



Dynamic network analysis of contact diaries[☆]

Jürgen Lerner^{a,*}, Alessandro Lomi^b, John Mowbray^c, Neil Rollings^c, Mark Tranmer^c

^a Department of Computer and Information Science, University of Konstanz, Box 67, 78457 Konstanz, Germany

^b University of the Italian Switzerland, Lugano, CH, Switzerland

^c School of Social and Political Sciences, University of Glasgow, UK

ARTICLE INFO

Keywords:

Duality of persons and groups
Two-mode networks
actor–event networks
Hypergraphs
Analysis of contact diaries
Longitudinal network models
Relational hyperevent models

ABSTRACT

Analyzing two-mode networks linking actors to events they attend may help to uncover the structure and evolution of social networks. This classic social network insight is particularly valuable in the analysis of data extracted from contact diaries where contact events produce — and at the same time are the product of relations among participants. Contact events may comprise any number of actors meeting at a specific point in time. In this paper we recall the correspondence between two-mode actor–event networks and hypergraphs, and propose relational hyperevent models (RHEM) as a general modeling framework for networks of time-stamped multi-actor events in which the diarist (“ego”) simultaneously meets several of her alters. RHEM can estimate event intensities associated with each possible subset of actors that may jointly participate in events, and test network effects that may be of theoretical or empirical interest. Examples of such effects include preferential attachment, prior shared activity (familiarity), closure, and covariate effects explaining the propensity of actors to co-attend events. Statistical tests of these effects can uncover processes that govern the formation and evolution of informal groups among the diarist’s alters. We illustrate the empirical value of RHEM using data comprising almost 2000 meeting events of former British Prime Minister Margaret Thatcher with her cabinet ministers, transcribed from contact diaries covering her first term in office (1979–1983).

1. Introduction

Interpersonal relations are shaped by, and at the same time shape membership of actors in informal groups (Breiger, 1974). Such latent group membership is often reflected in observable patterns of co-participation to social events (Davis et al., 1941; Freeman, 2003). Previous work has suggested a variety of analytic methods (Borgatti and Everett, 1997; Seidman, 1981; Everett and Borgatti, 1993; Everett et al., 2018) and statistical models (Wang et al., 2013; Koskinen and Edling, 2012; Snijders et al., 2013; Butts, 2008; Stadtfeld and Geyer-Schulz, 2011; Conaldi et al., 2012; Conaldi and Lomi, 2013; Lerner and Lomi, 2020; Chodrow and Mellor, 2020) for two-mode networks or – equivalently – hypergraphs. As we discuss more formally in Section 2.1, a hypergraph contains hyperedges that may link any number of nodes (Bretto, 2013). In contrast, edges in a graph link exactly two nodes.

Events typically have associated time points (or time intervals) and this is – almost by definition – true for meeting events transcribed from contact diaries whose entries have associated dates and/or times (Fu et al., 2013; Yen et al., 2016). A major shortcoming of most

previous statistical network models is that they either cannot cope with the fine-grained time information associated with meeting events – or they are designed for dyadic relational events, having a single sending node and a single receiving node. Building on the framework proposed in Lerner et al. (2019), in this paper we develop relational hyperevent models (RHEM) for networks of time-stamped multi-actor events transcribed from contact diaries. RHEM can cope with the fine-grained time information of meeting events, where often each event has its own unique time stamp – and at the same time can specify separate event intensities for all subsets of a given sample of actors. RHEM provide a general framework for likelihood-based inference of model parameters from observed sequences of time-stamped multi-actor events and enable researchers to test hypotheses on processes governing co-participation in social events.

The RHEM framework represents an important and transformative step for diary studies – a recognized area of inquiry in social research Bartlett and Milligan (2015), Stafford (2009). Appointment diaries, which are increasingly in electronic formats, contain multi-actor event data with fine-grained time information and hence lend

[☆] Financial disclosure: This work has been supported by the Deutsche Forschungsgemeinschaft (DFG) - 321869138 and Economic and Social Research Council, UK - ES/R009236/1.

* Corresponding author.

E-mail addresses: juergen.lerner@uni-konstanz.de (J. Lerner), alessandro.lomi@usi.ch (A. Lomi), John.Mowbray@glasgow.ac.uk (J. Mowbray), Neil.Rollings@glasgow.ac.uk (N. Rollings), Mark.Tranmer@glasgow.ac.uk (M. Tranmer).

<https://doi.org/10.1016/j.socnet.2021.04.001>

themselves particularly well to illustrate the applicability and empirical value of RHEM. According to Alaszewski (2006, p.1), a diary is “a document created by an individual who has maintained a regular, personal and contemporaneous record”. While appointment diaries regularly appear in the private papers of politicians, civil servants, artists (Hackett, 1989), and business executives, they have rarely been analyzed through the lens of longitudinal network models explaining co-participation of alters in meetings with the diarist. Instead, diary research in the social sciences and humanities has typically focused on the narrative content of unsolicited personal diaries, or on their solicited use as a research tool; for overviews, see Mehl and Conner (2012), Bartlett and Milligan (2015). Personal and contact diaries have been used mostly, albeit not exclusively, as sources of data for qualitative research (Alaszewski, 2006). By specifying RHEM for the study of appointment diaries, however, it will be possible to forge a new and more factual understanding of prominent public figures, based on the evolving structural dynamics of their interpersonal relationships. Findings from such studies would directly inform debates in contemporary history, political science, and elite studies, thus opening new avenues of research in those fields.

We illustrate the empirical value of RHEM in an analysis of event data extracted from former British Prime Minister (PM) Margaret Thatcher’s appointment diaries. While the objective of this paper is methodological, we illustrate how the RHEM framework could help to illuminate existing research debates on topics such as the nature of the policy process in the British core executive (Heffernan, 2005; Richardson, 2018), the development of Thatcherism (Williamson, 2015; Jackson and Saunders, 2012), and many others. We highlight the analytical potential of RHEM by presenting an exploratory analysis of Thatcher’s meetings with cabinet ministers during the term of the 1979 Parliament (May 1979 to June 1983). We hypothesize that participation in meetings – in which ministers can potentially influence the PM and/or can take influence on the PM’s interaction with other ministers – can reflect and establish latent competition and power differences among cabinet ministers. More specifically, we demonstrate how the following three research questions could be rigorously addressed with RHEM. Based on a binary actor-level covariate labeling “dry” ministers,¹ we test whether: (i) The PM has a preference (or reluctance) to meet with dry ministers and (ii) she has a preference (or reluctance) to meet with homogeneous groups of ministers that are either mostly dry or mostly non-dry. Furthermore, as an example of a structural effect in hyperevent networks we test whether (iii) there is a significant tendency for or against triadic closure in the PM meeting diary data. We conjecture that a negative closure effect – together with a tendency to partially repeat meetings – could lead to the emergence of stable structural holes (Burt, 1992) and might be a signal of competition and systematic power differences among ministers. We emphasize that it is not the goal of this paper to obtain conclusive evidence on these questions – but rather to illustrate how such questions could be tackled with RHEM.

2. Background

In this paper we propose statistical models for networks of time-stamped multi-actor events. Formally, the data we consider involve a sequence of events $E = (e_1, \dots, e_N)$, where each event $e = (t_e, h_e) \in E$ comprises the event time t_e and a set of participants $h_e \subseteq V_{t_e}$,

¹ Cabinet ministers who fully supported Thatcher’s free-market agenda, which prioritized controlling inflation over controlling unemployment, were labeled “dry”. On the other hand, ministers who were critical of her hard-line economic policies were labeled “wet”, because Thatcher perceived their inclination towards consensus politics and state intervention to be a weakness (Heppell, 2020). In our analysis we make a binary split between dry and non-dry ministers, where the latter category contains Wets but also ministers that are neither labeled dry nor wet.

which is a subset from a given sample of actors V_{t_e} at the event time. For example, in the case of contact diaries that we examine in the empirical case study of this paper, the actors are the participants to the various meetings recorded in the diary, that is, the actors are the diarist’s alters. In the following we recall in Section 2.1 the well-known equivalence between hypergraphs and two-mode actor–event networks (Seidman, 1981) and in Section 2.2 we discuss the applicability of related statistical network models to networks of time-stamped multi-actor events.

2.1. Hypergraphs and two-mode actor–event networks

The underlying data structure for RHEM is a hypergraph containing hyperedges that may connect any number of nodes (Bretto, 2013). Therefore, hypergraphs may be viewed as a generalization of graphs because every edge in a graph can only connect two nodes. Formally, a hypergraph $G = (V, H)$ comprises a set of nodes V (representing, e.g., the actors of a social network) and a set of hyperedges $H \subseteq \mathcal{P}(V)$, where each hyperedge $h \in H$ is a subset $h \subseteq V$ of any size (representing, e.g., the set of actors attending a social event). A relational hyperevent $e = (t, h)$ is a hyperedge $h \subseteq V$ (recording the participants to the event) together with a timestamp t (recording the time when the event takes place). In the data that we consider in this paper, the same hyperedge h (i.e., the identical set of participants) can experience more than one event at different points in time.

It is well known that a hypergraph (V, H) gives rise to a bipartite, two-mode actor–event network and vice versa (Seidman, 1981). The first set of nodes of the bipartite network (comprising the “actor nodes”) is identical with the node set V of the hypergraph, the second set of nodes (comprising the “event nodes”) is identical with the set of hyperedges H , and an actor node $v \in V$ is connected to an event node $h \in H$ in the bipartite graph if and only if the hyperedge h contains the actor v in the hypergraph. Fig. 1 displays a sequence of relational hyperevents (middle) and the associated two-mode actor–event network (left) and hypergraph (right). The one-mode projection (to the set of actors) of a two-mode network is a graph whose nodes are the actor nodes of the two-mode network and in which two nodes u and v are connected by an edge if and only if there is an event node in the two-mode network that is connected to both u and v – in other words, if u and v co-participate in a common event. Fig. 1 (right) displays the one-mode projection by dashed lines within the hyperedges. We recall that two-mode networks and hypergraphs are equivalent representations of the same data. In contrast, a one-mode projection does not uniquely represent a two-mode network or its associated hypergraph. This can be seen in the example from Fig. 1, where the triads $\{A, B, C\}$ and $\{D, E, F\}$ are identically connected in the one-mode projection but are structurally different in both the two-mode network and the hypergraph. Note that actors in the first triad experience one common event while actors in the second triad only have pairwise co-participation in different events.

