Simulating Relational Event Histories - Why and How

Rumana Lakdawala[1], Joris Mulder[1], and Roger Leenders[1,2]

[1]Tilburg University
[1]Jheronimus Academy of Data Science

Abstract

Many important social phenomena result from repeated interactions among individuals over time such as email exchanges in an organization, or face-to-face interactions in a classroom. Insights into the mechanisms underlying the dynamics of these interactions can be achieved through simulations of networks on a fine temporal granularity. In this paper, we present statistical frameworks to simulate relational event networks under dyadic and actor-oriented relational event models. These simulators have a broad applicability in temporal social network research such as model fit assessment, theory building, network intervention planning, making predictions, understanding the impact of network structures, to name a few. We show this in three extensive applications. First, it is shown why simulation-based techniques are crucial for relational event model assessment, for example to investigate how past events affect future interactions in the network. Second, we show how simulation techniques are important when building and extending theories about social phenomena such as understanding social identity dynamics using optimal distinctiveness theory. Third, we demonstrate how simulation techniques contribute to a better understanding of the longevity of network interventions.

1 Introduction

Our understanding of social interaction mechanisms between individuals has been enhanced by dynamic network-based approaches. A network approach allows for the representation of individuals by vertices and an event or interaction between the individuals by the edges between the vertices. In traditional, static network analysis, edges are assumed to be stable and their presence or absence is indicative for the state the network is in. In dynamic network analysis, edges (and vertices, for that matter) are allowed to appear and disappear. In most dynamic network applications, one observes a network repeatedly and then notes for each edge its presence or absence at observation time. In continuous time dynamic network analysis (of which this paper is an example), edges are observed to appear and dissolve in real time at any time. To distinguish between edges that are stable and edges that can be in constant flux, the latter are often termed "relational events." The measurement of network changes over time gives access to fine-grained temporal information

1

about the evolution of the network. This allows for investigation of the relationship between individual behaviour of individual nodes or dyads in the network, and collective behaviour of the network. Particularly the temporal information yields insight into how variation in individual level dynamics affect network level characteristics, which in turn further affect individual behaviour. Several statistical methods have been developed to study dynamic relational event or longitudinal social network data (Holland and Leinhardt, 1977; Snijders et al., 2010; Hanneke et al., 2010; Krivitsky and Handcock, 2014; Perry and Wolfe, 2013). Of particular interest to us is the relational event model introduced by Butts (2008), and Stadtfeld (2014) which allows for modelling of dynamic networks when the exact time-stamp of the events are available. Relational event models are a class of models that model the event rate in a dynamic network as a function of the network's past history and exogenous information available about the actors (nodal covariates) or pairs of actors (dyadic covariates) in the network. These models provide a flexible framework for network inference that allows for estimation of various network effects such as homophily, reciprocity, transitivity, and preferential attachment while jointly accounting for the fine-grained temporal nature of real-life social interaction.

Despite its usefulness to quantify the relative importance of various effects on the network evolution, it may still be difficult to truly understand the underlying social mechanisms based on a fitted relational event model. For example, it may be that several relational event models achieve similar fit to the data or the true underlying social mechanisms may be too complex to discover using a relational event model only. This is echoed in Box (1976) famous quote that "all models are wrong...". It is possible, however, to get a better understanding about the true underlying social mechanisms by further exploring what network structures look like under different relational event modeling settings via simulations. Simulations can provide an approximation of applied research settings as long as the simulations generate realistic networks. A major feature of the relational event-based approach to simulations, making it a suitable choice for simulating realistic datasets, is its flexibility in allowing for a wide range of network effects and covariates that can be utilized to generate datasets with many complex mechanisms in juxtaposition.

In this paper we introduce general frameworks to simulate relational event networks under the dyadic (Butts, 2008) and the actor oriented (Stadtfeld and Block, 2017) relational event model that can be utilized for simulating relational event sequences. The simulation frameworks can be used for various objectives:

(i) Assess goodness of fit: For a richer goodness of fit assessment, key network characteristics of the empirical relational event networks (such as degree distributions, density, triadic sub-structures, inter-event time distribution) can be compared to the simulated networks under the fitted model.

(ii) Evaluate network interventions: Relational event simulations can be used to evaluate network interventions, by investigating the temporal dynamics of social networks under various intervention scenarios. For instance, researchers can use simulations to explore how quickly networks respond to interventions, how long interventions need to be carried out to achieve desired outcomes, and whether the effects of interventions persist over time or fade away. Additionally, simulations can be used to study the effects of targeting specific actors in the network, and to identify the most effective intervention strategies for different network configurations.

(iii) Develop theory: Relational event simulations can be a powerful tool to develop social theories by providing a way to test and refine theoretical models in a controlled and systematic way. Simulations can be used to validate, test and extend existing theories as well as developing new theories about the social phenomena that drive the interactions in social networks. Moreover, simulations can help identify gaps in existing theories or data. For instance, if a simulated network does not match observed network characteristics, this could indicate that the theoretical model is incomplete or that data collection methods need to be revised. By identifying these gaps, researchers can develop more comprehensive theories and improve data collection methods.

(iv) Analyze the impact of network characteristics: By simulating networks with different characteristics (i.e. network size, combinations of endogenous and (or) exogenous effects, interaction effects, scaling of covariates, etc.) researchers can gain insight into how intricate combinations of the network settings affect social network dynamics.

(v) Make predictions: Relational event simulations fitted with predictive models can help us understand how relational sequences may evolve beyond the observation period, enabling us to make predictions about the most likely events in the near future or explore alternate scenarios for how network dynamics may develop over a longer period. In the latter case, it is worth noting the well-known statistical issue of the Hamill forecast (Hamill, 2001), where a forecast is "wrong" in specific instances but "correct on average". When a researcher marginally calibrates the model, each prediction of a collection of events may be unrealistic, but the model can still accurately predict the occurrence of events on average (Gneiting et al., 2007).

(vi) Evaluate model sensitivity and statistical power: Simulation-based analysis can be used to assess the sensitivity of relational event models to specific model parameters and violations of assumptions such as proportional hazards and conditional independence. Researchers can also perform extensive analysis of power, accuracy, and precision of relational event models (e.g., Schecter and Quintane, 2020) under idealized circumstances to test and benchmark novel extensions of the model where the truth conditions that generated the data are known.

This article has several goals. In Section 2, we introduce the dyadic and actor oriented relational event models and outline the general-purpose simulation frameworks for these models that can be customized in various ways using problem-specific information. Next we illustrate a range of simulation scenarios that showcase the broad potential of the simulation frameworks in an attempt to inspire and guide new research. In Section 3 we demonstrate how the simulations can be used to assess goodness of fit of relational event models using an email network from an organization. In Section 4, we describe how simulations can be used to develop theories about social phenomena that drive interactions and provide an example by simulating a well known social theory and test a boundary condition on this theory using simulations. In Section 5 we discuss how the simulation framework can be used to investigate outcomes of network-based interventions. We demonstrate how strategies for network-based interventions can be simulated and compared on an organizational network. Finally, we conclude with a discussion on limitations and future prospects of the simulation techniques in Section 6.

2 Simulating Relational Event Networks

A relational event can be thought of as an event in which a sending actor (e.g. a person, group of individuals or other entity) directs an action to a receiving actor (e.g. another person, group, organisation, etc.) in the form of a discrete instantaneous event at a certain time point. The specification for an observed event $e = (i, j, t)$ entails a sender $i \in \mathcal{A}$, a receiver $j \in \mathcal{A}$, and t, the time at which the event was observed, where $\mathcal{A}$ is a set of actors [1]. An event represents one step in the dynamic network, and a temporal sequence of relational events forms the relational event history (Butts and Marcum, 2017). In this paper, we describe networks in which the ties are directed (i.e an event from $i \to j$ is distinct from $j \to i$) and do not contain self loops (i.e from $i \to i$). However, extensions to undirected ties and networks with self-loops are straightforward.

In order to simulate a relational event network sequence[2] E in a time window $[0, \tau)$, every event e in the sequence E should be specified. This involves generating event times $t \in [0, \tau)$ at which each event occurs and the dyad (i, j) associated with each event. This process of generating event sequences can be described under the dyadic Relational Event Model (REM) of Butts (2008) as well as the actor-oriented Dynamic Network Actor Model (DyNAM) of Stadtfeld and Block (2017). We describe each of these simulation methods in detail in the following sections.

2.1 Simulation Framework 1: Relational Event Model (REM)

Consider a sequence of M events $E = \{e_1, e_2, \dots e_M\}$ such that. $t_m < t_{m+1}$. An event $e_m = (i_m, j_m, t_m)$, entails a sender $i_m \in \mathcal{A}$, a receiver $j_m \in \mathcal{A}$, and the time at which the event was observed, t_m. The set of possible events that can occur at any time is referred to as the risk set $\mathcal{R} \subseteq \{(i, j) : i, j \in \mathcal{A}\}$. The risk set may be fixed or vary with time endogenously, i.e particular events may give rise to the possibility of other events that were previously impossible, or the risk set may vary exogenously, where the possibility of occurrence of certain events is determined by exogenous factors such as availability of actors, location, or other circumstances specific to the research setting [3].

In REM, each dyad has its own rate of occurrence. A higher rate implies that the event involving that dyad is more likely to occur soon, and a lower rate implies that the occurrence of the event is more rare. The rate is modeled as a log linear function of the endogenous and exogenous statistics pertaining to that dyad along with the parameters β that represent the strength of these statistics to explain social interaction behavior in the network. The occurrence of events is modelled using a piece-wise constant hazard model. Under the piece-wise constant model, the rates are assumed to only change when an event occurs (anywhere in the network). Thus, as time progresses and events are observed, the rate is updated to reflect the new network history. The rate λ_{ij}^{dyadic} of a dyad (i, j) at time t is then specified in a log linear form as:

---

[1] Without loss of generality, we keep the set $\mathcal{A}$ fixed over time in this paper

[2] Relational Event Network, Relational Event Histories, and Relational Event Sequence are referred to analogously.

[3] Without loss of generality, we will keep the riskset constant in this paper.

$$\lambda_{ij}^{dyad}(t) = \begin{cases} \exp\{\beta^T X(i,j,E_t)\} & (i,j) \in \mathcal{R} \\ 0 & (i,j) \notin \mathcal{R} \end{cases} \tag{1}$$

where $X(i,j,E_t)$ is a vector of P statistics for the dyad (i,j) on the event sequence E_t until time t and $\beta \in \mathbb{R}^P$ is the vector of corresponding parameters. The statistics vector $X(i,j,E_t)$ can capture either endogenous network characteristics as a function of the past events in the history E_t before time t or exogenous actor or dyadic attributes. The notation for rate has a superscript 'dyad' to distinguish it from the rate under the actor-oriented model introduced in the next section.

