Received: 24 May 2019 | Revised: 29 October 2019 | Accepted: 4 November 2019
DOI: 10.1111/stan.12198

SPECIAL ISSUE ARTICLE

WILEY

Tempus volat, hora fugit: A survey of tie-oriented dynamic network models in discrete and continuous time

Cornelius Fritz | **Michael Lebacher** | **Göran Kauermann**

Department of Statistics, Ludwig-Maximilians-Universität München, Munich, Germany

Correspondence
Cornelius Fritz, Department of Statistics, Ludwig-Maximilians-Universitát München, Ludwigstr. 33, 80539 Munich, Germany.
Email: cornelius.fritz@stat.uni-muenchen.de

Funding information
European Cooperation in Science and Technology, Grant/Award Number: COST Action CA15109 (COSTNET); German Research Foundation (DFG), Grant/Award Number: KA 1188/10-1; Munich Center for Machine Learning (MCML)

Given the growing number of available tools for modeling dynamic networks, the choice of a suitable model becomes central. The goal of this survey is to provide an overview of tie-oriented dynamic network models. The survey is focused on introducing binary network models with their corresponding assumptions, advantages, and shortfalls. The models are divided according to generating processes, operating in discrete and continuous time. First, we introduce the temporal exponential random graph model (TERGM) and the separable TERGM (STERGM), both being time-discrete models. These models are then contrasted with continuous process models, focusing on the relational event model (REM). We additionally show how the REM can handle time-clustered observations, that is, continuous-time data observed at discrete time points. Besides the discussion of theoretical properties and fitting procedures, we specifically focus on the application of the models on two networks that represent international arms transfers and email exchange, respectively. The data allow to demonstrate the applicability and interpretation of the network models.

KEYWORDS
continuous-time, discrete-time, ERGM, event modeling, random graphs, REM, STERGM, TERGM

1 | INTRODUCTION

The conceptualization of systems within a network framework has become popular within the last decades; see Kolaczyk (2009) for a broad overview. This is mostly because network models provide useful tools for describing complex dependence structures and are applicable to a wide variety of research fields. In the network approach, the mathematical structure of a graph is utilized to model network data. A graph is defined as a set of nodes and relational information (ties) between them. Within this concept, nodes can represent individuals, countries, or general entities, whereas ties are connections between those nodes. Dependent on the context, these connections can represent friendships in a school (Raabe, Boda, & Stadtfeld, 2019), transfers of goods between countries (Ward, Ahlquist, & Rozenas, 2013), sexual relations between people (Bearman, Moody, & Stovel, 2004), or hyperlinks between websites (Leskovec, Lang, Dasgupta, & Mahoney, 2009) to name just a few. Given a suitable data structure for the system of interest, the conceptualization as a network enables analyzing dependencies between ties. A central statistical model that allows this is the exponential random graph model (ERGM, Robins & Pattison, 2001). This model permits the inclusion of monadic, dyadic, and hyperdyadic features within a regression-like framework.

Although the model allows for an insightful investigation of *within-network* dependencies, most real-world systems are typically more complex. This is especially true if a temporal dimension is added, which is relevant, as most systems commonly described as networks evolve dynamically over time. It can even be argued that most static networks are *de facto* not static but snapshots of a dynamic process. A friendship network, for example, typically evolves over time and influences like reciprocity often follow a natural chronological order.

Of course, this is not the first paper concerned with reviewing temporal network models. Goldenberg, Zheng, Fienberg, and Airoldi (2010) wrote a general survey covering a wide range of models. The authors laid the foundation for further articles and postulated a soft division of statistical network models into latent space (Hoff, Raftery, & Handcock, 2002) and p_1 models (Holland & Leinhardt, 1981), all originating in the Erdös-Rényi-Gilbert random graph models (Erdös & Rényi, 1959; Gilbert, 1959). Kim, Lee, Xue, and Niu (2018) give a contemporary update on the field of dynamic models building on latent variables. Snijders (2005) discusses continuous-time models and reframes the independence and reciprocity model as a stochastic actor-oriented model (SAOM; Snijders, 1996). Block, Koskinen, Hollway, Steglich, and Stadtfeld (2018) provide an in-depth comparison of the temporal ERGM (TERGM, Hanneke, Fu, & Xing, 2010) and the SAOM with special focus on the treatment of time. Furthermore, the ERGM and SAOM for networks that are observed at single time points are contrasted by Block, Stadtfeld, and Snijders (2019), deriving theoretical guidelines for model selection based on the differing mechanics implied by each model.

In the context of this compendium of articles, the scope is to give an update on the dynamic variant of the second strand of models relating to p_1 models. We therefore extend the summarizing diagram of Goldenberg et al. (2010), as depicted in Figure 1. Generally, we divide temporal models into two sections, by differentiating between discrete and continuous-time network models. This review paper will focus on tie-oriented models. Tie-oriented models are concerned with formulating a stochastic model for the existence of a tie in contrast to the actor-oriented approach by Snijders (2002), which specifies the model from the actor's point of view (Block et al., 2018). The dynamical actor-oriented model (DyNAM, Stadtfeld & Block, 2017) adopts this actor-oriented paradigm to event data. This type of model was formulated with a focus on social networks

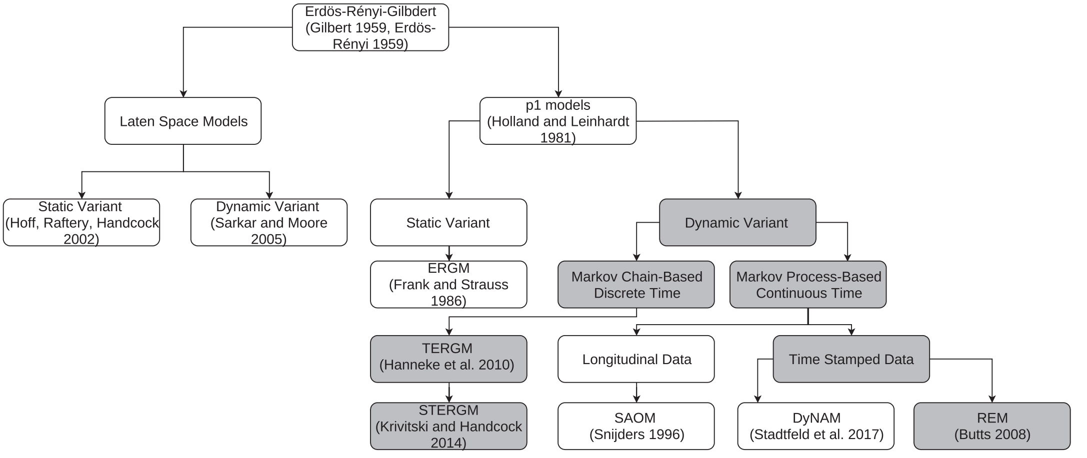

FIGURE 1 Tree diagram summarizing the dependencies between models originating in the Erdös-Rényi-Gilbert graph model; the models situated in a box with a grey background are discussed in this article. This graph is an update of Figure 6.1 in Goldenberg, Zheng, Fienberg, and Airoldi (2010). ERGM = exponential random graph model; TERGM = temporal exponential random graph model; STERGM = separable temporal exponential random graph model; SAOM = stochastic actor-oriented model; DyNAM = dynamical actor-oriented model; REM = relational event model

(Snijders, 1996). Tie-oriented models, on the other hand, can be viewed as more general because they are also applicable to nonsocial networks.

Statistical models for time discrete data rely on an autoregressive structure and condition the state of the network at time point t on previous states. This includes the TERGM and the separable TERGM (STERGM; Krivitsky & Handcock, 2014). There exists a variety of recent applications of the TERGM. White, Forester, and Craft (2018) use a TERGM for modeling epidemic disease outcomes and Blank, Dincecco, and Zhukov (2017) investigate interstate conflicts. In He, Dong, Wu, Jiang, and Zheng (2019), Chinese patent trade networks are inspected, and Benton and You (2017) use a TERGM for analyzing shareholder activism. Applications of STERGMs are given, for example, by Stansfield et al. (2019) that model sexual relationships and by Broekel and Bednarz (2018) that study the network of research and development cooperation between German firms.

In case of time-continuous data, the model regards the network as a continuously evolving system. Although this evolution is not necessarily observed in continuous time, the process is taken to be latent and explicitly models the evolution from the state of the network at time point $t - 1$ to t (Block et al., 2018). In this paper, we discuss the relational event model (REM, Butts, 2008) for the analysis of event data. Applications of the REM are manifold and range from explaining the dynamics of health behavior sentiments via Twitter (Salathé, Vu, Khandelwal, & Hunter, 2013), interhospital patient transfers (Vu, Lomi, Mascia, & Pallotti, 2017), online learning platforms (Vu, Pattison, & Robins, 2015), and animal behavior (Tranmer, Marcum, Morton, Croft, & de Kort, 2015) to structures of project teams (Quintane, Pattison, Robins, & Mol, 2013). Eventually, the REM is adapted to time-discrete observations of networks. That is, we observe the time-continuous developments of the network at discrete observation times only. Henceforth, we use the term time-clustered for this special data structure.

