MODELING FREQUENCY AND TYPE OF INTERACTION IN EVENT NETWORKS

Jürgen Lerner[1]
Margit Bussmann[2]
Tom A. B. Snijders[3]
Ulrik Brandes[4]

Abstract Longitudinal social networks are increasingly given by *event data;* i.e., data coding the time and type of interaction between social actors. Examples include networks stemming from computer-mediated communication, open collaboration in wikis, phone call data and interaction among political actors. In this paper, we propose a general model for networks of dyadic, typed events. We decompose the probability of events into two components: the first modeling the frequency of interaction and the second modeling the conditional event type, i. e., the quality of interaction, given that interaction takes place.

While our main contribution is methodological, for illustration we apply our model to data about political cooperation and confiicts collected with the Kansas Event Data System. Special emphasis is given to the fact that some explanatory variables affect the frequency of interaction while others rather determine the level of cooperativeness vs. hostility, if interaction takes place. Furthermore, we analyze if and how model components controlling for network dependencies affect findings on the effects of more traditional predictors such as geographic proximity or joint alliance membership. We argue that modeling the conditional event type is a valuable – and in some cases superior – alternative to previously proposed models for networks of typed events.
Keywords network analysis, statistical network models, event data, signed networks, structural balance theory

1 Jürgen Lerner is postdoctoral researcher in Computer Science at University of Konstanz; e-mail: *lerner@inf.uni-konstanz.de*

2 Margit Bussmann is professor of International Politics at University of Greifswald; e-mail: *margit.bussmann@uni-greifswald.de*

3 Tom A. B. Snijders is professor of Statistics at University of Oxford; e-mail: *tom.snijders@nuffield.ox.ac.uk*

4 Ulrik Brandes is professor of Computer Science at University of Konstanz; e-mail: *Ulrik.Brandes@uni-konstanz.de*

1 INTRODUCTION

More and more social network datasets encode interaction events (such as sending an email or co-authoring a scientific article) rather than relational states between actors (such as friendship or esteem). The increased availability of event data is especially due to the advent of automated data collection facilities. For instance, log-data of computer mediated communication (e.g., email, Usenet-groups, or social network services), open collaboration in wikis, or phone-call data naturally gives rise to event networks. In this paper we consider networks of political actors together with interaction events that are routinely observed and reported in the news.

We consider networks of dyadic, typed events where the type is a real number indicating the level of cooperativeness (if positive) or hostility (if negative). From a modeling point of view, the general research questions that we consider here are about the causes and effects of network interaction. When analyzing the causes of network events, the network is seen as a dependent variable and one seeks to answer questions like *what makes actor A interact more or less with actor B or what makes actor A engage in a specific type of interaction towards actor B*. When analyzing the effects of events, the network is seen as an explanatory variable and one seeks to answer *what results from interaction (of a certain type)*. Here we consider network events to be both the dependent and the explanatory variables; more specifically, we want to find out how past events (and externally given actor and dyad covariates) stochastically determine the frequency and type of future events.

The occurrence of events of specific types can be modeled in at least two distinct ways whose difference is crucial for this paper. For illustration, assume that we want to model event networks with two types of interaction, positive events encoding cooperation and negative events encoding hostilities, and that we want to test hypotheses about the causes of both types of events. The first way to do so is to adapt the model proposed by Butts (2008). In that model there are two different rate functions for the two types of events and the estimated parameters reveal which explanatory variables increase or decrease the frequency of cooperative or hostile events, respectively. The second possibility to model typed events is the one that we propose in this paper. In our model there is one rate function modeling the frequency of events of any type, and a type function modeling the conditional probability of cooperative versus hostile interaction, given that interaction occurs.[5] Thus, the estimated parameters in our model reveal what triggers:

1. an increase/decrease in the frequency of interaction;
2. positive vs. negative interaction, given that interaction occurs.

There are at least two benefits resulting from this alternative model for typed events. First, the conditional event type models are not restricted to a finite number of event types (i. e., to categorical event types) but can also deal with types characterized by continuous variables. Second, the results stemming from the conditional event type models provide additional information about the causes of events and may clarify seemingly counterintuitive findings that result from modeling the frequency of typed events separately. For instance, we demonstrate in the next section that a bivariate model for international relations suggests that countries have an increased probability of engaging in a militarized dispute with their alliance partners (compared to countries with which they do not share an alliance membership). On the other hand, an application of conditional event type models reveals that allies consistently show a tendency to engage in cooperative rather than conflictive interaction– under the precondition that they do interact.

The remainder of this article is structured as follows. Section 2 introduces a dataset on which we conduct an illustrative analysis, reports related previous results in international relations research and develops the exemplary hypotheses. Our newly-proposed model is described in Section 3 and results of the illustrative application of the model are given in Section 4. Section 5 concludes and indicates future research.

2 POLITICAL NETWORK ANALYSIS

Scholars of international politics increasingly realize the advantages of network analysis in various contexts (e. g., Maoz 2009; Hafner-Burton and Montgomery 2006). One approach within social network analysis, structural balance theory, is particularly well suited to addressing questions of cooperations and conflict between states.[6] A signed network (i. e., a network

5 Since we consider weighted events later in this paper, we model the conditional probability density for event weights, rather than the conditional probability of positive/negative events; the latter serves only for simplified illustration.

6 For a detailed description of structural balance theory see Heider (1946); Cartwright and Harary (1956).

with positive and negative ties) is *balanced* if every semi-cycle has an even number of negative ties.[7] Structural balance theory (SBT) claims that actors have a preference for balanced networks. Specifically, if two ties in a triplet of actors are present and the third tie is to be created then its sign is predicted by the following four rules resulting from SBT: "the friend of a friend is a friend," "the friend of an enemy is an enemy," "the enemy of a friend is an enemy," and "the enemy of an enemy is a friend."

The influence of common friends and enemies on a dyad in political networks has been analyzed in Maoz et al. (2007) and Crescenzi (2007). Maoz et al. (2007) compute the conditional probabilities of alliances and militarized interstate disputes (MIDs) between two countries, given that these satisfy the conditions of being (1) friends of enemies, (2) enemies of friends, and (3) enemies of enemies (the relations *friend* and *enemy* are derived from the alliance and MID relations, respectively). It turns out that all three preconditions increase both the probability of alliances and the probability of MIDs. Thus, the results simultaneously support and reject structural balance theory. Seen from a different angle, actors that are indirectly related via a third actor have a higher probability to interact–both positively and negatively. This result can be refined by applying our newly-proposed network model: later in this paper we show that actors that are (say) enemies of enemies have a higher probability of interaction but, given that they do interact, their relation has a tendency towards cooperation–clearly supporting SBT. In related work, Crescenzi (2007) defined a combined dyadic indicator that is positive if the two actors evaluate most other actors consistently (both positive or both negative), negative if they evaluate most other actors inconsistently, and (close to) zero if these effects cancel out. Crescenzi operationalized a test of SBT by estimating the influence of this indicator on the time it takes until the next MID in that dyad breaks out. Indeed, he found that dyads receiving a negative score have shorter waiting times until the next conflict. This provides support for the combined predictions of structural balance theory. In contrast to Maoz et al. (2007) and Crescenzi (2007), we analyze the effect of indirect relations on the conditional event type, rather than on the occurrence of ties. Thus, our model estimates the sign of a tie *(a, b)* only if *a* does interact with *b*.[8]

7 A *semi-cycle* is a sequence of actors $v_1, v_2, \ldots, v_{k+1} = v_1$, $k \geq 3$ where for all $i = 1, \ldots, k$ there is a tie from v_i *to* v_{i+1} or vice versa.

8 Neither Maoz et al. (2007) nor Crescenzi (2007) use/s daily event data, but rather data coding the yearly state of the world system on the country-level. Some researchers argue that yearly data is too coarse-grained to capture quick responses to hostility as they occurred, e. g., in the Israel/Palestine conflict (King and Lowe 2003, p. 617).

The analysis of event data (alternatively referred to as *time-to-event analysis, survival analysis, event history analysis, or lifetime analysis)* is an established research area; see Lawless (2003) for a general reference. Some recent papers analyze network dependencies among events that happen in dyads (e. g. Butts 2008; De Nooy 2008, 2011; Brandes et al. 2009; Stadtfeld 2010). Although event data analysis is common in political science (e. g., Box-Steffensmeier and Jones 1997), network dependencies are rarely considered there. Exceptions include Goldstein et al. (2001) who applied vectorautoregression to the dyadwise aggregated levels of cooperation/ conflict over short time-intervals and Hoff and Ward (2004) who estimate dependencies in networks constructed from event data by aggregating over the whole observation period. Our work differs from these references since we do not aggregate events over time-intervals but rather model the probability of each single event.