Under the piece-wise constant hazard model, the waiting time $\delta_m = t_m - t_{m-1}$ between subsequent events are assumed to be conditionally independent and are specified by an exponential distribution, i.e.,

$$p^{dyad}(\delta_m | E_{t_{m-1}}, \beta, X) \sim \text{Exponential}(\Lambda^{dyad}(t_{m-1})) \tag{2}$$

with the cumulative rate as distributional parameter, given by

$$\Lambda^{dyad}(t_{m-1}) = \sum_{(i,j) \in \mathcal{R}} \lambda_{ij}^{dyad}(t_{m-1}).$$

The probability of a dyad $(i,j) \in \mathcal{R}$ to be involved in the next observed event, follows a multinomial distribution where the probabilities are proportional to the dyadic rates:

$$p^{dyad}(i,j \mid t, \beta, X, E_t) = \frac{\lambda_{ij}^{dyad}(t)}{\Lambda^{dyad}(t)}. \tag{3}$$

In order to simulate relational event networks, the user needs to specify which endogenous or exogenous network statistics (e.g., Leenders et al., 2016) are included in the model and the magnitude of the corresponding parameters (based on theory, a specific dynamic of theoretical interest, or a fitted model). If preferred, it also possible to initialize simulations with a pre-defined starting sequence E_0. This would be useful if a user wants to predict the immediate future of an empirical network or to simulate using different models from the same event history. The inputs to the simulation algorithm play an important role in determining the characteristics of the sequence output. The number of actors N in the network determines the size of the riskset as $N(N-1)$ (when all dyads are at risk) and τ, the time until which the simulation algorithm is run, determines the number of events in the sequence. The parameters β capture the sign and magnitude of the effect of the statistics. Given these inputs, a relational sequence E can be simulated using Algorithm 1.

Algorithm 1: Simulation of a relational event sequence under dyadic model

    input : $\beta, \tau, \mathcal{R}, E_o$

    output: sequence E

1 **if** E_o *is empty* **then**

2   Initialize event sequence E with a random event e_1

3 **else**

4   Initialize event sequence $E = E_o$

5 **end**

6 **while** $t_m \leq \tau$ **do**

7   compute statistics matrix X on $E_{t_{m-1}}$

8   update event rates $\lambda_{ij}^{dyad}(t_{m-1} \mid E_{t_{m-1}}, \beta, X) \, \forall (i,j) \in \mathcal{R}$ using Equation (1)

9   sample a dyad $(i_m, j_m) \in \mathcal{R}$ from multinomial distribution in Equation (3)

10   sample inter-event time $\delta_m \sim \text{Exp}\left(\sum_{(i,j) \in \mathcal{R}} \lambda_{ij}^{dyad}(t_{m-1}) \right)$

11 $t_m = t_{m-1} + \delta_m$

12   append (i_m, j_m, t_m) to E

13 $m + +$

14 **end**

2.2 Simulation Framework 2: Dynamic Network Actor Model (DyNAM)

Instead of all dyads competing with each other to participate in the next event, under the DyNAM the actors compete with each other to becomes the next sender, where each actor has its own rate parameter to be the next sender. A higher rate implies that the actor is more likely to be the sender of the next event. The rate of the actor as a sender is specified similarly to the rate parameter in a dyadic REM using a log linear model of endogenous and exogenous sender statistics, which are now summarized in the statistics matrix X^s. The time of the next event then follows an exponential distribution,

$$p^{actor}(\delta_m | E_t, \beta, X^s) \sim \text{Exponential}(\Lambda^{sender}(t)),$$

where $\Lambda^{sender}(t) = \sum_{i \in A} \lambda_i^{sender}(t)$, with $\mathcal{R}^s$ denotes the set of actors that are at risk to become a sender, and the rate parameter of actor i to become a sender is defined by

$$\lambda_i^{sender}(t) = \exp\{\gamma^T X^s(i, E_t), \} \tag{4}$$

where γ denotes the vector of coefficients which quantify the importance of the respective statistic in X^s. Furthermore, the probability of actor i to become a sender is proportional to the respective rate parameters according to a multinomial distribution

$$p^{actor}(i \mid E_t, \gamma, X^s) = \frac{\lambda_i^{sender}(t)}{\Lambda^{sender}(t).} \tag{5}$$

Finally, the probability of actor j to become the receiver is modeled conditionally on the sender i using a multinomial distribution with probability

$$p^{actor}(j|i, t, E_t, \alpha, X^r) = \frac{\lambda_{j|i}^{receiver}(t)}{\Lambda_i^{receiver}(t)} \tag{6}$$

where the preference parameter of actor j to be come the receiver is given by

$$\lambda_{j|i}^{receiver}(t) = \exp(\alpha^T X_{ij}^r),$$

which is modeled in a similar manner as the rate parameter, and the summation is given by

$$\Lambda_i^{receiver}(t) = \sum_{j \in A \backslash \{i\}} \lambda_{j|i}^{receiver}(t).$$

The simulation framework of relational event sequences under the DyNAM is then summarized in Algorithm 2

Algorithm 2: Simulation of relational event sequence under actor-oriented model

input : $\gamma, \alpha, \tau, \mathcal{R}, E_o$

output: sequence E

1 **if** E_o *is empty* **then**
2 $\quad$ Initialize event sequence E with a random event e_1
3 **else**
4 $\quad$ Initialize event sequence $E = E_o$
5 **end**
6 **while** $t_m \leq \tau$ **do**
7 $\quad$ compute sender statistics matrices $X^s(E_{t_{m-1}})$
8 $\quad$ compute sender rates $\lambda_i(t_{m-1} \mid E_{t_{m-1}}, \gamma, X^s) \; \forall i \in \mathcal{A}$ using Equation (4)
9 $\quad$ sample inter-event time $\delta \sim \exp \left(\sum_{(i) \in \mathcal{A}} \lambda_i(t_{m-1} \mid E_{t_{m-1}}, \gamma, X^s) \right)$
10 $\quad$ sample a sender $i_m \in \mathcal{A}$ from categorical distribution in Equation (5)
11 $\quad$ compute receiver statistics matrix X^r
12 $\quad$ sample a receiver $j_m \in A \backslash \{i_m\}$ given sender i_m from categorical distribution in (6)
13 $\quad$ append $e_i = (i_m, j_m, t_m)$ to G
14 $\quad$ update β if time-varying
15 $\quad$ $m++$
16 **end**

The main differences between the dyadic relational event model and its actor-oriented counterpart (i.e., the DyNAM) can be understood from how they model the occurrence of relational events. In the dyadic REM,

the timing of the next event and the dyad that is observed are modeled using dyad specific rate parameters. Alternatively, in the actor-oriented model, the timing and the sender of the next event are modeled via actor (sender) specific rate parameters, and the next receiver is modeled conditionally on the sender (i.e., *chosen* by the sender) where all actors have separate rate parameters as potential receivers (Stadtfeld and Block, 2017). These differences in modelling are also reflected in the simulation frameworks. In the dyadic model, all dyads are competing to be sampled as the next observed event whereas in actor-oriented model, the actors first compete to become the sender and then the remaining actors compete to become the receiver that is chosen by the given sender. The two models also differ in how they conceive network change. In the dyadic model, the building block is the dyad and the model assumes the two participating actors in the dyad to actively determine whether and when they will interact. In the actor-oriented model, on the other hand, the agency is at the level of the sender, who determines when to get active. Therefore, the latter model is conceptually actor-driven rather than dyad-driven. In the simulation frameworks this difference translates to which statistics are used to determine the next sampled event. In case of the dyadic REM, dyadic statistics are used to sample the next event whereas in case of actor-oriented model, actor-focused statistics are used to sample the sender of the next event and dyadic statistics are used to make the choice of the receiver given the sender.

3   Assessing goodness-of-fit through simulations

In the following we will present three extensive applications that highlight why relational event simulators are essential for temporal social network research. We begin with assessing goodness of model fit through simulations.

Typical approaches to evaluating goodness-of-fit of relational event models involve a balance of parsimony and accuracy, i.e. choosing a model with fewest parameters and with satisfactory accuracy for predicting events. This usually involves using information criteria such as the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), or other likelihood-based goodness-of-fit measures. These criteria are useful as a relative measure of fit when comparing models but they do not give an indication of the model fitness in an absolute sense neither do they provide a substantive direction of misfit. For example, when comparing multiple models using general information criteria, the "best" model may still have a poor fit to the data in terms of capturing important network characteristics, however, and it may result in poor predictions. On the other hand, simulation-based methods can be used to assess model fit in an absolute sense by assessing whether important network characteristics in the data are also present in simulated data using the fitted model. If simulated sequences based on a fitted model bear little resemblance to the observed sequence (or misses important characteristics of the network dynamic), this generally suggests a poor fit of the model to the data. Misfit can occur due to inclusion of effects that are not operative in the data or due to exclusion of key effects or external factors that play an important role in the network dynamics. Inferences based on misfitted models are therefore (theoretically) unreliable, and simulation-based methods are needed to assess that the model results in a reasonable fit to the observed data sequence before drawing inferences.

Simulation-based methods for model assessment have also been extensively used in other directions of social network research. Hunter et al. (2008) introduced simulation-based goodness-of-fit tests for social networks under exponential random graph models (ERGMs) using structural indices such as degree distributions, edgewise shared partners, and geodesic distance distributions. Snijders and Steglich (2015) introduce a variety of structural fit indices to describe cross sectional data such as the size of the largest component, number of components, median geodesic distance, transitivity coefficient, variance of in or out degree divided by mean, correlation between in and out degrees, graph hierarchy and, least upperboundedness. Lospinoso and Snijders (2019) include higher order indices in the form of triad census to assess if the simulated networks accurately resemble triadic structures in the observed network. Wang et al. (2022) describe behavioral indices to assess fit for models that jointly model behaviour and network structure. Brandenberger (2019) considered the accuracy of event predictions to assess the goodness-of-fit of relational event models for political networks. Because relational event models not only model the network structure but the timing and order of events as well, goodness-of-fit tests would be incomplete without assessing the fit of candidate models on temporal indices. It is possible that the model fits data well within certain time periods but not in others or that events between certain dyads in the predicted dataset are further apart than in the observed network. In addition, it is possible to compute structural indices that respect the timing and order of events in relational event datasets, such as temporal degree, temporal reachability, and temporal betweenness (Fal-

zon et al., 2018). Nicosia et al. (2013) similarly define temporal connectedness, various temporal centrality indices and temporal shortest paths for time-respecting paths in dynamic networks. These temporal indices enhance the fit indices used for static or longitudinal networks, by also incorporating the timing information available from relational event datasets. The measures mentioned above are just some of the fit indices used in social network literature; there are abundant other measures that can also be considered. It is impossible to define an exhaustive set of characteristics for general use with the relational event model. Rather than searching for the best fitting model, we believe researchers should focus on a model that reproduces specific aspects of the data that are of theoretical or practical interest rather than applying a default set of statistics. Ultimately, the choice of the fit indices should depend on the specific phenomena the researcher intends to address in their analysis.