In reviewing dynamic network models, we assume a temporal first-order Markov dependency. To be more specific, this implies that the network at time point t only depends on the previous observation of the network. This characteristic is widely used in the analysis of longitudinal networks (Hanneke et al., 2010; Krivitsky & Handcock, 2014) and the resulting conditional independence among states of the network facilitates the estimation with an arbitrary number of time points. In that respect, it suffices to only include two observational moments for illustrative purposes because the interpretation and estimation with a longer series of networks is unchanged. This allows for a clear-cut comparison of the methods at hand.

This paper is structured as follows. In Section 2, we give basic definitions that are used throughout this paper and present the two data examples that will be analyzed as illustrative examples. After that, Section 3 introduces a time-discrete network model and Section 4, a time-continuous network model. They are applied in Section 5 on two data sets and Section 6 concludes. Additional results relating to the applications can be found in the Supplementary Material.

2 | DEFINITIONS AND DATA DESCRIPTION

2.1 | Definitions

This article regards directed binary networks, with ties representing directed relations between two nodes at a time point. The respective information can be represented in an adjacency matrix $Y_t = (Y_{ij,t})_{i,j=1,...,n} \in \mathcal{Y}$, where $\mathcal{Y} = \{Y : Y \in \{0,1\}^{n \times n}\}$ represents the set of all possible networks with n nodes. The entry (i,j) of Y_t is "1" if a tie is outgoing from node i to j in year t and "0," otherwise. Furthermore, the discrete time points of the observations of Y_t are denoted as $t = 1, \dots, T$. We restrict our analysis to two time points in both exemplary networks, which suffices for comparison. Hence, we set $T = 2$. In many networks, including our running examples, self-loops are meaningless. We therefore fix $Y_{ii,t} \equiv 0 \, \forall \, i \in \{1, \dots, n\}$ throughout the article. Furthermore, all subscripted temporal indices (Y_t) are assumed to take discrete values and all indices in brackets ($Y(t)$) continuous values. The temporal indicator t denotes the observation times of the network, and to notationally differ this from time-continuous model, we write $\tilde{t}$ for continuous time.

To sufficiently compare different models, we use two application cases. The first one represents the international trade of major weapons, which is given by discrete snapshots of networks that are yearly aggregated over time-continuous trade instances, that is, the time-stamped information is not observed. However, the second application, a network of email traffic, comes in time-stamped format, which can be aggregated to discrete-time observations.

2.2 | Data set 1: International arms trade

The data on international arms trading for the years 2016 and 2017 are provided by the Stockholm International Peace Research Institute (SIPRI, 2019). To be more specific, information on the exchange of major conventional weapons together with the volume of each transfer is included. In order to have a binary network representation, we discretize the data and set edges $Y_{ij,t}$ to "1" if country i sent arms to country j in t.

The left side of Table 1 gives some descriptive measures (Csardi & Nepusz, 2006) and Figure 2 visualizes the arms trade network using the software `Gephi` (Bastian, Heymann, & Jacomy, 2009). The density of a network is the proportion of realized edges out of all possible edges and is similar in both years, indicating the sparsity of the modeled network. Clustering can be expressed by the transitivity measure, providing the percentage of triangles out of all connected triplets. Reciprocity in a graph is the ratio of reciprocated ties and is similar in both years. As expressed by the high percentage of repeated ties, most countries seem to continue trading with the same partners.

TABLE 1 Descriptive statistics for the international arms trade network (left) and the European research institutions email correspondence (right)

		Arms trade network		Email network	
Time	t	2016	2017	Period 1	Period 2
Number of events		–	–	4, 957	2, 537
Number of nodes	n	180	180	88	88
Number of possible ties	$n(n-1)$	32,220	32,220	7,656	7,656
Density		0.021	0.020	0.123	0.087
Transitivity		0.195	0.202	0.407	0.345
Reciprocity		0.081	0.083	0.7	0.687
Repetition		–	0.641	–	0.574

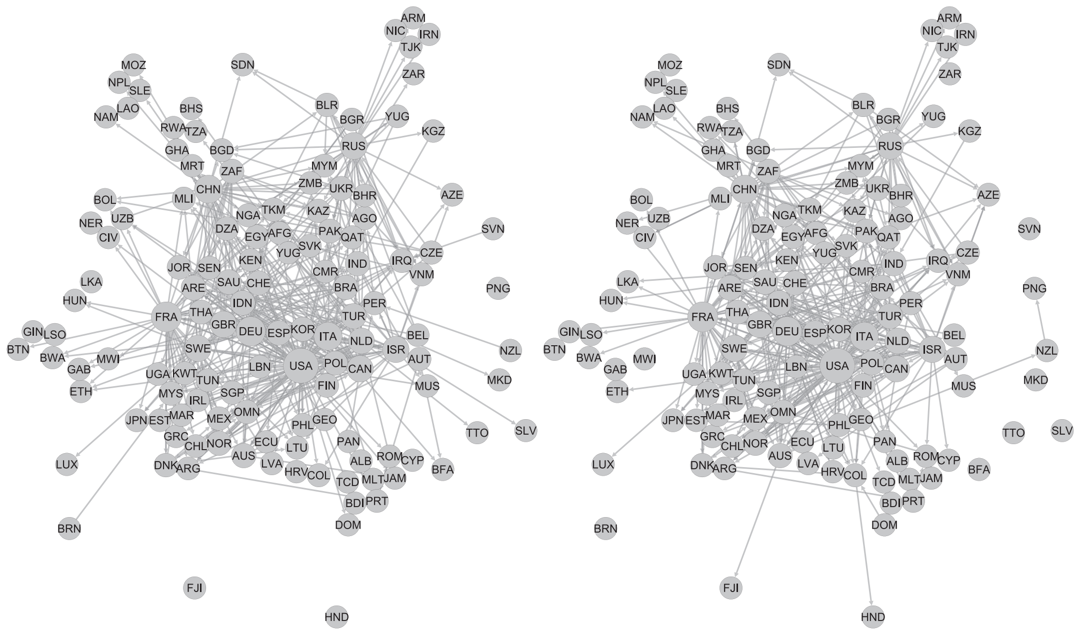

FIGURE 2 The international arms trade as a binary network in 2016 (left) and 2017 (right). Nodes that are isolated in both years are not depicted for clarity and the node size relates to the sum of the out- and in-degrees. The labels of the nodes are the ISO3 codes of the respective countries

Additionally, different kinds of exogenous covariates may be controlled for in statistical network models. In the given example, we use the logarithmic gross domestic product (GDP; World Bank, 2019) as monadic covariates concerning the sender and receiver of weapons. We also include the absolute difference of the so-called polity IV index (Center for Systemic Peace, 2017), ranging from zero (no ideological distance) to 20 (highest ideological distance), as a dyadic exemplary covariate. These covariates are assumed to be nonstochastic and we denote them by x_t. See the Supplementary Material for a list of all included countries and their ISO3 code.

2.3 | Data set 2: European research institution email correspondence

The second network under study represents anonymized email exchange data between institution members of a department in a European research institution (see Email EU Core, 2019; Paranjape,

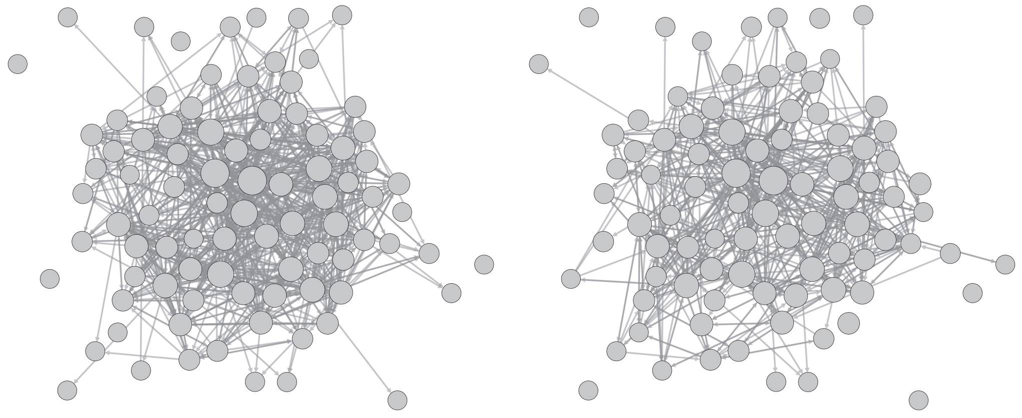

FIGURE 3 The European research institution email correspondence aggregated to a binary network divided into Period 1 (days 1–365, left) and Period 2 (days 1–365, right). The node size relates to the sum of the out- and in-degrees

Benson, & Leskovec, 2017). In this data set, we observe events $\omega = (i, j, \tilde{t})$ that represent emails sent from department member i to department member j at a specific time point $\tilde{t}$.

