relational event network analysis formalization developed by Brandes and colleagues (2009). For completeness, we reproduce the details of that formalization and described in terms of a communication network, here. In a relational event network model, interactions between each pair of individuals follow a Poisson process whose rate depends on the network structure of past interactions in the whole network (not just between to focal pair) as well as other covariates (such as time spent online by each person). The rate of communication for each pair is held constant between events. Once an event occurs between any pair, all pairs in the network (potentially) update their rate of communication. This model also accounts for right-censored observations of the many pairs that never interact, factoring in the time that elapsed without any interaction. Finally, the communication network uses weighted communication ties, representing an accumulated history of communication between two people.

In more detail, suppose that person a sends a chat c to person b at time t_c. At a later time t, the weight of the tie $w_t(a,b)$ from a to b will be equal to the number of all their past chats from a to b decayed by an exponential half-life parameter τ. Taking $C(a,b)$ as the set of all chats from a to b, the weighted volume of communication between two actors up to time t can be written as

$$w_t\,(a,b) = \sum\nolimits_{c\in C(a,b),t_c\le t} \frac{\ln 2}{\tau}\cdot \exp\left(-\,(t-t_c)\cdot\frac{\ln 2}{\tau}\right)$$

We refer to the whole network of ties $w_t(a,b)$ at a given time t by G_t. The rate of communication from person a to person b is determined by statistics derived from individual characteristics or from the structure of the network. We denote the k such statistics by s_i, and refer to the coefficients measuring their effects by θ_i. For example, if s_i is the statistic representing reciprocity (weighted communication from b to a), positive θ_i signifies that reciprocity increases the rate at which a chats to b. The rate of communication, $\lambda_t(a, b)$, from a to b at time t is then given by

$$\lambda_t(a,b) = \exp\left(\sum\nolimits_{i=1}^{k} \theta_i \cdot s_i(G_t; a, b)\right)$$

This rate is the parameter of an exponential waiting process determining time to the next communication event on the dyad. To keep the model tractable, this hazard of communication is held constant between events, so that t is fixed to the time of the last event in the whole network.

Combining the chat rates of all dyads in the network, we can calculate the likelihood of the observed history of chat events C by:

$$L = \prod_{c \in C}\left(\lambda_{t_c}(a_c, b_c) \cdot \exp\left(-\Delta t \cdot \lambda_{t_c}(a_c, b_c)\right) \cdot \prod_{a' \neq a_c, b' \neq b_c} \exp\left(-\Delta t \cdot \lambda_{t_c}\left(a', b'\right)\right)\right)$$

$$L = \prod\nolimits_{c \in C}\left(\lambda_{t_c}(a_c, b_c) \cdot \exp\left(-\Delta t \cdot \sum\nolimits_{a,b \in A} \lambda_{t_c}(a,b)\right)\right)$$

Here, A is the set of all individuals in the network and Δt is the time elapsed since the last event.

This formulation allows us to examine a rich set of possible mechanisms, looking not only at the local or global communication structure, but at individual characteristics and multiple modes of interaction as well.

CASE STUDY: *SECOND LIFE*

To demonstrate how relational event network analysis can be used to model and understand the evolution of an online communication network, we apply the technique to data drawn from the virtual world *Second Life*. Specifically, we use relational event network analysis to examine the processes that lead to more communication among friends than nonfriends. Several studies, including our own previous qualitative research within *Second Life,* have revealed that people spend a disproportionate amount of time communicating with friends online (Boellstorff, 2008; Foucault Welles, Rousse, Merrill and Contractor, 2014; de Nood & Attema, 2006). Although that observation is not especially surprising—indeed, it makes intuitive sense that people would communicate more with friends than nonfriends—when we consider the processes that may lead to more communication among friends, several alternative explanations emerge. These processes are discussed in detail, below. Relational event network modeling is uniquely able to evaluate the extent to which various *processes* individually and collectively contribute to the overall volume of communication among friends and nonfriends in *Second Life.*

Dyad Inertia

The most basic explanation for increased communication with friends is that people simply get into the habit of communicating with their friends. That is, because friends communicated more in the past, they continue to do so in the future (Hannan & Freeman, 1977). In a relational event network model, this effect is called *inertia* and it is measured by testing whether past communication predicts future communication for the same pair of people in the network (Brandes et al., 2009). This conditional pattern of communication is not often tested in communication research, and is an excellent example of the sort of process-based pattern that relational event network analysis can model directly. There are a number of reasons why dyad inertia may not operate in an online network as it does offline. For example, without the demands of face-to-face communication residents of *Second Life* may not feel compelled to continue communication as they would in the offline world or, similarly, residents who have communicated in the past may be co-present (logged into *Second Life*) less frequently than would be normal in offline settings. Nevertheless, we hypothesize (H1) that past communication will predict future communication within the network in general and (H2) past communication will predict future communication among pairs of friends.