3.1 General methodology

Suppose the goal is to assess the goodness-of-fit of a candidate model M_c for the observed sequence $\mathcal{E}$ on a set of fit indices $f(\mathcal{E})$. The fit indices can be a single value (in case the dynamic network is aggregated) or a vector of values measured during various time intervals across the observation period. The steps for evaluating goodness-of-fit using this approach are:

1. Identify relevant indices f of the observed sequence against which the goodness-of-fit will be evaluated.

2. Compute the corresponding indices on the observed sequence, i.e. compute $f(\mathcal{E}^{obs})$ on $\mathcal{E}^{obs}$.

3. Estimate the parameters $\hat{\beta}_{M_c}$ of the candidate model M_c.

4. Generate L sequences $E^{sim} = \{\mathcal{E}_1, \mathcal{E}_2 \ldots \mathcal{E}_L\}$ from the estimates $\hat{\beta}_{M_c}$ using the simulation frameworks described in Section 2.

5. Compute the fit indices $f(\mathcal{E}_l)$ for each of the simulated sequences $\mathcal{E}_l \in E^{sim}$.

6. The resulting distribution of L goodness-of-fit indices $f(\mathcal{E}_1), f(\mathcal{E}_2) \ldots f(\mathcal{E}_L)$ can be compared with $f(\mathcal{E}_o)$ visually (Hunter et al., 2008) or quantitatively (Lospinoso and Snijders, 2019; Chen and Onnela, 2019) to evaluate the fit of M_c on the observed data. If the computed fit indices (of their distributions) on the L sequences differ greatly from $f(\mathcal{E}^{obs})$, it can be concluded that M_c does not fit the data appropriately and further exploration is needed to determine where the model may be lacking.

In the context of relational event models, this approach to evaluating goodness-of-fit (i) allows researchers to test the model's fit across a set of interpretable structural and social indices, and (ii) guides the selection of network effects to be included in the model by pointing the researcher towards a direction in which the model may be misfit or lacking in specification.

This approach to evaluating the fit is very flexible in terms of the indices against which the fit can be evaluated. The indices can be global network properties such as network density, degree variances, or local properties such as reciprocated ties, or transitivity. They can be based on social behaviour, for example, homophily on specific attributes, subgroups in data, or proportion of actors in a network exhibiting behaviours

such as smoking, alcohol use, et cetera, or can be purely structural such as clustering or connected components. The flexibility also allows researchers to monitor these indices across all time points or networks aggregated over sub-intervals of the observation period.

3.2 Example: Assessing goodness-of-fit of an email dataset

We use a subset of publicly available email data from the Enron Corpus (Klimt and Yang, 2004) to demonstrate the evaluation of goodness-of-fit through simulations. Our purpose is not to provide a comprehensive analysis of goodness-of-fit as this depends on the research question, but to showcase the usefulness and flexibility of simulation techniques for assessing goodness-of-fit for relational event models. The Enron Corpus comprises time-stamped email exchanges between employees of the Enron corporation, a former energy, commodities, and services company. Email datasets are widely used in social network research, including relational event modeling, organization studies, and network science (Mulder and Leenders, 2019; Quintane et al., 2013; Kleinbaum, 2012; Hardin et al., 2015; Estévez-Mujica and Quintane, 2018). The Enron dataset has been used previously for analysis with relational event models (Perry and Wolfe, 2013; DuBois et al., 2013). For clarity of exposition, we consider a subset of the actors only belonging to either the Legal or Trading department. This subset consists of 60 employees and contains information on the gender, title, seniority, department, and division of these employees. We split the emails with more than one recipient as individual emails between the sender and each of the recipients. Further, we discard emails that contain more than four recipients in order to avoid including broadcast emails that are sent to a large number of receivers as they are covered by very different dynamics than emails exchanged in a smaller group size (Perry and Wolfe, 2013). A portion of the dataset from January 1, 2001 - August 30, 2001 containing approximately 5000 events is used to illustrate the application of simulation-based goodness-of-fit assessment. For this example, we do not aim to necessarily provide a well-fitting model of the data, but rather will provide a sensible model that might well coincide with an initial model of a researcher who wants to fit a relational event model to this data. This allows us to focus our attention on how simulation could guide the goodness-of-fit process and how it could point the researcher to the strengths and weaknesses of the fitted model.

We compare the goodness-of-fit for two REM models. Both models attempt to capture the communication dynamics in the data through endogenous network effects such as inertia, reciprocity, degree and participating shift effects. See Butts (2008) for the specification of these statistics. Additionally, the models include exogenous effects such as homophily, i.e. the tendency for actors to interact with other actors sharing the same attributes. The difference between the two models is that one of them accounts for the differences between the impact of events that occurred far in the past versus more recent events. In email interactions, in particular at a corporate organization, the recency of emails (received and sent) plays a crucial role in determining the dynamics of events. We incorporate this into the model following Perry and Wolfe (2013) by specifying an interval (δ_1, δ_2) in the past $(t - \delta_2, t - \delta_1)$ with $0 \leq \delta_1 < \delta_2 \leq t$ such that only events that occurred during the interval are counted towards the computation of the statistics. We consider intervals $(0, 1), (1, 7)$ and $(7, \inf)$ days. Therefore emails in the past day, week, and beyond are considered separately in computing endogenous statistics. Consider $A(t_1, t_2)$ to be an $n \times n$ adjacency matrix, where n is the

number of actors and each element of the adjacency matrix $a_{ij}(t_1, t_2)$ represents the count of events observed for the dyad (i, j) in the time interval $[t_1, t_2)$. The definitions of the endogenous statistics and their 'interval' counterparts can be formulated as:

inertia: $\qquad\qquad\qquad\qquad\quad X(i, j, t) = a_{ij}(0, t)$

interval inertia: $\qquad\qquad\qquad\; X(i, j, t, \delta_1, \delta_2 \mid 0 \leq \delta_2 < \delta_1 \leq t) = a_{ij}(t - \delta_1, t - \delta_2)$

reciprocity: $\qquad\qquad\qquad\qquad X(i, j, t) = a_{ji}(0, t)$

interval reciprocity: $\qquad\qquad\; X(i, j, t, \delta_1, \delta_2 \mid 0 \leq \delta_2 < \delta_1 \leq t) = a_{ji}(t - \delta_1, t - \delta_2)$

out-degree sender: $\qquad\qquad\; X(i, j, t) = \sum_{k \in \mathcal{A}/\, i} a_{ij}(0, t)$

interval out-degree sender: $\quad X(i, j, t, \delta_1, \delta_2 \mid 0 \leq \delta_2 < \delta_1 \leq t) = \sum_{k \in \mathcal{A}/\, i} a_{ij}(t - \delta_1, t - \delta_2)$

in-degree receiver: $\qquad\qquad X(i, j, t) = \sum_{k \in \mathcal{A}/\, j} a_{kj}(0, t)$

interval in-degree receiver: $\;\; X(i, j, t, \delta_1, \delta_2 \mid 0 \leq \delta_2 < \delta_1 \leq t) = \sum_{k \in \mathcal{A}/\, j} a_{kj}(t - \delta_1, t - \delta_2)$

To assess the goodness-of-fit of the models, we first start with the classical measures of goodness-of-fit. The Akaike information criterion (AIC) of the model without interval statistics is 29810 and that of with interval statistics is 28423. Further, the Bayesian information criterion (BIC) of the model with the interval statistics is 28573 and 29921 without. The results suggest that the model with the interval statistics fits the data relatively better than the model without. However, we do not learn about the overall fit and the adequacy of the two models in capturing key characteristics of interest. We also do not gain any insight into how we could potentially improve the fit of the models. A simulation-based goodness-of-fit approach can be utilized in addition to the AIC, BIC measures. To do this we need to select indices that we want the model to reproduce from the observed network. The indices that one would use in a full-fledged empirical study would depend on the research question at hand, the theory being employed, and the baseline against which the researcher wants to compare the model. The indices we have chosen to calculate for this example are meant to provide examples of different kinds of indices one might want to consider in an actual empirical study of this data.

In Figure 1 (a) and (b) we see the in-degree and out-degree distributions of observed and simulated sequences under the two models with and without interval statistics. Both models reproduce the degree distributions in the observed dataset poorly indicating that there might be a mis-fit in the models; therefore, if replicating the degree is important to the researcher, more statistics or attributes need to be included in the model that will attempt to better capture the degree distributions of the data. Figure 1 (e) and (f) report the interevent times for events exchanged in the legal and trading department of the organization. The model with the interval statistics is able to replicate the departmental inter-event times better than the model without interval statistics indicating that including the interval statistics did indeed improve this temporal aspect of the model. Figure 1 (c) shows the proportion of inter-departmental events exchanged in the network per week. We see the model with interval statistics again fits the observed data slightly better than the model without but the real dynamics is is not captured by either model. In Figure 1 (d) we report the proportion of reciprocated ties in the network computed every two weeks. The proportion of reciprocated ties is computed as the proportion of dyads that are symmetric. In our data this includes the dyads with actors who have sent and reciprocated