The data contain $n = 89$ persons and are recorded over 802 days. For this paper, we select the first two years and split them again into two years, labeled Period 1 and Period 2. Within the first period, $8,068$ events are recorded, and in the second period, $4,031$ events. We only regard one-to-one email correspondences; therefore, we exclude all group mails from the analysis. In the right column of Table 1, the descriptive measures for the two aggregated networks are given and in Figure 3, they are visualized. All descriptive statistics are higher in the email exchange network as compared to the arms trade network. In comparison to the arms trade network, the aggregated network is more dense with more than 10% of all possible ties being realized. In both years, the transitivity measure is relatively higher in both time periods. The high share of reciprocated ties is intuitive given that the network represents email exchange between institution members that may collaborate. No covariates are available for this network. See the Appendix for the visualization of the degree distributions of both applications.

3 | DYNAMIC EXPONENTIAL RANDOM GRAPH MODELS

3.1 | Temporal ERGM

The ERGM is among the most popular models for the analysis of static network data. Holland and Leinhardt (1981) introduced the model class, which was subsequently extended with respect to fitting algorithms and network statistics (see Lusher, Koskinen, & Robins, 2012; Robins, Pattison, Kalish, & Lusher, 2007). Spurred by the popularity of ERGMs, dynamic extensions of this model class emerged, pioneered by Robins and Pattison (2001) who developed time-discrete models for temporally evolving social networks. Before we start with a description of the model, we want to highlight that the TERGM and the STERGM are most appropriate for equidistant time points. That is, we observe the networks Y_t at discrete and equidistant time points $t = 1, \dots, T$. Only in this setting, the parameters allow for a meaningful interpretation. See Block et al. (2018) for a deeper discussion.

Hanneke et al. (2010) is the main reference for the TERGM, a model class that utilizes the Markov structure and, thereby, assumes that the transition of a network from time point $t-1$ to time point t can be explained by exogenous covariates as well as structural components of the present and preceding networks. Using a first-order Markov dependence structure and conditioning on the first network, the resulting dependence structure of the model can be factorized into

$$\mathbb{P}_\theta(Y_T, \dots, Y_2|Y_1, x_1, \dots, x_T) = \mathbb{P}_\theta(Y_T|Y_{T-1}, x_T)\dots \mathbb{P}_\theta(Y_3|Y_2, x_3)\mathbb{P}_\theta(Y_2|Y_1, x_2). \tag{1}$$

In the formulation above, the joint distribution is decomposed into yearly transitions from Y_{t-1} to Y_t. Furthermore, it is assumed that the same parameter vector θ governs all transitions. Often, this is an unrealistic assumption for networks observed at many time points because the generative process may change other time. Therefore, it can be useful to allow for different parameter vectors for each transition probability (i.e., $\theta_T, \theta_{T-1}, \dots$). In such a setting, the parameters for each transition can either be estimated sequentially (e.g., Thurner, Schmid, Cranmer, & Kauermann, 2019) or by using smooth time-varying effects (e.g., Lebacher, Thurner, & Kauermann, 2019).

Given the dependence structure (1), the TERGM assumes that the transition from Y_{t-1} to Y_t is generated according to an exponential random graph distribution with the parameter θ:

$$\mathbb{P}_\theta(Y_t = y_t|Y_{t-1} = y_{t-1}, x_t) = \frac{\exp\{\theta^T s(y_t, y_{t-1}, x_t)\}}{\kappa(\theta, y_{t-1}, x_t)}. \tag{2}$$

Generally, $s(y_t, y_{t-1}, x_t)$ specifies a p-dimensional function of sufficient network statistics, which may depend on the present and previous network, as well as on covariates. These network statistics can include static components, designed for cross-sectional dependence structures (see Morris, Handcock, & Hunter, 2008, for more examples). However, the statistics $s(y_t, y_{t-1}, x_t)$ explicitly allow temporal interactions, for example, delayed reciprocity

$$s_{\text{delrecip}}(y_t, y_{t-1}) \propto \sum_{i \neq j} y_{ji,t}\, y_{ij,t-1}. \tag{3}$$

This statistic governs the tendency whether a tie (i,j) in $t-1$ will be reciprocated in t. Another important temporal statistic is stability

$$s_{\text{stability}}(y_t, y_{t-1}) \propto \sum_{i \neq j} \left(y_{ij,t}\, y_{ij,t-1} + (1 - y_{ij,t})(1 - y_{ij,t-1})\right). \tag{4}$$

In this case, the first product in the sum measures whether existing ties in $t-1$ persist in t and the second term is one if nonexistent ties in $t-1$ remain nonexistent in t. The proportionality sign is used because in many cases the network statistics are scaled into a specific interval (e.g., $[0, n]$ or $[0, 1]$). Such a standardization is especially sensible for networks where the actor set changes with time. Additionally, exogenous covariates can be included, for example, time-varying covariates $x_{ij,t}$

$$s_{\text{dyadic}}(y_t, x_t) = \sum_{i \neq j} y_{ij,t}\, x_{ij,t}. \tag{5}$$

There exists an abundance of possibilities for defining interactions between ties in $t-1$ and t. From this discussion and Equation (2), it also becomes evident that, in a situation where the interest lies in the transition between two periods, a TERGM can be modeled simply as an ERGM, including lagged network statistics. This can be done for example by incorporating $y_{ij,t-1}$ as explanatory variable in (5), which is mathematically equivalent to the stability statistic (4). In the application we call this statistic repetition (Block et al., 2018).

Concerning the estimation of the model, maximum likelihood estimation appears to be a natural candidate due to the simple exponential family form (2). However, the normalization constant in the denominator of model (2) often poses an inhibiting obstacle when estimating (T)ERGMs. This can be seen by inspecting the normalization constant $\kappa(\theta, y_{t-1}, x_t) = \sum_{\tilde{y} \in \mathcal{Y}} \exp\{\theta^T s(\tilde{y}_t, y_{t-1}, x_t)\}$, which requires summation over all possible networks $\tilde{y} \in \mathcal{Y}$. This task is virtually infeasible, except for very small networks. Therefore, Markov Chain Monte Carlo (MCMC) methods have been proposed in order to approximate the logarithmic likelihood function (see Geyer & Thompson, 1992, for Monte Carlo maximum likelihood, and Hummel, Hunter, & Handcock, 2012, for its adaption to ERGMs). The article by Caimo and Friel (2011) provides an alternative algorithm that uses MCMC-based inference in a Bayesian model framework. Another approach is to employ maximum pseudolikelihood estimation (MPLE, Strauss & Ikeda, 1990) that can be viewed as a local alternative to the likelihood (van Duijn, Gile, & Handcock, 2009) but is often regarded as unreliable and poorly understood in the literature (Handcock, 2003; Hunter, Goodreau, & Handcock, 2008). However, the MPLE is claimed to be consistent and asymptotically efficient (Desmarais & Cranmer, 2012) and the biased standard errors can be corrected via bootstrap (Leifeld, Cranmer, & Desmarais, 2018). A notable special case arises if the network statistics are restricted such that they decompose to

$$s(y_t, y_{t-1}, x_t) = \sum_{i \neq j} y_{ij,t} \tilde{s}_{ij}(y_{t-1}, x_t), \tag{6}$$

with $\tilde{s}_{ij}$ being a function that is evaluated only at the lagged network y_{t-1} and covariates x_t for tie (i, j). With this restriction, we impose that the ties in t are independent, conditional on the network structures in $t-1$. This greatly simplifies the estimation procedure and allows to fit the model as a logistic regression model (see, e.g., Almquist & Butts, 2014) without the issues related to the MPLE.

A problem, that is very often encountered when fitting (T)ERGMs with endogenous network statistics is called degeneracy (Schweinberger, 2011) and occurs if most of the probability mass is attributed to network realizations that provide either full or empty networks. One way to circumvent these problems is the inclusion of modified statistics, called geometrically weighted statistics (Snijders, Pattison, Robins, & Handcock, 2006). Using the definitions of Hunter (2007), the geometrically weighted out-degree distribution (GWOD) controls for the out-degree distribution with one statistic, via

$$s_{\text{GWOD}}(y_t) = \exp\{\alpha_O\} \sum_{k=1}^{n-1} \left(1 - (1 - \exp\{-\alpha_O\})^k\right) O_k(y_t), \tag{7}$$

with $O_k(y_t)$ being the number of nodes with out-degree k in t and α_O being the weighting parameter. Correspondingly, the in-degree distribution is captured by the geometrically weighted in-degree distribution (GWID) statistic by exchanging $O_k(y_t)$ with $I_k(y_t)$, which counts the number of nodes with in-degree k, and α_O with α_I. While on the one hand, the weighting often effectively counteracts the problem of degeneracy, the statistics become more complicated to interpret. Negative values of the associated parameter typically indicate a centralized network structure.