Reciprocity

Beyond simple inertia, a second factor that may lead to more communication among friends is reciprocity. A number of social scientific studies have demonstrated that reciprocity is normative in many communication settings (e.g., Berger & Calabrese, 1975; Gouldner, 1960; Roloff, 1985). If a sender sends a message to a receiver, the receiver is very likely to send a message back to the sender in the future. A number of reasons have been given why this is true in the offline world, ranging from simple conversational turn-taking, to the development of trust and rapport, to a cognitive preference towards relational balance (Epstein, 1983; Sacks, Schegloff, & Jefferson, 1974; Wheeless, 1976).

The literature indicates that the norm of reciprocity is especially potent among friends. Because reciprocity is associated with trust and rapport, it has been noted to be especially important in the establishment and maintenance of friendship relationships in the offline world (Derlega, Wilson, & Chaikin, 1976; Hallinan, 1978). So, it may be the case that the increased volume of communication among friends originates in reciprocal exchanges. That is, reciprocity begets reciprocity among friends, resulting in more overall communication over time.

Although we expect reciprocity patterns to hold in the online world as they do offline, the anonymous nature of communication in *Second Life*, as well as the ability to ignore messages without specifically refusing to reciprocate may diminish the normative tendency towards reciprocal communication in *Second Life*. Nevertheless, we predict that (H3) communication ties will tend to be reciprocated over time, and (H4) that reciprocity among friend pairs will be more common than reciprocity in general.

Triadic Closure

A third factor that may increase the volume of communication among friends is triadic closure. Like reciprocity, a wide range of social scientific studies have observed triadic closure effects in a number of settings (e.g., Burt, 2005; Cartwright & Harary, 1956; Granovetter, 1973; Heider, 1958; Kossinets & Watts, 2006). Colloquially, this effect is described as "a friend of my friend is a friend," and triadic closure is achieved when two actors who share a common communication contact subsequently begin to communicate with one another. The reasons given for triadic closure are similar to those for reciprocity, with evidence suggesting that closed triads are associated with trust and that they are more cognitively pleasing (Burt, 2005; Granovetter, 1983; Krackhardt, 1992).

Further, there is evidence that triadic closure may be especially likely when at least two of the actors involved in the triad are friends (Espelage, Green, & Wasserman, 2007; Granovetter, 1983). If this is the case, it suggests that communication may be germane to friendship groups. That is, the increased volume of communication among friends is a function of the tendency for friends to communicate in groups, resulting in more overall communication with friends over time.

Although there is some evidence that triadic closure is relatively less common online than offline (Steinkuehler & Williams, 2006), we nevertheless hypothesize that (H5) triadic closure will predict communication within the network over time, and that (H6) closure among triads of friends will be more common than closure within the network in general.

Taken together, we expect that all of the hypotheses (H1-6) will contribute to explaining the patterns of communication that emerge among friends and nonfriends in the virtual world *Second Life*.

METHOD

To test the hypotheses outlined above, we used CMC data from the virtual world *Second Life*. The data were collected unobtrusively by Linden Labs, the company that owns *Second Life* and given to the authors for research purposes. *Second Life* is an immersive online virtual world where users (called "residents") interact with one another via avatars and can socialize, join groups, own land, and build a wide range of objects. Unlike other virtual worlds and online games that are plot-driven, *Second Life* has no overarching storyline or goal. Chiefly, residents join and use *Second Life* to socialize (Boellstorff, 2008; de Nood & Attema, 2006).

There are two main ways to socialize in *Second Life*. First, residents can designate other residents as "friends," which affords a variety of privileges, including being able to easily contact one another, see one another online, locate one another in *Second Life*, and use one another's virtual possessions, depending on the level of friendship access granted. Residents may also form and join groups. In *Second Life*, users form and join groups for a variety of reasons, including pursuing shared interests, engaging in role-playing, or starting businesses together. Although some groups require a financial commitment to join, most often, residents can join groups by simply expressing an interest. When a resident is a member of a group s/he can see a list of all other members and, depending on his/her privileges, send messages to and receive messages from that list. There is no requirement for residents to become friends with other group members, although they may be somewhat more likely to select group members as friends over other random *Second Life* residents for the simple fact that they have a mechanism for meeting and interacting with group members. Likewise, there is no requirement for residents to communicate with group members, although it is relatively easier to communicate with group members because of the contact list provided.