One possibility to analyze two-mode networks, or hypergraphs, is to consider the associated one-mode projection (Borgatti and Everett, 1997), which, however, has its drawbacks. First, as discussed above, one-mode projections do not uniquely represent the original two-mode data. Second, one-mode projections introduce structural artifacts such as high local density and clustering. For instance, a single event with 22 participants (the largest events in our illustrative data) introduces 231 actor–actor ties and 1,540 closed triangles in the one-mode projection. Thus, analyzing one-mode projections can bias findings – at least – towards “detecting” a tendency for triadic closure. For such reasons network analytic methods have been developed that operate directly on two-mode networks or – equivalently – hypergraphs. These include methods to compute centrality or clusters (Borgatti and Everett, 1997), role colorings (Everett and Borgatti, 1993), blockmodels (Doreian et al., 2004), information flows (Everett et al., 2018), or the structure of overlapping subsets (Foster and Seidman, 1982) in two-mode networks

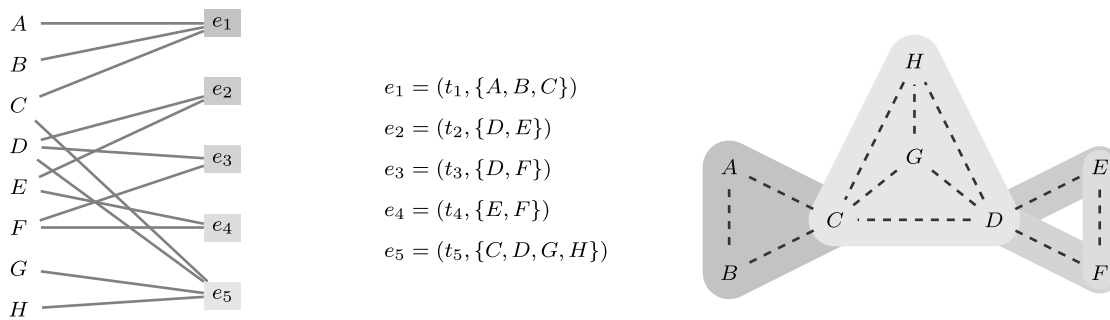


Fig. 1. Middle: Stylized example of five hyperevents e_1, \dots, e_5 occurring at event times $t_1 < \dots < t_5$. Participating actors are given by uppercase letters A, \dots, H . Left: Representation of the sequence of hyperevents as a two-mode actor–event network. Events are displayed as rectangular nodes ordered from top to bottom by their time (older events are displayed in a darker shade than younger events) and are connected to their participants by solid lines. Right: Representation of the sequence of hyperevents as a hypergraph. Events are represented by hyperedges displayed as convex hulls enclosing their participants. Dashed lines represent ties in the one-mode projection to the set of actors.

or hypergraphs. These network analytic methods have a different objective than the statistical models that we propose in this paper. Our framework allows analysts to conduct statistical tests whether certain characteristics of sets of actors (such as being homogeneous with respect to an actor-level covariate or being indirectly connected via third actors) increase or decrease the probability to experience common events (compare the three exemplary research questions formulated in Section 4). Thus, we need methods that estimate interaction probabilities for hyperedges (i. e., sets of actors) by comparing those on which events are observed with alternative hyperedges that could potentially have experienced events. We therefore turn our discussion to related statistical models defining probability distributions on a given space of two-mode networks (or hypergraphs).

2.2. Related statistical models for two-mode networks and relational event networks

Statistical models for 2-mode networks (or hypergraphs) have been proposed in different frameworks for (temporal) ERGM (Wang et al., 2013; Krivitsky and Handcock, 2014), SAOM (Koskinen and Edling, 2012; Snijders et al., 2013), relational event models (REM) (Butts, 2008; Conaldi et al., 2012; Lerner and Lomi, 2020; Valeeva et al., 2020), or by defining configuration models for hypergraphs that condition on observed degree sequences (Chodrow and Mellor, 2020). However, (T)ERGM, SAOM, or the configuration model are designed for networks of relational states and cannot cope with the fine-grained time information associated with contact events, where each event might have its own unique time stamp.

REM, on the other hand, are explicitly designed to model networks of relational events with fine-grained time information. However, REM typically specify event rates for dyads having one sending node and one receiving node – which is inappropriate for multi-actor interaction resulting from contact diaries requiring that event rates are specified for all subsets of actors (i. e., hyperedges) in the risk set. We point out that, in the case of meeting events transcribed from contact diaries, the “event nodes” in the two-mode actor–event representation are not exogenously given but arise – simultaneously with their ties – in an endogenous process. Event nodes are “active” in a single point in time in which they are created by the joint interaction among actors, simultaneously with all their incident ties. This marks a considerable difference to settings in which dyadic relational events connect single actors to existing nodes in the second mode: for instance, Valeeva et al. (2020) specify and estimate event rates for dyads comprising one director and one company board (one event recording that a director joins a board); Lerner and Lomi (2020) estimate dyadic REM for two-mode networks connecting Wikipedia users to the articles they contribute to (one event recording that a user edits an article). In contrast, events in our data do not “exist” over extended periods of time (event duration is very short compared to the observation period and

events do not overlap) in which actors could decide to join or leave; in our data, events record simultaneous interaction of a set of actors. The distinction whether events are associated with dyads or with groups of actors of any size is also reflected in the size of the risk set, that is, the set of instances that could experience a common event. If events are associated with dyads, the risk set can comprise pairwise combinations of nodes, leading to a maximal risk set size that is quadratic in the number of nodes. If events are associated with hyperedges, the maximal risk set size can be exponential in the number of nodes, since any subset can constitute the participant list of the next event. Moreover, if events are associated with groups of any size, models can test for more complex effects, such as subset repetition of order larger than two (see Section 3.2.1), that cannot be included in dyadic RHEM. Being related to our paper in a different way, Marcum and Butts (2015) propose REM for data extracted from diaries – but their events are about time-use data, rather than contact events.

As discussed in Lerner et al. (2019), some attempts to generalize REM to multi-actor interaction exist. Butts (2008) recommends to create “virtual” nodes representing subsets of elementary nodes to treat interaction with multiple senders and/or receivers. This can be a feasible approach, e. g., to represent the entire set of actors by a node that is the receiver of broadcast messages (Gibson, 2005; DuBois et al., 2013) – but it is easy to see that this strategy is restricted to a limited number of predefined sets of actors since there are exponentially many possible subsets. Following a different approach, Kim et al. (2018) propose the hyperedge event model for directed multicast events that have exactly one sender and any number of receivers (or conversely any number of senders and exactly one receiver). Their model specifies for every possible sender–receiver dyad (i, j) an intensity function λ_{ij} . These dyadic intensity functions then stochastically determine (1) who is the sender of the next event and (2) which is the set of receivers, given the chosen sender. The difference to RHEM proposed in Lerner et al. (2019), and elaborated for co-attendance data in this paper, is that in the model from Kim et al. (2018) the intensities are defined for dyads, while RHEM can define separate intensities for all hyperedges of any size. Bartolucci et al. (2018) define models for bipartite event networks including first-order effects (e. g., propensity of individual actors to participate in events) and second-order effects (e. g., propensity of pairs of actors to co-participate in events). However, generalizing their model to higher-order effects (like the ones that can be included in RHEM) seems to lead to inefficient estimation algorithms.

Another related model has been defined by Hoffman et al. (2020), who propose a model for the dynamics of face-to-face interaction, such as the formation and dissolution of informal conversation groups during social gatherings. This model maintains a two-mode network of actors and groups in which – at any point in time – each actor is connected to exactly one group (which may be a single-actor group representing an isolate). Actors who are currently isolates can decide to connect to existing groups; actors who are currently members of some multi-actor

group can decide to leave the group. Similar to REM for two-mode networks, discussed above, the model from Hoffman et al. (2020) seems not to apply to empirical data of the kind that we consider in this paper. The atomic observations in data that we gather from contact diaries comprise meeting events jointly experienced by groups of actors – rather than events in which individual actors join or leave ongoing meetings. Indeed, we consider the two models as complementary, where Hoffman et al. (2020) model individual actors’ decisions to join and leave groups, while we model events jointly experienced by groups of actors.

3. Relational hyperevent models for networks of time-stamped multi-actor events

In this section we develop relational hyperevent models for time-stamped multi-actor events extracted from contact diaries, building on the framework proposed in Lerner et al. (2019). Contact diaries (Fu, 2007) differ from the more common personal diaries (Bolger et al., 2003). The latter record personal histories that may or may not include systematic accounts of contacts with others. Personal diaries are typically retrospective, i. e., based on the diarist’s memories of time past. The former are specifically kept to record systematic information about relational events involving multiple alters simultaneously. Contact diaries are typically prospective, i. e., their objective is to organize and structure the diarist’s time in the future.

We assume that we have a time-varying population of actors that could potentially appear in the participant list of events at time t , denoted by V_t . Moreover, we assume that we are given a sequence of observed relational hyperevents $E = (e_1, \dots, e_N)$, where each event $e = (t_e, h_e) \in E$ comprises the event time t_e and a set of participants $h_e \subseteq V_{t_e}$. For instance, in our exemplary empirical data, discussed in Section 4, V_t is the set of all cabinet ministers at time t and an event $e = (t_e, h_e)$ represents a meeting with the PM (i. e., the diarist) that takes place at t_e and whose participants are the cabinet ministers listed in h_e .

3.1. General model framework

Following Lerner et al. (2019), RHEM specify an *event intensity* $\lambda(t; h)$ (also *event rate* or *hazard rate*) associated with hyperedges h , which intuitively is the expected number of events on h in a time interval of length one, starting at t (Lawless, 2003). Formally, if $n(t; h) = |\{e \in E : t_e \leq t \wedge h_e = h\}|$ is the number of events on hyperedge h up to and including time t , the intensity is defined by

$$\lambda(t; h) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{E}[n(t + \Delta t; h) - n(t; h)]}{\Delta t},$$

where $\mathbb{E}[n(t + \Delta t; h) - n(t; h)]$ is the expected number of events on h in the interval $[t, t + \Delta t]$.

Given a sequence of relational hyperevents $E = (e_1, \dots, e_N)$ and a point in time t , we denote by $G[E; t]$ the *network of past events* which is a function of all events from E that happen strictly before t (Brandes et al., 2009). We specify the likelihood of the sequence of hyperevents E based on the *Cox proportional hazard model* (Cox, 1972). We decompose the event rate $\lambda(t; h)$ into a time-varying baseline rate $\lambda_0(t)$ and a relative event rate $\lambda_1(t; h; \theta; G[E; t])$, conditional on a vector of *hyperedge statistics* $s(t; h; G[E; t]) \in \mathbb{R}^k$, being a function of the network of past events, and a vector of associated parameters $\theta \in \mathbb{R}^k$:

$$\lambda(t; h; \theta; G[E; t]) = \lambda_0(t) \cdot \lambda_1(t; h; \theta; G[E; t]), \tag{1}$$

$$\lambda_1(t; h; \theta; G[E; t]) = \exp(\theta \cdot s(t; h; G[E; t])). \tag{2}$$

The baseline rate λ_0 represents time-variation in the intensity of events in the whole network and is left unspecified. The partial likelihood based on the observed event sequence E is

$$L(\theta) = \prod_{e \in E} \frac{\lambda_1(t_e; h_e; \theta; G[E; t_e])}{\sum_{h \in R_e} \lambda_1(t_e; h; \theta; G[E; t_e])}, \tag{3}$$

where the $R_e \subseteq \mathcal{P}(V_e)$ are a suitable definition of the risk sets at the event times. As pointed out in Lerner et al. (2019), this framework for relational hyperevents is very close to some specifications of dyadic REM (Butts, 2008; Perry and Wolfe, 2013; Vu et al., 2015; Lerner and Lomi, 2020). The difference is that separate event rates are specified for all hyperedges in the risk set, rather than for all pairs of nodes as in dyadic REM, and that the statistics are functions of sets of any number of nodes, rather than being functions of dyads. The hyperedge statistics $s(t; h; G[E; t])$ can operationalize hypothetical effects explaining multi-actor interaction, for instance, by preferential attachment (popularity effects), prior shared activity (familiarity), closure, or covariate effects (e. g., homophily). The associated parameters – estimated by maximizing the likelihood function – provide statistical tests for such hypotheses.