Regarding statistics capturing clustering, the most common geometrically weighted triangular structure is called geometrically weighted edge-wise shared partners (GWESP) and builds on the number of two-paths that indirectly connect two nodes i and j given the presence of an edge (i, j):

$$s_{\text{GWESP}}(y_t) = \exp\{\alpha_S\} \sum_{k=1}^{n-2} \left(1 - (1 - \exp\{-\alpha_S\})^k\right) S_k(y_t), \tag{8}$$

where α_S is a weighting parameter. The number of edges with k shared partners ($S_k(y_t)$) is uniquely defined in undirected networks. If the edges are directed it must be decided which combination should form a triangle; see Lusher et al. (2012) for a discussion. As a default, the number of directed two-paths is chosen (Goodreau, Kitts, & Morris, 2009). Generally, a positive coefficient for GWESP indicates that triadic closure increases the probability of edge occurrence, and globally, a positive value for the associated parameter means more triadic closure as compared to a regime with a negative value (Morris et al., 2008).

3.2 | Separable TERGM

A useful improvement of the TERGM (2) is the STERGM proposed by Krivitsky and Handcock (2014). This model can be motivated by the fact that the stability term leads to an ambiguous interpretation of its corresponding parameter. Given that we include (4) in a TERGM and obtain a positive coefficient after fitting the model, it is not clear whether the network can be regarded as "stable" because existing ties are not dissolved (i.e., $y_{ij,t} = y_{ij,t-1} = 1$) or because no new ties are formed (i.e., $y_{ij,t} = y_{ij,t-1} = 0$). To disentangle this, the authors propose a model that allows for the separation of formation and dissolution.

Krivitsky and Handcock (2014) define the formation network as $Y^+ = Y_t \cup Y_{t-1}$, being the network that consists of the initial network Y_{t-1} together with all ties that are newly added in t. The dissolution network is given by $Y^- = Y_t \cap Y_{t-1}$ and contains exclusively ties that are present in t and $t-1$. Given the network in $t-1$ together with the formation and the dissolution network, we can then uniquely reconstruct the network in t because $Y_t = Y^+ \backslash (Y_{t-1} \backslash Y^-) = Y^- \cup (Y^+ \backslash Y_{t-1})$. Define $\theta = (\theta^+, \theta^-)$ as the joint parameter vector that contains the parameters of the formation and the dissolution model. Building on that, Krivitsky and Handcock (2014) define their model to be separable in the sense that the parameter space of θ is the product of the parameter spaces of θ^+ and θ^- together with conditional independence of formation and dissolution given the network in $t-1$:

$$\mathbb{P}_\theta(Y_t = y_t | Y_{t-1} = y_{t-1}, x_t) = \underbrace{\mathbb{P}_{\theta^+}(Y^+ = y^+ | Y_{t-1} = y_{t-1}, x_t)}_{\text{Formation Model}} \underbrace{\mathbb{P}_{\theta^-}(Y^- = y^- | Y_{t-1} = y_{t-1}, x_t)}_{\text{Dissolution Model}}. \tag{9}$$

The structure of the model is visualized in Figure 4. On the left-hand side, the state of the network Y_{t-1} is given, consisting of two ties (i, h) and (h, j). In the top network all ties that could possibly be formed are shown dashed and the actual formation in this example (i, j) is shown solid. On the bottom, the two ties that could possibly be dissolved are shown, and in this example, (h, j) persists while (i, j) is dissolved. On the right-hand side of Figure 4, the resulting network at time point t is displayed. Given this structure and the separability assumption (9), it is assumed that the formation model is given by

$$\mathbb{P}_{\theta^+}(Y^+ = y^+ | Y_{t-1} = y_{t-1}, x_t) = \frac{\exp\{(\theta^+)^T s(y^+, y_{t-1}, x_t)\}}{\kappa(\theta^+, y_{t-1}, x_t)}, \tag{10}$$

with $\kappa(\theta^+, y_{t-1}, x_t)$ being the normalization constant. Accordingly, the dissolution model can be defined. It becomes apparent how the STERGM is a subclass of the TERGM by inserting the

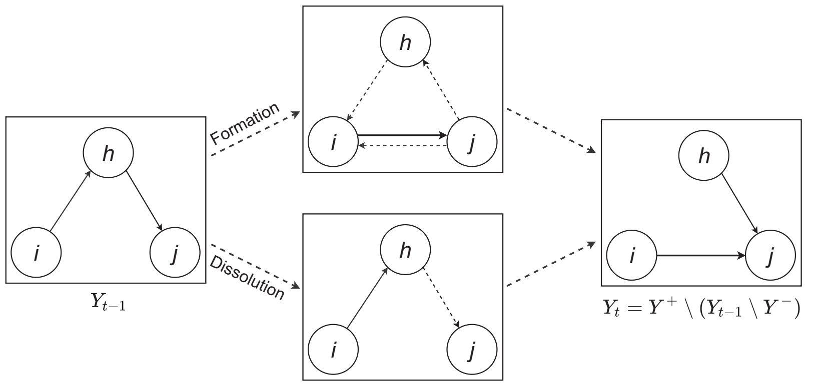

FIGURE 4 Conceptual representation, illustrating formation, and dissolution in the separable temporal exponential random graph model

separable models in (9) because

$$\mathbb{P}_\theta(Y_t = y_t | Y_{t-1} = y_{t-1}, x_t) = \frac{\exp\{(\theta^+)^T s(y^+, y_{t-1}, x_t)\}}{\kappa(\theta^+, y_{t-1}, x_t)} \frac{\exp\{(\theta^-)^T s(y^-, y_{t-1}, x_t)\}}{\kappa(\theta^-, y_{t-1}, x_t)}$$
$$= \frac{\exp\{\theta^T s(y_t, y_{t-1}, x_t)\}}{\kappa(\theta, y_{t-1}, x_t)}$$

with $\theta = (\theta^+, \theta^-)^T$, $s(y_t, y_{t-1}, x_t) = (s(y^+, y_{t-1}, x_t), s(y^-, y_{t-1}, x_t))^T$, and the normalization constant set accordingly.

For practical reasons, it is important to understand that the term dissolution model is somewhat misleading because a positive coefficient in the dissolution model implies that nodes (or dyads) with high values for this statistic are less likely to dissolve. This is also the standard implementation in software packages but can simply be changed by switching the signs of the parameters in the dissolution model.

The network statistics are used similarly as in a cross-sectional ERGM. In Krivitsky and Handcock (2014), they are called *implicitly dynamic* because they are evaluated either at the formation network y^+ or the dissolution network y^-, which are both formed from y_{t-1} and y_t. For example, the number of edges is separately computed now for the formation and the dissolution network, giving either the number of edges that newly formed or the number of edges that persisted. For example, reciprocity in the formation network is defined as

$$s_{\text{recip}}(y^+, y_{t-1}) = s_{\text{recip}}(y^+) \propto \sum_{i \neq j} y^+_{ji} y^+_{ij}, \tag{11}$$

and in case of the dissolution model, y^+ is simply exchanged with y^-. Similarly, edge covariates or the geometrically weighted statistics shown in Equations (5), (7), and (8) are now functions of y^+ or y^-, not y_t.

3.3 | Model assessment

In analogy to binary regressions models, the (S)TERGM can be evaluated in terms of their receiver-operator-characteristic (ROC) curve or precision-recall (PR), where the latter puts more emphasis on finding true positives (e.g., Grau, Grosse, & Keilwagen, 2015). A comparison between different models is possible using, for example, the Akaike information criterion (AIC, Claeskens & Hjort, 2008). Here, we want to highlight that the AIC fundamentally builds on the log likelihood, which in most realistic applications is only available as an approximation; see Hunter et al. (2008) for further discussion.

However, in statistical network analysis, it is often argued that suitable network models should not exclusively provide good predictions for individual edges but also be able to represent topologies of the observed network. The dominant approach to asses the goodness-of-fit of (S)TERGMs is based on sampling networks from their distribution under the estimated parameters and then comparing network characteristics of these sampled networks with the same ones from the observed network (Hunter et al., 2008). For this approach, it is recommendable to utilize network characteristics that are not used for specifying the model. For instance, models that include the GWOD statistic (7) may not be compared to its simulated values but against the out-degree distribution.

Hanneke et al. (2010) point out that for networks with more than one transition from $t-1$ to t available, it is possible to employ a “cross-validation-type” assessment of the fit. The parameters can be fit repetitively to all observed transitions except one hold-out transition. It is then checked, how well the network statistics from the hold-out transition period are represented by the ones sampled from the coefficients obtained from all other transitions.