TABLE 1
Sample Relational Event Data Format

Sender	*Receiver*	*Time*	*Type*
A	B	Day 1, 9:50 AM	Message
A	C	Day 1, 9:55 AM	Message
B	A	Day 1, 10:21 AM	Message
B	A	Day 1, 10:24 AM	Friend request
B	A	Day 1, 10:26 AM	Message

Data

At minimum, to use relational event network analysis, a researcher must have time-stamped data that indicate the sender and receiver of each relational event in a network of interest. Optionally, the data may also include weights for each relational event, such as the number of characters in a chat message (where the chat message is the relational event) and/or types for each relational event, such as a collaborative or adversarial message. This format is similar to a network edgelist, with the minimal addition of a time stamp for each event (edge) in the list. (See Table 1 for an example of how relational event data can be formatted.) In this case, the sample data contain the sender and receiver of each event, the time of the event, and the event type (message versus friend request).

Data used to generate the sample for this analysis included records of relational (chat) events, their time stamps, weights, and types. The data were drawn from complete server-side records of online behavior collected on an ongoing basis between April 1 and July 22, 2009, for 9,962,359 *Second Life* residents. Data included the following:

Groups

The complete dataset included a record of every group in *Second Life,* identified by anonymous group ID number. This list served as the sample space from which one group was drawn for analysis (see sampling procedure description, below).

Group Membership

The dataset also included a list of all of the members of each group. This list was used to generate the list of nodes in the network we analyzed.

Chat

The dataset contained complete, time-stamped records of private chat messages sent by *Second Life* residents. Data files contained records of the sender and receiver of each chat, a time stamp measured in tenths of seconds, and the total number of characters in each chat. The messages served as the relational events in our network, with the number of characters per message serving as the weights on the events.

In Second Life, there are a number of ways to communicate including general chat not directed at any particular individual (similar to posting in a chat room), announcements directed to particular groups (similar to sending an email to a listserv), and private chats directed at a particular individual (similar to an IM). To send a private chat, a resident can either click on the avatar of another resident who is co-present to initiate a chat, or send a message to a friend or group member (who can be co-present or not) by selecting the relevant avatar's name from the friend or group contact list.

An individual "chat" includes all of the characters typed prior to hitting the "return" key, and therefore chats may vary considerably in length, frequency, and synchronicity. Although private chats can only be exchanged between users logged into *Second Life,* it is not necessarily the case that chats are exchanged synchronously. For example, a user could type a long message of several sentences or more, and send it to another user to read later. Or, alternatively, two users may engage in a conversation where a series of short, contingent chats are exchanged quickly. Private chats are not necessarily reciprocal, and the resulting chat networks are directed.

Friendship

The final data field used in this analysis included records of each friendship within *Second Life,* as well as the in-game privileges set for each friendship. Friendships in *Second Life* are always reciprocal, but in-game privileges are directed. The records of friendship were used to generate the values on the relational events, where chats were valued "between friends" or "not between friends."

Sample and Network Specifications

The relational event network model is computationally intensive. Because the model takes into account possible interactions among all actors between every observed event, model estimation time increases geometrically in the size of the network and linearly in the number of events: with n actors and m events, the computational complexity is at least on the order of $O(mn^2)$, though computationally intense network statistics will increase this. The maximum likelihood estimation procedure is also difficult to parallelize because each event relies on the state of the network after the previous, and event statistic calculations potentially rely on the state of the whole network. The procedure cannot be easily split by regions of the network or by events. A network of about 100 actors and 10,000 events (such as the one we chose) takes about half an hour to run on a modern single-core machine, and the algorithm becomes prohibitively slow at perhaps 400–500 actors, using contemporary computing technology.

Because of these computational limitations on the ability to model event-based data, for this case study, we randomly sampled a single group from the data set for analysis. The sampled group contained 94 members connected by 16,706 private chat events. We generated two different networks representing the group. The first was a time-stamped chat network consisting of group members (the nodes) connected by private chats (the ties). This network was directed and each tie was weighted by the total number of chats sent.

The second network was a friendship network consisting of the 94 group members (the nodes) connected by 106 friendship relations (the ties). Because of the ambiguous nature of "friendship" online (boyd, 2007), we limited analyses involving friendship to include only those friendships that included trusting privileges in at least one direction (namely, the ability to access, take and/or delete *Second Life* belongings). Prior research using this same data set has indicated that these relationships share more of the network structural features associated with friendship in the offline world than so-called "friendships" without trusting privileges, and may therefore be a better conceptual proxy for friendship as we interpret the results (Burt, 2011). The friendship network was undirected and fixed over the time period of the analysis, including only the friendships with trusting privileges established prior to the first chat record.