In this paper we propose a general model for networks of dyadic typed events. With the increase in importance and availability of network event data we hope that this model will be applied to a variety of data sets. As an illustration, we apply it here to the publically-available data referred to as *Gulf data coded from full stories* from the Kansas Event Data System website (KEDS 2012).[9] This data set consists of events related to the Persian Gulf region for the period from April 15th, 1979 to March 31st, 1999. It includes more than 304,000 events among 202 unique actors. For the analysis done in this paper we exclude all non-state actors (such as ethnic groups or international organizations) yielding 168 actors and more than 217,000 events between them.

The KEDS (Schrodt et al. 1994) is a software tool that automatically extracts daily events from news reports. Events encode *who did when what to whom* and, thus, describe time-stamped, dyadic, typed interaction. Event types are classified using the World Event/Interaction Survey (WEIS) codes (McClelland 1976) and each event type is assigned a weight from the interval [–10, 10], where –10 stands for the most hostile and +10 for the most cooperative type of interaction (Goldstein 1992). These codings are explained in more detail in the following. Descriptive visualization, animation, and clustering of this data set can be found, e. g., in Brandes et al. (2006) and Brandes and Lerner (2008).

9 We have chosen the Gulf data set since mainly state actors are involved in this conflict; this implies that a consistent set of established alternative explanatory variables (introduced later) is available. Note that other data sets available from the KEDS website include many non-state actors.

Roughly spoken, the KEDS software (Schrodt et al. 1994) extracts triples of the form **`(subject, verb, object)`** from news wire reports. Each triple encodes the information that the **`subject`** performs an action (specified by the **`verb`**) towards the **`object`**. The subjects and objects are mapped to actors defined by the analyst. The following excerpt illustrates the coding of some of the actors included in the Gulf data:

```
AMERICA [USA]
CENTRAL INTELLIGENCE AGENCY [USA]
ISLAMIC_COUNTRIES [ARB]
ARAB_MONETARY_FUND [ARB]
GULF STATES [ARB]
```

Subjects and objects in news wire texts are interpreted as referring to specific actors. For instance, the general term **`AMERICA`** as well as the more specific **`CENTRAL INTELLIGENCE AGENCY`** are mapped to the same actor labeled **`USA`**. As another example, the tokens **`ISLAMIC COUNTRIES, ARAB MONETARY FUND,`** and **`GULF STATES`** (among others) are mapped to an actor labeled ARB. Actor ARB is, thus, an example of a non-state actor which is excluded from the network analyzed in this paper.

The following excerpt is an (incomplete) list of events that happened on August 10th, 1990 in the Gulf region.

```
900810      ARB    IRQ    012    RETREAT
900810      IRQ    USA    122    DENIGRATE
900810      IRQ    ARB    094    CALL FOR
900810      USA    IRQ    160    WARN
900810      USA    IRQ    051    PROMISE POLICY
900810      USA    IRQ    223    MIL ENGAGEME
```

For instance, the last event (dated **`900810`**; i. e., August 10th 1990) codes a military action (WEIS event type 223), initiated by actor **`USA`** and directed to the Iraq (**`IRQ`**). The text at the end of the line (**`MIL ENGAGEMENT`**) is a textual description of the event type (which is not needed in the analysis, since it is implied by the event type). In total there are more than 100 different types of events.

The WEIS event types are mapped to an established scale whose entries are referred to as *Goldstein weights* (Goldstein 1992) and indicate the level of cooperativeness (if positive) or hostility (if negative). Examples of weights associated with specific types are the following.

072	EXTEND MIL AID	8.3
054	ASSURE	2.8
160	WARN	-3.0
173	SPECIF THREAT	-7.0
223	MIL ENGAGEMENT	-10.0

Extending military aid is considered a highly cooperative action (weight equal to 8.3), whereas warnings are mildly hostile ($w = –3.0$), specific threats much more severe ($w = –7.0$), and military engagements are the most hostile type of events ($w = –10.0$).

It must be kept in mind that using the KEDS data entails some problems for the analysis. Since the data are generated on the basis of news reports we do not, strictly speaking, estimate the tendency to interact but rather the likelihood of interaction being reported in the news. We believe that, given that the interpretation takes account of this bias, the results still are meaningful.

Interaction among political actors is not only influenced by previous interaction but also by additional actor or dyad characteristics such as whether they share a common border or are members in the same military alliance. To control for different actor or dyad characteristics we rely on data from a frequently-used model of international militarized disputes (Oneal and Russett 2005). We will include several realist and liberal covariates, such as geographic adjacency, capability distribution, the countries' democracy scores and trade flows; however, we will pay special attention to the effect of military alliances. Research has not yet established whether military alliances reduce or increase the likelihood that a militarized dispute breaks out in a dyad (Bueno de Mesquita 1981; Bremer 1992; Oneal and Russett 2005; Kimball 2006). We will contribute to this debate by examining whether two countries that share an alliance membership generally interact more frequently, and if so, whether they behave more cooperatively or in a more hostile way towards each other. Including alliance membership as an explanatory variable is also quite illustrative from a methodological point of view. As it turns out, the positive influence of joint alliances on the conditional event type can be consistently validated–independent of which control variables we used. On the other hand, the positive relationship between alliances and the frequency of dyadic interaction that can be validated in a bivariate model diminishes, or even gets reversed, if we control for network dependencies and other covariates. This illustrates that the conditional event type is conceptually different from the absolute level of (friendly or hostile) interaction and emphasizes the need to control for network effects when testing associations among dyadic variables.

While the main contribution of this work is methodological, we present and test several hypotheses to illustrate and exemplify how our newly-proposed model can be applied in political science research and how it performs on empirical network event data. We have chosen the below-mentioned hypotheses since they illustrate different aspects of our model. Structural balance theory explicitly predicts that dyads are dependent. More precisely, interaction on a dyad *(a, b)* is claimed to depend on previous interaction on *(a, c)* and *(b, c),* for any third actor *c.* On the other hand, the hypotheses about the effect of alliances (H_5 and H_7) claim that interaction on *(a, b)* depends on a binary indicator *on the same dyad.* Although H_5 and H_7, thus, make no statement about dependencies among different dyads, we will see that controlling for network effects leads to different findings for some of these hypotheses. Thus, even if a particular research question is not about network dependencies these should nevertheless be tested and, if present, be included in the model.

Structural balance theory explains the *type* of events from *a* to *b* by the type of indirect relations via a third actor. More detailed, SBT predicts that actors behave:

H_1 cooperatively towards the friends of their friends;
H_2 hostile towards the friends of their enemies;
H_3 hostile towards the enemies of their friends;
H_4 cooperatively towards the enemies of their enemies.

Drawing on previous results on the effect of alliances, we hypothesize that events among allies are rather cooperative than hostile. Thus:

H_5 allies interact more cooperatively than non-allies. As hypotheses about event frequencies, we test the following two:

H_6 *Transitivity of activity:* the more actors *a* and *b* interacted (cooperatively or hostile) with common others, the higher the event rate on the dyad *(a, b).*

Finally, we hypothesize that alliances are only established among countries that, loosely speaking, have something to do with each other. Thus:

H_7 if actor *a* and *b* are allies then the event rate on the dyad *(a, b)* is higher than if *a* and *b* are not allies.

Note that the models we use later to test these hypotheses control for many more network dependencies which are, however, not of central interest for this paper and therefore not formulated as explicit hypotheses.

3 SPECIFICATION OF CONDITIONAL NETWORK EVENT TYPE MODELS

We assume that the occurrence of events and the type of events are dependent on previous events on the same or on other dyads. It is the goal of the analyst to test and/or control for such network dependencies and thereby to establish rules that govern the behavior of actors. As an example, if actors *a* and *b* both had frequent hostile interaction with common third actors (i. e., if they are enemies of enemies) then this may increase the probability that *a* and *b* interact cooperatively with each other. The model introduced in this section can be applied to perform statistical tests for such hypotheses. Note that a preliminary version of this model has been proposed in Brandes et al. (2009).

3.1 Model Overview

To model the probability of an observed sequence of events $E = (e_1,...,e_N)$, we assume that each event e_i is only dependent on events that happened earlier. To obtain a tractable model, we further assume that this dependence is completely captured by a dynamic network encoding the essential aspects of past interaction among actors. The past events (i. e., the events that happen before e_i) determine the *event network* G_{ei} and, given the state of G_{ei}, the next event e_i is assumed to be conditionally independent of all other events. The probability of e_i given G_{ei} is modeled parametrically so that the parameter estimates give the information which properties of G_{ei} increase/decrease the frequency of events and which properties of the network influence the conditional event type.