at least one email to each other in the past two weeks. Both models replicate the observed reciprocated ties poorly as they generate inter-departmental events that are fairly stable. This is expected as neither the parameters nor the attributes change over time in the fitted model which generates a stable sequence of interactions over time. Any researcher with data that changes considerably over time, should bear in mind to include time varying effects or attributes or other temporal aspects that may account for the deviation from the equilibrium. Thus we conclude that simulation-based goodness-of-fit assessment provides a fine-grained understanding of the fit of relational event models and allows researchers to evaluate potential areas where the model specification could be improved further.
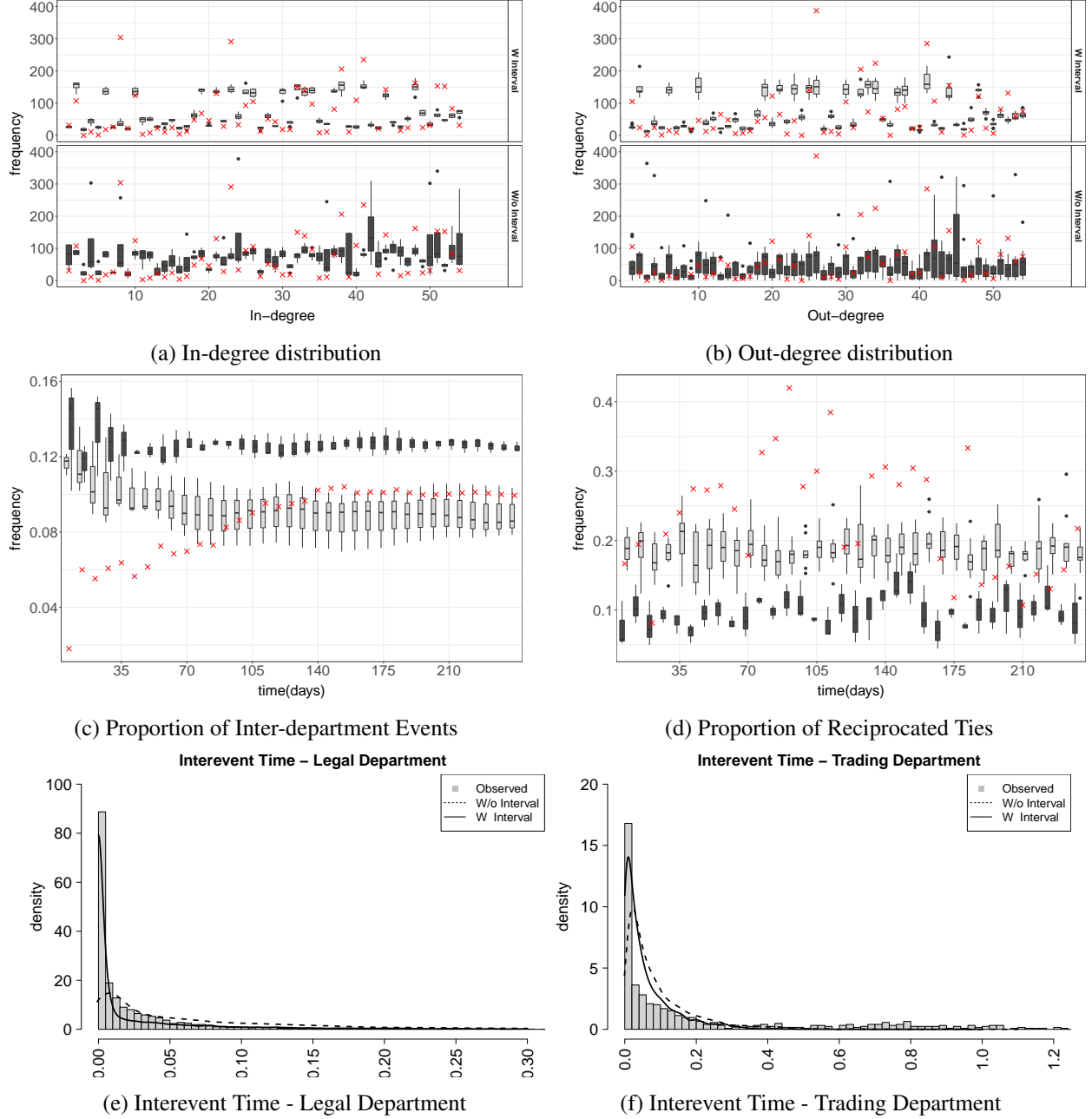
(a) In-degree distribution

(b) Out-degree distribution

(c) Proportion of Inter-department Events

(d) Proportion of Reciprocated Ties

(e) Interevent Time - Legal Department

(f) Interevent Time - Trading Department

Figure 1: Goodness-of-fit comparison. Each figure plots the values of the fit indices over 50 simulated sequences for each model.

4 Developing theories using simulations

The second important application of the relational event simulations we present here involves the development of social theories using simulations. Many social phenomenon arise from repeated interactions among individuals in a social network over time. Undoubtedly, researchers continue to gain insights into the drivers of repeated social interactions through empirical studies. The relational event model itself has contributed

greatly towards modelling the data arising out of such empirical research. While this is useful for estimating the magnitude of network effects or confirming differences between different groups in specific empirical settings, fitted REMs are typically based on theoretical ideas of the researcher and contribute to testing existing theories rather than being employed as a way to develop or refine (new) theory. Simulations can provide a powerful and flexible tool for (further) developing or expanding theories (Davis et al., 2007) regarding why, how, and when interactions occur in a social network. The proposed relational event simulation frameworks provide flexibility to combine multiple network effects such as transitivity, homophily, or preferential attachment in one theoretical model, allowing researchers to test and develop richer theories about social interaction dynamics. Below we first describe how we can (further) develop social network theories using relational event simulations, and subsequently, we present an extensive application of the methodology to better understand group formation using optimal distinctive theory.

4.1 Building, evaluating, and extending social network theories

Relational event simulations can be utilized to develop theories to a) study the emergence of social phenomena, b) evaluate boundary conditions of theories, and c) incorporate timing and dynamism in theories. We briefly elaborate on this below.

Study the emergence of social phenomena. Emergence is a fundamental concept in social sciences, referring to the way collective phenomena arise from the interactions of individuals (Kozlowski and Klein, 2000). Social phenomena can be considered emergent when they cannot be reduced to the behavior of individuals, but instead emerge from complex interactions among them. Relational event sequences can help explain elements of the emergent properties that occur on a global level but are a consequence of the events, their sequence, timing and patterns associated with event occurrence.

Simulations offer a powerful tool for studying the emergence of social phenomena (Epstein and Axtell, 1996; Gilbert and Troitzsch, 2005), especially those that are unexpected or counter-intuitive. By simulating relational event models, researchers can generate scenarios that test the boundaries of existing theories and develop new insights into the mechanisms underlying emergent phenomena. In this sense, simulations offer a natural way of describing complex dynamics as a combination of familiar network mechanisms. Relational event simulation frameworks are particularly useful in studying emergent properties because they can capture the dynamic nature of social phenomena over time. These frameworks allow researchers to model how the behavior of individuals within a social network can lead to collective outcomes that are not predictable from the behavior of any individual alone. Through relational event simulations, researchers can study the effects of different network structures, parameters, and mechanisms on the emergence of social phenomena. By leveraging the rich data available on relational events and using simulation techniques, researchers can gain new insights into the mechanisms underlying emergent phenomena and develop more accurate and effective theories.

Evaluate boundary conditions of theories Evaluating boundary conditions, or scenarios under which a theory makes sense is critical for developing and advancing theories. Unfortunately, little is known about the boundary conditions of most social science theories. Knowing the boundary conditions of a specific theory

15

can help guide setting up proper empirical experiments or the appropriate use of the theory in an empirical study. In our context, boundary conditions could refer to ranges of the parameter space of network effects, constraints on values of statistics, size of the network, or exogenous conditions that need to be present for a theory to be applicable. Also, knowing that a particular theory only can make valid predictions into the short future or only under fairly stable conditions informs the researcher how far into the future one can predict based on a fitted model or how much history must be taken into account in predicting future relational events in accordance with a specific theory. Simulations can be a powerful tool to explore the applicability of theories under various such conditions.

The approach here is to take a specific theory about social interaction that one intends to use in a research project and translate that into a relational event model. By simulating network interaction patterns across a range of parameter values or initial conditions, it can be assessed beyond which (combinations of) values the model starts to generate non-sensical or unrealistic dynamics such as the emergence of strongly separated subgroups, unrealistically high density, hyperactivity of actors with certain traits, odd speeding up and slowing down of interaction rates, etc. When appropriate characteristics of the simulated relational event network no longer pass the relevant sanity checks (either by being obviously non-sensical or by being quite removed from the dynamics of empirical reference data), it becomes clear that the theory that is mimicked by the relational event simulation no longer is realistic beyond these (combinations of) parameter values or initial conditions. This, then, provides the boundary within which a researcher would like to stay when using the theory to explain a phenomenon of interest.

Incorporate timing and dynamism. The proposed simulation frameworks allow researchers to incorporate dynamism when developing theories by including feedback loops that are realized through endogenous statistics, time-varying exogenous covariates when attributes of actors and dyads can vary with time, time-varying parameters that can simulate realistic change-points, and memory when the influence of past events on new events decays with time. One advantage of this is that a theorist can assess what happens when a static social theory (which almost all established social theories to date are) is reformulated in an explicitly dynamic, time-sensitive manner. Although many of our current social theories are built on inherently dynamic ideas and arguments, they are rarely formulated (and validated) as such. It is often not at all obvious how to turn established static social theories into time-sensitive theories, since we have very little knowledge of how fast interactions develop, how long it takes until routine sets in, or how long emergent properties last. Simulating several dynamic versions of current social theories can provide hints to a theorist as to how an explicitly time-sensitive and dynamic version of the theory could be developed and put to the test.

4.2 Group formation using Optimal Distinctiveness Theory

As an example, we present a simulation approach to represent the Optimal Distinctiveness Theory using relational event simulations. The Optimal Distinctiveness Theory of Brewer (1991) dictates that individuals have two fundamental and competing needs: 1) the need for inclusion, i.e., a desire to assimilate and interact with other individuals who share their social identity and 2) the need for distinctiveness from others in the actor's surroundings. Individuals prefer to be identified with social groups that are neither too inclusive

nor too distinctive, but are of optimal distinctiveness. These competing psychological mechanisms that operate at the actor level motivate the emergence of social groups that satisfy both these needs at the same time (Leonardelli et al., 2010). Drawing upon optimal distinctiveness theory, we design an actor-based simulation experiment to observe the emergence of groups and their dynamics based on actor's competing needs for optimal distinctiveness. We use an actor-orientated simulation approach, keeping in mind that the actor retains agency to decide if and when to send events to other individuals in the network in accordance with the theory. Our simulation model assumes that actors make decisions about when to send an event in an attempt to satisfy their preference for interactions that are optimally distinct. Each actor i has an attribute z_i (that reflects their fixed identity) and a preference for the optimal value of distinctiveness associated with that actor d_i^*. This value reflects the desired proportion of actors that i interacts with who do not share the same identity as i.

The distinctiveness of an actor d_i at any time t is quantified with a proportion of distinct events, i.e. the total number of incoming and outgoing events that the actor i sends and receives from other actors who have an attribute value different from i's divided by the total number of events sent and received by i from all other actors irrespective of the attribute:

$$d_i(t) = \frac{\sum\limits_{j \in \mathcal{A}/\{i\}} (a_{ij}(t) + a_{ji}(t)) \, \mathbb{I}[z_i \neq z_j]}{\sum\limits_{j \in \mathcal{A}/\{i\}} (a_{ij}(t) + a_{ji}(t))}, \tag{7}$$

where $a_{ij}(t)$ is the corresponding element of an $n \times n$ adjacency matrix $A(t)$ containing the total number of past events between actors at time t. Recall from Section 2.2 that the DyNAM model requires two separate statistics for the sender and receiver choice respectively. The distinctiveness aspect of an actor's interaction preferences is incorporated into the simulation framework as a combination of a sender dissatisfaction statistic X^s and a receiver choice statistic X^r.