Model Specification

To test the predictions outlined above, we used the specified chat and friendship networks and included the following network statistics in our model. As described above, the model can include a decay rate to model the importance of recent interaction. We used a half-life of 10 days for the chat network, and no decay for the friendship network (see *Model Selection* discussion below for details on how this half-life was selected). Because we measure directional communication, all network statistics focus on a resident (ego) and the partner that ego may be chatting to (alter). In the following, w_{ij} will indicate the weighted volume of past chats from resident i to partner j, and f_{ij} will indicate the value of the event, set to 1 if there is a friendship between i and j, and 0 otherwise.

Inertia

Inertia records the number of times a given ego has chatted with a given alter in the past. If the inertia hypothesis (H1) holds, we expect *Inertia* to be positive, indicating that past communication increases the rate at which two residents will communicate in the future.

$$Inertia\,(i,j) = w_{ij}$$

Inertia Between Friends

Friend Inertia records the number of times ego has chatted with alter in the past, if ego considers alter a friend. If the *Friend Inertia* hypothesis (H2) holds, we expect *Friend Inertia* to be positive, indicating that past communication increases the rate at which two friends will communicate in the future.

$$FriendInertia\,(i,j) = w_{ij} \cdot f_{ij}$$

Reciprocity

Reciprocity records the number of times a given alter has chatted with a given ego in the past. If the reciprocity hypothesis (H3) holds, we expect *Reciprocity* to be positive, indicating that ego will communicate with alter more often if alter has communicated with ego in the past.

$$Reciprocity\,(i,j) = w_{ji}$$

Reciprocity Between Friends

Friend Reciprocity records the number of times a given alter has chatted with a given ego in the past, if ego considers alter a friend. In other words, this is an interaction term between reciprocity and perceived friendship from ego to alter. If H4 holds, we expect *Friend Reciprocity* to be positive, indicating that friendship will promote reciprocal communication above and beyond what is seen between nonfriends.

$$FriendReciprocity\,(i,j) = w_{ji} \cdot f_{ij}$$

Triadic Closure

Triadic Closure records the volume of communications that both ego and alter have shared with any other resident in the network. Any resident that has communicated with both ego and alter contributes to this measure. If the triadic closure hypothesis (H5) holds, we expect *Triadic Closure* to be positive, indicating that the more both ego and alter have chatted with some other resident, the more often will ego chat with alter.

$$TriadicClosure\,(i,j) = \sqrt{\sum\nolimits_{k \neq i,j} (w_{ik} + w_{ki}) \cdot \left(w_{kj} + w_{jk}\right)}$$

Triadic Closure Among Friends

Friend Closure records the total number of friends that ego and alter share in common, if ego considers alter a friend. If H6 holds, we expect *Friend Closure* to be positive, indicating that if ego and alter share a spot in many friendship triangles, ego will be more likely to communicate with alter in the future.

$$FriendClosure\,(i,j) = f_{ij} \cdot \sqrt{\sum\nolimits_{k \neq i,j} (w_{ik} + w_{ki}) \cdot \left(w_{kj} + w_{jk}\right)}$$

Statistical Controls

We also include two controls to avoid biased measurements of the above statistics:

Ego Hub records the total volume of chats that ego participates in. This measures controls for the uneven tendency to participate in chatting, and the possibility that certain people communicate very frequently and/or broadcast information to the group.

$$EgoHub\,(i,j) = \sum\nolimits_{k \neq i} w_{ik} + w_{ki}$$

Alter Hub records the total volume of chats that alter participates in. This measure controls for the tendency of residents to seek information from or chat with popular alters.

$$AlterHub\,(i,j) = \sum\nolimits_{k \neq j} w_{jk} + w_{kj}$$

A constant term controls for the baseline level of communication within the network and captures the extent to which communication is random or determined by the network structures under analysis.

Model Selection

There has been little work on model selection and goodness of fit in relational event network models. In lieu of a tailor-made model selection criterion, we use the Akaike information criterion (AIC) (Burnham & Anderson, 2002) to select the optimal model specification. AIC penalizes the log-likelihood of the estimated model by correcting for k, the number of statistics estimated, the optimal model will minimize AIC.

$$AIC = 2k - 2\ln(L)$$

We used AIC to select both the statistics to estimate as well as the optimal half-life to use in the weighted communication network.