More formally, let $E = (e_1,...,e_N)$ be a sequence of events and let

$$\theta = (\theta^{(\lambda)}; \theta^{(\mu)}) = (\theta_1^{(\lambda)}, \dots, \theta_{k_\lambda}^{(\lambda)}; \theta_1^{(\mu)}, \dots, \theta_{k_\mu}^{(\mu)})$$

be the parameters of the model, where the rate *parameters* $\theta^{(\lambda)}$ stochastically determine the event frequency and the type *parameters* $\theta^{(\mu)}$ stochastically determine the event type, as we shall see later. The probability ensity function for an event sequence $E = (e_1,...,e_N)$ is

$$f(E;\theta) = f(e_1|G_{e_1};\theta) \cdot f(e_2|G_{e_2};\theta) \cdot \ldots \cdot f(e_N|G_{e_N};\theta) \ . \tag{1}$$

Here $f(e_i|G_{e_i};\theta)$ denotes the conditional probability density for the event e_i, given the network of past events G_{ei}.

For a given observed sequence of events $E = (e_1,...,e_N)$ the function $f(E; \theta)$ is the likelihood function when considered as a function of θ and hypothesis testing is operationalized by the maximum likelihood estimates, i. e., the values θ that maximize $f(E; q)$ (see Young and Smith 2005). The following sections provide details about the different components of our model.

3.2 Input Data

The input data we consider consists of sequences of dyadic, typed events $E = (e_1,...,e_N)$. A *(dyadic, typed)* event $e \in E$ is defined to be a tuple $e = (a_e, b_e, w_e, t_e)$, where:

- a_e is the *source* (initiator) of e;
- b_e is the *target* (addressee) of e;
- $w_e \in R$ is the *type,* coding the quality of the event e; and
- t_e is the time when e happens.

The source and the target of events are termed actors. Actors are, e. g., people, groups of people, organizations, or countries.

Time is given on some scale, e. g., by second, minute, hour, day, month, or year. In the KEDS data time is given by the day. Several events may happen during the same time unit. The event sequence is assumed to be in non-decreasing order with respect to time. The order of events that happen within the same time unit is considered as undefined. We note that for our analysis we do not need the absolute time t but rather the time difference Δt between events.

The type w_e of an event e (also referred to as its *weight*) characterizes the quality of e. In the exemplary application of this paper the weight w_e of an event e is a real number from the interval [–1, 1] obtained by dividing the Goldstein weights of KEDS events by ten. A positive weight indicates a cooperative event, a negative weight a hostile event, and the absolute value of event weights measures the magnitude of cooperativeness or hostility, respectively (so that this scale has a non-arbitrary zero indicating neutral events). In other applications, events may have other types, e. g., binary, multinomial, ordered multinomial, or event types might be multidimensional. While our model could be extended to these more general types of events this is not considered in this paper.

3.3 Explanatory Variable: The Network of Past Interaction

Given a sequence of events $E=(e_1,...,e_N)$ and a specific timepoint t (denoting the current time), the event network at time t (referred to as *network of past events* if t is implied) is a weighted graph $G_t=(A; W_t)$ defined as a function of the set of past events $E_{<t} = \{e \in E\ ;\ t_e < t\}$; i. e., the set of events that happen before t. Furthermore, the event network might encode (potentially time-dependent) actor, dyad, or network covariates that are not a function of previous events but that give additional information. For instance, in the case of political networks, such covariates might be the gross domestic product of a country (as an example of an actor covariate) or the geographical distance between countries (as an example of a dyad covariate). In our application these covariates are given as yearly data.

The components of $G_t=(A; W_t)$ are explained in the following. The set A consists of the actors that are involved in any event (thus we keep the set of actors fixed over time) and W_t is a vector-valued function mapping each dyad (a, b) to a value that characterizes the essential aspects of how a interacted with b in the past, i. e., before t. More formally, let $D = \{(i, j); i, j \in A, i \neq j\}$ be the set of all *dyads*. Then W_t is a function

$$W_t\colon D \to \mathbb{R}^d;\ (a,b) \mapsto (W_{t,1}(a,b),\ldots,W_{t,d}(a,b))\ ,$$

where $W_{t,i}(a, b) \in R$ denotes the real value in the i'th dimension, for $i = 1,...,d$.

In our concrete application, the network of past events $G_t=(A; w_t^+, w_t^-)$ is a weighted network with a two-dimensional weight function $W_t=(w_t^+, w_t^-)$, encoding past cooperative and past hostile interaction, respectively.[10] The value of cooperative/hostile interaction of a particular dyad (a, b) increases whenever a initiates a cooperative/hostile event e targeted at b. When the difference between the current time t and the event time t_e increases, the influence of e diminishes. The latter property is motivated by the assumption that actors forget (or forgive) cooperative and hostile actions. Assuming that the rate of forgetfulness or forgiveness is only dependent on the current weight, we obtain an exponentially decreasing influence of each event when time increases. More precisely, let $T_{1/2} \in R_{>0}$ be a given positive number

10 By a slight abuse of notation we denote the weight on a dyad by the same letter, namely w, as the weight of an event; this should not cause any confusion.

denoting the *halflife* of the influence of events. Then, the function w_t^+: D → $R_{\geq 0}$ is defined by

$$w_t^+(i,j) = \sum_{\substack{e:a_e=i,\ b_e=j,\\ w_e>0,\ t_e<t}} |w_e| \cdot \exp\left(-(t-t_e)\cdot\frac{\ln(2)}{T_{1/2}}\right)\cdot\frac{\ln(2)}{T_{1/2}}$$

and the function w_t^-: $D \to R_{\geq 0}$ is defined by[11]

$$w_t^-(i,j) = \sum_{\substack{e:a_e=i,\ b_e=j,\\ w_e<0,\ t_e<t}} |w_e| \cdot \exp\left(-(t-t_e)\cdot\frac{\ln(2)}{T_{1/2}}\right)\cdot\frac{\ln(2)}{T_{1/2}}$$

Thus, the value $w_t^+(i, j)$ is defined as a function of all weights of events $e = (a_e, b_e, w_e, t_e)$ that involve *i* as source $(a_e = i)$ and *j* as target $(b_e = j)$, that happen before the current time *t* $(t_e < t)$, and that have positive weight $(w_e > 0)$. Similarly, $w_t^-(i, j)$ is the sum over events with negative weight $(w_e < 0)$. How strongly an event *e* is counted at time *t* depends on the time-difference $t–t_e$. Each time this difference increases by $T_{1/2}$ the factor for w_e is halved. The choice of $T_{1/2}$ is dependent on whether the analyst is interested in short-term or long-term responses to previous events. Estimation of $T_{1/2}$ from empirical data is possible, but not considered in this paper. The last factor $\frac{\ln(2)}{T_{1/2}}$ is used to give $w_t^{\pm}(i, j)$ the interpretation of "aggregate weight per time unit." This interpretation is quite accurate if time units are small and only an approximation if time is given on a coarse scale.

Higher dimensional weight functions might additionally encode, e. g., the *conflict increase* defined as the difference of the current value of w_t^- minus its value Δ*t* time units in the past, i. e., $w_t^- - w_{t-Dt}^-$. Whether the level of conflict is currently increasing or decreasing might have a significant effect on future events–additional to the absolute level of conflict. The increase of conflict is, however, not used as an explanatory variable in this paper.

If *e* is an event in *E,* we sometimes write G_e for G_{te}. Note that G_e is only dependent on events that happen earlier than *e* (and not on events that happen in the same time unit as *e*).

3.4 Dependent Variable: The Next Event

In this section we model the probability density $f(e|G_e;\theta)$; i. e., the probability density of an event *e* dependent on the event network G_e, compare Eq. (1).