We make two further decisions in defining the model framework. The first decision – constraining the risk set R_e at the time of the event $e = (t_e, h_e)$ to those hyperedges that have the same size as h_e – is due to a very strong dependence of event intensity on hyperedge size. The second decision is taken to achieve computational tractability and consists of sampling from the risk sets.

Conditioning the model on the size of the observed hyperevents seems to be necessary – at least until we have found better ways to control for size – to avoid the confounding effect of hyperedge size on event intensity. To illustrate this, we note that the baseline probability to experience events typically depends strongly on the size of hyperedges. For instance, in our exemplary empirical data (compare Fig. 6) hyperedges of size one have a baseline event intensity that is about 50 times higher than the intensity on hyperedges containing exactly two actors and more than 500,000 times higher than the baseline intensity on hyperedges of size 11. Such examples illustrate that event intensities should not be compared between hyperedges of different size. Indeed, the preliminary analysis in Lerner et al. (2019) has shown that models with unconstrained risk sets yield results that strongly depend on how models control for the effect of hyperedge size on event intensity.

While we do not claim that this approach is preferable in all settings, we consider in this paper *conditional-size models* (Lerner et al., 2019), where the risk set at the time of an observed event $e = (t_e, h_e)$ consists of all hyperedges of size $|h_e|$, i. e., $R_e = \{h \subseteq V_{t_e} : |h| = |h_e|\}$. This implies that for each of the factors in Eq. (3) the hyperedges h appearing in the denominator have the same size as the hyperedge h_e in the numerator on which the event e has been observed. Thus, event intensities are only compared among hyperedges of the same size. The alternative hyperedges h can be obtained from the event hyperedge h_e by removing a subset of actors $h' \subseteq h_e$ of any size from h_e and then adding a set of new participants h'' of the same size as h' . Formally, the event hyperedge h_e is only compared with alternative hyperedges h satisfying $h = (h_e \setminus h') \cup h''$ for some $h', h'' \subset V_{t_e}$, where $h' \subseteq h_e$, $h'' \cap h_e = \emptyset$, and $|h'| = |h''|$. Thus, as discussed in Lerner et al. (2019), conditional-size models are consistent with the point of view that (groups of) actors *compete* for participation in meetings. We further discuss the implications of conditional-size models – and the need for future work on this issue – after having presented the empirical results in Section 5.

The second decision – taken to maintain computational tractability – is that the risk sets R_e at the event times are replaced by sampled risk sets \tilde{R}_e (Lerner et al., 2019), obtained via case-control sampling (Borgan et al., 1995). Sampling has become established in estimating approximate parameters for dyadic REM on large networks. For instance, Butts (2008) points out that estimating parameters via approximate likelihood functions, obtained by sampling, could make REM applicable to larger networks. Vu et al. (2015) specifically apply case-control sampling (Borgan et al., 1995), where all events and for each event a fixed number of randomly drawn “controls” (non-events from the risk set) are considered. Lerner and Lomi (2020) experimentally assess the reliability of case-control sampling for REM estimated on large networks. Following Lerner et al. (2019), we draw – for a given

number m of non-events per event – sampled risk sets \tilde{R}_e that contain the hyperedge h_e of the observed event $e = (t_e, h_e)$ and m non-event hyperedges drawn uniformly at random from $R_{t_e} = \{h \subseteq V_{t_e}; |h| = |h_e|\}$. This leads to the following sampled likelihood function:

$$\tilde{L}(\theta) = \prod_{e \in E} \frac{\lambda_1(t_e; h_e; \theta; G[E; t_e])}{\sum_{h \in \tilde{R}_e} \lambda_1(t_e; h; \theta; G[E; t_e])} \quad (4)$$

Given the values of the statistics $s(t_e; h; G[E; t_e])$ for all hyperedges h in the sampled risk sets \tilde{R}_e , maximum likelihood estimates for the parameters in Eq. (4) can be computed with standard statistical software. For the empirical analysis reported in this paper we compute hyperedge statistics with an extension of *eventnet*² (Lerner and Lomi, 2020) and use the R package *survival*³ (Therneau and Grambsch, 2013) to estimate parameters.

3.2. Model specification: network effects

Network effects included in the empirical model specification that we estimate in our illustrative analysis fall into three classes. First, we control for the tendency to repeat previous events, either in exact repetition with the identical set of participants or in partial repetition where a subset of participants of a previous meeting co-participates in a future meeting, potentially with yet other participants. Second, we introduce hyperedge statistics based on actor-level covariates, assessing first-order effects and homophily effects of the respective covariate. (In our exemplary case study we use the binary covariate labeling ministers as dry or non-dry.) Third, we define a hyperedge statistic for how strongly groups of actors have previously co-participated in meetings with common third actors to assess a tendency for or against triadic closure in meeting data. Technically, effects are added to the model by defining the vector of hyperedge statistics $s(t; h; G[E; t])$ in the specification of the relative event rate in Eq. (2). These statistics assign real numbers to hyperedges h at time t based on the network of past events $G[E; t]$.

3.2.1. Repetition and subset-repetition

A very basic effect in hyperevent networks is that events are repeated or partially repeated (Lerner et al., 2019). In all models reported in this paper we let the effect of past events decay over time in the way as it has been suggested in Brandes et al. (2009). For a given half life period $T_{1/2}$, we set the decay factor $w(\Delta t)$ for the time difference Δt to

$$w(\Delta t) = \exp\left(-\Delta t \cdot \frac{\log(2)}{T_{1/2}}\right).$$

The repetition statistic associated with a given hyperedge h at time t counts the number of previous events in $E_{<t} = \{e = (t_e, h_e) \in E; t_e < t\}$ whose set of participants h_e is identical with h , weighted by the respective decay factor. Formally, it is defined by

$$repetition(t; h; G[E; t]) = \sum_{e \in E_{<t}} w(t - t_e) \cdot \chi(h = h_e),$$

where χ is the indicator function that is one if the argument is true and zero else.

Often the set of participants of one event is not exactly repeated in another event – as it is required in the definition of *repetition* – but only partially and possibly in co-participation with yet other actors. We first define a time-varying indicator, denoted as *hypergraph degree*, assessing to what extent a given set of actors h co-participated in past events. This indicator counts the number of events in which all actors in h co-participated – weighted by the respective decay factor – and is formally defined by

$$hy.deg(t; h; G[E; t]) = \sum_{e \in E_{<t}} w(t - t_e) \cdot \chi(h \subseteq h_e).$$

Subset repetition is a family of hyperedge statistics parametrized by the order (also size or cardinality) of subsets. Formally, for a given integer $p \in \mathbb{N}$, subset repetition of order p is defined by

$$sub.rep^{(p)}(t; h; G[E; t]) = \sum_{h' \in \binom{h}{p}} hy.deg(t; h'; G[E; t]) / \binom{|h|}{p},$$

where $\binom{h}{p}$ denotes the set of all subsets of h that have exactly p elements. The formula above takes the average hypergraph degree over all subsets of size p of h . Thus, subset-repetition of order p can assess whether sets of p actors who have co-participated in previous events (potentially with other participants) are more or less likely to co-participate in future events (potentially with other participants).

Repetition and subset-repetition of various order is illustrated in Fig. 2. Note that subset-repetition of order one assigns the identical value to all three triads, subset-repetition of order two can distinguish between h_1 on one hand and h_2 and h_3 on the other hand but fails to distinguish between the latter two, subset-repetition of order three can recognize that h_3 has one previous joint event, in contrast to h_1 and h_2 .

We emphasize that subset-repetition statistics are not merely “control” variables, but may be able to uncover important aspects of participant selection based on familiarity of different order. Subset-repetition of order one accounts for individual past activity (i.e., the number of prior events of individual actors) and can model a tendency for or against preferential attachment. Subset-repetition of order two accounts for prior *shared* events on dyads (*dyadic familiarity*); subset-repetition of order three considers prior shared events on triads (*triadic familiarity*), and so on.

Although we do not have event types or weights in the empirical data analyzed in Section 4, we note that types or weights could be taken into account in the repetition or subset-repetition statistics in a straightforward way. Different statistics could count only past events of certain types and/or could add up the weights of past events. This would be very similar to the approach for typed and weighted dyadic events proposed in Lerner et al. (2013a).

3.2.2. Covariate effects (first-order and homophily effects)

We assume that we are given a binary, time-constant actor-level covariate $x: V \rightarrow \{0, 1\}$. (In the exemplary case study of this paper we use the binary covariate *dry*, where $dry(v) = 1$ indicates that v is a dry minister.) We define two hyperedge statistics modeling (1) a first-order effect of the covariate x , assessing whether actors v with $x(v) = 1$ are more likely to be among the participants of meetings than actors v' with $x(v') = 0$ and (2) a second-order effect on homophily with respect to x . Both statistics are independent of time and independent of the network of past events.

For a hyperedge h , the statistic *average-x* is the ratio of actors v with $x(v) = 1$ in h ; formally

$$avg.x(t; h; G[E; t]) = \sum_{v \in h} x(v) / |h|.$$

In our exemplary case study (with x operationalizing whether ministers are dry) this statistic can assess whether dry ministers do more often participate in meetings – or, from the other point of view, whether the PM has a preference to meet with Dries. If a positive parameter is associated with *avg.x*, then in the example given in Fig. 3, the right-most hyperedge would be predicted to have the highest event rate, followed by the hyperedge in the middle, and then the left-most hyperedge.

For a hyperedge h , let $h_x = \{v \in h; x(v) = 1\}$ denote the set of actors with $x = 1$ contained in it and let $\bar{h}_x = \{v \in h; x(v) = 0\}$ denote those with $x = 0$. The statistic *x-homogeneity* measures to what extent the larger of the two groups outnumbers the smaller, normalized by the size of h ; formally

$$x.hom(t; h; G[E; t]) = \frac{\max(|h_x|, |\bar{h}_x|) - \min(|h_x|, |\bar{h}_x|)}{|h_x| + |\bar{h}_x|}.$$

² <https://github.com/juergenlerner/eventnet>.

³ <https://CRAN.R-project.org/package=survival>.

