Article

Organizational Research Methods
1-24

DOI: 10.1177/1094428118825300
journals.sagepub.com/home/orm
SAGE

Stochastic Actor-Oriented Models for the Co-Evolution of Networks and Behavior: An Introduction and Tutorial

Yuval Kalish[1]

Abstract

Stochastic actor-oriented (SAO) models are a family of models for network dynamics that enable researchers to test multiple, often competing explanations for network change and estimate the extent and relative power of various influences on network evolution. SAO models for the co-evolution of network ties and actor behavior, the most comprehensive category of SAO models, examine how networks and actor attributes—their behavior, performance, or attitudes—influence each other over time. While these models have been widely used in the social sciences, and particularly in educational settings, their use in organizational scholarship has been extremely limited. This paper provides a layperson introduction to SAO models for the co-evolution of networks and behavior and the types of research questions they can address. The models and their underpinnings are explained in nonmathematical terms, and theoretical explanations are supported by a concrete, detailed example that includes step-by-step model building and hypothesis testing, alongside an R script.

Keywords

stochastic actor-oriented models, longitudinal models, social network analysis, RSiena

Researchers across diverse domains of science are increasingly considering social phenomena from a network perspective. This trend has manifested in a surge in research on social networks; in the Web of Science, for example, the number of studies on the topic of "social networks" has nearly tripled over the past decade (Borgatti, Mehra, Brass, & Labianca, 2009). Research in this vein is based on the notion that *actors* (individuals, teams, organizations, countries, etc.) are embedded in *ties* (social relations) and that these ties are important for individual and group outcomes because they constrain and provide actors with opportunities for action (Kilduff & Brass, 2010; Kilduff & Tsai, 2003).

[1]Coller School of Management, Tel Aviv University, Tel Aviv, Israel

Corresponding Author:
Yuval Kalish, Coller School of Management, Tel Aviv University, Tel Aviv 69978, Israel.
Email: yuvalk@tauex.tau.ac.il

Actors possess *attributes*—characteristics, attitudes, perceptions, personality traits, emotions, behaviors, and values—that influence how they "structure their ties," that is, select other actors with whom to connect (or from whom to disconnect; Emirbayer, 1997; Kalish & Robins, 2006; Klein, Lim, Saltz, & Mayer, 2004). The structuring of an actor's ties in turn shapes that actor's behavior, attitudes, perceptions, and emotions. For example, in a given team of individuals, team members who possess a given attribute—for example, strong perceptions of being psychologically safe—may be more likely than others to form friendships (ties) with other team members, and these friendships in turn may further enhance their perceptions of psychological safety (Schulte, Cohen, & Klein, 2012). Clearly, there is a temporal dimension inherent to this example: The attribute influences the tie structure, which later influences the attribute. Yet, traditional network analytic techniques cannot adequately accommodate the passage of time or model the reciprocal influence of attributes on network ties and vice versa. Accordingly, they are limited in their capacity to shed light on the mechanisms underlying organizational phenomena, which are dynamic by nature. It is necessary to overcome this limitation to bridge important knowledge gaps in organizational research. For example, because resources flow through ties and ties facilitate sense-making (cf. Kalish, Luria, Toker, & Westman, 2015), the capacity to link attributes and ties may provide the micro-macro link in organizational scholarship (Kilduff & Tsai, 2003) by explaining the emergence of various team characteristics such as shared perceptions of climate, leadership quality, leadership roles, emotional tone, group identity, and attitudes toward turnover (see also Lang, Bliese, & de Voogt, 2018).

The current article presents a network analysis approach that addresses the limitations outlined previously: stochastic actor-oriented (SAO) models for the co-evolution of networks and behavior. This family of models links changing individual attributes with evolving network structures and can thereby elucidate a broad spectrum of organizational phenomena—including but not limited to emergence of leadership and status hierarchies, development of shared perceptions, onboarding and socialization, emergence of communities, collective turnover, emergence of team climate, and team emotional tone. While SAO models have gained traction in the fields of sociology and education, their use in organizational contexts has been limited: As of 2018, of over 650 articles referring to SAO models, less than 6% examine inter- or intraorganizational contexts (compared with 39% in educational contexts). Herein, I aim to provide a nonmathematical, hands-on introduction to SAO models for the co-evolution of networks and behavior, oriented toward organizational researchers who do not typically work with these types of models. I further present and develop examples of research questions that these models can be used to address. Accordingly, this article complements prior works that provide a more statistical, mathematical, or methodological introduction to SAO models (Snijders, 1996, 2001, 2008; Snijders, Steglich, & Schweinberger, 2007; Snijders, van de Bunt, & Steglich, 2010; Veenstra & Steglich, 2012).

Stochastic Actor-Oriented Models: An Overview

SAO models are a family of models that express empirically observed changes in network ties and in many cases, changes in individual attributes as time-aggregated outcomes of a series of individual decisions (Steglich, Snijders, & Pearson, 2010). Consider the emergence of shared climate, thought to be the result of interactions between individuals (Lang et al., 2018). A situation in which interacting partners agree with one another with regard to climate might emerge because actors decide to change their own perceptions of climate after interacting with others with different perceptions (they "catch" the others' perceptions) or because members stop interacting with others who hold climate perceptions that differ from their own (they "de-select" others because of their different perceptions of climate). Each of these theoretical explanations can explain how interaction patterns are related to individual attributes dynamically and can therefore bridge the theoretical gap between individual perceptions and the emergence of higher level constructs. Unfortunately, most existing techniques cannot unpack these different explanations. SAO models overcome this limitation by (statistically) modeling individual decisions of actors. They are "actor-oriented" in the sense that they model change over time from the perspective of each actor, under the assumption that each actor has the option to adjust the structure of his of her network (by "deciding" whom to start, continue, or stop communicating with) or the level of a specific attribute (whether to increase or decrease the extent to which he or she perceives the team's climate as being positive). The decisions of a focal actor (referred to as *ego*) are assumed to be based on the structure of the network at the time of the decision as well as on ego's own attributes and the attributes of other actors (referred to as *alters*).

SAO models enable researchers to use data from a first measured timepoint to test whether a given set of hypothesized effects can produce the network structures and attribute levels measured at later timepoints. Central to these models is the idea that changes in network ties and actors' attributes occur continuously even though data on the state of the network and its actors are collected at discrete timepoints.[1] These models also assume that the difference between observed timepoints can be broken down into probabilistic, sequential small steps, called *ministeps*. At each ministep, a focal actor is randomly selected and has the opportunity to make a single decision. In a *network ministep*, the actor has the opportunity to modify one of his or her outgoing ties by creating a tie to a new actor, terminating an existing tie, or maintaining current ties. In an *attribute* (or behavioral) *ministep*, the actor can modify (increase, decrease, or maintain) his or her level of a given attribute.

To model these changes, the researcher constructs, on the basis of theoretical and empirical considerations, a set of "rules" that might drive an actor's decision to change a network tie or adjust the level of one of his or her attributes. Rules represent four broad categories: *Network evolution rules* determine how ties develop given the structure of ties in the previous timepoints. An example might be that individuals prefer to communicate with those who communicated with them in previous timepoints or that individuals prefer to communicate with those with whom many others communicated in previous timepoints. *Attribute evolution rules* determine how attributes evolve over time. An example might be that there is a general tendency for the attribute (negative affectivity, turnover intention, identification with a program) to increase over time. *Social selection rules* describe how ties develop in response to actors' attributes (and ties) in the previous timepoint. For these rules, actors' attributes are considered to be the drivers of network ties. An example might be that people who have higher levels of negative affectivity talk to fewer others over time or that people prefer to communicate with others whose levels of turnover intentions resemble their own. Finally, *social influence rules* describe how actors' attributes change in response to ties (and attributes) in the previous timepoint. For these rules, the network is considered to drive actors' attributes. For example, communicating with fewer others over time may increase a person's turnover intentions, or people may catch their network partners' levels of negative affectivity over time (emotional contagion).

The model estimates parameter values associated with each rule (called *effects* in SAO terminology) such that it identifies parameter values that could have brought the collected panel observations to follow one another. These parameter estimates can be used to test competing theoretical explanations of social processes underpinning network (and attribute) evolution. Importantly, the fact that actors' decisions are assumed to be based on a set of rules—given the current state of the network coupled with those actors' (and their alters') attribute levels—enables researchers to make strong (statistical) inferences regarding temporal causality in the relationships between network structures and attributes (Steglich et al., 2010), thereby allowing for a robust statistical test of emergence of specific phenomena over time (Schulte et al., 2012).

SAO models can be applied to diverse data structures of different degrees of complexity. In the current paper, I focus on SAO models for the co-evolution of network ties and behavior (Snijders et al., 2007, 2010)—models that assume that both networks and actor attributes change over time and influence each other. As such, these models enable researchers to unpack and model a wide range of theoretical questions, such as what drives the emergence of micro-climates within a team, emergence of consensus over time, leadership emergence, and emotional contagion. Table 1 presents some of the organizational research questions and specific hypotheses that have been addressed using SAO models for the co-evolution of networks and behavior. In what follows, for convenience, I refer to this family of models simply as SAO models. Readers who are interested in learning more about additional categories of stochastic actor-based models should refer to Tom Snijders's earlier work (e.g., Snijders, 1996, 2001, 2008; Snijders et al., 2007) and the excellent RSiena manual (Ripley, Snijders, Boda, Voros, & Preciado, 2018a).