Software

Relational event networks are still a recent entry to the modeling world. In fact, we know of only one off-the-shelf product that can estimate such models: the *relevent* package for *R* (Butts, 2008). The *relevent* package and its documentation are available for download on the Comprehensive R Archive Network (CRAN, http://cran.us.r-project.org/). Although *relevent* can accept custom networks and statistics, it does not handle weighted networks or estimate all of the statistics we describe here. We developed our own solver for the relational event network model to produce the results described here. Our implementation does incorporate weighted networks. The source code and documentation for our solver is available for download at: https://www.stanford.edu/~vashevko/pages/renm/

RESULTS

Descriptive Results

The sampled group used for this analysis had 94 residents connected by 16,706 chats and 106 friendship ties. In an undirected network of 94 nodes, there are a total of 4,371 possible ties ((n*(n-1))/2). Therefore, friendship ties represent 2.45% of all possible ties. Of the 16,706 total chats, 3,300 chats (19.75%) were exchanged between friends, and the remaining 13,406 chats (80.25%) were exchanged between nonfriends. This pattern confirms previous observations that a disproportionate amount of chatting occurs between friends in the network.

TABLE 2
Event-Based Network Analysis Results

Parameter	*Rate Estimate*	*Rate Multiplier*	*Standard Error*
Inertia	0.00156188**	1.8974	0.000045
Reciprocity	0.00054796**	1.2520	0.000045
Friend Reciprocity	0.00083864**	1.4105	0.000175
Triadic Closure	0.00052699**	1.2412	0.000025
Friend Closure	0.00014497**	1.0613	0.000031
Controls			
Ego Hub	0.00003133**	1.0129	0.000002
Alter Hub	−0.00001989**	0.9919	0.000003
Friend Inertia	−0.00136819**	0.5706	0.000169
Constant	−4.470289	0.0114	0.010954

Note. *28 days of data preloaded, half-life = 10 days. "Rate Multiplier" assumes the underlying statistic has mean level of 410.82. Constant "Rate Multiplier" is baseline rate of model.*

**indicates $p < 0.05$, **indicates $p < 0.01$*

Relational Event Network Analysis Results

Results of the relational event network modeling appear in Table 2. Coefficients on each term in the model can be understood as the logarithms of multipliers of a baseline rate of chatting. The constant term in the model gives the baseline rate of chatting between any two individuals in the network, and implies that two people chat at a rate of 0.01 messages per day (exp(-4.47) = 0.0114). All other coefficients multiply this baseline rate as a function of the state of the network. For example, if *a* sent a chat to *b*, the inertia coefficient implies that *a* has an 89.7% greater than baseline chance of sending another chat, while the reciprocity coefficient suggests that *b* has a 25.2% greater chance over baseline of replying. Because the data in this network are weighted by number of characters in each chat, the column "Rate Multiplier" in Table 2 interprets all coefficients assuming that the corresponding statistic is at the mean chat length of 410.82 characters.

The first hypothesis proposed that past communication would predict future communication within the network. This hypothesis was tested using the *Inertia* term in the model. The significant, positive coefficient on the *Inertia* term indicates that past chatting positively predicts future chatting within the network (89.7% more likely). The more an individual has chatted with someone in the past, the more likely he or she is to do so in the future. Therefore, the Inertia hypothesis (H1) is supported.

The second hypothesis proposed that past communication would predict future communication among friends in the network. The significant, negative coefficient on the *Friend Inertia* term indicates that the opposite may be true. Past communication is associated with a *reduced* likelihood of future communication among friends (42.9% less likely than baseline). Because the terms in the model are interdependent, it is most appropriate to interpret the *Friend Inertia* term in light of the general *Inertia* results for the whole network. Interpreted this way, we see that being friends has a dampening effect on inertia, but there is still a small inertia effect over baseline between friends (messages are 8.3% more likely than baseline when combined with *Inertia*). That is, the more someone has communicated with a friend in the past, the more likely s/he is to do so in the future, although the effect is substantially smaller between friends than between members of the network in general. Therefore, the Friend Inertia hypothesis (H2) is supported, however the effect is not strong.

The next hypothesis predicted that chatting would tend to be reciprocal among pairs of residents in the network. The significant positive coefficient on the *Reciprocity* term suggests that this is true, residents are 25.2% more likely than baseline to reciprocate chats over time. Therefore, the Reciprocity hypothesis (H3) is supported.

The fourth hypothesis predicted that chat reciprocity would be stronger among pairs of friends than among pairs of nonfriends. The significant positive coefficient on the *Friend Reciprocity* term suggests that pairs of friends do tend to reciprocate chat over time. Accounting for the general tendency towards reciprocity in the network, friends are an additional 41.1% more likely to reciprocate messages. Therefore, the Friend Reciprocity hypothesis (H4) is supported.

The fifth hypothesis predicted a triadic closure effect within the network, where past communication with a common partner predicts future communication between previously unconnected residents. The significant positive coefficient for *Triadic Closure* suggests that having a common chat partner in the past predicts the emergence of a future chat tie – if *a* sent a chat to *b* who sent a chat to *c, a* is 24.1% more likely than baseline to send a chat to *c*. That is, the third chat is more likely, as a function of the first two. Therefore, the Triadic Closure hypothesis (H5) is supported.