11 Note that w^- also maps to the non-negative numbers, since the absolute value $|w_e|$ is taken.

The first step is the decomposition of the probability density of events into a rate component and a conditional type component. Let $e=(a_e,b_e,w_e,t_e)$ be an event in the observed sequence E. The probability density of e, given the network of past events G_e, is decomposed into two factors:

$$f(e|G_e;\theta) = f_\lambda(a_e,b_e,t_e|G_e;\theta^{(\lambda)}) \cdot f_\mu(w_e|a_e,b_e,t_e;G_e;\theta^{(\mu)}) \ . \qquad (2)$$

Here $f_\lambda(a_e,b_e,t_e|G_e;\theta^{(\lambda)})$, called the *rate component,* is the probability density that the next event happens at time t_e and involves a_e as source and b_e as target. Likewise, $f_\mu(w_e|a_e,b_e,t_e;G_e;\theta^{(\mu)})$, called the *conditional type component,* is the conditional probability density that event e has type w_e, given that the next event involves a_e as source, b_e as target, and happens at time t_e. Defining

$$f_\lambda(E;\theta^{(\lambda)}) = \prod_{e\in E} f_\lambda(a_e,b_e,t_e|G_e;\theta^{(\lambda)})$$

and

$$f_\mu(E;\theta^{(\mu)}) = \prod_{e\in E} f_\mu(w_e|a_e,b_e,t_e;G_e;\theta^{(\mu)})$$

we can decompose the joint probability density into

$$f(E;\theta) = f_\lambda(E;\theta^{(\lambda)}) \cdot f_\mu(E;\theta^{(\mu)}) \ , \qquad (3)$$

where f_λ is the rate component of the joint probability density, modeling the occurrence of events, and f_μ is the conditional type component, modeling the distribution of event types.

Despite its simplicity, Eq. (2) is a key concept in the definition of conditional event type models. The main decision here is that a set of parameters $\theta^{(\lambda)}$ stochastically determines the occurrence of events (of any type) and a disjoint set of parameters $\theta^{(\mu)}$ stochastically determines the conditional type of an event, given that an event happens on a particular dyad. The decomposition in Eq. (2) implies that the rate parameters $\theta^{(\lambda)}$ and the type parameters $\theta^{(\mu)}$ can be estimated separately and, more importantly, the maximum likelihood estimates of $\theta^{(\mu)}$ are independent of the specification of the event frequency f_λ and, vice versa, the maximum likelihood estimates of $\theta^{(\lambda)}$ are independent of the specification of the distribution of the conditional event type f_μ.

The argument for this modeling decision has been given in the introduction: results about the determinants of the conditional event type give additional insights into the causes of typed events. Indeed, it will be shown in this paper that for our example, the effect of joint alliances on the frequency of dyadic events is very sensitive to the inclusion or exclusion of specific control variables. On the other hand, sharing an alliance membership consistently has a positive effect on the conditional event type – meaning that this association is robust to the inclusion or exclusion of control variables and is even robust whether or not we consider any network dependencies at all.

Next, we clarify the functional form of the density for the conditional event type (event quality) and the event rate (event frequency). We assume that the type w_e of an event *e* from actor *a* to actor *b* is dependent on the current state of the network G_e and the type parameters $\theta^{(\mu)}$.

For given parameters $\theta^{(\mu)} = (\theta_1^{(\mu)}, \ldots, \theta_{k\mu}^{(\mu)}, \sigma)$, the conditional distribution of the weight *w* of event $e = (a, b, w, t)$ is modeled as a normal distribution, leading to a likelihood which is the same as that of a linear regression model:

$$f_\mu(w|a,b;G_e;\theta^{(\mu)}) = \varphi_{\mu_{ab},\sigma^2}(w) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{[w-\mu_{ab}]^2}{2\sigma^2}\right) \quad . \qquad (4)$$

Here, the *expected event weight* $\mu ab = \mu_{ab}(G_e;\ \theta^{(\mu)})$ is postulated to be dependent on the parameters $(\theta 1^{(\mu)}, \ldots, \theta_{k\mu}^{(\mu)})$ and the values of various *statistics* $s_h(G_e;\ a, b)$, $h = 1, \ldots, k_\mu$ that characterize the network around *a* and *b* (see Sect. 3.5). More precisely, the expected event weight is modeled as a function

$$\mu_{ab}(G_e;\theta^{(\mu)}) = \sum_{h=1}^{k_\mu} \theta_h^{(\mu)} \cdot s_h(G_e;a,b) \quad . \qquad (5)$$

The maximum likelihood estimates of the weight parameters $\hat{\theta}_h^{(\mu)}$ reveal dependencies between characteristics of the network and later observed event weights. For instance, if a particular statistic $s_h(G_e;\ a, b)$ encodes how much *b* attacked *a* in the past, then a (significantly) negative value for the associated parameter $\hat{\theta}_h^{(\mu)}$ would imply that actors show a tendency to initiate hostile events towards attackers.

The modeling of the event frequency is slightly more complicated, since not the frequency itself but rather the waiting time between events is observed. However, estimating event frequencies from observed time-to-event data is a common task in lifetime analysis (Lawless 2003) from where we get the necessary methodology.

A key concept in modeling event times is the so-called hazard function (also called intensity function, or rate function). To simplify notation we assume here that event times are known exactly and that at any point in time there can happen at most one event; the likelihood function for the occurrence of events is later derived without this assumption (i. e., where time is only known up to some fixed precision and where more than one event may happen during the same time interval).

Let $(a, b) \in D$ be any dyad in the network, let t denote a point in time, and let $N_{ab}(t)$, defined by

$$N_{ab}(t) = |\{e \in E\,;\; a_e = a,\, b_e = b,\, t_e \leq t\}|,$$

denote the number of events that happen on the dyad (a, b) before or at time t. The function λ_{ab} mapping a time point t to

$$\lambda_{ab}(t) = \lim_{\Delta t \to 0} \frac{\mathbb{E}[N_{ab}(t + \Delta t) - N_{ab}(t)\,|\,G_t]}{\Delta t} \tag{6}$$

is called the *hazard function* for the dyad (a, b). (Here, the function E[•] denotes the expectation of the argument.) Intuitively, the hazard function can be interpreted as the expected number of events in a time interval of length one. Thus, $\lambda_{ab}(t)$ is also referred to as the *event rate* on the dyad (a, b) at time t. Note that, if at most one event can happen on (a, b), this definition is equivalent to the more usual definition of the hazard function as being the conditional probability density that *the* event happens at time t, given that *it* did not happen before (Lawless 2003). The definition given in Eq. 6 is preferable in our case, since it generalizes to repeated events. We assume that the hazard is a function of the current state of the network but that this function is time-invariant.[12] Similar to the event type, the event rate $\lambda_{ab}(G_t; \theta^{(\lambda)})$ is dependent on the rate parameters $\theta^{(\lambda)} = (\theta_1^{(\lambda)}, \ldots, \theta_{k\lambda}^{(\lambda)})$ and the values of various statistics $s_h(G_t; a, b)$, $h = 1, \ldots, k_\lambda$ that characterize the network around a and b. More precisely, the rate is specified to be a function

$$\lambda_{ab}(t) = \lambda_{ab}(G_t; \theta^{(\lambda)}) = \exp\left(\sum_{h=1}^{k_\lambda} \theta_h^{(\lambda)} \cdot s_h(G_t; a, b)\right) \,. \tag{7}$$

The exponential link function form ensures a positive event rate.

12 For more general hazard functions that have an explicit time dependency see Lawless (2003).

The maximum likelihood estimates for the rate parameters $\hat{\theta}_h^{(\lambda)}$ reveal dependencies between the network of past events and the frequency of future events. For instance, if a particular statistic $s_h(G_t; a, b)$ encodes how much b interacted with a in the past, then a significantly positive associated parameter $\hat{\theta}_h^{(\lambda)}$ would indicate that actors reciprocate activity. The rate parameters do not reveal whether responses are positive (more cooperation), neutral, or negative (more hostility).

The hazard function already determines the likelihood function for the occurrence of events which we derive next. We have to take into account that we have a fixed time precision (e. g., a day in our illustrative application) and that several events on the same or on different dyads may happen during the interval of timepoints that get the same timestamp. In the following, let the expression *time interval* always refer to an interval of length one consisting of the timepoints with the same timestamp, e. g., all timepoints within one day. Henceforward, we assume for sake of simplicity that the event network does not get updated during a time interval.

Let t denote a particular time interval and let, for any dyad $(i, j) \in D$,

$$n_{ij}(t) = |\{e \in E\,;\ a_e = i,\, b_e = j,\, t_e = t\}|,$$

denote the number of events that happen on (i, j) in the interval t. The use of time intervals implies that we go from a continuous-time model to a discrete-time model, where the number of events in interval t has a Poisson distribution with parameter $\lambda_{ij}(t)$; this is also the expected number of events in this time interval. For any dyad (i, j) the probability that exactly $n_{ij}(t)$ events happen on (i, j) in the interval t is

$$\frac{\lambda_{ij}(t)^{n_{ij}(t)} \exp(-\lambda_{ij}(t))}{n_{ij}(t)!} \quad ,$$

see, e. g., Lawless (2003). If t_1 denotes the first and t_N the last timestamp, the rate component of the joint probability density is[13]

$$f_\lambda(E; \theta^{(\lambda)}) = \prod_{t=t_1}^{t_N} \prod_{ij \in D} \frac{\lambda_{ij}(t)^{n_{ij}(t)} \exp(-\lambda_{ij}(t))}{n_{ij}(t)!} \quad . \tag{8}$$

13 Note that, if at most one event happened in any time interval, Eq. (8) is indeed identical to Eq. (2) in Butts (2008, p.163), although it is arranged in a different way.