Decision Frequency and Decision Rules: Rate and Objective Functions

As outlined previously, SAO models assume that whereas network and attribute data are collected at discrete timepoints, underlying time is continuous. At discrete, unobservable timepoints, a randomly selected actor gets an opportunity to change his or her outgoing ties (a network ministep) or level of the attribute (a behavioral ministep). This implies two separate subprocesses that need to be modeled. The first subprocess involves modeling the frequency of change: how often a ministep occurs. The *rate function* models expected opportunities for change on the network and on an attribute between every two consecutive measured timepoints. The second subprocess models how change occurs: what changes occur once an actor is given an opportunity to change his or her network or attribute (in a network or behavioral ministep, respectively). These changes are expressed by two interdependent mathematical functions termed *objective functions*, which model the course of action that the actor selects. The actor makes his or her decision by evaluating the outcome of each possible change (to his or her network or attribute, depending on the type of ministep) in terms of its effect on the value of his or her objective function and then choosing the optimal change (with a small amount of randomness). In a utility-based approach, the objective functions directly reflect the actors' utilities such that in making their decisions, actors seek to maximize the utility they derive.

Equation 1, f_i^{net}, denotes the objective function that actor i seeks to optimize in a network ministep. Equation 2, f_i^{beh}, represents the objective function that actor i seeks to optimize in an attribute (behavioral) ministep.

$$f_i^{net}(x,z) = \sum_k \beta_k^{net} s_{ik}^{net}(x,z) \tag{1}$$

$$f_i^{beh}(x,z) = \sum_k \beta_k^{beh} s_{ik}^{beh}(x,z) \tag{2}$$

In each equation, $f_i(x, z)$ denotes the value of the objective function for actor i for a given network state x and i's level of the attribute z. The value of f is dependent on a series of parameter values β_k, each of which is coupled to an effect, denoted s_{ik}. An effect represents a subgraph count in the network neighborhood of the focal actor (or a function of the attributes of actors sending or receiving ties) in a given ministep.[2] Mathematical specifications of frequently used effects are presented in Appendix B (available in the online version of the journal). The parameter value assigned to a given effect translates into probabilities for change in actor i's network and attributes and can be interpreted as the rules that govern change, or more simply as the attractiveness of the change—its cost or benefit to the actor.[3] Parameter values are estimated by an estimation technique called *method of moments* (Snijders, 2001). The outcome of this process is a set of parameter values (and standard errors) associated with effects that link evolving network ties and actor attributes and yield outcomes that are most consistent with the series of panel measurements.

Table 1. Example of Main Constructs, Exemplary Hypotheses, and Effects Used to Model the Co-Evolution of Networks and Attributes in Previous Research.

	Main Constructs		Effects Included in the Model												Hypotheses					
			Net. Func. Default		Net. Func Cluster		Net. Func. Degree Distribution			Net. Func. Attributes			Beh. Func. Default		Net. Func. – Attribute-Based Hypotheses (and Associated Effects)			Beh. Func. – Attribute Based Hypotheses (and Associated Effects)		
The Study	Network relationship	Main attribute	Outdegree	Reciprocity	TranTrip/GWESPP	3-cycle	InPop	OutPop	isolates	egoX	altX	simX	sahpe	Shape-squared	Network function – ego effect (egoX)	Network objective function – alter effect (altX)	Network objective function – similarity effect (simX)	Behavioral objective function – activity	Behavioral objective function – popularity	Behavioral objective function – average/total similarity ("contagion")
Checkley, Steglich, Angwin, and Endersby (2014)	Cooperation	Firm Performance	X	X	X		X			X	X	X	X	X	Increasing performance influences connectedness		High performers prefer to network with high performers	Increasing connectedness influences performance		Performance is contagious
Fransen, Delvaux, Mesquita, and Van Puyenbroeck (2018)	Leadership	Warmth/ competence/ gender	X	X	X	X	X	X		X	X	X	X	X		Warm individuals are perceived more as leaders				
Kalish and Luria (2016)	Communication	Perceived stress	X	X	X		X	X		X	X	X	X	X	The higher the stress, the fewer communication ties	People communicate less with others who have high level of stress	People communicate with others with similar levels of stress	The more communication ties, the lower the stress		Stress is contagious
de Klepper, Sleebos, van de Bunt, and Agneessens (2010)	Friendship	Military discipline	X	X	X					X	X	X	X	X						Students adjust their discipline to be more similar to their friends
de Kleppers, Labianca, Sleebos, and Agneesens (2017)	Control attempts	Status	X	X	X	X			X	X	X	X	X	X	Higher status leads to controlling more others	People do not attempt to control higher status others				
Schulte, Cohen, and Klein(2012)	Advice	Psychological safety	X	X	X					X	X	X	X	X	The more psychologically safe, the individual, the more friendships		Persons send friendships to others with similar levels of psych safety	The more friendships sent, the more safe the person is	The more friendships received, the higher safety climate perception	Over time, a person's psychological safety will be more similar to their friends
Schulte et al (2012)	Difficulty	Psychological safety	X	X	X					X	X	X	X	X	The more psychologically safe, the fewer difficult friendships	People with lower levels of psychological safety attract more difficult relationships from others		The more difficult relationships, the less safe the person is	The more difficult relationships received, the lower the psych. safety	

Note: Rate effects are not included in the table. Net. Func. = network objective fuction; Beh. Func = behavioral objective function.

Table 2. A Simplified Example of Data Structures Used for Stochastic Actor-Oriented Models.

Friendship Network Time 1					Friendship Network Time 2					Constant Attribute	Changing Attribute	
	A	B	C	D		A	B	C	D	Male	Negative Affectivity Time 1	Negative Affectivity Time 2
A	0	0	0	1	A	0	0	0	1	0	2	1
B	1	0	0	0	B	0	0	1	0	1	5	5
C	0	1	0	0	C	1	1	0	1	1	4	5
D	1	0	0	0	D	1	0	0	0	0	1	2

Basic Example

To better explain the data structure and the internal calculations of SAO models, this section presents a simple example concerning a network of friendship among individuals. Consider four actors, denoted A to D, who form a business unit. At two points in time, data were collected regarding the friendship relations among the actors (i.e., the network ties between them) and each actor's level of negative affectivity (NA), an attribute whose value is assumed to have the capacity to change over time. Data were also collected regarding a constant attribute, actors' gender. The two attributes (gender and NA) are referred to, in SAO terminology, as *covariates*, where NA is considered a changing covariate and gender is a constant covariate. Table 2 presents the information for each actor; a 1 in cell (i, j) indicates that actor i (ego) chose actor j (alter) as a friend at that observed timepoint.

We observe that at Time 1, Actor A, a female with relatively low levels of NA, nominated actor D, another female with relatively low levels of NA, as a friend and that this relationship was maintained at Time 2. Actor B, a male with high levels of NA, reported Actor A as a friend at Time 1, but this friendship tie was not maintained at Time 2. By Time 2, B had also created a new friendship tie, to Actor C (another male with relatively high levels of NA). Thus, by the end of Time 2, we observe that friendship ties reflect higher homophily on gender than they do at Time 1. One possible explanation for this outcome is that individuals tend to select friendships on the basis of similarity in gender. By the end of Time 2, we also observe that friendship ties reflect higher homophily in NA than they do in Time 1. What could be the mechanism underpinning this similarity in NA over time? It could be the case that people select as friends others who display similar levels of NA (a social selection effect similar to that observed in the case of gender). However, it is possible, for example, that people catch their friends' level of NA (a social influence explanation, also referred to as *emotional contagion*; Barsade, 2002). Note that both these processes reflect different rules by which over time, we observe emergent consensus in NA among people who are tied to each other.

An SAO model can unpack these two options by providing the net effect of each explanation in contributing to the observed future states of the network and the attribute. To achieve this goal, the model breaks the discrete network (friendship) measurements (Matrix 1 and 2) and changing attributes (NA1 and NA2) down into very small (unobserved) network and attribute ministeps, respectively.

Network ministep. In a network ministep, the SAO model randomly selects an actor. Let us assume the selected actor is Actor A. At Time 1, Actor A has a tie only with Actor D. Actor A now "considers" all her options: She can (a) create a new tie to actor B, (b) create a new tie to actor C, (c) drop her existing tie to D, or (d) do nothing. Her final decision is based on the value of the objective function (Equation 1) for each option. Let us assume for the sake of simplicity that the objective function comprises only two effects: (a) an outdegree effect, corresponding to the act of changing one's number of connections (establishing a new tie or terminating an existing tie), and (b) a reciprocity effect, corresponding to the act of reciprocating a tie or dropping an unreciprocated tie. Let us further assume that initial parameter values for those effects are automatically estimated by RSiena at –0.4 and +1.0, respectively.[4] The meaning of the negative parameter value for the outdegree effect is that each new tie "costs" the focal actor –0.4 (and dropping a tie entails a gain of +0.4), and the meaning of the positive parameter value for the reciprocity effect is that actors "gain" +1.0 by reciprocating ties sent to them (or dropping unreciprocated ties) or "lose" –1.0 by dropping a tie that was previously reciprocated.