4 | RELATIONAL EVENT MODEL

4.1 | Time-continuous event processes

The second type of dynamic network models results by comprehending network changes as a continuously evolving process (see Girardin & Limnios, 2018 as a basic reference for stochastic processes). The idea was originally introduced by Holland and Leinhardt (1977). In their work, changes in the network are not occurring at discrete time points but as a continuously evolving process, where only one tie can be toggled at a time. This framework was extended by Butts (2008) to model behavior, which is understood as a directed event at a specific time that potentially depends on the past. Correspondingly, the observations in this section are behaviors that are given as triplets $\omega = (i, j, \tilde{t})$ and encode sender i, receiver j, and exact time point $\tilde{t}$. This fine-grained temporal information is often called time-stamped or time-continuous; we adopt the latter name. Furthermore, we only regard dyadic events in this article, that is, a behavior only includes one sender and receiver.

The concept of behavior, hereinafter called event, generalizes the classical concept of binary relationships based on graph theory as promoted by Wasserman and Faust (1994). This event framework does not intrinsically assume that ties are enduring over a specific time frame (Butts, 2009; Butts & Marcum, 2017). For example in an email exchange network, sending one email at a specific time point is merely a brief event, which does not convey the same information as a durable relationship. Therefore, the time-stamped information cannot adequately be represented in a binary adjacency matrix without having to aggregate the relational data at the cost of information loss (Stadtfeld, 2012). Nevertheless, a friendship between actors i and j at a given time point can still be viewed as an event that has a one-to-one analogy to a tie in the classical framework.

The overall aim of relational event models (REMs, Butts, 2008) is to understand the dynamic structure of events conditional on the history of events (Lerner, Bussmann, Snijders, & Brandes, 2013). This dynamic structure, in turn, controls how past interactions shape the propensity of future events. To make this model feasible, we leverage results from the field of time-to-event analysis or survival analysis, respectively (see, e.g., Kalbfleisch & Prentice, 2002, for an overview).

The central concept of the REM can be motivated by the introduction of a multivariate time-continuous counting process

$$N(\tilde{t}) = (N_{ij}(\tilde{t}) | i, j \in \{1, \dots, n\}), \tag{12}$$

where $N_{ij}(\tilde{t})$ counts how often actors i and j interacted in $[0, \tilde{t})$. Note that we indicate continuous time $\tilde{t}$ with a tilde to distinguish from the discrete time setting with $t = 1, 2, \dots, T$ assumed in the previous section. Process (12) is characterized by an intensity function $\lambda_{ij}(\tilde{t})$ for $i \neq j$, which is defined as

$$\lambda_{ij}(\tilde{t}) = \lim_{dt \downarrow 0} \frac{\mathbb{P}(N_{ij}(\tilde{t} + dt) = N_{ij}(\tilde{t}) + 1)}{dt}.$$

This is the instantaneous probability of observing a jump of size "1" in $N_{ij}(\tilde{t})$, which indicates observing the event $(i, j, \tilde{t})$. Because we assume that there are no self-loops $\lambda_{ii}(\tilde{t}) \equiv 0 \; \forall \; i = 1, \dots, n$ holds.

4.2 | Time-continuous observations

Butts (2008) introduced the REM to analyze the intensity $\lambda_{ij}(\tilde{t})$ of process (12) when time-continuous data on the events are available. He assumed that the intensity is constant over time but depends on time-varying relational information of past events and exogenous covariates. Vu, Hunter, Smyth, and Asuncion (2011) extended the model by postulating a semiparametric intensity similar to Cox (1972):

$$\lambda_{ij}(\tilde{t} | N(\tilde{t}), x(\tilde{t}), \theta) = \lambda_0(\tilde{t}) \exp\{\theta^T s_{ij} \left(N(\tilde{t}), x(\tilde{t})\right)\}, \tag{13}$$

where $\lambda_0(\tilde{t})$ is an arbitrary baseline intensity, $\theta \in \mathbb{R}^p$ is the parameter vector, and $s_{ij}(N(\tilde{t}), x(\tilde{t}))$ is a statistic that depends on the (possibly time-continuous) covariate process $x(\tilde{t})$ and the counting process just prior to $\tilde{t}$.

Generally, similar statistics as already introduced in Section 3 can be included in $s_{ij}(N(\tilde{t}), x(\tilde{t}))$. Solely, the differing level of the model needs to be accounted for because model (13) takes a local time-continuous point of view to understand the relational nature of the observed events. This necessitates defining the statistics $s_{ij}(N(\tilde{t}), x(\tilde{t}))$ from the position of specific ties, in contrast to the globally defined statistics $s(y_t, y_{t-1}, x_t)$ in (2). To give an example, the tie-level version of reciprocity for the event (i,j) is defined as

$$s_{ij,\text{reciprocity}}(N(\tilde{t}), x(\tilde{t})) = \mathbb{1}(N_{ji}(\tilde{t}) > 0),$$

where $\mathbb{1}(\cdot)$ is the indicator function. It only regards, whether already having observed the event (j, i) prior to $\tilde{t}$ has an effect on $\lambda_{ij}(\tilde{t} | N(\tilde{t}), x(\tilde{t}), \theta)$, in comparison to the network level version (3) of delayed reciprocity that counted all reciprocated ties between the networks y_t and y_{t-1}.

Degree statistics can be specified as either sender- or receiver-specific. If we, for example, want to control for the out-degree of the sender, the corresponding tie-oriented statistic is

$$s_{ij,SOD}(N(\tilde{t}), x(\tilde{t})) = \sum_{h=1}^{n} \mathbb{1}(N_{ih}(\tilde{t}) > 0).$$

The in-degree of the receiver can be formulated accordingly.

Clustering in event sequences may be captured by different types of nested two-path configurations. For instance, the tie-oriented version of directed two-paths, henceforth called transitivity, is given by

$$s_{ij,TRA}(N(\tilde{t}), x(\tilde{t})) = \sum_{h=1}^{n} \mathbb{1}(N_{ih}(\tilde{t}) > 0)\mathbb{1}(N_{hj}(\tilde{t}) > 0).$$

The inclusion of monadic and dyadic exogenous covariates becomes straightforward by setting $s_{ij,\text{dyadic}}\left(N(\tilde{t}), x(\tilde{t})\right)$ equal to the covariate values of interest. Because the effect of a past event at time δ, say, on a present event at time $\tilde{t}$ may vary according to the elapsed time $\tilde{t} - \delta$, Stadtfeld and Block (2017) introduced windowed effects, which only regard events that occurred in a prespecified time window, for example, a year. We will come back to this point in the next section.

If time-continuous observations are available, each dimension of the observed counting process is conditionally independent given the past. This, in turn, enables the construction of a likelihood, which can subsequently be maximized. Assuming that Ω is the set of all observed events and $\mathcal{T}$ is the interval of observation, the likelihood can be written as:

$$\mathcal{L}(\theta) = \prod_{(i,j,\tilde{t})\in\Omega} \lambda_{ij}(\tilde{t}|N(\tilde{t}), x(\tilde{t}), \theta) \exp\left\{ -\int_{\mathcal{T}} \sum_{k,h=1}^{n} \lambda_{kh}(u|N(u), x(u), \theta)du \right\}. \tag{14}$$

It is straightforward to maximize the likelihood in the case of a parametric baseline intensity $\lambda_0(t)$, for example, Butts (2008) assumes $\lambda_0(t) = \gamma_0$. Alternatively, Butts (2008) analyzed events with ordinal temporal information. In this setting, the likelihood is equal to the partial likelihood introduced by Cox (1972) for estimating parameters of semiparametric intensities as in (13). Letting U_t denote the set of all possible events that could have occurred at time point t but did not, the partial likelihood for continuous event data is defined as

$$\mathcal{PL}_{\text{cont}}(\theta) = \prod_{(i,j,\tilde{t})\in\Omega} \frac{\lambda_{ij}(\tilde{t}|N(\tilde{t}), x(\tilde{t}), \theta)}{\sum_{(k,h)\in U_{\tilde{t}}} \lambda_{kh}(\tilde{t}|N(\tilde{t}), x(\tilde{t}), \theta)}. \tag{15}$$

Consecutively, $\Lambda_0(t) = \int_0^t \lambda_0(u)du$ can be estimated with a Nelson Aalen estimator (see Kalbfleisch & Prentice, 2002, for further details on the estimation).

When dealing with large amounts of event data, the main obstacle is evaluating the sum over the intensities of all possible ties in (15) (Butts, 2008). One exact option is to trade a longer running time for a slimmer memory footprint by means of a coaching data structure. Vu et al. (2011) exploit this by saving prior values of the sum and subsequently changing it event-wise by elements of $U_{\tilde{t}}$ whose covariates changed. Alternatively, Vu et al. (2015) proposes approximate routines that utilize case-control sampling and stratification for the Cox model (Langholz & Borgan, 1995). More precisely, the sum is only calculated over a sampled subset of possible events in addition to stratification. Lerner and Lomi (2019) go one step further and sample events out of Ω for the calculation of $\mathcal{PL}(\theta)$ in (15).