Taking into account that in a given time interval there may be many dyads on which no event happens, and letting $D_{act}(t)$ denote the set of active dyads at time t, i. e., dyads on which at least one event happens, we can rewrite this as

$$f_\lambda(E;\theta^{(\lambda)}) = \prod_{t=t_1}^{t_N} \left(\prod_{ij \in D_{act}(t)} \frac{\lambda_{ij}(t)^{n_{ij}(t)}}{n_{ij}(t)!} \right) \cdot \exp\left(-\sum_{ij \in D} \lambda_{ij}(t) \right) \quad . \qquad (9)$$

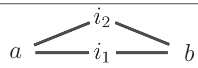
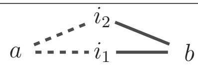
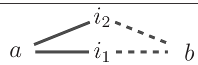
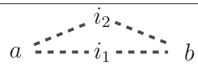
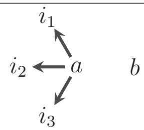
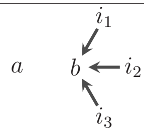
Note that the second product is over all dyads that are active at time t, while the sum is over all dyads, including inactive ones. If it is known that during a time interval t there cannot be an event on some dyads, then these have to be left out from the summation. Thereby one can, for instance, address situations where the set of actors changes over time: if actor a is not in the network at time t, then all dyads having a as source or as target have to be left out from the summation in the normalizing constant $e^{-\sum_{ij\in D}\lambda_{ij}(t)}$.

3.5 Network Statistics

The general model outlined so far can be applied to test many hypotheses concerning the interplay between network structure and the frequency and quality of dyadic events. The specialization is done by using various statistics in Eqs. (5) and (7). The particular statistics that we define below are similar to those of previous statistical models for cross-sectional (Robins et al. 2007) or longitudinal networks (Snijders 2005). The statistics are illustrated in *Table 1.*

Note that some of these statistics are used to test the hypotheses on structural balance theory, while others mostly serve to control for certain trivial regularities (e. g., inertia). A control statistic that always has to be taken to obtain meaningful results is the *constant* statistic, defined by **`constant`** $(G_t; a, b) = 1$. The constant statistic just controls for possible deviation from zero of the average event weight (similar argument for the rate). The function **`constant`** $(G_t; a, b)$, as well as the ones whose definition follows, correspond to the statistics $s_h\ (G_t; a, b)$ in Eqs. (5) as well as (7). First we propose some further statistics for use in Eq. (5) to specify the distribution of the conditional event type; subsequently statistics for use in Eq. (7) are proposed, specifying the event frequency.

Table 1: Illustration of some statistics explaining the tie from actor a towards actor b at time t. A $\pm$ sign indicates that there is a version for positive and for negative weights. The *symmetric* positve/negative weight on a dyad i, j is defined by $w^{\pm}_{t,\mathrm{sy}}(i,j) = w^{\pm}_t(i,j) + w^{\pm}_t(j,i)$. Dashed lines indicate negative ties.

name	formula	$a \longrightarrow b$ depends on
`inertia`$^{\pm}$	$w^{\pm}_t(a,b)$	$a \longrightarrow b$
`reciprocity`$^{\pm}$	$w^{\pm}_t(b,a)$	$a \longleftarrow b$
`friendOfFriend`	$\sqrt{\sum_{i\in A} w^{+}_{t,\mathrm{sy}}(a,i)\cdot w^{+}_{t,\mathrm{sy}}(i,b)}$	
`friendOfEnemy`	$\sqrt{\sum_{i\in A} w^{-}_{t,\mathrm{sy}}(a,i)\cdot w^{+}_{t,\mathrm{sy}}(i,b)}$	
`enemyOfFriend`	$\sqrt{\sum_{i\in A} w^{+}_{t,\mathrm{sy}}(a,i)\cdot w^{-}_{t,\mathrm{sy}}(i,b)}$	
`enemyOfEnemy`	$\sqrt{\sum_{i\in A} w^{-}_{t,\mathrm{sy}}(a,i)\cdot w^{-}_{t,\mathrm{sy}}(i,b)}$	
`activitySource`$^{\pm}$	$\sum_{i\in A} w^{\pm}_t(a,i)$	
`popularityTarget`$^{\pm}$	$\sum_{i\in A} w^{\pm}_t(i,b)$	

The most simple model would assume that actors just continue to act in the way they did in the past. For instance, if actor *a* initiated many hostilities targeted at actor *b,* the dyad *(a, b)* is likely to be a hostile one in the future. This effect is controlled for by the two statistics capturing the *inertia* of positive respectively negative events, defined by

$$\texttt{\textbf{inertia}}^{+}(G_t; a, b) = w_t^{+}(a, b)\ \texttt{\textbf{inertia}}^{-}(G_t; a, b) = w_t^{-}(a, b)\ .$$

The type parameter associated with **`inertia`**$^{+}$ is expected to be positive (events from *a* to *b* are more cooperative if *a* cooperated with *b* in the past) and the parameter associated with **`inertia`**$^{-}$ is expected to be negative (events from a to b are more hostile if a fought b in the past).

A non-trivial, but very reasonable, network effect is that actors reciprocate, i. e., actor a adapts its events towards actor b in accordance to how b treated a in the past. This is captured for positive and negative events by the two statistics

$$\texttt{reciprocity}^{+}(G_t; a, b) = w_t^{+}(b, a) \quad \texttt{reciprocity}^{-}(G_t; a, b) = w_t^{-}(b, a).$$

A positive estimate for the type parameter associated with **`reciprocity`**$^{+}$ would imply that actors reward cooperation; a negative estimate for the type parameter associated with **`reciprocity`**$^{-}$ would imply that actors retaliate when receiving hostilities.

Structural balance theory predicts that the relation of two actors a and b is dependent on whether they have common friends or foes. In the following we take it as an indication of friendship if two actors cooperate (in either direction) and as an indicator that they are enemies if they exchange hostilities. Let $w_{t,sy}^{+}(i, j) = w_t^{+}(i, j) + w_t^{+}(j, i)$ denote the *symmetrized positive weight* on adyad (i, j) and let $w_{t,sy}^{-}(i, j) = w_t^{-}(i, j) + w_t^{-}(j, i)$ denote the *symmetrized negative weight*. To test the four hypotheses of structural balance theory, we define four statistics measuring to what extent b is with respect to a a friend of a friend, a friend of an enemy, an enemy of a friend, and an enemy of an enemy, respectively.

$$\begin{aligned}
\texttt{friendOfFriend}(G_t; a, b) &= \sqrt{\sum_{i \in A} w_{t,\mathrm{sy}}^{+}(a, i) \cdot w_{t,\mathrm{sy}}^{+}(i, b)} \\
\texttt{friendOfEnemy}(G_t; a, b) &= \sqrt{\sum_{i \in A} w_{t,\mathrm{sy}}^{-}(a, i) \cdot w_{t,\mathrm{sy}}^{+}(i, b)} \\
\texttt{enemyOfFriend}(G_t; a, b) &= \sqrt{\sum_{i \in A} w_{t,\mathrm{sy}}^{+}(a, i) \cdot w_{t,\mathrm{sy}}^{-}(i, b)} \\
\texttt{enemyOfEnemy}(G_t; a, b) &= \sqrt{\sum_{i \in A} w_{t,\mathrm{sy}}^{-}(a, i) \cdot w_{t,\mathrm{sy}}^{-}(i, b)}
\end{aligned}$$

The square root expresses the assumption that a second (third, fourth, etc.) actor who is an enemy of a and a friend of b has a decreasing marginal effect on how strongly a perceives b as a friend of an enemy (compare Snijders et al. 2010). According to the hypotheses developed before, the type parameters associated with **`friendOfFriend`** and **`enemyOfEnemy`** are predicted to be positive and those associated with **`friendOfEnemy`** and **`enemyOfFriend`** are predicted to be negative.[14]

As a matter of fact, some actors are more active than others, some do rather initiate hostile events (aggressive actors), and others are more cooperative. Likewise, some actors are typical targets of hostilities, while others tend to experience cooperation. To control for such differences in actors' network *positions* or *roles,* we introduce a set of statistics dependent on the degree of actors. These statistics vary in three dimensions: (1) outdegree (activity) vs. indegree (popularity); (2) positive vs. negative weight; and, (3) whether we want to analyze the influence of these degree statistics on the initiator of events (source) or on the addressee of events (target). Together we obtain eight different statistics; two of them are defined below, the others are implied by analogy. The activity of the source actor with respect to positive events is defined to be

$$\texttt{activitySource}^{+}(G_t; a, b) = \sum_{i \in A} w_t^{+}(a, i) \ ,$$

while the popularity of the target actor with respect to negative events is defined to be

$$\texttt{activitySource}^{+}(G_t; a, b) = \sum_{i \in A} w_t^{+}(a, i) \ ,$$

For instance, a positive estimate for the type parameter associated with **`activitySource`**$^{+}$ would imply that actors *a* who initiate a lot of cooperation (towards any actor *i*), are more likely to cooperate with (rather than fight) the particular actor *b* who is the target of the next event.