The final hypothesis predicted that the triadic closure effect would be stronger among friends. We predicted that if two residents have a common friend with whom they chat, they are very likely to chat in the future. The significant positive result on *Friend Closure* supports this hypothesis, although the effect size is somewhat small. Accounting for the general tendency towards triadic closure in the network, triadic closure among friends is 6.1% more likely. Therefore, the Friend Closure hypothesis (H6) is supported.

Significant effects on the two Hub control variables (*Ego Hub* and *Alter Hub*) suggest that there are certain hubs in the network that broadcast information without otherwise engaging in the chat patterns tested in the hypothesis tests. Including this variable in the model controls for the effect of these hubs, and allows for a more accurate examination of the remaining hypothesized patterns.

DISCUSSION

In this article, we advocate for the use of relational event network modeling for studying online communication networks. Unlike other network analytic techniques that focus on static networks, relational event network modeling allows researchers to model dynamic network processes. Relational event network analysis leverages the detailed event data available for many online communication networks and integrates both time-ordering and the conditional interdependence of networked communication in order to test hypotheses about how communication processes affect network structure over time.

We demonstrated one application of relational event network modeling using an online chat network sampled from the virtual world *Second Life.* Our case study focused on unpacking the processes that lead to a disproportionate amount of communication among friends online. We discovered that inertia, reciprocity, and triadic closure all work in concert to generate more communication among friends than nonfriends. Among these, the strongest effect was reciprocity.

This suggests that increased communication among friends does not stem from people *starting* more conversations with friends than nonfriends (indeed, we observed a dampening effect on inertia among friends over time). Instead, the overall volume of communication increases because friends are more likely than nonfriends to *reciprocate* messages, leading conversations to emerge and resulting in more overall communication between friends than between nonfriends.

This finding reinforces the relational value of reciprocal communication among friends noted elsewhere in the CMC literature (Gilbert & Karahalios, 2009; Peter, Valkenburg, & Schouten, 2005). It also begins to paint a picture of the communication processes associated with the emergence of friendship online. Before friendship is established, communication tends to be rather one-sided. People send many messages out, occasionally reciprocating or closing triads, but more often sending repeat messages to the same person, perhaps in the hopes of eliciting a response. When friendship exists, conversations are the norm; people tend not to follow up with unresponsive friends, instead focusing communication on responsive partners, or less often, in responsive groups. Although we currently have only between-subjects evidence of these different patterns, future studies using relational event network analysis to examine the evolution of communication networks as friendships emerge could lend further evidence to this claim.

Although preliminary, these time-conditioned insights represent important advances in understanding the processes that generate online communication networks. Historically, CMC research has focused on stable network structures and used inferential statistics to explain how those network structures may have emerged. In contrast, relational event network modeling directly observes the processes that generate stable network structures, and distinguishes between the relative impact of different processes that lead to the to the same overall network structures. Doing so allows researchers to move beyond discussing how CMC networks *are,* and instead focus on how they *came to be.*

This opens up a wealth of new research opportunities for CMC researchers. For example, rather than observing that there are certain influential nodes in a CMC network, relational event network analysis allows us to determine how those nodes became influential. Instead of noting that there are clusters in a network, relational event network analysis allows us to determine why they formed. As demonstrated in this paper, rather than identifying different levels of communication between different types of users, relational event network analysis also allows us to identify the patterns that generated these differences.

Naturally, as with any exploratory research, the results of our case study should be interpreted with caution. *Second Life,* although suitable for this methodological demonstration, is one virtual world with a relatively unique culture and set of norms around communication. Generalization of the communication patterns observed in this case study to other online communities and/or offline interactions would not be appropriate. Further, the present analysis is necessarily limited by the type of data we had about *Second Life.* Notably, while we can draw conclusions about how patterns of communication evolve within the network, without records of content exchanged, we do not offer insight into how specific communication messages, topics, themes, etc. influence network evolution.

Nevertheless, we hope that computer-mediated communication researchers can leverage the examples provided here to more precisely and accurately deal with the wealth of information available about online communication. In the future, we plan to analyze additional network samples using the same techniques, and to further investigate the interdependence of chatting and friendship. Specifically, we plan to examine how chat patterns vary before and after the emergence of friendship ties to further clarify the preliminary results noted above. These and other relational event network analyses offer great potential to yield insights about online communication patterns never before accessible using alternate network analytic or sociometric techniques.

ACKNOWLEDGMENTS

The authors would like to thank Linden Labs for providing the data used in this research.

FUNDING

This research was supported by grants from the National Science Foundation (IIS-084-1583) and the Army Research Institute (W91WAW-08-C-0106). We gratefully acknowledge all our sponsors. The findings presented here do not in any way represent, either directly or through implication, the policies of these organizations.