Let us consider Actor A's choices again: If she creates a new tie to Actor B, the value of her objective function will be $-0.4 + 1.0 = 0.6$ because she has created a new tie to an alter (B), and that new tie is now reciprocal (B already has a tie to A). The value of a new tie to Actor C will be –0.4 because A will have created a new (nonreciprocated) tie. The termination of a tie to Actor D will yield the objective function value of $+0.4 + -1.0$ because A "frees" a tie yet damages her reciprocity level. Finally, doing nothing (maintaining her tie to D) will elicit a value of 0. Each of these values translates into a probability according to the exponential transformation, $e^{0.6}$, $e^{-0.4}$, $e^{-0.6}$, $e^{0} = 1.82$, 0.67, 0.54, and 1, for each of the four options, respectively. This suggests that at this network ministep, and given the parameter values of –0.4 and +1 for outdegree and reciprocity, respectively, Actor A is most likely to create a new tie to B.

In the next network ministep, the model selects another actor, and that actor makes a decision according to the same process. The model continues to randomly select actors in this way such that each actor has, on average, the number of ministeps as estimated by the rate function; however, for the sake of simplicity, let us assume that our model stops after two ministeps. In the second ministep of our example, the model selects Actor C, who, given the parameter values for the objective function, also decides to sever his existing tie to Actor B. Thus, at the end of these two ministeps, Actor A has created a tie to Actor B, and Actor C has terminated his tie to Actor B. Yet, in the matrix on the right (Table 2, representing the next measured timepoint), this is clearly not the case since Actor A has no tie to Actor B and Actor C has maintained his tie to Actor B.

As a result, parameter values are recalibrated so that the actors' actions better fit the observed outcomes at the next timepoint. Estimation of the parameter values for the objective functions takes place in an iterative process using the method of moments estimation technique (Snijders, 2001), as noted previously, such that the final parameter values are the ones that best describe the decision rules with which the network progresses from the matrix on the left to that on the right in Table 2.

Adding effects that are dependent on actor attributes to the objective function (Equation 1) works in a similar way. Each actor attribute receives an initial weight and is weighted in the (network) utility function. Let us assume, for example, that initial parameter estimates for females are –0.1, +0.2, and +0.3 for each additional tie sent to anyone, each additional tie received, and for forming a tie to an actor of similar gender, respectively.[5] This suggests that for females (compared with males), there is an additional cost associated with forming a new friendship tie (–0.1), females gain more from receiving friendship ties (from either males or females, +0.2), and all actors (male or female) prefer to send ties to others of similar gender (+0.3). The (network) objective function (Equation 1) now contains five effects associated with initial parameter values (β): outdegree (–0.4), reciprocity (+1.0), sender effect for females (–0.1), receiver effect for females (+0.2), and similarity effect on gender (tie sent from females to females or from males to males, +0.3).

Let us now examine the value of the (network) objective function for the focal actor, D, a female, forming a new friendship tie to either Actor C, a male, and compare it to the value of the objective function for her forming a new friendship tie to Actor C*, a female who has exactly the same network relationships as Actor C. The value of the objective function for any actor forming a new tie is –0.4 (outdegree). Since D is a female, there is an additional cost associated with her forming new ties (–0.1, the sender effect for females). Since the gender similarity effect is weighted as +0.3, everything else being equal, Actor D is 1.35 ($e^{0.3} = 1.35$) more likely to form a tie to Actor C* (a fellow female) compared with Actor C (a male).

Attribute ministep. A similar process operates for the attribute ministeps (modeled by the behavioral objective function, Equation 2). Recall that the attribute ministep models how attributes change over time due to previous levels of the attribute and network ties. Let us assume that the randomly selected actor is Actor A and that Equation 2 only includes the *average alter* effect, an effect that captures the average level of the attribute among the people ego is connected to (in our case, the average level of negative affectivity among ego's connections).[6] Let us further assume that the initial parameter value for the average alter effect is +1.1. The effect coupled with the parameter value means that when comparing two actors who are otherwise equal in all respects, if the friends of the first actor are (on average) 1 point higher on NA compared with the friends of the second actor, the first actor is three times ($e^{1.1} = 3.0$) more likely than the second actor to increase his or her level of NA (rather than maintain it). In other words, given an initial parameter of (+1.1), we would expect actors to strive to decrease the difference in the level of the attribute between themselves and the average level of their alters.

Let us now consider Actor A, whose initial level of NA is 2. Recall that at Time 1, A is connected only to Actor D, whose initial NA level is 1. With an initial average similarity parameter value (+1.1), the implication is that Actor A will tend to decrease her level of NA over time (to move closer to her average alter's level of NA), which is precisely what happens. Another actor is then randomly selected (say, Actor C with an initial NA level of 4 and an average alter level of 5). Actor C is also expected to adjust his attribute level to match the average level of his alters over time; indeed, we observe that Actor C shifts to an NA level of 5 at Time 2. These observations support the initial choice of parameter value for the average alter effect.

Overview of SAO Model Implementation

RSiena (Simulation Investigation for Empirical Network Analysis as implemented in the R environment; Ripley et al., 2018b) is a program that implements SAO models in the R environment. In what follows, we provide an overview of the steps involved in specifying and implementing an SAO model in RSiena.

An RSiena script contains three parts: The first part prepares data structures, the second specifies effects for the model, and the third tests for goodness of fit (GOF) of the obtained model. The heart of RSiena is in the second part, the modeling of the co-evolution of networks and attributes. An SAO model can incorporate diverse types of effects, and the researcher must select which ones to include in his or her model in accordance with the characteristics of the network and the question the researcher is studying. Modeling the co-evolution of networks and attributes may seem like an ominous task. Table 1 presents the effects used to model different networks and attributes according to the objective function that was modeled (network or behavioral) and the purpose of the effect. It also includes examples of specific research hypotheses presented in the corresponding papers and the effects used to directly test those hypotheses. An examination of Table 1 reveals that it is possible to model diverse networks using similar effects.

Selection of effects is typically done in a few iterative steps, described (and exemplified) in the following. The first step is to adequately model the (network) objective function without accounting for attributes. Effects incorporated in this step are considered necessary controls because they model the temporal evolution of the network independently of actor attributes. They pertain to default effects, effects that model clusterability, and effects that model the degree distribution (Table 1). The second step is to add actor attributes to the network (objective) function of the model. Three effects are typically added, some of which may already relate to specific hypotheses the researcher has (these effects are elaborated in Table 1, "Network Function–Attributes" columns). Once the network evolution part is modeled adequately, including actor attributes, researchers start modeling the second (behavioral) objective function. Two effects are incorporated automatically, and additional effects are specified by the researcher and typically pertain to additional research hypotheses (see Table 1, "Behavioral Function–Attributed-Based Effects" columns for hypotheses and the effects that model them). At each step, the researcher examines model adequacy using convergence and GOF (as described in the following). Once an adequate model is found, the researcher interprets the results.

Modeling Steps in RSiena: An Empirical Example

In this section, I present an illustrative worked out example, including all the steps necessary for successfully modeling the example. I first present the context, walk the reader through the steps that are necessary to adequately model the data, and provide the relevant RSiena script.

Context

A researcher is interested in examining collective turnover—the tendency of similarly dissatisfied employees to leave their organizations together (Feeley, Hwang, & Barnett, 2008; Felps et al., 2009; Krackhardt & Porter, 1986). The researcher posits, following Bartunek, Huang, and Walsh (2008), that friendships play a key role in collective turnover because they provide a route by which sense-making occurs. Specifically, the researcher hypothesizes that because friends discuss their dissatisfactions, ideas, and goals with each other, turnover intention may be contagious (social influence explanation, Hypothesis 1), that people befriend one another because they hold similar attitudes towards turnover (social selection explanation, Hypothesis 2), and that over time, people who are thinking of quitting their job select fewer friends (Hypothesis 3) and are chosen by fewer others as friends (Hypothesis 4). Note that Hypothesis 1 assumes that network structure drives the level of the attribute and is therefore modeled by effects in the behavioral objective function, whereas Hypotheses 2, 3, and 4 suggest that the level of the attribute drives the network structure and are therefore modeled by effects in the network objective function.

To test the hypotheses, the researcher examines the friendship networks and turnover intentions among individuals in a newly formed team. Participants included 32 respondents, 18 of whom were male. Questionnaires were administered at three points in time during the first year of the team's activity: 1 month, 7 months, and 11 months post-formation. At each timepoint, participants completed a questionnaire asking which other participants they considered to be friends (network questionnaire) and a questionnaire examining their turnover intentions (attribute questionnaire). Specifically, in the network questionnaire, each participant was given a list of all others in the team and was requested to identify those whom they considered as friends. In the attribute questionnaire, participants completed a single item taken from the turnover intention scale (Bothma & Roodt, 2013): "How often have you considered leaving your job in the previous month?" Reponses were on a 7-point Likert scale ranging from *never* to *all the time*. All participants agreed to complete the questionnaires. Appendix A (available in the online version of the journal) provides the complete RSiena script (see also Ripley et al., 2018a, for a more comprehensive illustrative script). Lines in Appendix A (available in the online version of the journal) and in the excerpts that follow are numbered (in brackets). Supplement 1 (available in the online version of the journal) provides the output from the script. Lines and excerpts are numbered (in brackets, with the prefix *r*). Supplement 2 (available in the online version of the journal) provides the (network and attribute) data used for the example.