Constant or slowly changing actor and dyad characteristics may additionally or alternatively explain the behavior of political actors. The following statistics characterize a dyad *(a, b)* by various covariates. The statistics are taken from a standard model for non-directed dyad analysis used in the study of international relations (Oneal and Russett 2005). The binary variable **`allies`**$(G_t; a, b)$ is one if and only if *a* and *b* have at least one common joint alliance membership. Hypothesis H_5 predicts that the conditional type parameter associated with

14 Clearly, by symmetrizing the friend and enemy relations we lose some information, since the direction of ties might cause different behavior. For instance, it might be possible that actors fight those who attack their friends but are indiferent to those who are attacked by their friends. If we distinguish all combinations of signs and directions of the two ties that indirectly relate a with b, we obtain 16 statistics for structural balance. This refinement is, however, not considered in this paper.

allies is positive. Statistic **lnCapRatio**$(G_t; a, b)$ is the logarithmized ratio of the capability score of the more powerful actor divided by the score of the less powerful actor.[15] Previous results have shown that a preponderance of national capabilities is related to less conflict within the dyad (Hegre 2008). The binary statistic **minorPowers**$(G_t; a, b)$ is one if and only if neither *a* nor *b* is a major power. The polity score of the less democratic actor in the dyad gives the value of the statistic **polityWeakLink**$(G_t; a, b)$. Since previous results have reported that democracies show a tendency to not fight each other (Russett and Oneal 2001), we expect a positive type parameter associated with **polityWeakLink.** The binary statistic **contiguity**$(G_t; a, b)$ is one if and only if *a* and *b* share a land border or a sea border less than 400 miles long and the variable **lnDistance**$(G_t; a, b)$ is the logarithmized distance between the capitals of *a* and *b*. The statistic **lnTrade**$(G_t; a, b)$ is the logarithm of the average of trade going from *a* to *b* and from *b* to *a* and **lnJointIGO**$(G_t; a, b)$ is the logarithmized number of joint memberships in intergovernmental organizations of *a* and *b*.

In principle, the statistics defined above can also be taken for the specification of the event rate in Eq. (7). However, structural balance theory makes no predictions about whether actors interact more or less frequently with, say, the friends of their enemies; it is just predicted that the tie is likely to be a hostile one. For this reason we argue that the rate is better specified by statistics that ignore the sign of previous interaction as defined in the following, leaving out the argument $(G_t; a, b)$.

$$\begin{aligned}
\mathbf{inertia} &= \mathbf{inertia}^{+} + \mathbf{inertia}^{-} \\
\mathbf{reciprocity} &= \mathbf{reciprocity}^{+} + \mathbf{reciprocity}^{-} \\
\mathbf{triangle} &= \mathbf{friendOfFriend} + \mathbf{friendOfEnemy} \\
&\quad + \mathbf{enemyOfFriend} + \mathbf{enemyOfEnemy} \\
\mathbf{activitySource} &= \mathbf{activitySource}^{+} + \mathbf{activitySource}^{-}
\end{aligned}$$

The definition of the statistics **activityTarget, popularitySource,** and **popularityTarget** is analogous to **activitySource.** The rate parameters associated with these statistics reveal dependencies between particular aspects of the network of past events and future event frequencies. For instance, a positive rate parameter associated with **reciprocity** would imply that if *b* interacted a lot with *a* (positively or negatively) then the current event frequency on the dyad *(a, b)* is typically higher than without this precondition; a negative rate parameter points to a decreased event rate.

15 The capability score of a country is a composite measure taking into account, among other things, demographic, economic, and military strength.

3.6 Parameter Estimation

The maximization of the log-likelihood of the type parameters, see Eq. (4), is very efficient and can be done by simply solving the usual OLS system of equations. The maximization of the log-likelihood of the rate parameters (see Eq. (8)) can be done numerically, e. g., by using the established Newton-Raphson algorithm. The computation of the rate parameters is much more time-consuming since the normalization constant in Eq. (9) contains a sum over all pairs of actors. The computation can be speeded up by approximating this sum via sampling over a sffiucient number of pairs; compare Butts (2008). The results in this paper, however, are computed without such an approximation.

4 RESULTS AND DISCUSSION

In this section we report and discuss the estimated parameters on the Gulf network, restricted to state actors. The half life parameter $T_{1/2}$ is set to 30 days.

4.1 Conditional Type Parameters

Table 2 shows the estimated conditional type parameters for three models, the first built from the 16 signed network effects, the second built from the covariate statistics, and finally a combined model which includes network and covariate effects. The log-likelihood of the null model M_0 (no effects except the constant offset) is –143,964, its BIC is 287,940, its AIC is 287,930. The maximized likelihood and information criteria for the other models are listed in Table 2. With respect to the information criteria the network model is better than the covariate model, but the combination of networks and covariates yields a strong further model improvement.[16]

16 Note that the models including covariates are estimated using a slightly smaller set of events due to dropping events with missing values; thus the network-only model can only be compared with the null model and the covariate model only with the joint model, according to the information criteria.

Table 2 *Estimated conditional type prameters and standard errors (in brackets)*

statistic	event network model	covariate model	combined model
`inertia`$^+$	**0.214 (0.012)***	·	**0.192 (0.012)***
`inertia`$^-$	*-0.085 (0.003)**	·	*-0.071 (0.003)**
`reciprocity`$^+$	**0.124 (0.014)***	·	**0.075 (0.014)***
`reciprocity`$^-$	*-0.082 (0.004)**	·	*-0.052 (0.004)**
`friendOfFriend`	**0.246 (0.027)***	·	**0.138 (0.027)***
`enemyOfFriend`	*-0.206 (0.014)**	·	*-0.119 (0.015)**
`friendOfEnemy`	*-0.224 (0.015)**	·	*-0.137 (0.015)**
`enemyOfEnemy`	**0.113 (0.008)***	·	**0.057 (0.008)***
`activitySource`$^+$	**0.051 (0.003)***	·	**0.009 (0.004)***
`activitySource`$^-$	*-0.008 (0.001)**	·	0.001 (0.002)
`activityTarget`$^+$	**0.040 (0.004)***	·	-0.006 (0.004)
`activityTarget`$^-$	0.002 (0.002)	·	**0.013 (0.002)***
`popularitySource`$^+$	-0.008 (0.005)	·	**0.023 (0.005)***
`popularitySource`$^-$	**0.003 (0.001)***	·	*-0.007 (0.002)**
`popularityTarget`$^+$	*-0.020 (0.005)**	·	0.005 (0.005)
`popularityTarget`$^-$	**0.004 (0.001)***	·	*-0.005 (0.002)**
`lnCapRatio`	·	**0.002 (0.001)***	*-0.007 (0.001)**
`allies`	·	**0.118 (0.003)***	**0.106 (0.003)***
`polityWeakLink`	·	$3.2E^{-4}(1.8E^{-4})$	*-0.001* $(1.9E^{-4})$*
`minorPowers`	·	**0.097 (0.003)***	**0.042 (0.004)***
`lnTrade`	·	**0.028 (0.001)***	**0.017 (0.001)***
`contiguity`	·	*-0.093 (0.003)**	*-0.060 (0.003)**
`lnDistance`	·	**0.011 (0.001)***	**0.013 (0.001)***
`lnJointIGO`	·	*-0.097 (0.003)**	*-0.076 (0.003)**
`constant`	*-0.082 (0.001)**	*-0.017 (0.008)**	-0.002 (0.008)
#events	217 479	200 886	200 886
log-likelihood	-136 669	-137 844	-134 441
ll-ratio to M_0	7 295	6 120	9 523
#params	17	9	25
BIC	273 547	275 798	269 187
AIC	273 372	275 706	268 932

The estimated parameters for the “network-only model” are reported in the first column in *Table 2*. The hypotheses derived from structural balance theory (H_1 to H_4) are fully supported by our analysis. The parameter associated with **`friendOfFriend`** is significantly positive, supporting H_1. This implies that actors have a tendency to cooperate with the friends of their friends: the more *b* is connected by past cooperative events to a third actor *c* who, in turn, is connected by past cooperative events to *a,* the higher (i. e., more cooperative) is the average weight of future events from *a* to *b.* The parameter associated with **`friendOfEnemy`** and with **`enemyOfFriend`** are significantly negative, supporting H_2 and H_3 respectively. This implies that actors have a tendency to fight the friends of their enemies as well as the enemies of their friends. Finally, the parameter associated with **`enemyOfEnemy`** is significantly positive as predicted by hypothesis H_4. This implies that actors in the Gulf conflict have a tendency to cooperate with the enemies of their enemies.