Specifically, we operationalize, a sender i's dissatisfaction statistic $X^s(i, \mathcal{E}_t)$ as the absolute difference between the current distinctiveness and the optimal value of distinctiveness the actor strives for $|d_i(t) - d_i^*|$. A positive effect of the distinctiveness statistics implies that actors who are most dissatisfied with their current level of distinctiveness are most likely to be activated as the next sender in an attempt to adjust their distinctiveness level. We assume that for any actor, the direction of deviation from their optimal distinctiveness, i.e towards more or less distinctiveness does not influence the absolute level of dissatisfaction. Moreover, dissatisfied actors with a lower distinctiveness than their optimal level ($d_i(t) < d_i^*$) are likely to reach out to actors with attribute levels distinct from their own. Similarly, dissatisfied actors with a higher distinctiveness than their optimal level ($d_i(t) > d_i^*$) are likely to reach out to actors with the same attribute level as their own. If all the actor's are equally satisfied with their current distinctiveness level, an equilibrium will be maintained where all actors are equally likely to be sampled as the next sender (Leonardelli et al., 2010).

For the receiver choice model of the DyNAM, a receiver choice statistic can be defined as follows:

$$X^r(j|i, \mathcal{E}_{t_m}) = \begin{cases} \mathbb{I}[z_i = z_j] & d_i(t) > d_i^* \\ \mathbb{I}[z_i \neq z_j] & d_i(t) < d_i^* \end{cases} \tag{8}$$

In case $d_i(t) > d_i^*$, i.e., when sender i's interactions are more distinct than desired, the receiver statistic assigns a value of 1 for potential receivers with attribute the same as that of actor i. Similarly in case $d_i(t) < d_i^*$, i.e., when actor i's interactions are more assimilated than desired, the receiver statistic assigns a value of 1 to receivers with attributes distinct from i's. Otherwise the statistics is set to 0. We further include an interaction of the above-defined statistic with an inertia receiver choice statistic in our model to ensure that receivers who share a greater volume of past events with a sender are more likely to be selected by the sender.

In our simulations, we assume that all actors have the same value of optimal distinctiveness d^*. We simulate networks consisting of 30 actors with a binary attribute that represents their fixed identity $z = \{1, 0\}$ for varying values of d^*. The proportion of actors with each attribute is equal. To identify emergent groups formed as a result of the simulations, we use the Louvain community detection algorithm Blondel et al. (2008). This method of community detection essentially partitions nodes based on a greedy modularity optimisation. The groups identified from this method are considered emergent because they arise from interactions and decisions made by actors on the actor level which lead to emergent group formation on a network level. Figure 2 depicts the network plots for simulated sequences with different d^* values. When $d^* = 0$ as illustrated in Figure 2 (a), actors in the simulation strongly prefer homophilic interactions, resulting in groups containing nodes of one attribute type. Actors with opposing attributes tend to cluster together rather than interact with actors of a different attribute.

For $d^* = 0.3$, shown in Figure 2 (b), a mixed grouping is observed, with some larger groups containing both attribute types and some groups being homogeneous. Actors of opposing attributes interact with each other more often, and the distance between them is shorter compared to the plot for $d^* = 0$. Figure 2 (c) for $d^* = 0.5$ shows a mixed grouping with equal proportion of squares and circles in the groups as well as a roughly random communication structure. The actors form a highly connected network, and actors with opposing attributes are roughly as close to each other as actors with the same attribute. In Figure 2 (d) where $d^* = 1$, there is a strong preference for distinctiveness among actors in their interactions, leading to a plot with mixed groupings. However, it is worth noting that the interactions within each group are bipartite, meaning that squares are only strongly associated with circles and vice versa.

These network plots pass our sanity checks, indicating the specific way we translated the optimal distinctiveness theory into the relational event model is sound and does not produce unexpected results. We can further test the boundary conditions of the theory through simulations. In the previous simulations, the attributes were equally distributed between the actors; however, this may not often be realistic. There may be an imbalance in the proportion of actors for instance when dealing with integration of minority communities or under-represented sections of society at a university. To understand the effects of proportions of actors with different attributes on the network under the optimal distinctiveness theory, we simulated networks with 50
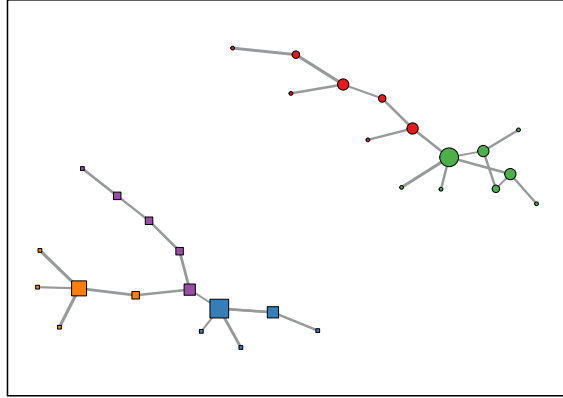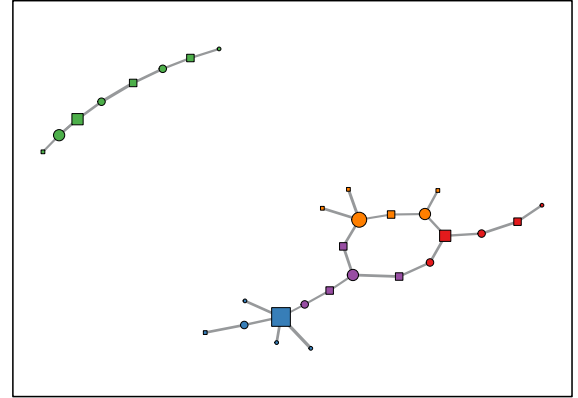
(a) $d^* = 0$
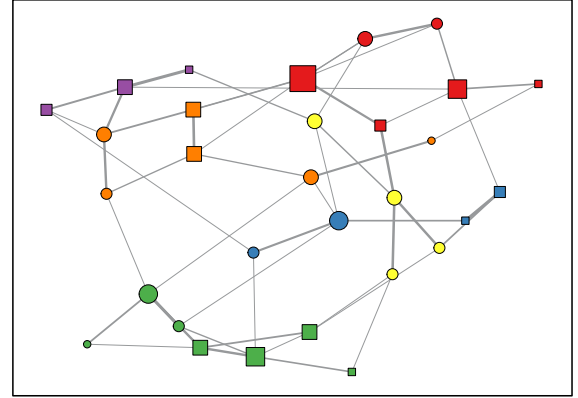
(b) $d^* = 0.3$
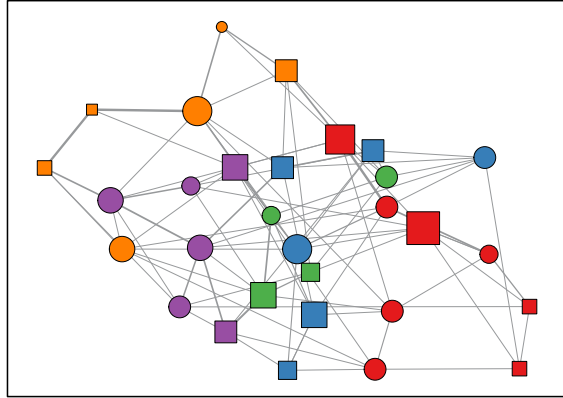
(c) $d^* = 0.5$

(d) $d^* = 1$

Figure 2: Simulated Network plots for different optimal distinctiveness d^* values. Square nodes represent nodes with $z = 1$ and circular nodes represent nodes with attribute value $z = 2$. The colours indicate the emergent groups to which the nodes were assigned based on the community detection algorithm. The size of nodes corresponds to the degree of the actors, and the width of the edges corresponds to the total volume of events exchanged in either direction between the two nodes.

actors, and varied the proportion of the minority within the population. The simulations were repeated 50 times for each combination of proportion of minority actors $p \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ and for each d^* values in the interval $[0, 1]$.
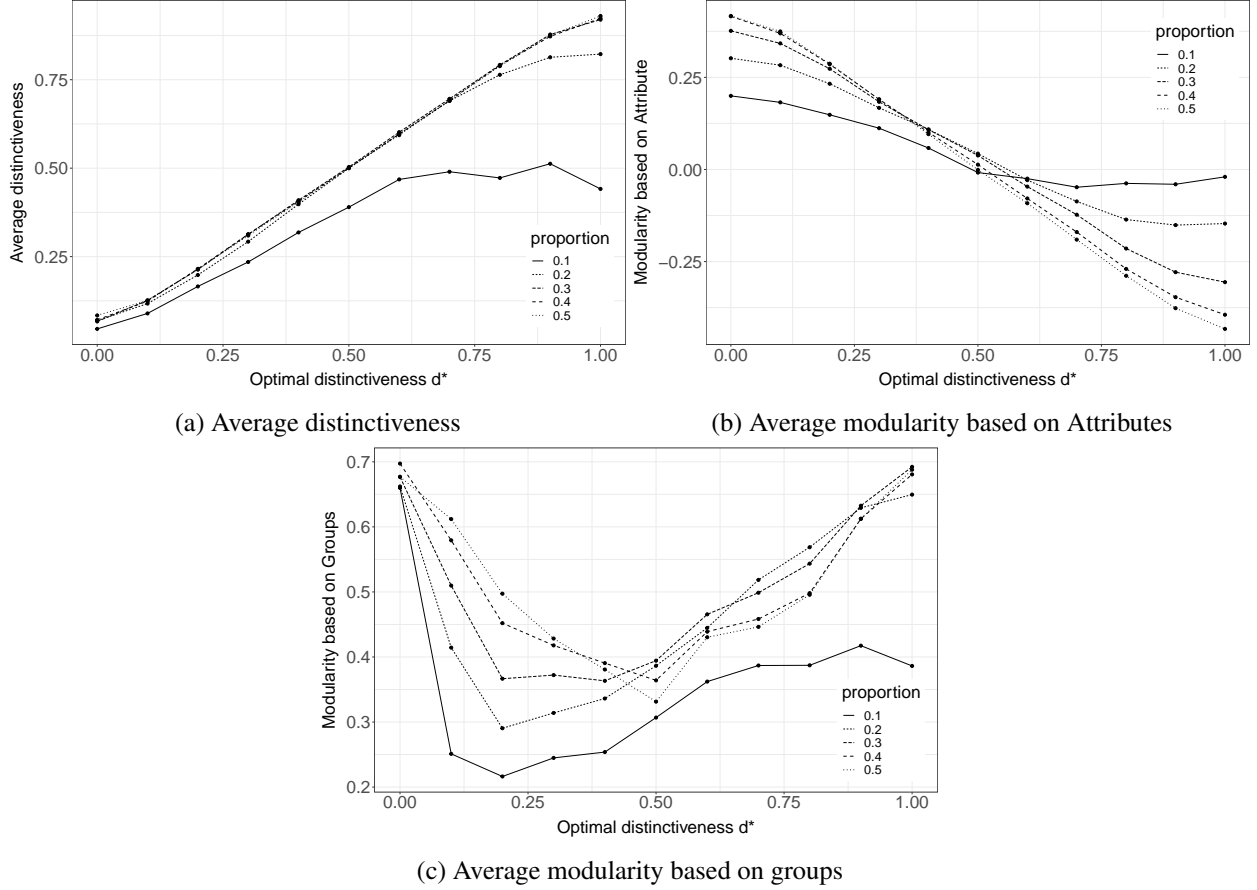

(a) Average distinctiveness

(b) Average modularity based on Attributes

(c) Average modularity based on groups

Figure 3: (a) Average distinctiveness of actors, (b) Average modularity computed based on the partitioning of nodes by actor attribute and, (c) Average modularity computed based on the partitioning of nodes by the detected groups at the end of 50 simulation runs for each value of optimal distinctiveness $d^* \in [0, 1]$ and for each proportion of minority attribute.