Step 1. Data Preparation

The first step in using RSiena involves data collection, data entry, creation of R objects, and an examination of the suitability of the data to modeling using SAO models. Data collection for network analytic studies is beyond the scope of this manuscript. Briefly, however, researchers must first identify a relevant population and conduct a census using the full network method. Two types of data need to be collected: the relationships between all actors in the population (collected either with a survey or using available information such as email exchanges or attendance at meetings) and actor attributes (again, either collected via self-completed questionnaires or through existing data). Interested readers are referred to Robins's (2015) comprehensive book.

Step 1a. Data preparation and data considerations. Data preparation necessitates preparing text files that represent network data (i.e., a complete set of data on network ties among actors) and attribute data (attributes of each actor) observed for the same set of actors at several (at least three) discrete timepoints. Networks should be binary (i.e., either a tie exists, or it does not) and directed (i.e., a tie between actor *i* and *j* does not necessitate a tie between *j* and *i*; it is possible to model nondirected networks as well; see RSiena manual [Ripley et al., 2018a] for details) and of a reasonable size (ideally containing at least eight actors). The network data are represented as matrices, in which rows and columns represent actors and cell (*i*, *j*) is a binary indicator of the existence (or lack thereof) of a relationship between actors *i* and *j* at time *t*. It is possible to specify some ties that are impossible (*structural zeros*)[7] and deal with changes in the composition of the network over time (e.g., addition or deletion of nodes; see the RSiena manual [Ripley et al., 2018a] for more details).

Attribute data in RSiena may contain dummy (e.g., gender) or ordinal (e.g., age bracket, personality) covariates; attributes on an ordinal scale should be limited in terms of the number of values they can take (preferably less than 10 and ideally 2 to 5). A distinction is drawn between constant covariates (e.g., gender, personality traits) and varying covariates (e.g., level of stress, salary, turnover intention). Dyadic covariates, representing the attributes of pairs of actors (e.g., reporting lines), can also be accommodated but are not the focus of this article.

Step 1b. Creation of RSiena objects. After inputting all data files into RSiena, the user creates RSiena objects by specifying the role of each object. Each object is classified either as a (network or attribute) dependent variable (*SienaDependent*) or an independent variable; in the latter case, the object is further classified as either a constant (*coVar*) or changing (*varCovar*) covariate. An RSiena data object is formed by linking the structures with the *SienaDataCreate* command.

[1]	library(RSiena)	#calls the RSiena library in R
[2]	setwd("C:/Users/User/Desktop/ORM SAOM")	#identifies the working directory
[3]	friend1 <- as.matrix(read.table("MH-friend1.dat"))	#reads the data files (suffix.dat)
[4]	friend2 <- as.matrix(read.table("MH-friend2.dat"))	#into R
[5]	friend3 <- as.matrix(read.table("MH-friend3.dat"))	
[6]	quit <- as.matrix(read.table("attribute - quit.dat"))	#reads the turnover intention attribute file into R
[7]	gender <- as.matrix(read.table("attribute - gender.dat"))	#read the gender attribute into R.

(continued)

(continued)

[8] friendship <- sienaDependent(array(c(friend1, friend2, friend3), dim=c(32, 32, 3)))	# creates a dependent network object called "friendship" that specifies the ordering of the networks and their dimensions (32×32 people, 3 timepoints)
[9] gender<- coCovar(gender[,1])	#creates gender as a constant attribute
[10] quit<- sienaDependent (quit, type="behavior")	#creates *quit* as a changing covariate that serves as both an independent variable (driving the friendship network) and a dependent variable (the outcome of the network; the dual role of *quit* is specified by the sienaDependent, type="behavior" command, which also indicates to RSiena that the model is one for the co-evolution of networks and behavior
[11] MyData <- sienaDataCreate(friendship, quit, gender)	#creates the full RSiena object (MyData) by adding the networks and attribute objects

At the end of this step, we have created an RSiena data object, which contains the networks and attributes, such that both the networks (*friendship*) and the attribute (*quit*) serve as both independent and dependent variables in the temporal dynamics.

Step 1c. Create an initial effects object. Now that all data objects are specified, we need to create the initial effects object using the *getEffects* command. Two effects are so important for longitudinal network modeling that RSiena adds them automatically to any model; they do not need to be specified manually. These effects are the outdegree effect, which reflects changes in the number of ties in the network, and the reciprocity effect, which models the tendency to reciprocate ties sent in a previous timepoint.

[12] MyEff <- getEffects(MyData)	#creates an initial effects object, which includes outdegree and reciprocity (as well as the rate parameters and shape parameters) – additional effects will be added to this object later.

Step 1d. Print an initial report to check adequacy of running SAO model. While not necessary, it is good practice to print a preliminary report that provides some descriptive statistics for each timepoint and change between timepoints using the *print01Report* command. The main purpose of the printed report is to get a "feel" for the data and—more importantly—check (using the Jaccard coefficients) that enough change exists in the data to allow for longitudinal modeling.

[13] print01Report(MyData, MyEff, modelname = 'init')	#prints the initial report from the data object and given the effects object

The printed report (presented in full in Supplement 1, available in the online version of the journal; excerpts in the following appear with corresponding line numbers prefixed with *r*) first provides an analysis of the input data, including number of observations, number of actors, and the specification of the variables in the analysis (r4-r9). Next, the report describes the various objects, including the network objects (r19-r42) and the various attributes (r43-r75). For the turnover intention variable, *quit*, we observe that the range for responses is between 1 and 6 (r47), and there seems to be an increase in level of turnover intention over time, as exemplified by the increasing means across time (from 3.375 to 4.125; r52-r54).

```
[r44] Reading dependent actor variables.
[r45] ————————————————————————————————————————————————
[r46] 1st dependent actor variable named quit.
[r47] Maximum and minimum rounded values are 1 and 6.
<<snip>>
[r52] Means per observation:
[r53] observation  1      2      3      overall
[r54] quit         3.375  3.844  4.125  3.781
```

Next, the report provides some internal calculations used for centering of the variables (r67-r75). Finally, it provides change indicators for all variables across the observed timepoints (r90-r93). For example, between the first and second timepoints (r92), 29 new ties were formed, 45 existing ties were terminated, and 29 ties were maintained.

```
[r90] Tie changes between subsequent observations:
[r91] periods  0 => 0  0 =>1  1=> 0  1=>1  Distance  Jaccard  Missing
[r92] 1==>2    889     29     45     29    74        0.282    0 (0%)
[r93] 2 ==>3   900     34     29     29    63        0.315    0 (0%)
```

Of particular importance are the Jaccard coefficients, indicating the extent of change between observations. For the modeling of network evolution, a moderate level of change in both network and attributes is necessary; Jaccard coefficients ought to be above 0.2, ideally higher than 0.3, and lower than 0.7. Jaccard coefficients in the current example are adequate.

Given the aforementioned values, the software calculates initial parameter values for the outdegree parameter (r103), the rate parameter for each pair of consecutive timepoints (r101-102), and the changing covariate, *quit* (r104-r120), together with the initial value for the shape parameter (r125).

Step 2. Build Increasingly Complex Models Until Adequate Fit Is Achieved

As previously stated, modeling often starts with getting the network evolution part of the model to adequately fit the observed panel data. The researcher starts by adding "pure" network effects to the (network) objective function (Equation 1). Only then are attributes added to the network objective function, and in the last stage, effects are added to the behavioral objective function (Equation 2). As described previously, two effects are added automatically to the network objective function—the outdegree effect and the reciprocity effect (alongside the rate effects). The researcher then adds effects until an adequate fit is achieved in the following order: First, effects that model clustering in the network are usually added, and the model's fit is assessed. If adequate fit is not achieved, effects that model the degree distribution are added, and fit is assessed again. Finally, if adequate fit is still not achieved, effects that pertain to the correlation between in- and outdegree distribution are added. Typically, these three types of effects suffice to achieve adequate fit (Table 1). The RSiena manual (Ripley et al., 2018a) provides additional model fit troubleshooting instructions.

Step 2a. Run the model and check for t-ratios for convergence. Let us run this initial (very basic, overly simplistic) model, which includes only the default effects.

[14]	Model1 <- sienaModelCreate(useStdInits = TRUE, projname = 'Model1-results')	#creates a model object
[15]	Model1.results <- siena07(Model1, data=MyData, effects=MyEff, batch=FALSE, verbose=FALSE, returnDeps=TRUE)	#creates a results object that includes the results of the estimated model. Model is estimated using the *Siena07* command. The data and effects objects (as well as some default output requests) are specified.
[16]	Model1.results	#prints the results object

Table 3. Parameter Estimates, Standard Errors, and Convergence t Ratios for Model 1.