The second column of Table 2 shows the parameters estimated for a "covariates-only model." The main purpose of the politically relevant covariates is checking whether the network effects reported above are merely side-effects of certain actor or dyad characteristics. We do not wish to debate here the extent to which the conclusions obtained from this data analysis may be regarded as generalizable tests of IR theories: by restricting the data to a focused region, the Gulf, we have selected actors which have specific characteristics with respect to democracy, trade, geographic closeness, etc. The parameter associated with **`allies`** (binary variable encoding whether the two actors have at least one alliance or not) is significantly positive, supporting hypothesis H_5. This implies that events tend to be more cooperative among allied actors. As we will see in the following, this result is robust to the inclusion of network effects and also to the exclusion of all other covariate statistics. Other parameters, which are not related to our hypotheses, are not discussed here.

Finally, we estimate a model in which network effects and covariate effects are combined (third column in Table 2). It is remarkable that controlling for covariates does not change the signs of any parameter associated with **`inertia`** (positive and negative), **`reciprocity`** (positive and negative), and the four structural balance effects **`friendOfFriend, friendOfEnemy, enemyOfFriend,`** and **`enemyOfEnemy.`** In particular, the support for of structural balance theory (hypotheses H_1 to H_4) is robust and not just the side-effect of actor characteristics. There are some changes of parameters associated with covariates that are not related to our hypotheses. The effect of alliances on the conditional event type, however, remains significantly positive; i. e., allies interact more cooperatively if they interact. This gives further evidence that the validation of hypothesis H_5 is robust and not just the result of uncontrolled network dependencies.

To further test the robustness of the effect of alliances on the conditional event type we fit a bivariate model that contains only the alliance statistic as predictor (and a constant offset). In this model we obtain a significantly positive type parameter equal to 0.112 (0.002) for allies and a constant of –0.132 (0.001). In a different model built from the 16 network statistics (as above), the constant, and the alliance statistic (without any other covariate) we estimate a significantly positive type parameter 0.077 (0.002) for **`allies`.** Thus, the effect of alliances on the conditional event type is indeed consistently positive.

4.2 Event Frequency

Table 3 shows the estimated rate parameters for three models, the first built from the seven unsigned network effects, the second built from the covariate statistics, and finally the joint model including both network and covariate effects. The log-likelihood of the null model M_0 (no effects except the constant offset) is –1,629,098, its BIC is 3,258,208 and its AIC is 3,258,198. The maximized likelihood and information criteria for the other models are listed in Table 3. The information criteria show that here the covariate model is better than the network model, but the combination of networks and covariates yields a strong further model improvement.

Table 3 *Event rate parameters and standard errors*

statistic	event network model	covariate model	combined model
`inertia`	-0.114 (0.002)	·	-1.9E^{-4} (0.002)
`reciprocity`	-0.090 (0.003)	·	0.042 (0.003)
`triangle`	0.506 (0.002)	·	0.348 (0.003)
`activitySource`	0.202 (0.001)	·	0.161 (0.001)
`activityTarget`	0.168 (0.001)	·	0.118 (0.001)
`popularitySource`	0.094 (0.001)	·	0.073 (0.001)
`popularityTarget`	0.131 (0.001)	·	0.119 (0.001)
`lnCapRatio`	·	-0.289 (0.002)	-0.225 (0.002)
`allies`	·	0.064 (0.006)	-0.223 (0.006)
`polityWeakLink`	·	-0.137 (0.001)	-0.122 (0.001)
`minorPowers`	·	-2.726 (0.007)	-1.970 (0.007)
`lnTrade`	·	0.062 (0.001)	0.142 (0.001)
`contiguity`	·	1.362 (0.006)	1.310 (0.007)
`lnDistance`	·	-0.287 (0.002)	-0.343 (0.002)
`lnJointIGO`	·	1.344 (0.005)	1.313 (0.005)
`constant`	-6.774 (0.002)	-6.964 (0.017)	-7.530 (0.016)
#events	217,479	200,886	200,886
log-likelihood	-1 302 411	-1 271 416	-1 029 255
ll-ratio to M_0	326 687	357 682	599 843
#params	8	9	16
BIC	2 604 920	2 542 941	2 058 705
AIC	2 604 838	2 542 850	2 058 542

The rate parameters estimated for the "network-only model" are reported in the first column in Table 3. The parameter associated with **`triangle`** is significantly positive, supporting hypothesis H_6. This implies that activity is transitively closed: the more *b* is connected by past (cooperative or hostile) events to a third actor *c* who, in turn, is connected by past events to *a,* the higher is the frequency of future events from *a* to *b.*

The rate parameters resulting from the "covariate-only model" are reported in the second column in Table 3. The rate parameter associated with **`allies`** is significantly positive. Thus, allied actors interact more frequently than non-allied actors, giving support to hypothesis H_7. It is interesting to consider this result together with the conditional type parameter associated with the **`allies`** statistic (see Table 2): allies interact more, and if they interact, their interaction is rather friendly. This might nevertheless lead to the observation that the absolute (rather than conditional) probability of conflictive interaction is higher among allies (compare Bremer 1992).

Finally, we report the event frequency parameters in the combined model (third column in Table 3). The sign of the rate parameter associated with the **`triangle`** statistic did not change when covariates are included. In particular, the validation of hypothesis H_6 (transitivity of activity) is robust to the inclusion of covariates. There is only one change in the signs of the covariate parameters when the model is augmented by network effects: the parameter associated with **`allies`** changes from significantly positive to significantly negative. Thus, the finding "allies interact more", related to hypothesis H_7, is sensitive to whether we control for past interaction on the dyad (**`inertia`** and **`reciprocity`** statistics), activity and popularity of the actors, and indirect relations (**`triangle`** statistic). Controlling for these network dependencies, allies do not interact with higher frequency.

For comparison, we estimate a bivariate event frequency model that includes only the **`allies`** statistic and, as always, the **`constant`**. The rate parameter associated with the **`allies`** statistic in this model is 1.600 (0.005) and for **`constant`** we get a parameter of –6.700 (0.002). Thus, while allies consistently interact more cooperatively if they interact, the finding that allies interact more frequently is only a side-effect of ignored variables. Indeed, without any control variables we obtain the result that allies interact about $e^{1.6} = 4.95$ times as much as non-allies. Controlling for the effects of other covariates, this factor is reduced to (approximately) $e^{0.064} = 1.07$. Finally, controlling for covariates and network effects, a joint alliance membership even diminishes the rate of interaction by a factor of $e^{-0.223} = 0.8$.

It is noteworthy that some covariates influence the frequency of interaction while others rather have an effect on the conditional event type. For instance, in contrast to alliances, geographic adjacency (operationalized by the `contiguity` statistic) consistently increases the frequency of interaction, independent of whether we control for network dependencies or not. The effect of `contiguity` on the conditional event type is consistently negative so that–considering their frequent interaction–adjacent countries indeed are "dangerous dyads" (Bremer 1992). A different behavior can be seen for the `lnCapRatio` statistic (logarithmized capability ratio). While its effect on the event rate is consistently negative (unequal capability decreases interaction frequency), its effect on the conditional event type changes when we control for network dependencies: ignoring these leads to a seemingly positive effect (more cooperation when the capability ratio increases). In contrast, controlling for network effects yields a negative association between capability ratio and the conditional event type. The last-mentioned findings also emphasize the need to control for statistical dependency among dyadic observations: ignoring these might lead to spurious associations that vanish or even become reversed if network effects are taken into account.