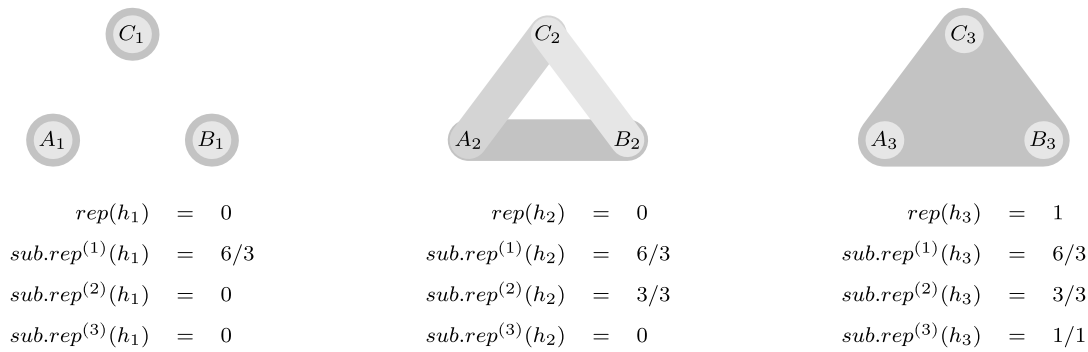


Fig. 2. Stylized example illustrating repetition and subset-repetition effects. Previous events on the three hyperedges $h_i = \{A_i, B_i, C_i\}$ for $i = 1, 2, 3$ are displayed as gray-shaded convex hulls enclosing their respective participants. *Left*: each of the three members of h_1 has participated in two single-actor events. *Middle*: each unordered pair of the three members of h_2 has co-participated in one two-actor event. *Right*: all three members of h_3 have co-participated in one three-actor event and, in addition, each of its members has participated in one single-actor event.

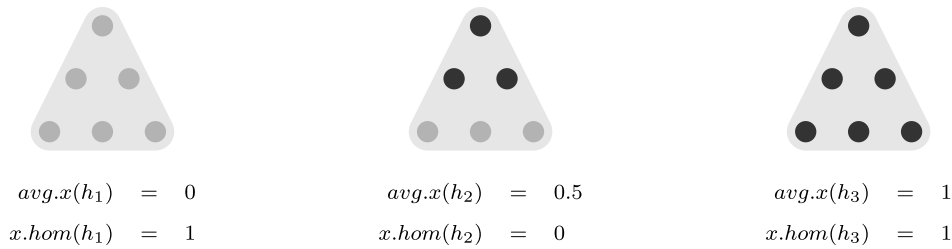


Fig. 3. Stylized example illustrating average- x and x -homogeneity for a binary actor-level covariate $x : V \rightarrow \{0, 1\}$. The image shows three hyperedges, each containing six actors, displayed as small circles. Actors v with $x(v) = 1$ are displayed in darker shade than those taking the value zero in the covariate x . *Left*: all actors take the value $x = 0$ and the hyperedge is maximally homogeneous. *Middle*: half of the actors take the value $x = 1$ and the hyperedge is maximally inhomogeneous (homogeneity is zero). *Right*: all actors take the value $x = 1$ and the hyperedge is maximally homogeneous.

If $|h|$ is odd, then it is not possible that the two groups have the same size, so that zero (minimally homogeneous) would not be attainable. To correct this, we take the value $\frac{x.hom(t; h; G[E; t]) - 1/|h|}{1 - 1/|h|}$, instead of $x.hom(t; h; G[E; t])$, if h has an odd number of participants. With this correction, the statistic can attain the extremal values zero and one.⁴

In our exemplary case study (with x operationalizing whether ministers are dry), the statistic x -homogeneity can assess whether meetings reveal a homophily effect with respect to the dry/non-dry characteristic of ministers – or, from the another point of view, whether the PM has a preference to meet dry ministers separately from non-dry ministers. If a positive parameter is associated with $x.hom$, then in the example given in Fig. 3, the right-most and the left-most hyperedge would be predicted to have a higher event rate than the hyperedge in the middle. If $x.hom$ has a negative parameter, then the “mixed” hyperedge in the middle would be predicted to have a higher event rate than the two “pure” hyperedges.

3.2.3. Closure

The closure statistic measures to what extent the members of a hyperedge h have co-participated in previous events with common third actors w . In contrast to subset-repetition of order three, these third actors w may be outside of the focal hyperedge h and different members of h may have co-participated with w in different past events. Formally,

closure is defined by⁵

$$closure(t; h; G[E; t]) = \sum_{\{u,v\} \in \binom{h}{2} \wedge u \neq v} \min[hy.deg(t; \{u, w\}; G[E; t]), hy.deg(t; \{v, w\}; G[E; t]) / \binom{|h|}{2}$$

In the formula above, we iterate over all combinations $(\{u, v\}, w)$, such that $\{u, v\} \in \binom{h}{2}$ is an unordered pair of actors who are both members of the focal hyperedge h and w is any actor different from u and v . For each of these triples, we compare the hypergraph degrees of the two sets $\{u, w\}$ and $\{v, w\}$ and add up the minimum of these two values. The resulting sum is divided by the number of unordered pairs within h . As an example of such a triple of actors $(\{u, v\}, w)$, consider $(\{E, F\}, D)$ in Fig. 4 at the time of the event $e_4 = (t_4, \{E, F\})$. Both members of the focal hyperedge $h_4 = \{E, F\}$ are connected by one past event to the third actor D . Thus, disregarding any decay in the influence of past events, $closure(t_4; h_4) = 1$.

Understanding the implication of the closure effect in hyperevent networks – and its interplay with subset-repetition of various order – is challenging. For illustration we use again Fig. 4 which recalls five hyperevents (assumed to be observed in the past) and defines two additional hyperedges, $h = \{A, B, H\}$ and $h' = \{D, G, H\}$, on which events could happen at a time point $t > t_5$.

A superficial look at the one-mode projection (given by the dashed lines) seems to suggest that the five observed events reveal a tendency

⁵ The closure statistic in our paper would have been denoted by $closure^{(1,1)}$ in Lerner et al. (2019) who define closure statistics of varying order — but do not estimate any closure effect in their empirical analysis. Another difference is that Lerner et al. (2019) propose a different normalization by also dividing over all possibilities of third actors; we deviate from this to be more consistent with usual definitions of closure in networks of dyadic events.

⁴ If $|h| = 1$, we set $x.hom(t; h; G[E; t]) = 0$.

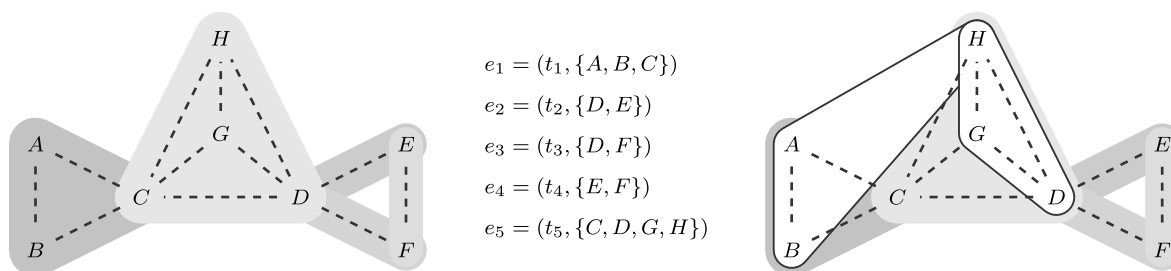


Fig. 4. Illustration of closure on the data from Fig. 1, recalled in the hypergraph (left) and in the sequence of hyperevents (middle). The hypergraph on the right displays two additional hyperedges $h = \{A, B, H\}$ and $h' = \{D, G, H\}$, on which events could potentially occur, as white convex hulls with dark borders. An event on h would point to closure while an event on h' could be explained by subset-repetition. See the text for additional explanation.

for triadic closure, since the one-mode projection contains many closed triangles. However, a more detailed inspection of the actual hyperevents tells us that just one of the five events (namely e_4 connecting actors E and F) has a closure statistic different from zero (E and F have co-participated in past events with actor D). None of the other four events closes any two-path that was present before the respective event. Most triangles in the one-mode projection are actually created by hyperevents of size three or larger. Thus, a first insight is that the existence of densely connected groups, local clustering, or an over-representation of closed triangles in hyperevent networks does not give any evidence for a positive closure effect but can alternatively be explained by events of size three or larger — or by the tendency to (partially) repeat such events.

Next we consider the two hyperedges $h = \{A, B, H\}$ and $h' = \{D, G, H\}$, on which events could potentially occur in the next time point and discuss whether such events would provide evidence for a closure effect in a model that controls for subset-repetition of various order. A hypothetical event on the hyperedge $h = \{A, B, H\}$ would give evidence for closure. Indeed A , B and H have previously co-participated each in one event with the common third actor C . Subset-repetition, on the other hand, does not appear to explain a hypothetical event on h . We note that A , B and H individually have participated in only one event each so that they are rather inactive actors, compared to others in the same network. Among the three unordered pairs in $\{A, B, H\}$ there is just one that has one previous joint event so that subset-repetition of order two on the hyperedge h is equal to $1/3$ and therefore is also below average (which is $12/\binom{8}{2} = 0.43$ in this network). Finally $\{A, B, H\}$ have never jointly co-participated in any event, so that sub-repetition of order three (or higher) could not explain an event on h either. Indeed, closure seems to be the only satisfying explanation for a hypothetical event on h – beyond random chance.

This is very different for a hypothetical event on $h' = \{D, G, H\}$. While the closure statistic on h' has a positive value (actually, it is equal to two), a hypothetical event on h' could alternatively be explained by subset-repetition. Indeed, the triad $\{D, G, H\}$ has co-participated in a joint event before and subset-repetition of order two and three on h' are above average, both being equal to one. Thus, assuming that we have positive subset-repetition effects, an event on h' would provide no evidence for closure. (It is obvious that from observing just a handful of events, as in this stylized example, we could not statistically separate the effects of closure from subset-repetition. However, in our exemplary empirical data we have thousands of events and – as it will turn out – we can actually separate these effects.)

Closure and its interplay with subset-repetition has implications for the macro-structure of hyperevent networks on one hand and for the existence of stable structural holes on the other. First we note that repetition and subset-repetition leads to a reinforcement of densely connected groups of actors. For instance, the data from Fig. 4(left) suggests the dense groups $\{A, B, C\}$, $\{C, D, G, H\}$, and $\{D, E, F\}$. Subset-repetition and repetition predicts that these groups will be re-inforced.

Crucially, these groups are overlapping in two “bridging” actors C and D , respectively. These two actors are surrounded by structural holes which could point to a position of power (Burt, 1992). Given such a precondition, the future evolution of the network crucially depends on whether we have a positive or a negative closure effect. A positive closure effect – which could lead for instance to an event on $h = \{A, B, H\}$ – would imply that overlapping dense groups have a tendency to merge over time. In turn, this would close structural holes and C and D would potentially lose their power positions. In contrast, a negative closure effect would prevent, for instance, events on $h = \{A, B, H\}$. In turn, overlapping dense groups would typically not merge and structural holes would remain open. In conclusion, a negative closure effect – in combination with positive (subset-)repetition – could explain overlapping but stable dense groups in dynamic co-participation networks or, from another point of view, is a way to sustain structural holes and thus power positions of actors. We take up this point again when discussing the empirical results on the closure effect in Section 5.