Figure 3 (a) reports the average distinctiveness at the end of the simulations for the simulated event histories for the corresponding p and $d*$ values. The results indicate that the average distinctiveness of the actors at the end of the simulations could converge to the optimal value when the proportions of the minority are greater than 10%. However, when $p = 0.1$ the actors do not achieve their desired distinctiveness. When a minority is too small in size relative to the population, the actors are unsatisfied in their need for distinctiveness particularly when d^* is high.

Figure 3 (b) presents the modularity values (Newman, 2006) computed based on the partitioning of nodes by actor attribute at the end of the simulations. The value of the modularity reflects a measure of the strength of the partitions based on actor attributes, i.e. when modularity is high actors with the same attributes tend to cluster together and have a denser communication structure. The results show that with increasing d^*,

20

the modularity values decrease. Indicating that at $d^* = 0$ i.e. the actors in the simulation preferred only homophilic interactions, the simulated network indeed has high modularity. Whereas at $d^* = 1$, when the preference for distinct interactions is at its peak, the modularity value is at its lowest. The decrease in modularity from $d^* = 0$ to $d^* = 1$ is less severe for lower proportions of minorities.

Figure 3 (c) on the other hand, presents modularity values computed based on partitioning of nodes, based on the groups detected by the community detection algorithm. When $d^* = 0.5$ the modularity is at its lowest because actors don't have a preference for distinctiveness. However on each extreme i.e $d^* = 1$ or 0 the clustering is high because actors have a strong preference for interacting with actors with specific attributes leading to tighter group formation. The modularity values for proportion $= 0.5$ is fairly symmetric around $d^* = 0.5$ however for lower proportion, the symmetry is altered with the minimum shifted to lower d^* values. A possible explanation of this is that when d^* is higher and proportion of minority actors present in the network are lower, the clustering within a group is not strong due to the majority actors fulfilling their desired for high distinctiveness from minority actors outside the group. The minimum at lower d^* and lower p values occurs when the groups are formed around minority actors, thus the majority actors are able to fulfill their distinctiveness criteria within their groups.

Our purpose was to illustrate through a simple example *how* an interested researcher could articulate theories using the REM simulation framework. The results indicate how individual preferences of actors, in this case their desired distinctiveness, led to emergent group formation on the network level. We further tested a simple boundary for the theory by varying the proportion of minorities in the network. We see that dissatisfaction across actors is greater when the size of the minority group is skewed toward a very small minority presence compared to societies with larger minority groups, where actors are able to achieve their desired optimal distinctiveness. Future work remains to be done to investigate the impact of including more than two attributes or when varying the optimal distinctiveness in the population where certain actors have a higher desire for distinctiveness than others. Further we assumed that the direction of dissatisfaction does not influence who is likely to be the next sender. We did not distinguish between actors dissatisfied with their distinctiveness and actors dissatisfied with their assimilation as long as the absolute value of dissatisfaction was equal. In reality, it is possible that these two cases may have a different impact on the probability of the next sender. Future work is needed to evaluate this. In addition, we also assumed that all actors had free choice and no restrictions in their communications, however, in reality spatial or cultural restrictions could influence the interactions of actors in different ways and needs to be taken into account.

5 Planning network interventions via simulations

Network interventions describe "the process of using social network data to accelerate behavior change or improve organizational performance" (Valente, 2012). Interventions involve the use of social network data to bring about a behaviour change or to influence social dynamics in a network towards desirable outcomes. Previous work in network interventions focuses on adoption of behaviour and transmission of the behaviour across neighbours of nodes in a network often using diffusion models (Valente, 2005, 2012, 2017; Badham et al., 2021). These models assume that behaviour is propagated in a network and once a node adopts a behaviour, it can influence it's neighbour to adopt that behaviour. The edges between actors in such a network can represent dissemination of information, opinions, ideas, goods or even diseases. The variables of interest in diffusion models are often the proportion of nodes that have adopted a desirable behaviour or have acquired certain knowledge. However, in relational event models, the variable of interest is the rate of interaction that describes who interacts with whom and when. REM or DyNAM can be utilized to model interventions in evolving interaction networks where the intervention aims to influence the interaction dynamics itself.

The relational event simulation framework can be availed to predict or better understand the potential impact of the intervention outcomes over time. This can be of great practical use, for instance by allowing managers to try out several interventions "in-silico" before deciding which one(s) to try out "in-practice"; the effects of competing interventions can be assessed without spending much money or other resources and without overhauling an organization based on a mere managerial hunch. Rather, a manager can get a fairly good idea of which intervention is likely to work, how it should be implemented, and what the pitfalls are that need to be monitored when the intervention is implemented in real life.

Ideally, the simulation is initialised from an estimated model based on an empirical observed network. This allows a realistic simulation of the intervention based on empirically derived conditions and effect estimates. Adams and Schaefer (2016) explored how the same intervention can have different results based on the initial conditions of the network. This is also applicable in the REM context. For example, consider an intervention that is conducted for a while and aims to promote inter-departmental collaboration in an organizational network. If the network is characterized by high tendencies towards inertia, the inter-department interactions that were originally triggered by the intervention (that only ran for a limited time) may become sustained into the long term due to the network's tendency for inertia. On the other hand, if inertia is only low (or negative) in this network, it is possible that, once the intervention ends, the network reverts back to its original state. Having an estimate of the effects operational in the empirical network before-hand would be highly beneficial in planning interventions. Simulations could then be carried out to evaluate the outcomes of planned interventions that explore the implications of a manipulated set of effects or initial conditions.

5.1 Types of network interventions

In order to effectively simulate realistic interventions using the simulation framework, it is important to understand the types of interventions that can be simulated in the relational event context and how the

mathematical modification to the model translates to implementation in practise. To be able to effectively translate an intervention into the REM or DyNAM model, it is important to define which aspects of the model the intervention modifies. To facilitate this, we distinguish between three types of REM interventions: i) Actor Attributes, ii) Network Effects, and iii) Composition - based on the locus of change in the model.

- Attributes: Interventions may involve modifications to attributes of actors or relationship between actors (dyadic attributes) to answer questions such as: can changing the layout of desks lead to a change in communication patterns of students in a classroom? What is the effect of changing the gender or hierarchy distribution on the communication in an organisation? The analytical translation of such interventions are straightforward because they involve simply modifying the actor or dyadic attributes while keeping other aspects of the model the same as pre-intervention.

- Network Effects: Interventions of this type focus on modifying network effects of the model itself by either changing the magnitude of network effects (by adjusting the corresponding parameters) or by removing (or adding) network effects that may (or may not) be operative in the social network previously. For example, consider an intervention that is carried out by organising social events designed to increase mixed-gender contact opportunities among university students. The analytical translation of this intervention could be a temporary increase in the magnitude of the gender heterophily effect in the duration of the organised events. Simulating events beyond the intervention duration could reveal the long-term effects of the designed intervention on ties between the students. Another example of this type of intervention is reduction in popularity of a smoker as students in a university are made aware of the harmful effects of smoking. The translation of this effect could be a decrease in magnitude of the effect for the smoker popularity statistic (interaction of the smoker attribute and the in-degree of the receiver).

- Structure: Interventions that change the structure of a social network involve changes to size of the network by addition or removal of actors, the riskset by restricting which pairs of actors may interact, and change in the composition and grouping of actors. Examples of structural changes include addition of employees when companies merge or acquire one another, removal of a key leader in an organisation, and reassignment of teams and managers when new projects are started. The analytical translation of these interventions are straightforward and often involve making alterations to the riskset. In case of re-assignment of teams for instance, dyads with employees from different teams or locations may no longer be considered in the riskset.

The three categories above are not disjoint, as interventions in practise can involve multiple categories. For example, consider an intervention that involves the addition of sport activities in a university. Students and staff from different departments would now have an opportunity to interact with each other during these events. The analytical translation of such an intervention involves all the three categorizations defined above. Actor attributes that represent whether or not individual actors participate in sports and the type of sport if they do, are added to the model. A new network effect can be added for homophily interactions that may take place between actors attending the same sports sessions. Additionally, the structure of the riskset may

also be altered to include dyads from different departments that participate in the same sports that may not previously be considered at risk of interaction.

5.2 Evaluating the Persistence of Intervention Outcomes

In this example, our aim is to explore the longevity of outcomes of different interventions through simulations. With simulations, it is possible to assess how a network would react to competing interventions and assists in improving intended network interventions towards a desirable outcome. The intervention in this example is intended to increase the inter-department collaboration between employees of two departments in an organization. The number of inter-departmental emails is used as a proxy for the collaboration across departments. This intervention example is a simplified version of an actual intervention study that we conducted for the R&D department of a large consumer goods company in Western Europe. To increase the inter-department communication, the organisation intended to intervene by providing the employees with incentives to increase their inter-department communication (e.g., by organising new multi-department projects and organising joint meetings). This intervention is modelled in our example by temporarily increasing the departmental heterophily effect β_{het} in the model. Our interest is in evaluating what the effect is of the intervention after it ends and how long the increased heterophily effect remains in the network after the intervention has ended and the network has returned to its "normal" dynamic. Our assumption is that the intervention by itself will not structurally affect employee interaction norms and preferences. Therefore, after the intervention, we assume that the "pre-intervention" model is again applicable to the interaction among the employees (but the interaction histories will have been altered by the interventions).