	Estimate	Standard Error	Convergence t Ratio
Network dynamics			
1. rate constant friendship rate (period 1)	5.7975	(1.1531)	0.0156
2. rate constant friendship rate (period 2)	4.6983	(1.0576)	0.0245
3. eval outdegree (density)	–2.2702	(0.1533)	–0.0133
4. eval reciprocity	2.5913	(0.2559)	–0.0343
Behavior dynamics			
5. rate quit (period 1)	1.8320	(0.5091)	0.0919
6. rate quit (period 2)	1.5399	(0.4808)	–0.0383
7. eval behavior quit linear shape	0.4544	(0.1748)	0.0381
8. eval behavior quit quadratic shape	–0.1665	(0.0874)	0.0625

Printing the *model1.results* object provides the output shown in Table 3. We first examine t ratios for convergence. The t ratios for convergence provide an indication that parameter estimates are stable, namely, that they converge across simulations. The t ratio (not to be confused with a t statistic, used for hypothesis testing) examines to what extent do estimated parameter values differ from parameter estimates simulated in (typically 1,000) simulated runs of network evolution. Ideally, we want the t ratio to be 0, indicating that simulated parameter values are exactly the same as estimated parameter values. Convergence is said to be excellent when the t ratios for convergence are lower than 0.1 (in absolute value) for all effects. The RSiena manual (Ripley et al., 2018a, Section 6.3) provides solutions on how to identify and troubleshoot convergence issues. The first (and easiest) solution is to add the argument *prevAns=Model1.results* to line [15] in the previous example. This tells RSiena to start estimation again, beginning with the previous estimation results. In our case (Table 3), all t ratios for convergence are smaller than $|0.1|$, indicating good convergence. For instructive purposes, I now turn to interpreting the output obtained for the initial model, noting that we are not yet certain that the model adequately fits the data. The following subsection, Step 3, elaborates on the interpretation of the output of the complete model, incorporating all relevant effects; this step is taken after both convergence is reached and GOF of the model is examined and found to be adequate.

The output of the initial model is divided into the two objective functions: the network objective function (top panel in Table 3, termed *network dynamics*) and behavior objective function (bottom panel in Table 3, termed *behavior dynamics*). We see the rate functions for both periods on both functions. On average, each actor was selected 5.8 times between Timepoints 1 and 2 and 4.7 times between and Timepoints 2 and 3 and was given an opportunity to change his or her network (of course, the actual number of changes in the network is lower than the number of opportunities to change because some actors may "choose" to maintain their networks). The outdegree parameter is negative, indicating that over time, the density of the friendship network decreases. More importantly, results indicate that reciprocity is statistically significant and positive (estimate = 2.59, $SE = 0.26, p < .01$), indicating that over time, there is a significant tendency to reciprocate friendship choices sent in the previous timepoint.

Turning to the behavioral objective function, after describing the rate functions, two effects that are added by default are estimated: The linear shape parameter is positive and statistically significant (its estimate is more than 1.96 times its standard error in absolute terms), indicating that over time, turnover intentions become stronger (there is a steady, linear increase on the attribute over time). The quadratic shape parameter is also statistically significant and negative, indicating that over time, there is a tendency toward a unimodal distribution of turnover intention.

Step 2b. Assess the model's GOF. Once satisfactory convergence has been reached, the researcher assesses the model's GOF to additional effects that were not directly modeled. Examples might include the in- and outdegree distribution and the geodesic distance distribution—effects that describe the global structure of the network over time. A model is said to be a good fit if the values of these additional statistics across multiple simulations are close to (fall within the 95% confidence interval [CI] of) the values observed in the data. We will examine the distribution of indegree [17], outdegree [18], and geodesic distances [19] over time, all considered customary indicators of GOF.[8]

[17] gofi <- sienaGOF(Model1.results, IndegreeDistribution, verbose=TRUE, join=TRUE, varName="friendship")	#creates three new objects that take the results object and examine the distributions of expected in- and outdegree distributions and the geodesic distance distribution.
[18] gofo <- sienaGOF(Model1.results, OutdegreeDistribution, verbose=TRUE, join=TRUE, varName="friendship")	
[19] gofgeo <- sienaGOF(Model1.results, GeodesicDistribution, levls=1:12, verbose=TRUE, join=TRUE, varName="friendship")	#specifies the distribution for levels 1 to 8 (default) or 1 to 12 (customized) on the friendship network.
[20] plot(gofi)	#plots the objects
[21] Plot(gofo)	
[22] Plot(gofgeo)	

Plotting these three variables [20-22] provides the following graphs (Figure 1). The (red) line connecting the (red) squares indicates the observed value of the statistic. The violin surrounding the line provides a 95% CI of the estimated statistics. For the indegree distribution, the red square falls within the 95% CI for expected values, indicating good fit, with an overall p value of .07—the model's expected indegree distribution over time is not significantly different from the observed indegree distribution over time. The model captures the outdegree distribution well (p = .839). Our very basic model, however, clearly does not capture the distribution of geodesic distances over time well.

Step 2c. Improve the model's GOF by adding additional effects. Failure to achieve adequate model GOF is resolved by adding additional effects. Typically, effects are added in the following order: (a) effects that model clusterability in the network, (b) effects that model the dynamics of in- and/or outdegree distributions, and (c) effects that model the relationship between in- and outdegree distribution.

Effects that model clusterability. Beyond the two basic effects specified in the mode object by default, researchers would typically add an effect that controls for the tendency for network

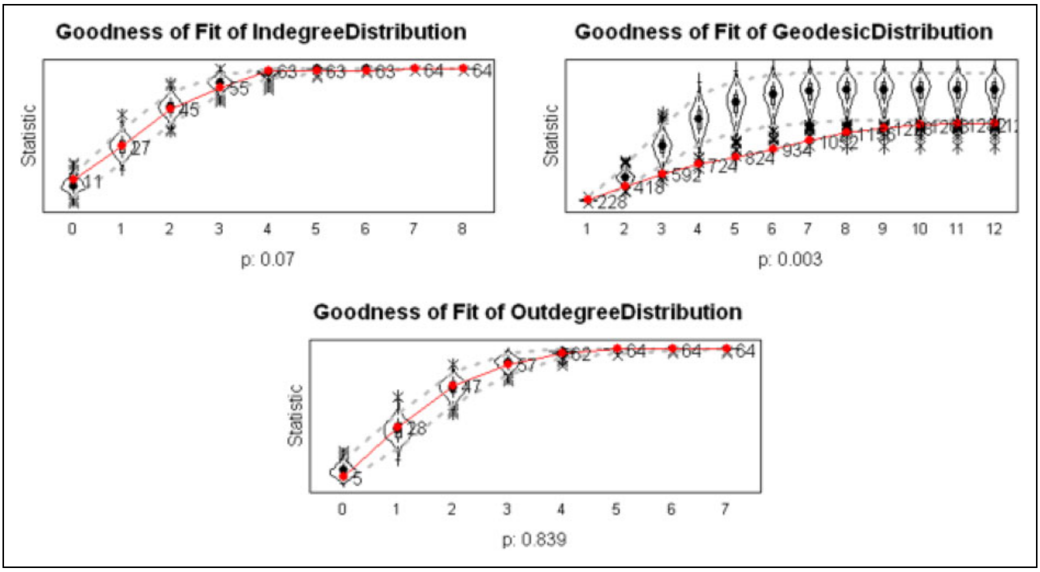

Figure 1. Goodness-of-fit plots for Model 1.

closure—the tendency for clustering to appear in networks. Multiple such effects are specified in RSiena; one or two of them are typically used, and the decision regarding which effect to use is typically based on theoretical considerations. Thus, for example, transitivity is typically found in friendship, leadership, and advice networks and should be captured with the effects denoted *GWESPFF* or *transTrip*. As a rule of thumb, it is best to start by adding the *GWESPFF* effect. The *GWESPFF* effect was initially developed for exponential random graph models and is more fully explained in the RSiena manual and in Tom Snijders's work (Snijders, Pattison, Robins, & Handcock, 2006). Briefly, it examines the number of times ego chose an alter for which there exists both a direct relationship and an indirect relationship through additional alters such that each additional alter who is connected to ego's alters is given a lower weight in the calculation of the statistic. As such, it is a weighted aggregate of the *transTrip* effect. If the model's fit is adequate, the researcher can start adding actor attributes to the network objective function. If it is not, it may also be necessary to add the *3cycle* effect, which controls for cycles in the network (particularly relevant for advice or trade networks).

```
[24]   MyEff <- includeEffects(MyEff, gwespFF)   #adds the GWESPFF effect to the effects object (MyEff)
```

As shown in Figure 2, the addition of the effect that controls for clustering improves our model: It captures the outdegree distributions well, and the distribution of geodesic distances has improved, but it is still not adequate.