Numerous extensions of this model that build on already well-established methods in social networks and time-to event analysis have been proposed. Perry and Wolfe (2013) used a stratified Cox model in (13). Stadtfeld, Hollway, and Block (2017) adopted the SAOM to events. DuBois and Smyth (2010) and DuBois, Butts, and Smyth (2013) extended the stochastic block model for time-stamped relational events. Furthermore, DuBois, Butts, McFarland, and Smyth (2013) adopted a Bayesian hierarchical model to event data when information is only available in smaller groups.

4.3 | Time-clustered observations

Generally, the approach discussed above requires time-continuous network data, meaning that we observe the precise time points of all events. To give an instance, in the first data example, this means that we need the exact time point $\tilde{t}$ of an arms trade between countries i and j. Often, such exact time-stamped data are not available and, in fact, trading between states can hardly be stamped with a single time point $\tilde{t}$. Indeed, we often only observe the time-continuous network process at discrete time points $t = 1, \ldots, T$. In such setting, we may assume a Markov structure in that we do not look at the entire history of the process $N(\tilde{t})$ but just condition the intensity (13) on the history of events from the previous observation $t-1$ to $\tilde{t}$. Technically, this means that $N(t)$ is adapted to $\tilde{Y}(\tilde{t}) := N(\tilde{t}) - N(t-1)$ and $x(\tilde{t})$ for $\tilde{t} \in [t-1, t]$. We then reframe (13) as

$$\lambda_{ij}(\tilde{t}|\tilde{Y}(\tilde{t}), x(\tilde{t}), \theta) = \lambda_0(\tilde{t}) \exp\{\theta^T s_{ij}(\tilde{Y}(\tilde{t}), x(\tilde{t}))\}. \quad (16)$$

In other words, we assume that the intensity of events between $t-1$ and t does not depend on states of the multivariate counting process (12) prior to $t-1$. For this reason, all endogenous statistics introduced in Section 4.2 are now evaluated on $\tilde{Y}(\tilde{t})$ instead of $N(\tilde{t})$. This is a reasonable assumption, if one is primarily interested in short-term dependencies between the individual counting processes. It enables a meaningful comparison to the models from Section 3 that assume an analog discrete Markov property. However, we want to emphasize that this dependence structure is not vital to inferential results.

If we observe the continuous process at discrete time points, it is inevitable that we observe time-clustered observations, meaning that two or more events happen at the same time point. Under the term tied observations, this phenomenon is well known in time-to-event analysis and treated with several approximations. One option is the so-called Breslow approximation (see Breslow, 1974; Peto, 1972). Let therefore

$$O_t = \{(i,j) \mid N_{ij}(t) - N_{ij}(t-1) > 0\},$$

where element (i,j) is replicated $N_{ij}(t) - N_{ij}(t-1)$ times in O_t, that is, if an event between i and j occurred multiple times in the interval from $t-1$ to t, then (i,j) appears respective times in O_t. Given that we have not observed the exact time point of an event, we also get no information on the baseline intensity $\lambda_0(\tilde{t})$ in (13) for $\tilde{t} \in [t-1, t]$ so that the model simplifies to a discrete choice model structure (see, e.g., Train, 2009), which resembles the partial likelihood (15) and is defined as

$$\mathcal{PL}_{\text{clust}}(\theta) = \prod_{t=1}^{T} \frac{\prod_{(i,j)\in O_t} \exp\{\theta^T s_{ij}(\tilde{Y}(t), x(t))\}}{\left(\sum_{(k,h)\in U_t} \exp\{\theta^T s_{kh}(\tilde{Y}(t), x(t))\}\right)^{n_t}}, \quad (17)$$

where $n_t = |O_t|$. Alternatively, one can replace the denominator in (17) by considering all possible orders of the unobserved events in O_t giving the average likelihood as introduced by Kalbfleisch and Prentice (2002). Because this can be a combinatorial and, hence, numerical challenge, random sampling of time point orders among the time-clustered observations can be used with subsequent averaging, which we call Kalbfleisch–Prentice approximation (see Kalbfleisch & Prentice, 2002). Further techniques to deal with unknown time ordering are augmenting the clustered events into possible paths of ordered events and adapting the maximum likelihood estimation proposed for the SAOM by Snijders, Koskinen, and Schweinberger (2010) or using random sampling of the ordering. This can be legitimized in cases where we may assume independence among events happening in one year because the events take a long time to materialize (Snijders, 2017).

4.4 | Model assessment

In comparison to the assessment for models operating in discrete time, widely accepted methods dealing with relational event data are scarce. The proposals either stem from time-to-event analysis or regard link prediction, which is the task of predicting the most likely next event given the history of past events (Liben-Nowell & Kleinberg, 2007). One example of the former option is the usage of *Schönfeld residuals* by Vu, Asuncion, Hunter, and Smyth (2011) to check the assumption of proportional intensities, which is central to semiparametric models as the one proposed by Cox (1972). For the latter approach, we need to define a predictive measure that quantifies how well the next event is predicted. Vu et al. (2011) proposed the recall measure that estimates the percentage of test events, which are in the list of K most likely next events according to a given model. Evaluating this percentage for different values of K permits a visualization of the predictive capabilities of the model. The strength of the predicted intensity allows the ordering of events according to the probability of being observed next. If we model the propensity of time-clustered events that represent binary adjacency matrices, one can alternatively adopt the analysis of the ROC and PR curve introduced in Section 3.3.

5 | APPLICATION

When it comes to software, there exist essentially three main `R` packages that are designed for fitting TERGMs and STERGMs. Most important is the extensive `statnet` library (Goodreau, Handcock, Hunter, Butts, & Morris, 2008) that allows for simulation-based fitting of ERGMs. The library contains the package `tergm` with implemented methods for fitting STERGMs using MCMC approximations of the likelihood. However, currently the package `tergm` (version 3.5.2) does not allow for fitting STERGMs with time-varying dyadic covariates for more than two time periods jointly. The package `btergm` (Leifeld et al., 2018) is designed for fitting TERGMs using either maximum pseudolikelihood or MCMC maximum likelihood estimation routines. In order to obtain Bayesian Inference in ERGMs, the package `bergm` by Caimo and Friel (2014) can be used. Besides implementations in `R`, the stand-alone program `PNet` (Wang, Robins, & Pattison, 2006) allows for simulating, fitting, and evaluating (T)ERGMs. In order to ensure comparable estimates, we estimate the TERGM, as well as the STERGM, with the `statnet` library, using MCMC-based likelihood estimation techniques. We use the package `ergm` and include delayed reciprocity and the repetition of previous ties as dyadic covariates. The STERGM is fitted using the `tergm` package.

Marcum and Butts (2015) implemented the R package `relevent` (version 1.0-4) to estimate the REM for time-stamped data. It was followed by the package `goldfish` (version 1.2) by Stadtfeld and Hollway (2018) for modeling event data with precise and ordinal temporal information with an actor- and tie-oriented variant of the REM. Furthermore, it is highly customizable in terms of endogenous and exogenous user terms and will be used in the following applications.

We want to remark that the STERGM coefficients are implicitly dynamic, whereas in the TERGM, all network statistics except the lagged network and delayed reciprocity terms are evaluated on the network in t. All covariates of the REM are continuously updated and the intensity at time point $\tilde{t} \in [t-1, t]$ only depends on events observed in $[t-1, \tilde{t})$. Like the building period proposed by Vu et al. (2011), the events in $t-1$ are only used for building up the covariates and not directly modeled. Moreover, the coefficients of the REM affect the intensity of a specific event on the tie level. For the (S)TERGM, the estimates can be interpreted as the effect of a global statistic $s(y_t, y_{t-1}, x_t)$ on the probability of observing the network y_t. They also allow for a tie-level interpretation based on the so-called change statistics (Cranmer & Desmarais, 2011). Due to no compositional changes, we did not scale any statistics.

5.1 | Data set 1: International arms trade

The results obtained for the arms trading data section are displayed in Table 2. For a detailed interpretation of effects focusing on political, social, and economic aspects, we refer to the relevant literature (e.g., Thurner et al., 2019). Here, we want to comment on a few aspects only. While we do not have time stamps for the arms trades, the longitudinal networks can still be viewed as time-clustered observations enabling the techniques from Section 4.3.