4. Illustrative case study: Thatcher meetings

We seek to establish the empirical value of RHEM in an analysis of empirical data sourced from Margaret Thatcher’s appointment diaries. While the contribution of our paper is methodological, we provide in this section a brief overview of the empirical setting from which the exemplary data stems. Indeed, to demonstrate that RHEM can be fruitfully applied in empirical social network analysis, we illustrate how relevant research questions could be addressed and how findings could advance existing empirical research. While details of the empirical setting are necessarily context-specific, we emphasize that RHEM provide a general model framework for time-stamped multi-actor events from contact diaries.

4.1. Background on the illustrative empirical setting

Network studies are firmly established in political science research (Victor et al., 2017; Ward et al., 2011). The majority are based on cross-sectional designs, but dynamic models such as temporal exponential random graph models (TERGM) and stochastic actor-oriented models (SAOM) have also been applied to longitudinal network data; for an overview, see Desmarais and Cranmer (2017). Longitudinal studies using time-stamped relational event data are increasingly adopted also in the political sciences (Lerner et al., 2013a; Stadtfeld et al., 2017; Brandenberger, 2019) – but models are typically for dyadic events and do not directly apply to multi-actor events transcribed from contact diaries. This is a severe limitation for political science since appointment diaries exist widely and contain masses of information about the working lives of individuals, particularly leaders and managers. Indeed, the empirical data on which we illustrate our models are sourced from the appointment diaries of former British PM Margaret Thatcher, which are publicly accessible and cover a significant portion of her

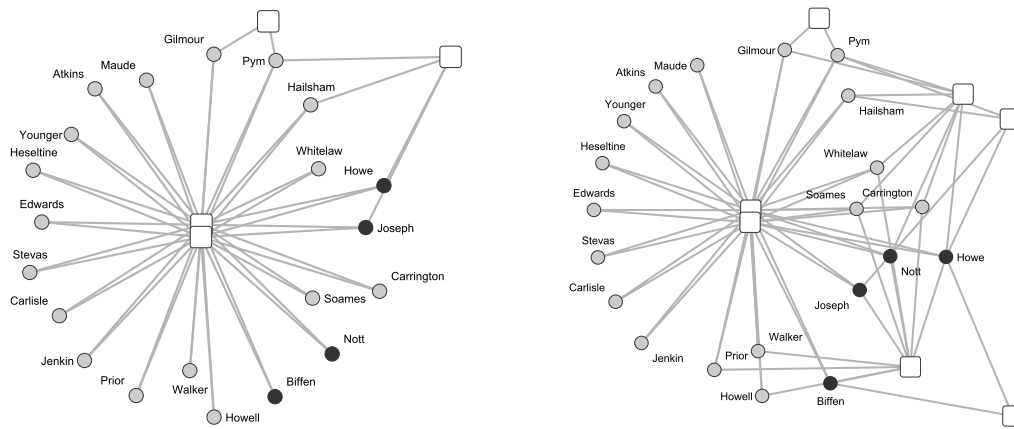


Fig. 5. Co-attendance networks at selected points in time. Circles represent ministers (circles with a dark shade represent dry ministers), squares represent events, and lines connect events to their participants. *Left*: The network after the fourth multi-actor event. *Right*: The network after the seventh multi-actor event.

premiership (Margaret Thatcher Foundation, 2019). Such data makes it possible to study the British core executive using advanced network analytic methods.

Many research themes have emerged in the literature that relate either directly or indirectly to the executive branch of the British government. For example, there are long-standing debates on whether policy implementation is driven by a prime ministerial or collectivist process (Bennister and Heffernan, 2012; Burch and Holliday, 2004; Burnham and Jones, 1993; Byrne and Theakston, 2019; Heffernan, 2003), and investigations into different leadership styles (Bowles et al., 2007; Kaarbo and Hermann, 1998; Theakston, 2011). Our case study illustrates how RHEM can add to this literature by specifically focusing on power dynamics within Thatcher’s early cabinets. When she entered Downing Street in May 1979, Thatcher faced a lack of internal support from several members of her conservative party – known as the “Wets” – who were critical of her hard-line economic policies and considered her to be an untried extremist with a limited shelf life as PM (Cannadine, 2016, p. 29). To increase the chances of implementing her domestic policy agenda, therefore, Thatcher placed her small band of cabinet supporters – i. e., the Thatcherites, or “Dries” – into key economic positions. Then, at subsequent cabinet reorganizations, she took the step of using promotions almost exclusively for policy ends, to gradually shift the balance of power in her favor (King, 1985, p. 132).

4.2. Orienting questions

It is within this context that we demonstrate the potential of the RHEM framework. To illustrate how RHEM could be applied in empirical research, we pose exemplary questions to explore the power balance between the Dries and Non-dries (i. e., Wets and other ministers not labeled as dry) in Thatcher’s early cabinets, and how it impacted her approach to ministerial meetings. Our first question relates to a first-order effect of the actor-level covariate *dry* on the propensity to participate in events. Essentially, it seeks to determine whether the PM preferred to meet with dry ministers over non-dry ministers:

RQ1 Does the chance of a group of ministers jointly meeting the PM increase as the proportion of dry ministers in that group increases?

Our second question also explores the general composition of Thatcher’s ministerial meetings. It addresses a second-order effect of the same covariate to test whether ministerial meetings tend to be homogeneous (mostly dry or mostly non-dry ministers) or heterogeneous (a balance of dry and non-dry ministers) in nature:

RQ2 Does the chance of a group of ministers jointly meeting the PM increase if the group is homogeneous, that is, if it is composed mostly of dry or mostly of non-dry ministers?

Our final question tests for the structural effect of triadic closure in the data. Positive triadic closure would imply that ministers who co-attended meetings with common third ministers have an increased probability to co-attend meetings themselves. Negative triadic closure, on the other hand would imply that brokers (i. e., those who bridge between different groups) typically keep their power positions since their different contacts are kept separate:

RQ3 Does the chance of a group of ministers jointly meeting the PM increase if they have previously co-attended meetings with common third ministers?

The models that we fit later in this paper to address these research questions control for other basic effects in networks of time-stamped multi-actor events, such as the tendency to repeat meetings with the identical or overlapping participant lists. We note that a tendency *against* triadic closure could point to the existence of stable structural holes and in turn to competition or power differences among cabinet ministers. This aspect has been already discussed in Section 3.2.3 and we will discuss it again in light of our results in Section 5.

4.3. Illustrative empirical data

Our illustrative analysis is based on a sequence of 1,989 meeting events of Margaret Thatcher with her cabinet ministers, which took place between 5th May 1979 and 8th June 1983, i. e., Thatcher’s full first term. These events are sourced from the PM’s appointment diaries, which are publicly accessible and cover a significant portion of her premiership (Margaret Thatcher Foundation, 2019). The cabinet consisted of 21 ministers (in addition to the PM), with the exception of eight months from September 1981 when there were 22 ministers. Cabinet composition – that is, the population of actors – changed in four cabinet reorganizations with a total of eight ministers exiting and eight ministers entering cabinet. Individual meeting events in the data are listed by the minute and do not overlap in time. The participant lists of events (i. e., their hyperedges) comprise all attending ministers *except the PM*, since she attended every meeting by definition. Meetings, therefore, involve one, several, or all cabinet ministers. In many cases, other non-cabinet members were involved in meetings. These have been removed from the data to focus explicitly on the PM’s interactions with ministers. In this paper we use a binary actor-level covariate that labels ministers as “dry” (that is, supportive of the PM’s economic agenda) or “non-dry”. Of the 29 ministers who featured in cabinet

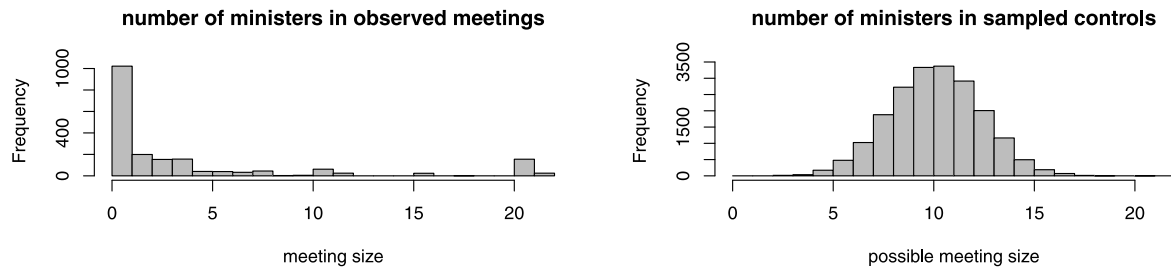


Fig. 6. *Left*: Histogram of sizes (i.e., number of cabinet ministers) of observed meetings in the illustrative empirical data. *Right*: Histogram of sizes of non-event subsets (i.e., groups of ministers that could have experienced a meeting event) sampled uniformly from the set of all cabinet members at the event times.

during Thatcher’s first administration, eight were Dries.⁶ Note that the category of “non-dries” contains ministers considered Wets (those who were opposed to Thatcherism) and those that were considered neither dry nor wet. This covariate does not vary over time.

In Fig. 5 we visualize the co-attendance network after selected points in time (to avoid visual clutter we only display the networks after the first few events in the sequence of almost 2,000 meetings). We display only events with two or more attending participants (*multi-actor events*), since meetings with only one attending minister result in uninformative event nodes of degree one. The layout has been computed by an algorithm for dynamic network visualization proposed by Brandes and Mader (2011) and implemented in the visone⁷ software (Baur et al., 2001). Note that this is an offline algorithm in which node positions for the network at time t take into account network data from time points $t' > t$. We visualize the co-attendance networks for the first 14 multi-actor events in the Appendix in Fig. 7.

The number of sets of actors that could potentially experience joint events depends exponentially on the size of the population. In our illustrative empirical analysis we get $2^{21} \approx 2.1$ million different sets for cabinets with 21 members and for cabinets with 22 members we get $2^{22} \approx 4.2$ million different sets of ministers that could potentially constitute the participant list of any meeting. (Strictly speaking we have to subtract one from these numbers since the empty subset, i.e., a meeting with no participating minister, would not be considered an event in our data.) The distribution of meeting sizes in our empirical data is given in Fig. 6 (*left*). This distribution of observed event sizes is contrasted in Fig. 6 (*right*) with the distribution of the sizes of subsets that are drawn uniformly at random from the set of all cabinet ministers at event times. (The latter are draws from the binomial distribution $f(k) = \binom{n}{k}$ with $n = 21$ or $n = 22$, respectively.) It can be seen that small meetings are over-represented in the empirical data, meetings of intermediate size are very infrequent compared to the large number of subsets of intermediate size, and large meetings are also over-represented – although not to the same degree as small events. As discussed above, the strong dependency of event frequency on subset size is the main reason to consider in this empirical study only RHEM with size-constrained risk sets. We discuss this aspect again after having presented the results. The largest size-constrained risk set in our empirical data contains more than 700,000 hyperedges and results from meetings with 11 participating ministers from a cabinet of size 22.

5. Results

We estimate RHEM parameters from the empirical PM meeting data by maximizing the sampled likelihood given in Eq. (4). For each observed event $e = (t_e, h_e)$ we sample 100 non-event hyperedges from the size-constrained risk set R_{t_e} (i.e., sets of cabinet ministers at t_e

⁶ Dries in the cabinet are: Geoffrey Howe, John Biffen, John Nott, Keith Joseph, Leon Brittan (from Jan-81), Nigel Lawson (from Sept-81), Cecil Parkinson (from Sept-81), and Norman Tebbit (from Sept-81).