REFERENCES

Ahn, Y.-Y., Han, S., Kwak, H., Moon, S., & Jeong, H. (2007). *Analysis of topological characteristics of huge online social networking services*. Paper presented at the ACM 16th International Conference on World Wide Web, Banff, Canada.

Amichai-Hamburger, Y., & Vinitzky, G. (2010). Social network use and personality. *Computers in Human Behavior, 26*(6), 1289–1295. doi:http://dx.doi.org/10.1016/j.chb.2010.03.018

Arnaboldi, V., Conti, M., Passarella, A., & Pezzoni, F. (2012). *Analysis of ego network structure in online social networks*. Paper presented at the Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on on Social Computing (SocialCom), Amsterdam, The Netherlands.

Berger, C. R., & Calabrese, R. J. (1975). Some expolrations in initial interaction and beyond: Toward a developmental theory of interpersonal communication *Human Communication Research*, *1*(2), 99–112. doi:10.1111/j.1468-2958.1975.tb00258.x

Boellstorff, T. (2008). *Coming of age in second life*. Princeton, NJ: Princeton University Press.

boyd, d. m. (2007). Why youth (heart) social network sites: The role of networked publics in teenage social life. In D. Buckingham (Ed.), *MacArthur foundation series on digital learning: Youth, identity and digital media volume*. Cambridge, MA: MIT Press.

Brandes, U., Lerner, J., & Snijders, T. A. B. (2009, July 20–22). *Networks evolving step by step: Statistical analysis of dyadic event data*. Paper presented at the IEEE International Conference on Advances in Social Network Analysis and Data Mining, Athens, Greece.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: A practical information-theoretic approach*: New York, NY: Springer.

Burt, R. S. (2005). *Brokerage and closure: An introduction to social capital*. New York, NY: Oxford University Press.

Burt, R. S. (2011). *Structural holes in virtual worlds* (Working Paper). Chicago, IL: University of Chicago.

Butts, C. T. (2008). A relational event framework for social action. *Sociological Methodology*, *38*(1), 155–200. doi:10.1111/j.1467-9531.2008.00203.x

Cartwright, D., & Harary, F. (1956). Structural balance: A generalization of Heider's theory. *Psychological Review*, *63*(5), 277–293.

de Nood, D., & Attema, J. (2006). *Second life, the second life of virtual reality*. The Hague, the Netherlands: EPN–Electronic Highway Platform.

de Nooy, W. (2008). Signs over time: Statistical and visual analysis of a longitudinal signed network. *Journal of Social Structure*, *9*(1). Retrieved from http://www.cmu.edu/joss/content/articles/volume9/DeNooy/

Derlega, V. J., Wilson, M., & Chaikin, A. L. (1976). Friendship and disclosure reciprocity. *Journal of Personality and Social Psychology*, *34*(4), 578–582.

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends": Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, *12*(4), 1143–1168.

Epstein, J. L. (1983). Examining theories of adolescent friendships. In J. L. Epsten (Ed.), *Friends in school: Patterns of selection and influence in secondary schools* (pp. 39–62). New York, NY: Academic Press.

Espelage, D. L., Green, H. D., Jr., & Wasserman, S. (2007, Winter). Statistical analysis of friendship patterns and bullying behavior among youth. *New Directions for Child and Adolescent Development*, *118*, 61–76. doi:10.1002/ed.201

Foucault Welles, B., Rousse, T., Merrill, N., & Contractor, N. (2014). Virtually friends: An exploration of friendship claims and expectations in immersive virtual worlds. *Journal of Virtual Worlds Research*, *7*(2), article 7024.

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

Gilbert, E., & Karahalios, K. (2009). *Predicting tie strength with social media*. Paper presented at the Proceedings of the 27th International Conference on Human Factors in Computing Systems, Boston, MA.

Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, *25*(2), 161–178.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Granovetter, M. S. (1983). The strength of weak ties: A network theory revisited. *Sociological Theory*, *1*, 201–233.

Hallinan, M. T. (1978). The process of friendship formation. *Social Networks*, *1*(2), 193–210. doi:10.1016/0378-8733(78)90019-9

Hannan, M. T., & Freeman, J. (1977). The population ecology of organizations. *American Journal of Sociology*, *82*(5), 929–964.

Hanneke, S., Fu, W., & Xing, E. P. (2010). Discrete temporal models of social networks. *Electronic Journal of Statistics*, *4*, 585–605.

Haythornthwaite, C. (1996). Social network analysis: An approach and technique for the study of information exchange. *Library & Information Science Research*, *18*(4), 323–342.

Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: John Wiley.