TABLE 2 Arms trade network: Comparison of parameters obtained from the TERGM (first column), STERGM (formation in the second column, dissolution in the third column), and REM (fourth column)

	TERGM	STERGM		REM	
		Formation	Dissolution		
Repetition	3.671***	–	–	2.661***	
	(0.132)	–	–	(0.143)	
Edges	−15.632***	−17.186***	−16.987***	–	
	(1.809)	(2.168)	(3.587)	–	
Reciprocity	−0.258	−0.620	−0.058	−0.109	
	(0.306)	(0.436)	(0.619)	(0.181)	
In-degree (GWID)	−1.823***	−2.106***	−0.412	0.060**	In-degree receiver
	(0.278)	(0.379)	(0.442)	(0.015)	
Out-degree (GWOD)	−3.220***	−4.126***	−0.326	0.010**	Out-degree sender
	(0.304)	(0.462)	(0.533)	(0.004)	
GWESP	0.050	0.076	0.150	0.010	Transitivity
	(0.066)	(0.071)	(0.126)	(0.029)	
Polity score (absolute difference)	−0.024*	−0.028*	−0.016	−0.016	
	(0.010)	(0.014)	(0.017)	(0.009)	
log(GDP) sender	0.313***	0.394***	0.323***	0.395***	
	(0.048)	(0.054)	(0.088)	(0.039)	
log(GDP) receiver	0.165***	0.135*	0.327***	0.192***	
	(0.043)	(0.054)	(0.087)	(0.032)	
Log likelihood	−949.833	−675.327	−258.425		
AIC	1,917.666	1,366.654	532.849		
$\sum$ AIC	1,917.666	1,899.503			

Note. Standard errors in brackets and stars according to p values smaller than 0.001 (***), 0.05 (**), and 0.1 (*). The decay parameter of the geometrically weighted statistics is set to log(2) and the Kalbfleisch–Prentice approximation was used with 100 random orderings of the events to find the estimates of the REM. TERGM = temporal exponential random graph model; STERGM = separable temporal exponential random graph model; REM = relational event model; GWID = geometrically weighted in-degree distribution; GWOD = geometrically weighted out-degree distribution; GWESP = geometrically weighted edge-wise shared partners; GDP = gross domestic product; AIC = Akaike information criterion.

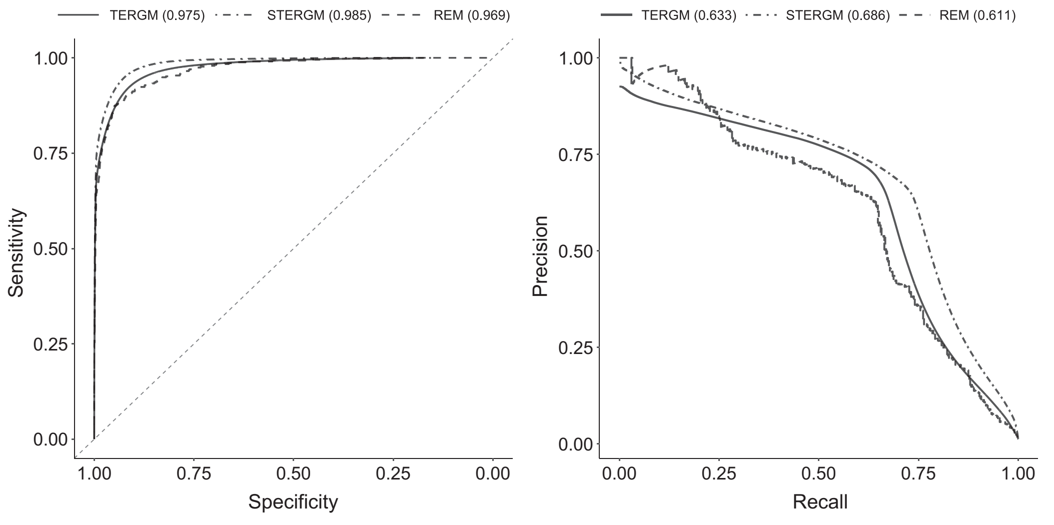

FIGURE 5 Arms trade network: ROC (left) and PR (right) curves from the TERGM (dotted line), STERGM (dot-dashed line), and REM (solid line). The AIC values of the respective curves are indicated in brackets. ROC = receiver-operator-characteristic; PR = precision-recall; TERGM = temporal exponential random graph model; STERGM = separable temporal exponential random graph model; REM = relational event model; AIC = Akaike information criterion

Both the TERGM (column 1) and the REM (column 4) identify the repetition of previous ties as a driving force in the dynamic structure of the network. Degree-related covariates, which are GWID and GWOD in the (S)TERGM and the in- and out-degrees in the REM, capture centrality in the network. The coefficients of the GWID and GWOD are negative and have low *p* values in the TERGM. This stands in contrast to the STERGM, where these effects are only pronounced in the formation model (column 2); however, they are insignificant effects in the dissolution model (column 3). Hence, these effects suggest a centralized pattern in the formation network, which is also captured by the TERGM. In the REM, an analogous pattern can be detected because a higher in-degree of the receiver increases the respective intensity; thus; spurs trade relations. Similar interpretations hold for the out-degree of the sender. Overall, countries that have a high out-degree are more likely to send weapons and countries with a high in-degree to receive them, which again results in a centralized network structure as indicated by the estimates in the TERGM and STERGM. This example illustrates how seemingly contrasting coefficients of the (S)TERGM and REM can still imply a similar interpretation in terms of the implied global network characteristics.

Lastly, consistent effects among the models were also found for the exogenous covariates. Consider, for instance, the coefficient of the logarithmic GDP of the importing country. The TERGM assigns a significantly higher probability to observe in-going ties to countries with a high GDP just like the REM. However, disentangling the model towards formation and dissolution, we see strongly significant coefficients in the dissolution model, whereas the effect for the formation model is weakly significant.

Based on the independence assumption in (9), we can sum up the two AIC values and see that the AIC value of the STERGM is smaller than of the TERGM. Furthermore, the ROC and PR curves of all three models are shown in Figure 5. Again, both measures indicate that the STERGM provides a slightly improved fit when compared to the TERGM and REM. The results of the simulation-based model assessment of the (S)TERGM can be found in the Supplementary Material.

5.2 | Data set 2: European research institution email correspondence

As already indicated by the descriptive statistics in Table 1, the email network seems to be driven by three major structural influences: repetition, reciprocity, and transitive clustering. The estimates from Table 3 demonstrate that all models were able to identify these forces.

According to the REM (column 4), the event network of email traffic in the research institution is not centralized and primarily based on collaboration between coworkers. We can draw those conclusions from insignificant estimates of degree-related statistics and highly significant estimates regarding reciprocity and repetition. In the TERGM (column 1), we find a positive and significant effect of GWID, whereas no effect can be found in the STERGM (columns 2 and 3). The estimates of repetition and reciprocity in the REM and TERGM are very pronounced. For instance, the estimates of the REM imply that a reciprocated event is 19.6 times more likely than an event with the same covariates only not being reciprocated. Interestingly, the STERGM detects a lower effect of GWESP in the formation and dissolution than the TERGM. The effect of the delayed reciprocity in the TERGM is less relevant than reciprocity in the formation and dissolution model. This strongly differing effect size results from the mathematical formulation of the statistics given in Equations (3) and (11).

Contrasting the AIC values of the TERGM and STERGM shows that the dynamic structure of the email network is again better explained by the STERGM. In the Supplementary Material, we

TABLE 3 Email exchange network: Comparison of parameters obtained from the TERGM (first column), STERGM (formation in the second column, dissolution in the third column), and REM (fourth column)

	TERGM	STERGM		REM	
		Formation	Dissolution		
Repetition	1.367***	–	–	2.27***	
	(0.107)	–	–	(0.084)	
Edges	−5.755***	−4.853***	−2.237***	–	
	(0.237)	(0.247)	(0.224)	–	
Reciprocity	0.398***	2.498***	2.586***	1.655***	
	(0.112)	(0.157)	(0.226)	(0.075)	
In-degree (GWID)	1.060**	1.349*	0.709	−0.004	In-degree receiver
	(0.333)	(0.648)	(0.415)	(0.003)	
Out-degree (GWOD)	0.031	−0.411	−0.369	−0.0001	Out-degree sender
	(0.312)	(0.431)	(0.397)	(0.003)	
GWESP	1.560***	0.655***	0.429***	0.070***	Transitivity
	(0.110)	(0.111)	(0.086)	(0.008)	
Log likelihood	−1,723.732	−1,000.506	−505.431		
AIC	3,459.464	2,011.012	1,020.862		
$\sum$ AIC	3,459.464	3,031.874			

Note. Standard errors in brackets and stars according to p values smaller than 0.001 (***), 0.05 (**), and 0.1 (*). Decay parameter of the geometrically weighted statistics is set to log(2). TERGM = temporal exponential random graph model; STERGM = separable temporal exponential random graph model; REM = relational event model; GWID = geometrically weighted in-degree distribution; GWOD = geometrically weighted out-degree distribution; GWESP = geometrically weighted edge-wise shared partners; GDP = gross domestic product; AIC = Akaike information criterion.

give the results of model assessment for the (S)TERGM and REM, as well as an application of the TERGM and STERGM to multiple time points.