Before testing the interventions using simulations, we first estimate the coefficients of a model based on interactions before the intervention, for a period spanning 243 days using a standard REM. The REM model contains various endogenous effects (represented by β_{endo}) such as inertia, reciprocity, participation shifts and degree effects. Exogenous effects (represented by β_{exo}) for the gender, seniority and sub-domain are also included. An additional effect for inter-departmental heterophily (β_{het} is also included and is the locus for the intervention.

The model fits the data reasonably well and we assume that it captures the way in which the company's employees tend to interact well. In the following, we will compare nine versions of the intervention to increase the inter-departmental homophily (by increasing β_{het}). The interventions differ in the length of time they run and in the strength of the intervention and we aim to evaluate the persistence of the intervention effects after the intervention has been carried out. Figure 4 graphically shows two examples. Intervention I lasts from t_1 to t_2 and for this period we increase β_{het} (the parameter for the heterophily effect) to β_2. Alternatively, intervention II runs longer than intervention I (from t_1 to t_2') and has a different heterophily parameter $\beta_{het} = \beta_3$.

Using the pre-intervention sequence as the initialization, we simulate the relational events in the network for the length of time the intervention runs and following the intervention, we further simulate the relational sequence for 50 weeks after each intervention. We consider three durations of intervention $\{28, 42, 56\}$ days and three inter-departmental heterophily effect sizes $\{0, 0.25, 0.5\}$. Therefore, we generate relational event
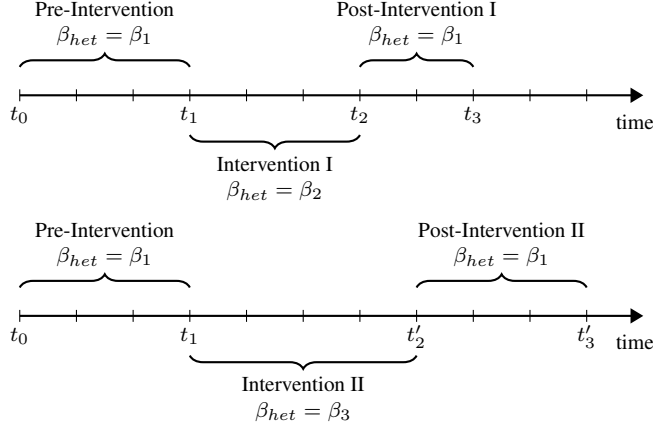
Figure 4: Diagrammatic representation of two interventions

sequences under $3 \times 3 = 9$ different interventions. Since the simulations have a stochastic component, we generate 50 relational event sequences for each intervention and compute the proportion of interdepartmental events in each. For each intervention, we report the median proportion in the 50 simulated sequences in Figure 5. The proportion of inter-department events increased most for interventions held for the longest period and for the strongest effect size. The effects of the interventions on the inter-departmental communication lasted 38 days for the case of $\beta_{het} = 0$, 91 days for the case of $\beta_{het} = 0.25$ and 208 days when $\beta_{het} = 0.5$ when the duration of the intervention period was 28 days. Similarly, the effects of the intervention last 61 days for $\beta_{het} = 0$, 247 days for $\beta_{het} = 0.25$, and 283 days for $\beta_{het} = 0.5$ when intervention duration was 42 days. Further, the impact of the intervention lasted 72 days for $\beta_{het} = 0$, 307 days for $\beta_{het} = 0.25$, and persist beyond our simulation period when $\beta_{het} = 0.5$ when intervention duration was 56 days. We conclude that relational event simulations can provide a useful tool in planning and designing network interventions as well as to evaluate the persistence of interventions outcomes over time.

5.3 Evaluating Intervention Targeting Strategies

Our second example concerns the situation where one would like to know whom to aim a specific intervention at. For example, in organisational networks, it may be prohibitively costly and complicated to intervene directly in everyone's work. Rather, it may often be effective to target an intervention at a limited set of employees only, acknowledging that their changed behaviour will subsequently affect that of their coworkers, causing the intervention effect to spread across the network naturally. Of course, the effectiveness of this will usually depend on which set of employees are targeted by the intervention. This is what we will show an example of here. Again, the intervention is intended to increase inter-departmental collaboration.

Keeping the duration of the intervention constant, we consider different targets of the intervention. Table 1 provides an overview of the nine different sets of targets that we consider. The "random targets" intervention is used as a baseline to compare the other, more specifically targeted, interventions against. Interventions 2-5 use the endogenous activity of actors as a criteria for target selection. Interventions 6-9 are based on exogenous attributes of the actors. We also include a scenario when no intervention is carried out.
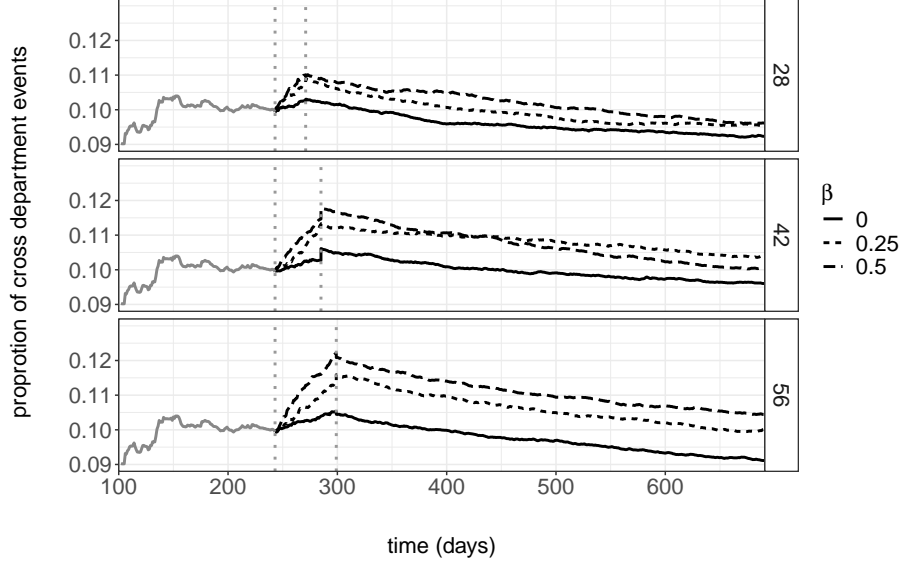
25

Figure 5: Median proportion of inter-departmental events simulated after an artificial intervention. The vertical dotted lines indicate the the intervention period. Top panel shows the simulations for an intervention lasting 4 weeks i.e 28 days in duration. Middle panel with duration 6 weeks i.e 42 days and Bottom panel with 8 weeks i.e 56 days.

The interventions in our example are carried out by modifying the department-heterophily parameter β_{het} only for the selected targets. Specifically, the departmental heterophily parameter during the intervention remains the same as its value pre-intervention ($\beta_{het}^{pre} = -0.58$) for dyads where the sender does not belongs to the target group of actors $\mathcal{Q}$. However, the heterophily parameter is increased to a higher value ($\beta_{het}^{intv} = 0.5$) for dyads where the sender belongs to the target group of actors. Thus, the heterophily effect is split into two groups with two different parameter values.

$$\beta_{het}(i,j) = \begin{cases} \beta_{het}^{pre} & i \notin \mathcal{Q}^K \\ \beta_{het}^{intv} & i \in \mathcal{Q}^K \end{cases}$$

In addition to comparing the different strategies for target selection, we also vary the fraction of actors (K) that are selected for each strategy (from the set of 60 employees). Figure 6(a) shows the reported median proportion of inter-department events simulated for the endogenous selection strategies, random selection, and no intervention for K = 10%, 20%, and 30%. For $K = 10\%$ and 20% actors selected based on having the lowest inter-department degree had the greatest impact on the proportion of inter-departmental communication while, selecting targets based on the remaining strategies impacted the proportion of inter-departmental somewhat equally as selecting random targets. When the $K = 30\%$, the low inter-departmental degree strategy is again leading followed by the high centrality. The remaining targeting strategies impacted the proportion of inter-departmental events somewhat equally. Figure 6(a) depicts the comparison among the exogenous attribute based strategies. Targeting seniors or females lead to the greatest impact for $K = 30\%$ whereas the strategies were roughly tied when fewer fraction of actors were selected i.e $K = 10\%$. In all

26

| Targets	Description
1. Random	Each actor has equal chance of selection.
2. Highest Centrality	Actors with the highest betweenness centrality are selected.
3. Lowest Centrality	Actors with the lowest betweenness centrality are selected.
4. Highest Inter-department Out-degree	Actors with the highest number of inter-departmental outgoing events are selected.
5. Lowest Inter-department Out-degree	Actors with the lowest number of inter-departmental outgoing events are selected.
6. Senior	Actors designated as 'Senior' in the dataset (based on their tenure at the organization) are selected.
7. Junior	Actors designated as 'Junior' in the dataset (based on their tenure at the organization) are selected.
8. Male	Actors designated as 'Male' in the dataset are selected.
9. Female	Actors designated as 'Female' in the dataset are selected.
10. No Intervention	No intervention was carried out.

Table 1: Overview of measures used to select intervention targets

cases targeting junior employees had least effect.