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, *24*(3), nihpa54860.

Hurme, T.-R., Veermans, K., Palonen, T., & Järvelä, S. (2008). Exploring changes in network structures during online discussions. In *Proceedings of the 8th International Conference for the Learning Sciences* (Volume 1, pp. 382–389).

Keegan, B., Gergle, D., & Contractor, N. (2012). *Do editors or articles drive collaboration? Multilevel statistical network analysis of wikipedia coauthorship*. Paper presented at the ACM Conference on Computer Supported Cooperative Work (CSCW2012), Seattle, WA.

Kossinets, G., & Watts, D. J. (2006). Empirical analysis of an evolving social network. *Science*, *311*(57), 88–90. doi:10.1126/science.1116869

Krackhardt, D. (1992). The strength of strong ties: The importance of Philos in organizations. In N. Nohrida & R. G. Eccles (Eds.), *Networks and organizations: Structure, form and action* (pp. 216–239). Boston, MA: Harvard Business School Press.

Krivitsky, P. N., & Handcock, M. S. (2014). A separable model for dynamic networks. *JRSS Series B-Statistical Methodology*, *76*(1), 29–46. doi:10.1111/rssb.12014 JAN 2014

Kumar, R., Novak, J., & Tompkins, A. (2010). Structure and evolution of online social networks. In Yu, P. S., Han, J., & Faloutsos, C. (Eds.) *Link mining: Models, algorithms and applications*. (pp. 337–357) New York, NY: Springer.

Lazer, D., Pentland, A. S., Adamic, L. A., Aral, S., Barabasi, A. L., Brewer, D., . . . van Alstyne, M. (2009). Life in the network: The coming age of computational social science. *Science*, *323*(5915), 721–723.

Leskovec, J., Huttenlocher, D., & Kleinberg, J. (2010). *Signed networks in social media*. Paper presented at the 28th International Conference on Human Factors in Computing Systems (CHI2010), Atlanta, GA.

Lewis, K. (2011). The co-evolution of social network ties and oline privay behavior. In S. Trepte & L. Reinecke (Eds.), *Privacy online* (pp. 91–110). Berlin, Germany: Springer-Verlag.

Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., & Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using Facebook.com. *Social Networks*, *30*(4), 330–342. doi:10.1016/j.socnet.2008.07.002

Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P., & Bhattacharjee, B. (2007). *Measurement and analysis of online social networks*. Paper presented at the Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement, San Diego, CA.

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

Peter, J., Valkenburg, P. M., & Schouten, A. (2005). Developing a model of adolescent friendship formation on the Internet. *CyberPsychology and Behavior*, *8*(5), 423–430.

Ratan, R., Chung, J.-E., Shen, C., Poole, M. S., & Williams, D. (2010). Schmoozing and smiting: Trust, social institutions and communication patterns in an MMOG. *Journal of Computer Mediated Communication*, *16*(1), 93–114.

Roloff, M. E. (1985). Communication and reciprocity within intimate relationships. In M. E. Roloff & G. R. Miller (Eds.), *Interpersonal processes: New directions in communication research* (vol. 14, pp. 11–38). Thousand Oaks, CA: Sage.

Rosen, D., Barnett, G. A., & Kim, J. H. (2011). Social networks and online environments: When science and practice co-evolve. *Social Network Analysis and Mining*, *1*(1), 27–42.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–725.

Shen, C., Monge, P., & Williams, D. (2012). Virtual brokerage and closure: Network structure and social capital in a massively multiplayer online game. *Communication Research*. doi:10.1177/0093650212455197

Shumate, M., & Palazzolo, E. T. (2010). Exponential random graph (p*) models as a method for social network analysis in communication research. *Communication Methods and Measures*, *4*(4), 341–371. doi:10.1080/19312458.2010.527869

Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. *Sociological Methodology*, *31*(1), 361–395.

Stefanone, M. A., Kwon, K., & Lackaff, D. (2011). The value of online friends: Networked resources via social network sites. *First Monday*, *16*(2). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/3314/2763

Steinkuehler, C., & Williams, D. (2006). Where everybody knows your (screen) name: Online games as “third places.” *Journal of Computer Mediated Communication, 11*(4), article 1.

Thelwall, M. (2008). Social networks, gender, and friending: An analysis of MySpace member profiles. *Journal of the American Society for Information Science and Technology*, *59*(8), 1321–1330.

Wang, P., Robins, G., & Pattison, P. (2006). LPNet: A program for the simulation and estimation of longitudinal exponential random graph models. *University of Melbourne*.

Wheeless, L. R. (1976). Self-disclosure and interpersonal solidarity: Measurement, validation and relationships. *Human Communication Research*, *3*(1), 47–61.