⁷ <http://visone.ethz.ch/>.

Table 1

Cox proportional hazard models for event intensities associated with hyperedges (sets of ministers). All three models are estimated on 1,989 events and 101,246 observations (the number of observations is the number of events plus the number of sampled controls).

Repetition	0.38 (0.02)***	0.37 (0.02)***	0.38 (0.02)**
Sub.rep ⁽¹⁾	0.07 (0.04) ^a	0.10 (0.04)**	0.07 (0.04) ^a
Sub.rep ⁽²⁾	2.34 (0.13)***	2.37 (0.14)***	2.36 (0.14)***
Sub.rep ⁽³⁾	2.33 (0.14)***	2.31 (0.14)***	2.30 (0.14)***
closure	−3.01 (0.30)***	−3.11 (0.30)***	−3.10 (0.30)***
Avg.dry	0.05 (0.02) ^a		0.04 (0.02) ^a
Hom.dry		−0.12 (0.05) ^a	−0.11 (0.05) ^a
AIC	9,983.14	9,982.50	9,979.74

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ^a $p < 0.1$.

of the same size as h_e). If the size-constrained risk set has less than 100 elements (which happens for instance for observed events of size one), we take the full risk set. We set the half life for the decay of past events to 30 days. Thus, an event counts close to one right after the event, one month later it counts 1/2, and so on.⁸ Effects in the hyperevent model therefore assess the impact of *recent* events. Before estimating the parameters, we standardize statistics to mean equal to zero and standard deviation equal to one. Since statistics have no natural units, this facilitates the interpretability of the relative effect sizes. Table 1 reports estimated parameters of three models. All models include repetition, subset-repetition of order one, two, and three, and the closure statistic. The first model additionally includes the statistic average-dryness of sets of ministers, the second model includes dry-homogeneity, and the third contains both covariate effects. Below, we discuss the results mostly independent of the specific empirical context. Recall that it is not the objective of this paper to actually draw conclusive empirical insight, but rather to present a model that can be fruitfully applied to this and related empirical settings.

Repetition and subset-repetition. The parameter associated with *repetition* is significantly positive. Thus, there is a tendency to repeat (recent) meetings with the identical set of participants. We also find a positive effect of subset-repetition of order one (in two of the models, this effect is only significant at the 10%-level). Thus, actors who (recently) participated in more meetings, are more likely to be included among the participants of future events. This can be seen as a preferential attachment effect where “popular” actors (i.e., those who were more frequent participants in the past) accumulate future interaction at a higher rate. Moreover, we find a positive effect of subset-repetition of order two and three. Thus, sets of actors containing dyads (or triads) that previously co-participated in joint meetings experience future meetings at a higher rate. This can be understood as a *familiarity* effect:

⁸ In preliminary tests we found that models with time decay consistently have a better fit than models without time decay. Moreover, for checking robustness, we repeated our analysis with a half life set to 7 days and again with a half life set to 90 days (not reported in this paper). Qualitatively, results have been very similar to those reported below.

Table 2

Cox proportional hazard models for event intensities associated with hyperedges (sets of ministers). All three models are estimated on 1,989 events and 101,246 observations.

Repetition	0.39 (0.02)***	0.45 (0.02)***	0.38 (0.02)***
Sub.rep ⁽¹⁾	0.12 (0.04)***	−0.04 (0.04)	0.07 (0.04) ^a
Sub.rep ⁽²⁾		2.83 (0.14)***	2.36 (0.14)***
Sub.rep ⁽³⁾			2.30 (0.14)***
Closure	4.23 (0.19)***	−1.41 (0.30)***	−3.10 (0.30)***
Avg.dry	0.06 (0.02)**	0.06 (0.02)**	0.04 (0.02)*
Hom.dry	−0.05 (0.04)	−0.14 (0.05)**	−0.11 (0.05)*
AIC	10,724.31	10,257.10	9,979.74

****p* < 0.001, ***p* < 0.01, **p* < 0.05, ^a*p* < 0.1.

actors have a tendency to co-attend meetings with others they have recently met before. This effect is not restricted to dyadic familiarity, where it just matters whether actors have *pairwise* met before, but also applies to triadic familiarity. For instance, in the example given in Fig. 2 the hyperedge h_3 would be predicted to have a higher event rate than the hyperedge h_2 and h_2 would be predicted to have a higher rate than h_1 . As discussed above, subset-repetition and repetition can reinforce dense groups of actors, or local clustering.

Covariate effects: dry and homogeneous meetings. The parameter associated with the average-dryness statistic – instantiating the statistic “average- x ” in this case study – is significantly positive. Thus, dry ministers are more likely to be among the participants of meetings than non-dry ministers. This finding could be interpreted in the sense that the dry ministers are in more powerful positions (having more opportunities to meet the PM), or – from another angle – that the PM has a preference to meet dry ministers. Moreover, we find a significantly negative parameter associated with dry-homogeneity. Thus, there seems to be a tendency to mix dry and non-dry ministers when assembling the participants of meetings. Taking these effects together, and looking at the example given in Fig. 3, we would expect a tendency for meetings somewhere “right of the middle” (i. e., towards the “drier” hyperedges) but not at the extreme right. These two covariate effects keep their sign when both are included in the same model. We note that there are fewer dry ministers than non-dry ministers in the cabinets (i. e., population of actors) and therefore, if this covariate had no effect on meeting frequency, then by chance alone we would expect more Non-dries than Dries among the participants of observed meetings. Thus, the negative effect of dry-homogeneity – which pushes hyperedge composition towards balanced meetings with about the same number of dry and non-dry ministers – tends to also increase the number of dry meeting attendees over random chance.

Closure effects. We consistently find a negative closure effect. Thus, looking at the example given in Fig. 4, the hyperedge h , which closes several open two-paths but could not be well explained by subset-repetition, would be predicted a rather low probability to experience meetings. As discussed in Section 3.2.3, a negative closure effect – together with positive subset-repetition – can explain overlapping but stable (i. e., non-merging) dense subgroups and the maintenance of stable structural holes. This, in turn could point to actors that maintain power positions by bridging structural holes (Burt, 1992).

Interplay between subset-repetition and closure. How does the estimated negative closure effect depend on whether we control for subset-repetition or not? To shed light on this question, we fit models, reported in Table 2, that control only for subset-repetition of order one (i. e., for past activity of individual actors), that control for subset-repetition of order one and two, and compare these to the model including subset-repetition up to order three.

The crucial difference is between the leftmost model, which controls only for individual activity, and the two others which also control for prior shared events of dyads or triads, respectively. If we control only for past individual participation in meetings, but fail to control for past co-participation, we spuriously estimate a *positive* closure effect. This

can be well explained with the example given in Fig. 4. The hyperedge h' closes several triangles – but could alternatively be explained by subset-repetition of order two or three. However, a model that fails to control for prior shared events misses the difference between events on the hyperedge h (which would point to closure) and events on h' (which would give no evidence for closure since they can be explained by subset-repetition or order two or higher). Since events on hyperedges like h' are apparently frequent in our data (due to subset-repetition), the first model in Table 2 falsely estimates a positive closure effect.

Explanatory power of estimated RHEM. We might ask how well the fitted models succeed in “recognizing” the true observed events among their associated alternatives. Here we do not consider out-of-sample prediction, since we believe that picking the right event out of 4 million alternatives (or up to 700,000 for conditional-size models) is a fairly impossible task. We also argue that RHEM – as we specify and apply them in this paper – are made for testing hypotheses in dynamic co-attendance networks, rather than to foresee the future. This view is consistent with the observation in Block et al. (2018) that in longitudinal network studies, models with the best performance in out-of-sample prediction often have limited relevance for inferential network modeling. Therefore, we fit RHEM to the whole observed data and then compare – separately for each observed event $e = (t_e, h_e)$ – the implied relative event rate on the event hyperedge h_e with the implied rate on all associated controls sampled from the risk set R_{t_e} . The relative event rate of a hyperedge h' at the time t_e (implied by a fitted model, that is, with given parameters θ) is

$$rel.rate(h') = \frac{\lambda_1(t_e; h'; \theta; G[E; t_e])}{\sum_{h \in \bar{R}_{t_e}} \lambda_1(t_e; h; \theta; G[E; t_e])}$$

compare Eq. (4). We report results for the largest model; see the right-most column of Table 1.

This model seems to succeed fairly well in assigning high relative event rates to hyperedges of observed events. More than 39% of all observed events are assigned the highest event rate among all *sampled* alternatives (note that more than one hyperedge can be assigned the same maximum event rate). Moreover, we compute for each observed event $e = (t_e, h_e)$ the percentile of the associated relative rate $rel.rate(h_e)$ in the vector of values $(rel.rate(h))_{h \in \bar{R}_{t_e}}$. For comparison, under random guessing we would expect percentiles equal to 0.5. The model seems to perform reasonably well also from that perspective: the median percentile over all observed events is more than 0.95, implying that the “typical” event hyperedge is surpassed by less than 5% of its associated alternatives. Note that this is nevertheless consistent with our view that predicting the correct event hyperedge is a difficult task, since 5% of 4 million (or up to 700,000 in conditional-size models) is still a large number. We think that developing more sophisticated methods to assess the fit of RHEM is a relevant task for future work.

5.1. Limitations of the empirical case study

We recall that the contribution of this paper is methodological, that the empirical case study has been included for illustrating the usefulness of RHEM, and that it is not our objective here to draw conclusive insights into Thatcher’s government. In fact, the concrete model specification that we applied to our empirical data is quite representative for basic RHEM and is almost independent of the concrete empirical setting. It seems plausible that (subset-)repetition, closure, and covariate effects (main effect and homophily) are present in most instances of relational hyperevent networks – although the nature of the covariate will most likely be different in other settings. We avoided to include effects that are specific to the concrete setting. For that reason, we discuss some of the limitations of our case study and how these limitations could be tackled in future work that seeks to draw empirical insights.