5 CONCLUSION

We propose a general model for the dynamics of networks that are given as sequences of dyadic, typed events. Our model exploits the time information of event data and can test and control for potentially complex dependencies among dyadic observations. The most distinctive feature of our model is that we decompose the joint probability of typed events into two components, the first modeling the frequency of events of any type and the second modeling the conditional type of events, given that events happen. Thus, our model is an alternative to previously proposed models that estimate the rate of events of different types separately. This distinction indeed turned out to be crucial when tackling substantive research questions. For instance, previous work showed that enemies of enemies have a higher probability of engaging in conflict and in cooperation (Maoz et al. 2007), which simultaneously rejects and supports structural balance theory. In contrast, we showed that the conditional event type among enemies of enemies is pushed towards cooperation, which clearly supports SBT. Thus, our main methodological conclusion is the following: when we want to test a typical hypothesis in political network analysis, such as *does condition X lead to more or less conflict,* we have to clarify first whether we mean the absolute level of conflictive interaction or rather the tendency of conflict vs. cooperation, given that interaction occurs.

The separation of the joint probability density into its rate component and conditional type component provides deeper insight into how specific predictors increase or decrease the probabilities of events of a given type. For instance, alliance ties increase the probability of dyadic conflict in a bivariate model without any control variables. This result is refined by our model: alliances push the conditional event type towards cooperation–independent of which control variables we use. Thus, given that two actors do interact, their interaction is more friendly and less conflictual if they are allies. On the other hand, the finding that allies interact more frequently is only due to uncontrolled variables; if we control for network effects and other covariates, alliances even decrease the event frequency. Thus, alliance ties co-occur with frequent interaction but, apparently, they do not cause it.

We emphasize that the concrete empirical findings that have been included for illustration in this paper should be treated with caution for at least two reasons. First, the KEDS data includes dyadic events only if they have been reported in the news. Thus, issues such as news bias or media fatigue might influence the results in a systematic way. This is, strictly spoken, not a problem with our model but rather of the particular data. When more and more event data become available the same model can be applied to data sets that (hopefully) suffer less from news bias. Second, the data we analyzed included events that are restricted to the Gulf region–thereby selecting actors that have specific characteristics with respect to democracy scores, geographic proximity, trade relations, capability, etc. An issue for future work is to repeat the analyses on global datasets that are not restricted to specific regions, e. g., King and Lowe (2003); O'Brian (2010). Otherwise, it remains unclear which of the observed patterns are special to the Gulf conflict and which ones are universal in the dynamics of international relations.

The model itself leads also to several possibilities for extension and refinement. Obviously, the decomposition of the probability density of typed events into a rate component and a conditional type component, as in Eq. (2), is not restricted to one-dimensional real-valued event types. More general event types (such as binominal, multinomial, ordered multinomial, or multidimensional types) would just require the adaptation of the specification of the distribution for the conditional event type and the adaptation of the estimation procedure for the type parameters. Developing and applying models for more general event types is a promising area for future research.

REFERENCES

Box-Steffensmeier, Janet M.-Bradford S. Jones (1997), "Time is of the essence: Event history models in political science," *American Journal of Political Science* Vol. 41, No. 4, pp. 1414–1461.

Brandes, Ulrik -Daniel Fleischer -Jürgen Lerner (2006), "Summarizing dynamic bipolar conflict structures," *IEEE Transactions on Visualization and Computer Graphics* Vol. 12, No. 6, pp. 1486–1499.

Brandes, Ulrik -Jürgen Lerner (2008), "Visualization of conflict networks," in: Kauffmann, Mayeul, ed., *Building and Using Datasets on Armed Conflicts,* IOS Press, pp. 169–188.

Brandes, Ulrik -Jürgen Lerner -Tom A. B. Snijders (2009), "Networks evolving step by step: Statistical analysis of dyadic event data," in *Proc. 2009 Intl. Conf. Advances in Social Network Analysis and Mining,* IEEE Computer Society, pp. 200–205.

Bremer, Stuart A. (1992), "Dangerous dyads: Conditions affecting the likelihood of interstate war, 1816–1965," Journal of Conflict Resolution Vol. 36, No. 2, pp. 309–341.

Bueno de Mesquita, Bruce (1981), *The War Trap,* Yale University Press.

Butts, Carter T. (2008), "A relational event framework for social action," *Sociological Methodology* Vol. 38, No. 1, pp. 155–200.

Cartwright, Dorwin -Frank Harary (1956), "Structural balance: A generalization of Heider's theory," *The Psychological Review* Vol. 63, No. 5, pp. 277–293.

Crescenzi, Mark J. C. (2007), "Reputation and interstate conflict," *American Journal of Political Science* Vol. 51, No. 2, pp. 382–396.

De Nooy, Wouter (2008), "Signs over time: Statistical and visual analysis of a longitudinal signed network," *Journal of Social Structure* Vol. 9, No. 1.

De Nooy, Wouter (2011), "Networks of action and events over time. A multilevel discrete-time event history model for longitudinal network data," *Social Networks* Vol. 33, No. 1, pp. 31–40.

Goldstein, Joshua S. (1992), "A conflict-cooperation scale for WEIS international events data," *Journal of Conflict Resolution* Vol. 36, No. 2, pp. 369–385.

Goldstein, Joshua S. -Jon C. Pevehouse -Deborah J. Gerner -Shibley Telhami (2001), "Reciprocity, triangularity, and cooperation in the Middle East, 1979-97," *Journal of Conflict Resolution* Vol. 45, No. 5, pp. 594–620.

Hafner-Burton, Emilie M. -Alexander H. Montgomery (2006), "Power positions: International organizations, social networks, and conflict," *Journal of Conflict Resolution* Vol. 50, No. 1, pp. 3–27.

Hegre, Havard (2008), "Gravitating toward war: Preponderance may pacify, but power kills," *Journal of Conflict Resolution* Vol. 52, No. 4, pp. 566–589.

Heider, Fritz (1946), "Attitudes and cognitive organization," *The Journal of Psychology* Vo. 21, pp. 107–112.

Hoff, Peter D. -Michael D. Ward (2004), "Modeling dependencies in international relations networks," *Political Analysis* Vol. 12, No. 2, pp. 160–175.

KEDS (2012), *"http://eventdata.psu.edu/data.dir/gulf.html,"* Web-page, last accessed on September 20th, 2012.

Kimball, Anessa L. (2006), "Alliance formation and conflict initiation: The missing link," *Journal of Peace Research* Vol. 43, No. 4, pp. 371–389.

King, Gary Will Lowe (2003), "An automated information extraction tool for international conflict data with performance as good as human coders: A rare events evaluation design," *International Organizations* Vol. 57, pp. 617–642.

Lawless, Jerald F. (2003), *Statistical Models and Methods for Lifetime Data,* Wiley.

Maoz, Zeev (2009), "The effects of strategic and economic interdependence on international conflict across levels of analysis," *American Journal of Political Science* Vol. 53, No. 1, pp. 223–240.

Maoz, Zeev -Lesley G. Terris -Ranan D. Kuperman -Ilan Talmud (2007), "What is the enemy of my enemy? Causes and consequences of imbalanced international relations, 1816–2001," *Journal of Politics* Vol. 69, No. 1, pp. 100–115.

McClelland, Charles A. (1976), "World Event/Interaction Survey Codebook (ICPSR 5211)" .

O'Brian, Sean P. (2010), "Crisis early warning and decision support: Contemporary approaches and thoughts on future research," *International Studies Review* Vol. 12, pp. 87–104.

Oneal, John R. – Bruce Russett (2005), "Rule of three, let it be? When more really is better," *Conflict Management and Peace Science* Vol. 22, No. 4, pp. 293–310.

Robins, Garry -Tom A. B. Snijders -Peng Wang -Mark Handcock -Philippa Pattison (2007), "Recent developments in exponential random graph (p*) models for social networks," *Social Networks* Vol. 29, No. 2, pp. 192–215.

Russett, Bruce–John R. Oneal (2001), *Triangulating peace -democracy, interdependence, and international organizations,* W.W. Norton & Company.

Schrodt, Phillip A. -Shannon G. Davis -Judith L. Weddle (1994), "Political science: KEDS–a program for the machine coding of event data," *Social Science Computer Review* Vol. 12, No. 3, pp. 561–588.

Snijders, Tom A. B. (2005), "Models for longitudinal network data," in: Carrington, Peter J. -John Scott -Stanley Wasserman, eds., *Models and Methods in Social Network Analysis,* Cambridge University Press.

Snijders, Tom A. B. – Gerhard G. van de Bunt – Christian E. G. Steglich (2010), "Introduction to actor-based models for network dynamics," *Social Networks* Vo. 32, pp. 44–60.

Stadtfeld, Christoph (2010), "Who Communicates With Whom? Measuring Communication Choices on Social Media Sites," in *Proc. IEEE 2nd Intl. Conf. Social Computation (SocialCom),* pp. 564–569.

Young, G. A. and Smith, R. L. (2005), *Essentials of Statistical Inference,* Cambridge University Press.