Effects that model actors' in- and outdegrees (and their correlation). If adequate fit is still not achieved after incorporation of effects that model clusterability, researchers should add effects that model actors' in- and outdegrees over time. These effects are typically added when degrees are of theoretical importance (e.g., because they represent social status—as in friendship, advice, and/or leadership nomination networks) or because there is high dispersion in degrees such that degree

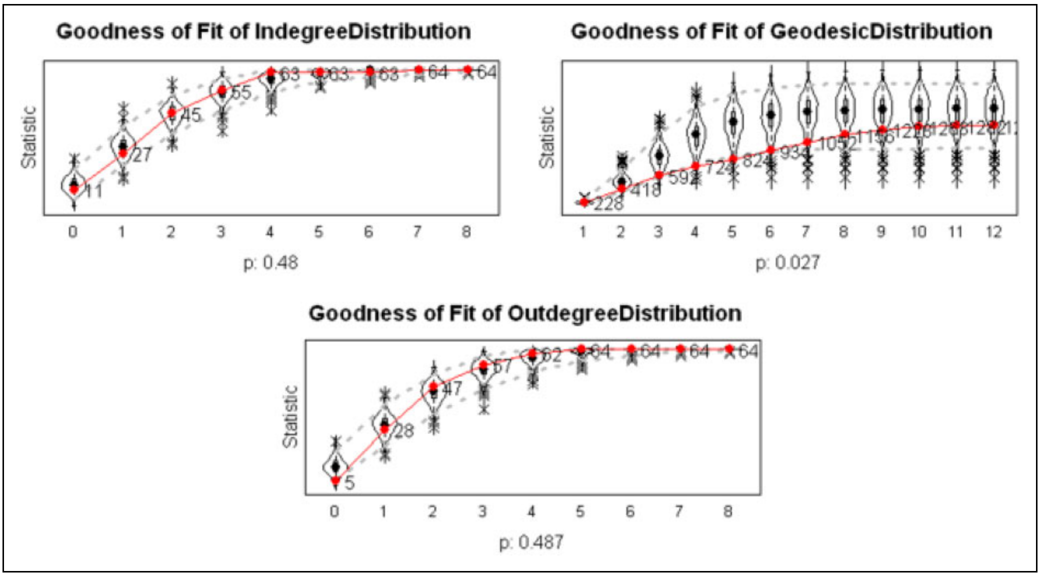

Figure 2. Goodness-of-fit plots for Model2 .

distribution on GOF plots does not adequately capture the observed degree distribution. Eight such effects are available in RSiena, but the one most often used is the *inPop* effect (see Table 1), which models the tendency of actors who were popular at previous timepoints to become more popular over time (the Matthew effect; Merton, 1968). If the network contains many disconnected actors, it might be important to add the *isolate* effect.

If adequate GOF is still not reached after directly modeling actors' in- and/or outdegrees, it might be necessary, in light of theoretical or empirical considerations, to add an effect that directly models the relationship between actors' indegrees and outdegrees (e.g., *outPop*, reflecting the tendency of actors who select multiple others to be selected more over time). Addition of these effects typically improves GOF on both degree distribution and geodesic distance distribution.

Returning to our example, because adequate GOF has not yet been reached (with the model containing outdegree, reciprocity, and GWESPFF), we will add the *inPop* effect. The effect is not significant (not shown; this is not surprising given that the indegree distribution is already adequately modeled). Perhaps we need to add an effect that controls for the relationship between actors' in- and outdegrees—this effect may assist in modeling the more problematic geodesic distribution. A researcher might argue that above and beyond reciprocity, the more actors choose others as friends at earlier timepoints, the more popular they become at later timepoints. Let us test whether the *outPop* effect should be included in the model (by fixing it to 0 and testing whether the assumption that the effect is 0 is valid [26]) by using the score test (Schweinberger, 2012).

```
[25] MyEff <- setEffect(MyEff, outPop, fix = TRUE,      #tests whether we should include the outPop
         test = TRUE, initialValue = 0)                    Effect by fixing it to 0. This is the way to test for the
                                                           inclusion of effects in RSiena.
```

summary(model2.results) provides the following result:

Generalized score test <c>
Testing the goodness of fit of the model restricted by
eval: outdegree – popularity = 0.0000

$\chi^2 = 15.6541$ $df = 1$ p value < .0001
one-sided (normal variate): –3.9565

The score-test suggests that we should include the *outPop* effect in our model (chi-square = 15.651, $df = 1$, $p < .05$). Using the model with outdegree, reciprocity, clustering, and an effect controlling for the correlation between in- and outdegrees and rerunning GOF diagnostics (not shown) suggests that the model now fits the data well.

With dozens of possible effects specified in RSiena and the built-in ability to create user-specified effects, modeling can be an ominous task. However, at the end of the day, modeling the network evolution part of the model is relatively simple and straightforward (Table 1).

Step 2d. Add actor attribute effects to the network objective function. Having controlled for (endogenous) network evolution effects, the researcher can now proceed to include effects related to actor attributes—some of these might already pertain to the researchers' hypotheses. Adding attributes to the network objective function typically involves adding three additional effects, which link an actor's level on a given attribute with the evolution of the actor's ties (see Table 1). The *covariate-similarity* effect models the tendency of actors to create ties to others with similar levels on a given attribute. This is an effect that directly measures social selection (or homophily). The *covariate-ego* and *covariate-alter* effects model the preferential tendency of people who are higher on an attribute to select more others or to be selected more often by others over time, respectively. Table 1 clearly reveals the types of hypotheses that are represented by these effects. For example, the covariate-ego effect, *egoX*, may represent hypotheses such as actors with higher status control more others over time (de Klepper, Labianca, Sleebos, & Agneessens, 2017) or actors with higher levels of stress will communicate with fewer others (Kalish et al., 2015). In our illustrative example, the *quit-ego* effect measures the tendency for people with higher levels of turnover intention to form more friendship ties over time—and is a direct test of Hypothesis 3. We expect a significant negative parameter estimate for this effect.

The covariate-alter effect, *altX*, measures the tendency of people with higher levels of the attribute to be selected more (less) over time. Examples for such hypotheses might be that warm (Fransen, Delvaux, Mesquita, & Van Puyenbroeck, 2018) and intelligent (Kalish & Luria, 2016) individuals are selected more as leaders over time. In our example with turnover intention, the *quit-alter* effect measures the tendency for people to befriend others who have higher levels of turnover intentions over time. We expect a significant negative parameter estimate for this effect as support for Hypothesis 4.

The covariate-similarity effect, *simX*, measures the tendency for people to form ties with others with similar levels of the attribute. As such, it is a direct test of homophily. It has been used to test hypotheses such as high-performing firms cooperate with other high-performing firms (Checkley, Steglich, Angwin, & Endersby, 2014) or people seek advice from others with similar levels of psychological safety (Schulte et al., 2012). In our example, a significant positive parameter estimate would support Hypothesis 2, stating that people choose each other as friends because of similarity in their levels of turnover intentions.

We added the same three effects for gender as controls.

[30] MyEff <- includeEffects(MyEff, egoX, altX, simX, interaction1="quit") [31] MyEff <- includeEffects(MyEff, egoX, altX, simX, interaction1="gender")	#includes the three attribute effects for both attributes. We are only testing hypotheses for the three effects for "*quit*". The effects for gender are added as controls.

Step 2e. Model the behavioral objective function. In modeling ego's choice regarding the level of an attribute (Equation 2), by default, RSiena includes two effects that control for the distributional shape of the attribute over time: the *shape* effect, which models the tendency for an attribute level to shift over time toward the midpoint of the range, and the *quadratic shape effect*, which models whether over time the values on the attribute shift toward a unimodal or a bimodal distribution. In addition to these effects, researchers typically incorporate two types of effects into the behavioral utility function. When researchers are interested in examining social influence, they should include the average similarity effect (*avSim*) or one of its theoretical variants, which model the tendency of actors to become more similar on the attribute to their alters over time. Examples might include students adjust their discipline to be more similar to their friends over time (de Klepper, Sleebos, van de Bunt, & Agneessens, 2010) or firm performance is contagious (Checkley et al., 2014).

When researchers are interested in examining how changes in levels of connectivity affect actors' attributes, they often include the indegree effect (*indeg*) or outdegree effect (*outdeg*), which examines how changes in an actor's popularity or expansiveness over time influence his level of a given attribute. Examples (Table 1) might include increases in firm connectivity influence firm performance over time (Checkley et al., 2014) or the fewer others a person communicates with, the higher his or her level of stress (Kalish et al., 2015).

In our turnover intention example, we include the *avSim* effects, which directly model social influence (Hypothesis 1)—the idea that people catch their friends' average level of the attribute. A significant positive parameter value for this effect would suggest the presence of social influence in the network.

[32] MyEff <- includeEffects(MyEff, name="quit", avSim, interaction1="friendship")	#includes the average similarity effect on the quit attribute and the friendship network.

Because we are now modeling the evolution of both the network and the attribute (turnover intention), we will add an additional indicator to evaluate GOF; namely, does the model capture the distribution of actors' attribute levels over time?

[33] gofbeh <- sienaGOF(Model4.results, BehaviorDistribution, levls=1:6, verbose=TRUE, join=TRUE, varName="quit")	#creates a GOF results object for the distribution of the attribute (*quit*) over time, and presents all 6 levels of the attribute.

Step 3. Interpret the Output

Assuming that convergence has been reached and GOF is adequate, it is now possible to interpret the output. Figure 3 provides GOF diagnostics for the final model. All GOF statistics are well above the required 0.05. Parameter estimates, standard errors, and t statistics for convergence are presented in Table 4.