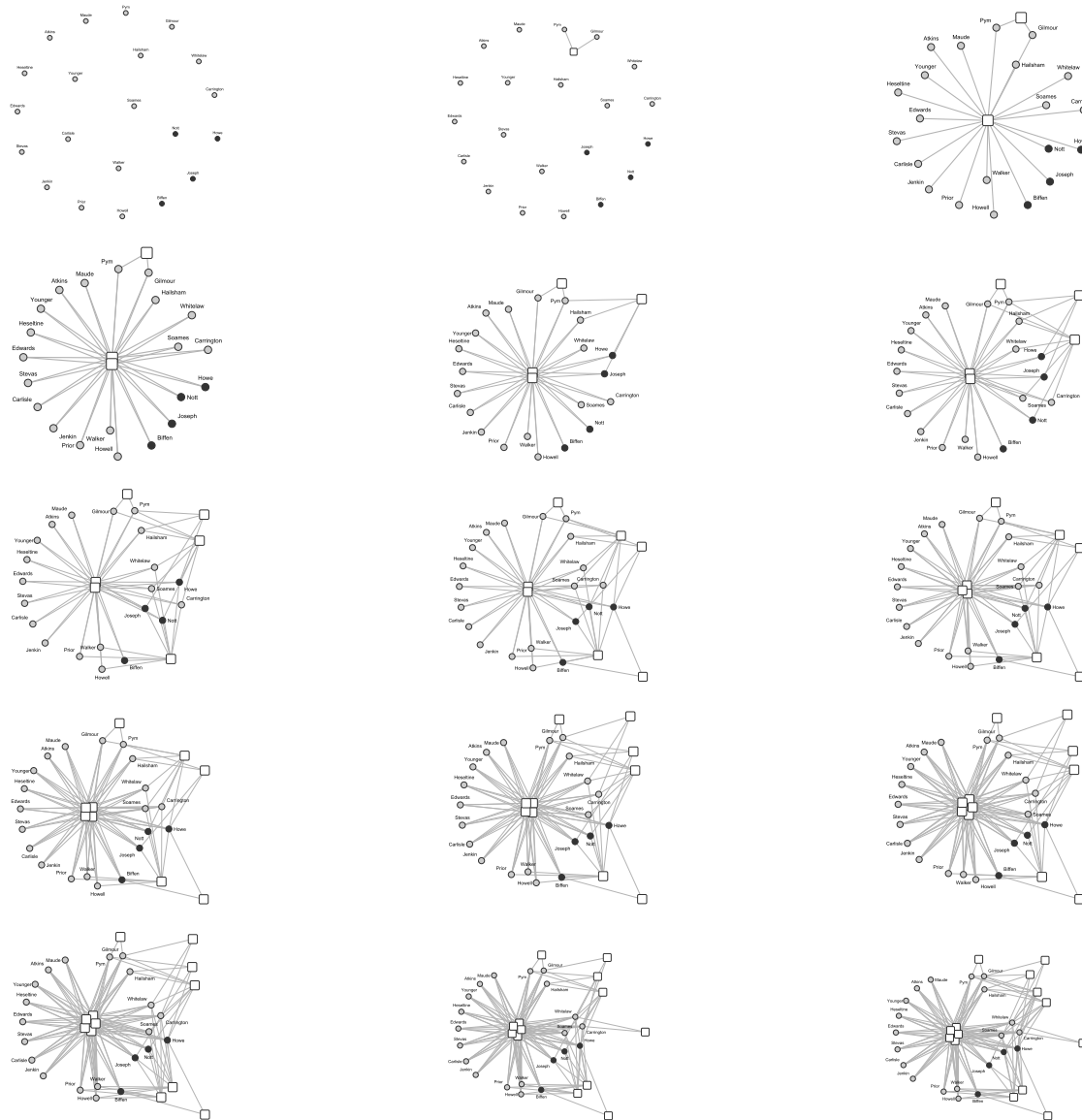


Fig. 7. Co-attendance networks at the time of the first 14 multi-actor events in the Thatcher meeting data. Circles represent ministers (circles with a dark shade represent dry ministers), squares represent events, and lines connect events to their participants.

In our study we considered only one covariate – “dry”, which intuitively identifies ministers supporting Thatcher’s political agenda – and we did not distinguish any type of meetings. A more sophisticated analysis might consider ministerial roles, membership in committees, or restricted access to some meetings, e.g., due to national security interests. Some meeting events are “formal” cabinet meetings or cabinet committee meetings which, in the medium term, have static attendee lists, while other events correspond to “ad-hoc” meetings with unconstrained participant lists. There are a number of possibilities to extend RHEM in order to control for such constraints. Ministerial roles or committee membership could be included via covariate effects and it is possible to specify separate models for, e.g., formal and ad-hoc meetings, or to interact effects with dummy variables indicating the type of meetings. If attendee lists of meetings are constrained – either by mandatory participation or restricted access – then this could be incorporated by specifying the risk set accordingly: if it is known that, say, minister *A* necessarily has to attend a meeting event $e = (t_e, h_e)$, then the risk set at time t_e should comprise only hyperedges that contain *A*; if it is known that minister *B* must not attend such a meeting

(for instance, due to national security interests), then the associated risk set should comprise only hyperedges not containing *B*. Since such constraints could explain latent dense groups or separation between groups of actors, not including them in the model might distort findings on effects such as repetition, homogeneity, or closure.

Another issue worth mentioning is our approach to condition the risk sets to the observed event size. We argue that in our empirical data the event rate associated with hyperedges depends strongly on hyperedge size, with factors up to 500,000. Conditional-size RHEM seem to be a way to avoid comparison of hyperedges with such different baseline rates. The decision to condition on hyperedge size, however, has implications for the interpretation of results. For instance, a finding such as the preference for meeting dry ministers in a conditional-size RHEM does not imply a tendency to add more and more dry ministers to the set of attendees (since this would change the size of the meeting event) but rather a tendency to add dry ministers whilst removing the same number of non-dry ministers, thereby increasing the overall proportion of dry attendees at the meeting event. Thus, the point of view in conditional-size models is that actors compete for participation

in meetings. Another implication is that hypotheses concerning event size (such as, a preference for having large or small meetings) cannot be tested with conditional-size RHEM. It is worth noting that the approach to condition hyperedge size cannot be justified by substantive arguments in our empirical data, as it seems unlikely that the PM can only meet with groups of ministers of a given size at a given point in time. We consider conditional-size RHEM rather as a way to prevent that hyperedges of observed events are compared with alternative hyperedges of different size that are known to have very different baseline intensities. For robustness checks we estimated RHEM with unconstrained risk sets, controlling for a linear and quadratic effect of hyperedge size on meeting intensity — similar to the approach taken by Lerner et al. (2019). Some of these findings were consistent with those made via conditional-size models: we found positive effects for repetition, subset-repetition of order two, and average-dryness and a negative closure effect. However, subset-repetition of order one and three, as well as dry-homogeneity, were no longer significant in RHEM with unconstrained risk sets. Moreover, findings made via unconstrained models are more sensitive to the inclusion or exclusion of effects. In summary, while some of the empirical findings seem to be robust, we argue that conditional-size models are a way to filter out the strong dependency of event rates on hyperedge size. An alternative to this approach would be to develop RHEM with unconstrained risk sets that better control for the effect of size. In general, the question whether to condition RHEM on observed event sizes needs future work (also see the discussion below).

6. Conclusion

Relational event models enable the analysis of time-stamped social interaction data that frequently represent the micro-structure of social networks (Pallotti et al., 2020). Because of their emphasis on temporally ordered event sequences, available relational event models are ill-suited as models for clusters of simultaneous events like those recorded in contact diaries where one person (the diarist) typically meets simultaneously multiple alters at a specific point in time giving rise to a cluster of dyadic events with the same time-stamp. This one-to-many and many-to-one “multicasting” situation is more common than our empirical illustration might suggest. For example, situations where ego interacts with multiple alters simultaneously are commonly encountered in empirical studies of conversation (Gibson, 2005) and in the analysis of venture capital investment syndicates (Sorenson and Stuart, 2008).

The main methodological contribution of this paper was to extend REMs by proposing RHEM as a general statistical framework for the analysis of time-stamped events involving multiple actors simultaneously. We have shown how RHEM may be used to estimate event frequencies associated with all subsets (i. e., hyperedges) of a given set of actors. By specifying appropriate hyperedge statistics, it is possible to rigorously test a wide range of hypotheses about how characteristics of a group may influence the probability that group members experience common events. Examples of characteristics of a group that may be of theoretical or empirical interest include its socio-demographic composition and its embeddedness in the network of prior interaction events.

We have exemplified the empirical applicability of RHEM using data extracted for the contact diaries of Ms. Margaret Thatcher. Our illustrative analysis revealed relevant patterns in the structure and dynamics of hyperevent networks. We discussed that an over-representation of closed triangles in one-mode projections of co-attendance networks does *not* necessarily give evidence for triadic closure — but can in a more straightforward manner be explained by “large” events (i. e., events involving three or more actors), or by the tendency to (partially) repeat such events. In our empirical analysis we rather find a tendency *against* triadic closure, given that we control for repetition and subset-repetition. We argued that negative triadic closure can explain the

existence of overlapping but stable (i. e., non-merging) dense groups and, in turn, the existence of actors occupying stable broker positions. In contrast, *positive* triadic closure would predict that overlapping groups tend to merge over time, thereby closing structural holes.

In closing, it may be useful to repeat that while we demonstrated the value of RHEM in a specific empirical context, the model itself is general and may be applied to data produced by a variety of social settings. RHEM apply to network data taking the form of a sequence of relational hyperevents, if the following scope conditions hold. Event times have to be available at a sufficient level of resolution to allow considering events as conditionally independent of each other, given the network of previous events (Lerner et al., 2013b). The framework allows simultaneous events, if conditional independence can be credibly assumed. Moreover, the approach to specify event rates at a given time as a function of all previous events assumes that decisions regarding participant lists are done at event time — an assumption that might be violated if events (with fixed set of attendees) are scheduled significantly in advance. A possible way to deal with such situations would be to specify event rates dependent on past events at the time of planning the event. Another scope condition deserving notice is that event participants have to be extracted from a larger set of “possible contacts” containing a population of known actors. This set itself may vary over time, but the identities of the actors at risk of experiencing an event must be known at the time of observed events.

Further developing RHEM opens up several avenues for future work. As we discussed above, the approach to condition RHEM on observed event sizes may not be appropriate in all settings. An alternative would be to control for the effect of event size on relative rates by including (functions of) hyperedge size in the model statistics. A preliminary analysis suggests that quadratic polynomials of hyperedge size do not seem to be sufficient — at least not for the empirical data we considered in this paper. Future work might explore fixed effects for (some) event sizes or interacting other explanatory variables with size. Another direction for advancing RHEM methodology would be to further develop methods for model selection and assessment of model fit. For instance, we might compare statistics of possible events predicted by a RHEM with those of observed events to assess whether models are likely to generate plausible events.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. Visualization of co-attendance networks for the first 14 events

In Fig. 7 we visualize the co-attendance network before and after the first 14 multi-actor events. To avoid visual clutter we choose the earliest time points in the sequence of almost 2000 events. We display only events with two or more attending participants (*multi-actor events*), since events with only one attending minister result in uninformative event nodes of degree one. The layout has been computed by an algorithm for dynamic network visualization proposed by Brandes and Mader (2011) and implemented in the visone⁹ software (Baur et al., 2001). Note that this is an offline algorithm in which node positions for the network at time t take into account network data from time points $t' > t$.

⁹ <http://visone.ethz.ch/>.