The simulation results indicate that targeting a smaller set of actors rather than the entire network indeed can lead to desirable results in that network. However, the selection of the intervention targets can impact the results considerably. In our example, we found that targeting juniors does not produce compelling results rather seniors in the organization should be targeted for promoting inter-departmental communication. Further targeting actors who have a lower inter-departmental communication rate can prove to be more impactful than actors with higher inter-department communication or even higher centrality. Hence, when designing an intervention in practise, simulations can be an effective tool to assess which individuals to target in order to achieve more efficient results.
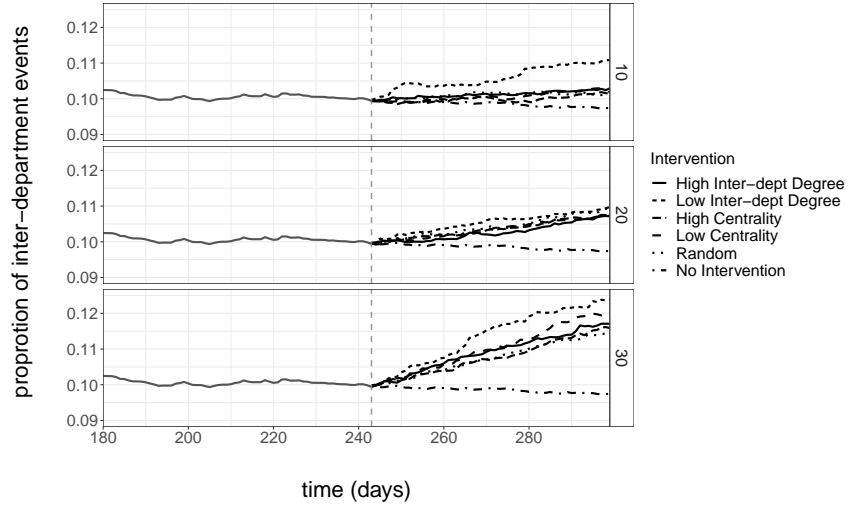
(a) Intervention targets based on endogenous attributes
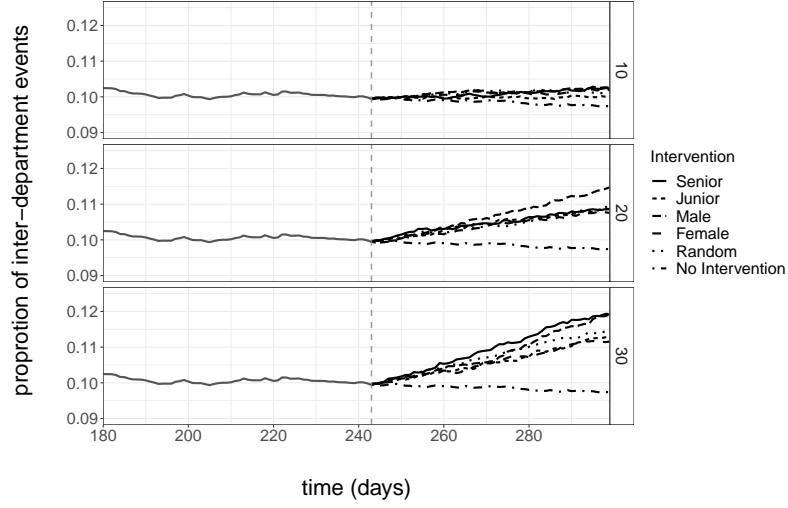
(b) Intervention targets based on exogenous attributes

Figure 6: Median proportion of inter-departmental events simulated for all the intervention strategies. (a) Intervention targets are selected on endogenous attributes, and (b) targets are selected on exogenous attribute. The panel in each sub-figure shows the results from top to bottom, K = {10% ,20% and, 30%}

6 Discussion

In this paper, we presented a flexible framework for simulating dynamic networks. Relational event modeling simulations make it possible to study complex interaction dynamics in isolation as well as in combination with other mechanisms that may operate in a dynamic network. Such simulations are effective for a variety of purposes. In this paper, we have explored the reasons why simulating relational event histories is such a valuable tool for social scientists and researchers across disciplines. They can be used for goodness of fit analyses of models fitted on empirical data. They can provide insights into building theories, testing assumptions, evaluating alternative explanations and hypotheses about the operation of interaction mechanisms in dynamic networks. Furthermore, they can be used to evaluate consequences of network interventions,identify intervention points, and inform policy decisions.

The simulation frameworks provide a practical means of simulating dynamic networks that resemble real dynamic networks. This approach while promising, also has certain limitations. First, the task of specifying an appropriate simulation model is not trivial and without enough theory, understanding, or extensive exploration, simulated networks can be prone to mis-specification. Simulations from a mis-specified model may result in unrealistic networks such as networks with unrealistically high (or low) connectivity. However, improving the model specification can help circumvent this problem. Second, the simulations we outlined in our applications only provide control over the local mechanisms of interaction. These local 'rules' give rise to global network level properties indirectly. However, our simulations do not give direct control over network level properties. For example, if a network of a specific density or degree distribution is desired, then the methods proposed in this paper may not be helpful as the density or degree distribution of a relational event network may only be partially dependent on the effects used to simulate relational event sequences. Other simulation models may be more suited towards these tasks.

Worthwhile directions of future research include relational event simulations methods that take into account heterogeneity in network data, for instance through multilevel models or latent variable models; simulations that incorporate more complex shapes of memory; the possibility to allow network effects to changes gradually or periodically (for example to mimic seasonal behavior); and the development of more advanced goodness of fit indices. The simulation objectives and examples discussed in this article along with the easy to use open source R package [package name redacted for double blind review] can provide the reader with a tool that opens up new avenues of research and shortens the design time for future attempts for simulating time-stamped relational event data.

Acknowledgement

References

Adams, j. and Schaefer, D. R. (2016). How Initial Prevalence Moderates Network-Based Smoking Change: Estimating Contextual Effects with Stochastic Actor Based Models. *Journal of health and social behavior*, 57(1):22–38.

Badham, J., Kee, F., and Hunter, R. F. (2021). Network structure influence on simulated network interventions for behaviour change. *Social Networks*, 64:55–62.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008.

Box, G. E. P. (1976). Science and Statistics. *Journal of the American Statistical Association*, 71(356):791–799. Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1976.10480949.

Brandenberger, L. (2019). Predicting Network Events to Assess Goodness of Fit of Relational Event Models. *Political Analysis*, 27(4):556–571.

Brewer, M. B. (1991). The Social Self: On Being the Same and Different at the Same Time. *Personality and Social Psychology Bulletin*, 17(5):475–482.

Butts, C. T. (2008). A Relational Event Framework for Social Action. *Sociological Methodology*, 38(1):155–200.

Butts, C. T. and Marcum, C. S. (2017). A Relational Event Approach to Modeling Behavioral Dynamics. *arXiv:1707.09902 [stat]*, pages 51–92.

Chen, S. and Onnela, J.-P. (2019). A Bootstrap Method for Goodness of Fit and Model Selection with a Single Observed Network. *Scientific Reports*, 9(1):16674.

Davis, J. P., Eisenhardt, K. M., and Bingham, C. B. (2007). Developing Theory Through Simulation Methods. *Academy of Management Review*, 32(2):480–499.

DuBois, C., Butts, C., and Smyth, P. (2013). Stochastic blockmodeling of relational event dynamics. In *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, pages 238–246. PMLR. ISSN: 1938-7228.

Epstein, J. M. and Axtell, R. L. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. The MIT Press.

Estévez-Mujica, C. P. and Quintane, E. (2018). E-mail communication patterns and job burnout. *PLoS ONE*, 13(3).

Falzon, L., Quintane, E., Dunn, J., and Robins, G. (2018). Embedding time in positions: Temporal measures of centrality for social network analysis. *Social Networks*, 54:168–178.

Gilbert, G. N. and Troitzsch, K. G. (2005). *Simulation for the Social Scientist*. Open University Press, Maidenhead, England ; New York, NY, 2nd ed edition.

Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268.

Hamill, T. M. (2001). Interpretation of Rank Histograms for Verifying Ensemble Forecasts. *Monthly Weather Review*, 129(3):550–560.

Hanneke, S., Fu, W., and Xing, E. P. (2010). Discrete temporal models of social networks. *Electronic Journal of Statistics*, 4(none):585–605.

Hardin, J. S., Sarkis, G., and URC, P. C. (2015). Network Analysis with the Enron Email Corpus. *Journal of Statistics Education*, 23(2):null.

Holland, P. W. and Leinhardt, S. (1977). A dynamic model for social networks. *The Journal of Mathematical Sociology*, 5(1):5–20.

Hunter, D. R., Goodreau, S. M., and Handcock, M. S. (2008). Goodness of Fit of Social Network Model. *Journal of the American Statistical Association*, 103(481):248–258.

Kleinbaum, A. M. (2012). Organizational misfits and the origins of brokerage in intrafirm networks. *Administrative Science Quarterly*, 57(3):407–452.

Klimt, B. and Yang, Y. (2004). Introducing the Enron Corpus. In *CEAS*.

Kozlowski, S. W. J. and Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions*, pages 3–90. Jossey-Bass/Wiley, Hoboken, NJ, US.

Krivitsky, P. N. and Handcock, M. S. (2014). A Separable Model for Dynamic Networks. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, 76(1):29–46.

Leenders, R. T. A. J., Contractor, N. S., and DeChurch, L. A. (2016). Once upon a time: Understanding team processes as relational event networks. *Organizational Psychology Review*, 6(1):92–115.

Leonardelli, G. J., Pickett, C. L., and Brewer, M. B. (2010). Optimal Distinctiveness Theory. In *Advances in Experimental Social Psychology*, volume 43, pages 63–113. Elsevier.

Lospinoso, J. and Snijders, T. A. (2019). Goodness of fit for stochastic actor-oriented models. *Methodological Innovations*, 12(3):2059799119884282.

Mulder, J. and Leenders, R. T. A. J. (2019). Modeling the evolution of interaction behavior in social networks: A dynamic relational event approach for real-time analysis. *Chaos, Solitons & Fractals*, 119:73–85.

Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 103(23):8577–8582.

Nicosia, V., Tang, J., Mascolo, C., Musolesi, M., Russo, G., and Latora, V. (2013). Graph Metrics for Temporal Networks. *arXiv:1306.0493 [physics]*, pages 15–40.

Perry, P. O. and Wolfe, P. J. (2013). Point process modelling for directed interaction networks. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(5):821–849.

Quintane, E., Pattison, P. E., Robins, G. L., and Mol, J. M. (2013). Short- and long-term stability in organizational networks: Temporal structures of project teams. *Social Networks*, 35(4):528–540.

Schecter, A. and Quintane, E. (2020). The Power, Accuracy, and Precision of the Relational Event Model. *Organizational Research Methods*, page 1094428120963830.

Snijders, T. A., van de Bunt, G. G., and Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1):44–60.

Snijders, T. A. B. and Steglich, C. E. G. (2015). Representing micro–macro linkages by actor-based dynamic network models. *Sociological Methods & Research*, 44(2):222–271.

Stadtfeld, C. (2014). *Events in Social Networks : A Stochastic Actor-Oriented Framework for Dynamic Event Processes in Social Networks*. KIT Scientific Publishing.

Stadtfeld, C. and Block, P. (2017). Interactions, Actors, and Time: Dynamic Network Actor Models for Relational Events. *Sociological Science*, 4.

Valente, T. W. (2005). Network Models and Methods for Studying the Diffusion of Innovations. In Scott, J., Carrington, P. J., and Wasserman, S., editors, *Models and Methods in Social Network Analysis*, Structural Analysis in the Social Sciences, pages 98–116. Cambridge University Press, Cambridge.

Valente, T. W. (2012). Network Interventions. *Science*, 337(6090):49–53.

Valente, T. W. (2017). Putting the network in network interventions. *Proceedings of the National Academy of Sciences*, 114(36):9500–9501.

Wang, C., Butts, C. T., Hipp, J., and Lakon, C. M. (2022). Model Adequacy Checking/Goodness-of-fit Testing for Behavior in Joint Dynamic Network/Behavior Models, with an Extension to Two-mode Networks. *Sociological Methods & Research*, 51(4):1886–1919.