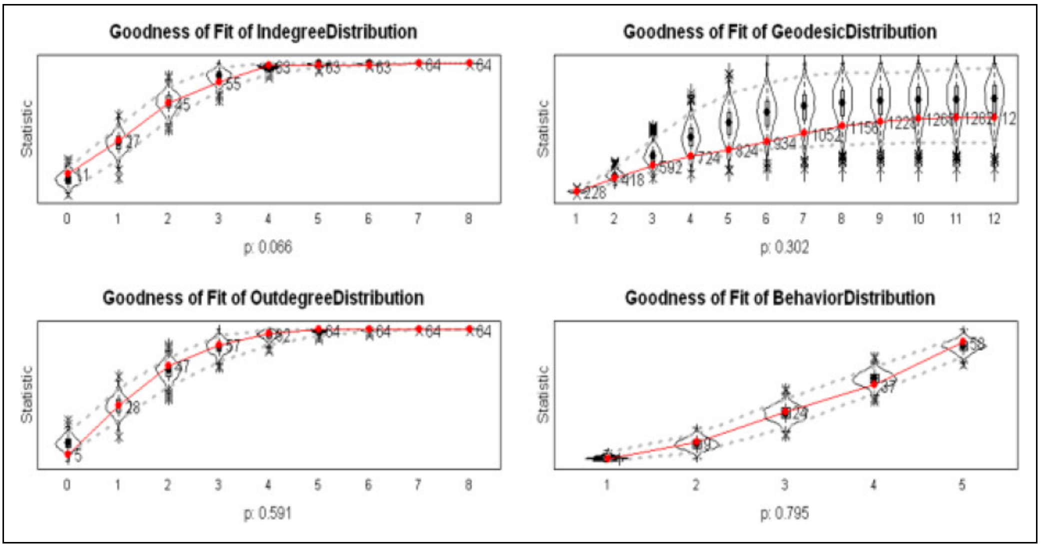

Figure 3. Goodness-of-fit plots for final model.

Table 4. Parameter Estimates, Standard Errors, and Convergence *t* Ratios for the Final Model.

	Estimate	Standard Error	Convergence *t* Ratio
Network dynamics			
1. rate constant friendship rate (period 1)	6.3187	(1.4402)	–0.0220
2. rate constant friendship rate (period 2)	4.5908	(0.9391)	0.0367
3. eval outdegree (density)	–2.3348	(0.2985)	–0.0118
4. eval reciprocity	3.0950	(0.4512)	0.0193
5. eval GWESP I -> K -> J (69)	3.5061	(0.6328)	–0.0413
6. eval outdegree - popularity	–0.4479	(0.1601)	–0.0034
7. eval gender alter	–0.0336	(0.2380)	–0.0078
8. eval gender ego	–0.0038	(0.2772)	–0.0459
9. eval gender similarity	0.2248	(0.2002)	–0.0749
10. eval quit alter	–0.0719	(0.1588)	0.0352
11. eval quit ego	0.1049	(0.1906)	0.0433
12. eval quit similarity	2.3262	(1.0414)	–0.0070
Behavior dynamics			
13. rate rate quit (period 1)	2.0494	(0.7097)	0.0016
14. rate rate quit (period 2)	1.5827	(0.5344)	–0.0414
15. eval behavior quit linear shape	0.5703	(0.2937)	0.0569
16. eval behavior quit quadratic shape	–0.0579	(0.1071)	–0.0475
17. eval behavior quit average similarity	6.2381	(3.8873)	0.0456

Hypothesis 1 suggests that people catch their friends' turnover intentions over time. The average similarity effect (effect 17) does not reach (two-sided) statistical significance—the hypothesis is not supported, but we note a trend in the predicted direction. Hypothesis 2 suggested that people select others as friends because of similarity in turnover intentions. The *quit* similarity effect (effect 12) is significant and positive as expected. Thus, Hypothesis 2 is supported. Hypotheses 3 and 4 suggested that people with higher levels of turnover intentions select fewer others as friends (Hypothesis 3, effect 11) and are selected by fewer others as friends (Hypothesis 4, effect 11) over time. Neither effect was statistically significant; that is, neither hypothesis was supported.

Additional interesting effects emerge from our analysis: There is indeed evidence for significant reciprocity and transitivity in the network over time: Over time, people select those who selected them as friends (effect 4), and people want their friends' friends to become their own friends (effect 5). People who are expansive in their choice of friends become less popular over time (effect 6): Being overly inclusive in (nonreciprocated) friendship harms a person's popularity. As for gender, it does not seem to affect network evolution: We observe no effect for gender either on being selected as a friend over time (effect 7), selecting more others as friends over time (effect 8), or selecting similar others over time (effect 9).

Taken together, these (fictional) results may explain previous results on collective turnover (Feeley et al., 2008; Felps et al., 2009). Results suggest that in the network under investigation, both selection and (possibly) influence effects may be at play, strengthening each other: Individuals with higher levels of turnover intentions select similar others as friends, and there is some indication that they may then catch their friends' levels of turnover intention. Individuals with low levels of turnover intention also select similar others, and they too catch their friends' (lower) levels of turnover intention. These effects are observed even after controlling for the tendency to reciprocate friendship choices and more importantly, for the tendency of the friendship network to form clusters. Without SAO models for the co-evolution of networks and behavior, it would have been exceedingly difficult to obtain such a complex, detailed analysis of human behavior. It should be noted that because the network objective function models change in $n \times (n - 1)$ network ties, whereas the behavioral function models n attributes, it is generally easier to find statistically significant effects that pertain to the former compared with the latter due to the increased power of the former analysis. The implication is that larger samples are needed to examine social influence as compared with social selection. Owing to the difficulty in collecting complete network and attribute data in large samples (this would require every person to complete a questionnaire on every other person), researchers often aggregate results from multiple small groups either by using structural zeros to highlight impossible ties between groups (cf. Kalish et al, 2015) or preferably, meta-analyzing results from multiple small groups using the RSiena meta-analysis option. The utilization of multiple groups also addresses the replicability issue that is often a concern is network analytic studies (Forbes, Wright, Markon, & Krueger, 2017).

Concluding Remarks

Organizations are characterized by change: Employees change over time in their commitment, performance, behaviors, and attitudes. Organizations and countries change in their levels of democracy, attitudes toward governance mechanisms, and hostility. The structures of relationships also change over time: Employees befriend new people and sever existing friendships; countries form and alter trade agreements. Only recently has an effective technique been developed for modeling these changes in networks and attributes and unpacking the complex dependencies between them: SAO models for the co-evolution of networks and behavior (Snijders et al., 2007, 2010). This paper provided a tutorial introduction to these models.

Notably, though these powerful models have been in use for quite some time in the fields of sociology and education, their potential has yet to be fully realized in organizational scholarship. Applications of SAO models for co-evolution of networks and actor attributes in applied organizational settings include research examining the co-evolution of advice and leadership networks and (in)civility (Porath, Gerbasi, & Schorch, 2015), communication networks and perceived stress (Kalish et al., 2015), trust networks and job satisfaction (Agneessens & Wittek, 2008), interfirm networks and governance orientation (Benton, 2016), interfirm collaborative networks and firm performance (Checkley et al., 2014), status networks and peer control attempts (de Klepper et al., 2017), information sharing networks and evidence-based practice (Mercken, Saul, Lemaire, Valente, & Leischow, 2015), friendship networks and unethical behavior (Zuber, 2015) and advice, and friendship and difficult tie networks and psychological safety (Schulte et al., 2012).

I propose that SAO models can—and should—be considered more often by organizational scholars. They allow researchers to model the temporal dynamics of networks and attributes and the complex interrelationships between them. Researchers who are interested predominantly in the dynamics of networks can use these models to explain emergence of status, leadership, or power hierarchies. In particular, these models can empirically address theoretical questions such as: When and how does shared leadership occur? When does centralization of power emerge? And when and how do "bad apples" emerge (Carter, DeChurch, Braun, & Contractor, 2015)? Likewise, these models might be useful to scholars interested in identifying the mechanisms—and specifically, the attributes and network relationship patterns—that drive the evolution of attributes such as group climate, norms, emotions, and attitudes (Lang et al., 2018). Researchers interested in examining how attributes drive network relationships might benefit from using SAO models to answer questions such as which attributes drive cooperation and which attributes drive helping behavior. Regardless of the specific research question involved, I propose that SAO models can provide researchers with a fine-grained analysis of multiple social and psychological processes and their interactions with each other and the manner in which these processes generate higher level, emergent structures. As such, these models may be an important step in bridging the micro-macro gap in organizational scholarship (Kilduff & Tsai, 2003). Since networks may—or may not—coincide with formal structures (Krackhardt & Hanson, 1993), utilizing SAO models may help reconcile the widely divergent results obtained when attempting to aggregate from the individual level to the team level (Kalish, 2013).

The SAO models discussed herein can be expanded in various directions (not included in the scope of this tutorial); such extensions include SAO models for multiple networks, whereby the evolution of one type of tie can depend on another type of tie; SAO models for the co-evolution of two-mode networks, whereby actors are tied to objects and both evolve simultaneously; and models that accommodate different weights on the effect of tie formation and the effect of tie dissolution. To implement an SAO model in RSiena, it is necessary to have some familiarity with the R framework, which may pose a challenge to some researchers. However, given this tutorial and the excellent RSiena manual (Ripley et al., 2018a), coupled with the immense value that SAO models can offer organizational scholarship, I hope that the use of SAO models will increase in this domain.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Supplemental Material

Supplemental material for this article is available online.