References

- Alaszewski, A., 2006. *Using Diaries for Social Research*. Sage.
- Bartlett, R., Milligan, C., 2015. *What Is Diary Method?* Bloomsbury Academic.
- Bartolucci, F., Mira, A., Peluso, S., 2018. Marginal models with individual-specific effects for the analysis of longitudinal bipartite networks. *arXiv preprint arXiv:1810.08778*.
- Baur, M., Benkert, M., Brandes, U., Cornelsen, S., Gaertler, M., Köpf, B., Lerner, J., Wagner, D., 2001. Visone software for visual social network analysis. In: *International Symposium on Graph Drawing*. Springer, pp. 463–464.
- Bennister, M., Heffernan, R., 2012. Cameron as prime minister: The intra-executive politics of Britain's coalition government. *Parliam. Aff.* 65, 778–801.
- Block, P., Koskinen, J., Hollway, J., Steglich, C., Stadtfeld, C., 2018. Change we can believe in: Comparing longitudinal network models on consistency, interpretability and predictive power. *Soc. Netw.* 52, 180–191.
- Bolger, N., Davis, A., Rafaeli, E., 2003. Diary methods: Capturing life as it is lived. *Annu. Rev. Psychol.* 54, 579–616.
- Borgan, Ø., Goldstein, L., Langholz, B., 1995. Methods for the analysis of sampled cohort data in the Cox proportional hazards model. *Ann. Stat.* 1749–1778.
- Borgatti, S.P., Everett, M.G., 1997. Network analysis of 2-mode data. *Soc. Netw.* 19, 243–270.
- Bowles, N., King, D.S., Ross, F., 2007. Political centralization and policy constraint in british executive leadership: Lessons from American presidential studies in the era of sofa politics. *Br. Politics* 2, 372–394.
- Brandenberger, L., 2019. Predicting network events to assess goodness of fit of relational event models. *Political Anal.* 27, 556–571.
- Brandes, U., Lerner, J., Snijders, T.A., 2009. Networks evolving step by step: statistical analysis of dyadic event data. In: *Proc. Intl. Conf. Advances in Social Network Analysis and Mining (ASONAM)*. IEEE Computer Society, pp. 200–205.
- Brandes, U., Mader, M., 2011. A quantitative comparison of stress-minimization approaches for offline dynamic graph drawing. In: *International Symposium on Graph Drawing*. Springer, pp. 99–110.
- Breiger, R.L., 1974. The duality of persons and groups. *Soc. Forces* 53, 181–190.
- Bretto, A., 2013. *Hypergraph Theory*. Cham: Springer.
- Burch, M., Holliday, I., 2004. The Blair government and the core executive. *Gov. Oppos.* 39, 1–21.
- Burnham, J., Jones, G.W., 1993. Advising Margaret Thatcher: The prime minister's office and the cabinet office compared. *Political Stud.* 41, 299–314.
- Burt, R.S., 1992. *Structural Holes: The Social Structure of Competition*. Harvard University Press.
- Butts, C.T., 2008. A relational event framework for social action. *Sociol. Methodol.* 38, 155–200.
- Byrne, C., Theakston, K., 2019. Understanding the power of the prime minister: structure and agency in models of prime ministerial power. *Br. Politics* 14, 329–346.
- Cannadine, D., 2016. *Margaret Thatcher: A Life and Legacy*. Oxford University Press.
- Chodrow, P., Mellor, A., 2020. Annotated hypergraphs: models and applications. *Appl. Netw. Sci.* 5, 9.
- Conaldi, G., Lomi, A., 2013. The dual network structure of organizational problem solving: a case study on open source software development. *Soc. Netw.* 35, 237–250.
- Conaldi, G., Lomi, A., Tonellato, M., 2012. Dynamic models of affiliation and the network structure of problem solving in an open source software project. *Organ. Res. Methods* 15, 385–412.
- Cox, D., 1972. Regression models and life-tables. *J. R. Stat. Soc. Ser. B* 34, 187–202.
- Davis, A., Gardner, B., Gardner, M., 1941. *Deep South*. The University of Chicago Press.
- Desmarais, B.A., Cranmer, S.J., 2017. Statistical inference in political networks research. In: Victor, J.N., Montgomery, A.H., Lubell, M. (Eds.), *The Oxford Handbook of Political Networks*, Oxford University Press, pp. 203–220.
- Doreian, P., Batagelj, V., Ferligoj, A., 2004. Generalized blockmodeling of two-mode network data. *Soc. Netw.* 26, 29–53.
- DuBois, C., Butts, C.T., McFarland, D., Smyth, P., 2013. Hierarchical models for relational event sequences. *J. Math. Psychol.* 57, 297–309.
- Everett, M.G., Borgatti, S.P., 1993. An extension of regular colouring of graphs to digraphs, networks and hypergraphs. *Soc. Netw.* 15, 237–254.
- Everett, M.G., Broccatelli, C., Borgatti, S.P., Koskinen, J., 2018. Measuring knowledge and experience in two mode temporal networks. *Soc. Netw.* 55, 63–73.
- Foster, B.L., Seidman, S.B., 1982. Urban structures derived from collections of overlapping subsets. *Urban Anthropol.* 177–192.
- Freeman, L.C., 2003. Finding social groups: A meta-analysis of the southern women data. In: Breiger, R., Carley, C., Pattison, P. (Eds.), *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*. National Research Council, The National Academies Press, Washington, DC, pp. 39–97.
- Fu, Y.-c., 2007. Contact diaries: Building archives of actual and comprehensive personal networks. *Field Methods* 19, 194–217.
- Fu, Y.-c., Ho, H.-C., Chen, H.M., 2013. Weak ties and contact initiation in everyday life: Exploring contextual variations from contact diaries. *Soc. Netw.* 35, 279–287.
- Gibson, D.R., 2005. Taking turns and talking ties: Networks and conversational interaction. *Am. J. Sociol.* 110, 1561–1597.
- Hackett, P., 1989. *The andy warhol diaries*. In: Warner Books.
- Heffernan, R., 2003. Prime ministerial predominance? Core executive politics in the UK. *Br. J. Politics Int. Relat.* 5, 347–372.
- Heffernan, R., 2005. Exploring (and explaining) the british prime minister. *Br. J. Politics Int. Relat.* 7, 605–620.
- Heppell, T., 2020. The ideological composition of the parliamentary conservative party from thatcher to may. In: *Thatcherism in the 21st Century*. Springer, pp. 15–34.
- Hoffman, M., Block, P., Elmer, T., Stadtfeld, C., 2020. A model for the dynamics of face-to-face interactions in social groups. *Netw. Sci.* 1–22.
- Jackson, B., Saunders, R., 2012. *Making Thatchers Britain*. Cambridge University Press.
- Kaarbo, J., Hermann, M.G., 1998. Leadership styles of prime ministers: How individual differences affect the foreign policymaking process. *Leadersh. Q.* 9, 243–263.
- Kim, B., Schein, A., Desmarais, B.A., Wallach, H., 2018. *The hyperedge event model*. *arXiv preprint arXiv:1807.08225*.
- King, A., 1985. Margaret Thatcher: The style of a prime minister. In: *The British Prime Minister*. Springer, pp. 96–140.
- Koskinen, J., Edling, C., 2012. Modelling the evolution of a bipartite network—Peer referral in interlocking directorates. *Soc. Netw.* 34, 309–322.
- Krivitsky, P.N., Hancock, M.S., 2014. A separable model for dynamic networks. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 76, 29–46.
- Lawless, J.F., 2003. *Statistical Models and Methods for Lifetime Data*. Wiley.
- Lerner, J., Bussmann, M., Snijders, T.A., Brandes, U., 2013a. Modeling frequency and type of interaction in event networks. *Corvinus J. Sociol. Soc. Policy* 4, 3–32.
- Lerner, J., Indlekofer, N., Nick, B., Brandes, U., 2013b. Conditional independence in dynamic networks. *J. Math. Psychol.* 57, 275–283.
- Lerner, J., Lomi, A., 2020. Reliability of relational event model estimates under sampling: how to fit a relational event model to 360 million dyadic events. *Netw. Sci.* 8, 97–135.
- Lerner, J., Tranmer, M., Mowbray, J., Hancean, M.-G., 2019. *REM beyond dyads: relational hyperevent models for multi-actor interaction networks*. *arXiv preprint https://arxiv.org/abs/1912.07403*.
- Marcum, C.S., Butts, C.T., 2015. Constructing and modifying sequence statistics for relevant using informr in R. *J. Stat. Softw.* 64, 1.
- Margaret Thatcher Foundation, 2019. *Large scale document archive*. Retrieved from <https://www.margaretthatcher.org/archive>.
- Mehl, M., Conner, T. (Eds.), 2012. *Handbook of Research Methods for Studying Daily Life*. Guilford Press.
- Pallotti, F., Weldon, S.M., Lomi, A., 2020. Lost in translation: Collecting and coding data on social relations from audio-visual recordings. *Social Networks*.
- Perry, P.O., Wolfe, P.J., 2013. Point process modelling for directed interaction networks. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 75, 821–849.
- Richardson, J., 2018. The changing british policy style: From governance to government? *Br. Politics* 13, 215–233.
- Seidman, S.B., 1981. Structures induced by collections of subsets: A hypergraph approach. *Math. Soc. Sci.* 1, 381–396.
- Snijders, T.A., Lomi, A., Torló, V.J., 2013. A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Soc. Netw.* 35, 265–276.
- Sorenson, O., Stuart, T.E., 2008. Bringing the context back in: Settings and the search for syndicate partners in venture capital investment networks. *Adm. Sci. Q.* 53, 266–294.
- Stadtfeld, C., Geyer-Schulz, A., 2011. Analyzing event stream dynamics in two-mode networks: An exploratory analysis of private communication in a question and answer community. *Soc. Netw.* 33, 258–272.
- Stadtfeld, C., Hollway, J., Block, P., 2017. Dynamic network actor models: Investigating coordination ties through time. *Soc. Methodol.* 47, 1–40.
- Stafford, F.P., 2009. Timeline data collection and analysis: Time diary and event history calendar methods. In: *Calendar and Time Diary Methods in Life Course Research*. pp. 13–30.
- Theakston, K., 2011. Gordon Brown as prime minister: Political skills and leadership style. *Br. Politics* 6, 78–100.
- Therneau, T.M., Grambsch, P.M., 2013. *Modeling Survival Data: Extending the Cox Model*. Springer Science & Business Media.
- Valeeva, D., Heemskerk, E.M., Takes, F.W., 2020. The duality of firms and directors in board interlock networks: A relational event modeling approach. *Soc. Netw.* 62, 68–79.
- Victor, J.N., Montgomery, A.H., Lubell, M. (Eds.), 2017. *The Oxford Handbook of Political Networks*, Oxford University Press.
- Vu, D., Pattison, P., Robins, G., 2015. Relational event models for social learning in MOOCs. *Soc. Netw.* 43, 121–135.
- Wang, P., Pattison, P., Robins, G., 2013. Exponential random graph model specifications for bipartite networks—A dependence hierarchy. *Soc. Netw.* 35, 211–222.
- Ward, M.D., Stovel, K., Sacks, A., 2011. Network analysis and political science. *Annu. Rev. Political Sci.* 14, 245–264.
- Williamson, A., 2015. *Conservative Economic Policymaking and the Birth of Thatcherism*. Palgrave Macmillan, pp. 1964–1979.
- Yen, T.-J., Fu, Y.-c., Hwang, J.-S., 2016. Alters as species: predicting personal network size from contact diaries. *Soc. Netw.* 45, 78–88.