Notes

1. Formally, the model assumes that changes in network ties and attributes occur following a Markov process, that is, the conditional probability distribution of future states of the network (and attributes) depends on the past states only as a function of the present states. In other words, the current state of the network at time *t* probabilistically determines the future state of the networks (see Snijders, van den Bunt, & Steglich, 2010)

2. For example, the reciprocity effect simply counts the number of reciprocated ties that the focal actor has (defined as $s_i^{net}(x) = \sum_j x_{ij}x_{ji}$).
3. More formally, parameter values are log-odds ratio for choosing between two alternatives: the result of changing the tie (or attribute) and the result of maintaining the current network (or attribute). For example, a positive reciprocity effect means that over time, there is a tendency to reciprocate ties—it is beneficial for actors to reciprocate ties sent to them in previous timepoints. This is further explained in the simplified example, and see Ripley, Snijders, Boda, Voros, and Preciado (2018a, p. 67) for a more comprehensive explanation.
4. Estimation of parameter values is conducted in three phases. In the first phase, initial values are estimated by the RSiena program through an initial inspection of data. In the second phase, multiple runs (typically 1,000) are conducted, and the program searches for parameter values where deviations between generated and observed statistics across time average out to 0. In the third phase, (now final) parameter values are held constant, and standard errors are computed. A more comprehensive introduction to the estimation procedure in RSienacan be found in Snijders (2001).
5. The calculations presented in the following example are a simplification of the calculations that stochastic actor-oriented (SAO) models perform. In reality, variables are centered prior to any calculations. This oversimplified example is provided to assist the reader in understanding the internal calculations of the model. The interested reader should refer to Snijders et al. (2010).
6. There are additional basic effects related to attributes that are incorporated into Equation 2 by default; they are referred to as the *shape* and the *quadratic shape* effects. For simplicity's sake, we will not consider them in this discussion. See the following section for a more comprehensive discussion of these effects.
7. Structural zeros are often used to represent impossible ties between actors who are not allowed to interact with each other because they belong to different groups or teams. As such, they are one solution often used to aggregate results from multiple groups (cf. Kalish, Luria, Toker, & Westman, 2015). An alternative, preferable solution is to run a meta-analysis of RSiena results (see the RSiena manual [Ripley et al., 2018a], section 11.2 for more details).
8. For geodesic distances, researchers need to run an additional script that formalizes the function. The function, presented in Appendix A (available in the online journal), can be used "as is." An additional distribution that is often useful to troubleshoot GOF issues is the triad census.

References

Agneessens, F., & Wittek, R. (2008). Social capital and employee well-being: Disentangling intrapersonal and interpersonal selection and influence mechanisms. *Revue Francaise de Sociologie*, *49*(3), 613-637.

Barsade, S. G. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, *47*(4), 644-675.

Bartunek, J. M., Huang, Z., & Walsh, I. J. (2008). The development of a process model of collective turnover. *Human Relations*, *61*(1), 5-38. doi:10.1177/0018726707085944

Benton, R. A. (2016). Corporate governance and nested authority: Cohesive network structure, actor-driven mechanisms, and the balance of power in American corporations. *American Journal of Sociology*, *122*(3), 661-713. doi:10.1086/689397

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(5916), 892-895.

Bothma, C. F. C., & Roodt, G. (2013). The validation of the turnover intention scale. *SA Journal of Human Resource Management*, *11*(1), 1-12.

Carter, D. R., DeChurch, L. A., Braun, M. T., & Contractor, N. S. (2015). Social network approaches to leadership: An integrative conceptual review. *Journal of Applied Psychology*, *100*(3), 597-622. doi:10.1037/a0038922

Checkley, M., Steglich, C., Angwin, D., & Endersby, R. (2014). Firm performance and the evolution of cooperative interfirm networks: UK venture capital syndication. *Strategic Change*, *23*(1-2), 107-118.

de Klepper, M. C., Labianca, G., Sleebos, E., & Agneessens, F. (2017). Sociometric status and peer control attempts: A multiple status hierarchies approach. *Journal of Management Studies*, *54*(1), 1-31.

de Klepper, M. C., Sleebos, E., van de Bunt, G. G., & Agneessens, F. (2010). Similarity in friendship networks: Selection or influence? The effect of constraining contexts and non-visible individual attributes. *Social Networks*, *32*(1), 82-90. doi:10.1016/j.socnet.2009.06.003

Emirbayer, M. (1997). Manifesto for relational sociology. *American Journal of Sociology*, *103*(2), 281-317.

Feeley, T. H., Hwang, J., & Barnett, G. A. (2008). Predicting employee turnover from friendship networks. *Journal of Applied Communication Research*, *36*(1), 56-73. doi:10.1080/00909880701799790

Felps, W., Mitchell, T. R., Hiekman, D. R., Lee, T. W., Holtom, B. C., & Harman, W. S. (2009). Turnover contagion: How coworkers' job embeddeness and job search behaviors influence quitting. *Academy of Management Journal*, *52*(3), 545-561.

Forbes, M. K., Wright, A. G. C., Markon, K. E., & Krueger, R. F. (2017). Evidence that psychopathology symptom networks have limited replicability. *Journal of Abnormal Psychology*, *126*(7), 969-998.

Fransen, K., Delvaux, E., Mesquita, B., & Van Puyenbroeck, S. (2018). The emergence of shared leadership in newly formed teams with an initial structure of vertical leadership: A longitudinal analysis. *Journal of Applied Behavioral Science*, *54*(2), 140-170.

Kalish, Y. (2013). Harnessing the power of social network analysis to explain organizational phenomena. In J. M. Cortina & R. S. Landis (Eds.), *Modern research methods for the study of behavior in organizations* (pp. 99-135). New York, NY: Routledge.

Kalish, Y., & Luria, G. (2016). Leadership emergence over time in short-lived groups: Integrating expectations states theory with temporal person-perception and self-serving bias. *Journal of Applied Psychology*, *101*(10), 1474-1486. doi:10.1037/apl0000126

Kalish, Y., Luria, G., Toker, S., & Westman, M. (2015). Till stress do us part: On the interplay between perceived stress and communication network dynamics. *Journal of Applied Psychology*, *100*(6), 1737-1751.

Kalish, Y., & Robins, G. L. (2006). Psychological predispositions and network structure: The relationship between individual predispositions, structural holes and network closure. *Social Networks*, *28*(1), 56-84.

Kilduff, M., & Brass, D. J. (2010). Organizational social network research: Core ideas and key debates. *Academy of Management Annals*, *4*, 317-357.

Kilduff, M., & Tsai, W. (2003). *Social networks and organizations*. London: Sage.

Klein, K. J., Lim, B., Saltz, J. L., & Mayer, D. M. (2004). How do they get there? An examination of the antecedents of centrality in team networks. *Academy of Management Journal*, *47*(6), 952-963.

Krackhardt, D., & Hanson, J. (1993). Informal networks: The company behind the chart. *Harvard Business Review*, *71*(4), 104-111.

Krackhardt, D., & Porter, L. W. (1986). The snowball effect—Turnover embedded in communication-networks. *Journal of Applied Psychology*, *71*(1), 50-55.

Lang, J. W. B., Bliese, P. D., & de Voogt, A. (2018). Modeling consensus emergence in groups using longitudinal multilevel methods. *Personnel Psychology*, *71*(2), 255-281.

Mercken, L., Saul, J. E., Lemaire, R. H., Valente, T. W., & Leischow, S. J. (2015). Coevolution of information sharing and implementation of evidence-based practices among North American tobacco cessation quitlines. *American Journal of Public Health*, *105*(9), 1814-1822.

Merton, R. K. (1968). The Mathew effect in science. *Science*, *159*(3810), 56-63.

Porath, C. L., Gerbasi, A., & Schorch, S. L. (2015). The effects of civility on advice, leadership, and performance. *Journal of Applied Psychology*, *100*(5), 1527-1541.

Ripley, R. M., Snijders, T. A. B., Boda, Z., Voros, A., & Preciado, P. (2018a). *Manual for RSIENA*. Oxford, UK: University of Oxford, Nuffiled College.

Ripley, R. M., Snijders, T. A. B., Boda, Z., Voros, A., & Preciado, P. (2018b). RSiena package for R (Version 4. 0). [Computer software].

Robins, G. (2015). *Doing social network research*. London: Sage.

Schulte, M., Cohen, N. A., & Klein, K. K. (2012). The co-evolution of network ties and perceptions of psychological safety. *Organization Science*, *23*(2), 564-581.

Schweinberger, M. (2012). Statistical modeling of network panel data: Goodness-of-fit. *British Journal of Statistical and Mathematical Psychology*, *63*, 263-281.

Snijders, T. A. B. (1996). Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology*, *21*(1-2), 149-172.

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

Snijders, T. A. B. (2008). Longitudinal methods of network analysis. In J. Scott (Ed.), *Encyclopedia of complexity and system science* (pp. 5998-6013). Dordrecht, the Netherlands: Springer Verlag.

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

Snijders, T. A. B., Steglich, C., & Schweinberger, M. (2007). Modeling the co-evolution of networks and behavior. In K. van Montfort, H. Oud, & A. Satorra (Eds.), *Longitudinal models in the behavioral and related sciences* (pp. 41-71). Mahwah, NJ: Lawrence Erlbaum.

Snijders, T. A. B., van de Bunt, G. G., & Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, *32*(1), 44-60. doi:10.1016/j.socnet.2009.02.004

Steglich, C., Snijders, T. A. B., & Pearson, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology*, *40*, 329-393.

Veenstra, R., & Steglich, C. (2012). Actor-based model for network and behavior dynamics: A tool to examine selection and influence processes. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 598-618). New York, NY: Guilford Press.

Zuber, F. (2015). Spread of unethical behavior in organizations: A dynamic social network perspective. *Journal of Business Ethics*, *131*(1), 151-172.

Author Biography

Yuval Kalish is adjunct professor at Coller School of Management, where he also serves as academic director of Executive MBA programs. His research focuses primarily on the interplay between social networks and individual attributes in explaining organizational and social phenomena.