6 | CONCLUSION

6.1 | Further models

Snijders (1996) formulated a two-stage process model operating in a continuous-time framework. The dynamics are considered to evolve according to unobserved microsteps. At first, a sender out of all eligible actors gets the opportunity to change the state of all his outgoing ties. Consecutively, the actor needs to evaluate the probability of changing the present configuration with each possible receiver, which entails each actor's knowledge of the complete graph whenever he has the possibility to toggle one of his ties. Lastly, the decision is randomly drawn relative to the probabilities of all possible actions. In general, the SAOM is a well-established model for the analysis of social networks, which was successfully applied to a wide array of network data, for example, in sociology (Agneessens & Wittek, 2012; de Nooy, 2002), political science (Bichler & Franquez, 2014; Kinne, 2016), economics (Castro, Casanueva, & Galán, 2014), and psychology (Jason, Light, Stevens, & Beers, 2014). Estimation of this model variant is predominantly carried out with the R package `RSiena` (Ripley, Boitmanis, & Snijders, 2013).

Another notable model that can be regarded as a bridge between the ERGM and continuous-time models is the longitudinal ERGM (LERGM; Koskinen, Caimo, & Lomi, 2015; Snijders & Koskinen, 2013). In contrast to the TERGM, the LERGM assumes that the network evolves in microsteps as a continuous-time Markov process with an ERGM being its limiting distribution. Similar to the SAOM, the model builds on randomly assigning the opportunity to change, followed by a function that governs the probability of a tie change. This model is still tie-oriented, meaning that dyadic ties instead of actors are chosen and then have the option to change the current network.

6.2 | Summary

In this article, we emphasize tie-oriented dynamic network models. Comparisons between these models can be drawn on the level at which each implied generating mechanism works and how time is perceived. The overall aim in the TERGM is to find an adequate distribution of the adjacency matrix Y_t conditioning on information of previous realizations of the network. In the separable extension, the aim remains unchanged, only splitting Y_t into two subnetworks that include all possible ties that were and were not present in Y_{t-1} separately. While the (S)TERGM proceeds in discrete time, the REM tackles modeling the intensity on the tie level in continuous time conditional on past events. Therefore, the (S)TERGM takes a global and REM a local point of view. Even though this results in substantially different interpretations of the estimates, they can still be related to one another by focusing on how the global effects of the (S)TERGM can be emergent from tie-level effects treated in the REM. Furthermore, our discussion exposed how the model assessment of the REM focuses on analogies with time-to-event analysis, that is, looking at the adequate behavior of the model on the tie-level, and the (S)TERGM on simulation-based procedures, which regards the capabilities of the model to capture global characteristics.

Finally, the typically required data structure differs between the model classes. While the time-discrete generating mechanism of the (S)TERGM naturally processes network data that were observed at discrete time points, the REM is based on time-continuous data. Of the two analyzed data sets, the international arms trade network represents the former data structure, whereas the email traffic data are an example of the latter. By extending the REM to time-clustered observations and aggregating events to binary adjacency matrices, a meaningful comparison between the STERGM, TERGM, and REM is enabled.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their careful reading and constructive comments. The project was supported by the European Cooperation in Science and Technology (COST Action CA15109 [COSTNET]). We also gratefully acknowledge funding provided by the German Research Foundation (DFG) for project KA 1188/10-1: *International Trade of Arms: A Network Approach*. Furthermore, we like to thank the Munich Center for Machine Learning (MCML) for the funding.

ORCID

Cornelius Fritz https://orcid.org/0000-0002-7781-223X

REFERENCES

Agneessens, F., & Wittek, R. (2012). Where do intra-organizational advice relations come from? The role of informal status and social capital in social exchange. *Social Networks*, *34*(3), 333–345.

Almquist, Z. W., & Butts, C. T. (2014). Logistic network regression for scalable analysis of networks with joint edge/vertex dynamics. *Sociological Methodology*, *44*(1), 273–321.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154

Bearman, P., Moody, J., & Stovel, K. (2004). Chains of affection: The structure of adolescent romantic and sexual networks. *American Journal of Sociology*, *110*(1), 44–91.

Benton, R. A., & You, J. (2017). Endogenous dynamics in contentious fields: Evidence from the shareholder activism network, 2006–2013. *Socius: Sociological Research for a Dynamic World*, *3*.

Bichler, G., & Franquez, J. (2014). *Conflict cessation and the emergence of weapons supermarkets* (pp. 189–215). Cham, Switzerland: Springer International Publishing.

Blank, M., Dincecco, M., & Zhukov, Y. M. (2017). Political regime type and warfare: Evidence from 600 years of european history. Available at SSRN: https://dx.doi.org/10.2139/ssrn.2830066.

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. *Social Networks*, *52*, 180–191.

Block, P., Stadtfeld, C., & Snijders, T. (2019). Forms of dependence: Comparing SAOMs and ERGMs from basic principles. *Sociological Methods & Research*, *48*(1), 202–239.

Breslow, N. (1974). Covariance analysis of censored survival data. *Biometrics*, *30*(1), 89–99.

Broekel, T., & Bednarz, M. (2018). Disentangling link formation and dissolution in spatial networks: An application of a two-mode STERGM to a project-based R&D network in the German biotechnology industry. *Networks and Spatial Economics*, *18*(3), 677–704.

Butts, C. (2008). A relational event framework for social action. *Sociological Methodology*, *38*(1), 155–200.

Butts, C. (2009). Revisiting the foundations of network analysis. *Science*, *325*, 414–416.

Butts, C., & Marcum, C. (2017). A relational event approach to modeling behavioral dynamics. In *Group processes data-driven computational approaches*. Cham, Switzerland: Springer.

Caimo, A., & Friel, N. (2011). Bayesian inference for exponential random graph models. *Social Networks*, *33*(1), 41–55.

Caimo, A., & Friel, N. (2014). Bergm: Bayesian exponential random graphs in R. *Journal of Statistical Software*, *61*(2), 1–25.

Castro, I, Casanueva, C., & Galán, J. L. (2014). Dynamic evolution of alliance portfolios. *European Management Journal*, *32*(3), 423–433.

Center for Systemic Peace (2017). Polity IV Annual Time-Series, 1800-2015. Version 3.1. http://www.systemicpeace.org

Claeskens, G., & Hjort, N. L. (2008). *Model selection and model averaging*. Cambridge, UK: Cambridge University Press.

Cox, D. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *34*(2), 187–220.

Cranmer, S., & Desmarais, B. (2011). Inferential network analysis with exponential random graph models. *Political Analysis*, *19*(1), 66–86.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*(5), 1–9.

Desmarais, B. A., & Cranmer, S. J. (2012). Statistical mechanics of networks: Estimation and uncertainty. *Physica A: Statistical Mechanics and Its Applications*, *391*(4), 1865–1876.

DuBois, C., Butts, C., McFarland, D., & Smyth, P. (2013). Hierarchical models for relational event sequences. *Journal of Mathematical Psychology*, *57*(6), 297–309.

DuBois, C., Butts, C., & Smyth, P. (2013). Stochastic blockmodeling of relational event dynamics. In C. M. Carvalho & P. Ravikumar (Eds.), *Proceedings of the 16th International Conference on Artificial Intelligence and Statistics* (pp. 238–246).

DuBois, C., & Smyth, P. (2010). Modeling relational events via latent classes. *Proceedings of the 16th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 803–812.

van Duijn, M. A., Gile, K. J., & Handcock, M. S. (2009). A framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential family random graph models. *Social Networks*, *31*(1), 52–62.

Email EU Core. (2019). Email EU core temporal. Department 3. https://snap.stanford.edu/data/email-Eu-core-temporal-Dept3.txt.gz

Erdös, P., & Rényi, A. (1959). On random graphs I. *Publicationes Mathematicae Debrecen*, *6*, 290.

Frank, O., & Strauss, D. (1986). Markov Graphs. *Journal of the American Statistical Association*, *81*(395), 832–842.

Geyer, C. J., & Thompson, E. A. (1992). Constrained Monte Carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *54*(3), 657–699.

Gilbert, E. N. (1959). Random graphs. *The Annals of Mathematical Statistics*, *30*(4), 1141–1144.

Girardin, V., & Limnios, N. (2018). *Applied probability* (2nd ed.). Heidelberg, Germany: Springer.

Goldenberg, A., Zheng, A. X., Fienberg, S. E., & Airoldi, E. M. (2010). A survey of statistical network models. *Foundations and Trends in Machine Learning*, *2*(2), 129–233.

Goodreau, S. M., Handcock, M. S., Hunter, D. R., Butts, C. T., & Morris, M. (2008). A statnet tutorial. *Journal of Statistical Software*, *24*(9), 1.

Goodreau, S. M., Kitts, J. A., & Morris, M. (2009). Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography*, *46*(1), 103–125.

Grau, J., Grosse, I., & Keilwagen, J. (2015). PRROC: Computing and visualizing precision-recall and receiver operating characteristic curves in R. *Bioinformatics*, *31*(15), 2595–2597.

Handcock, M. S. (2003). *Assessing degeneracy in statistical models of social networks* (Working Paper No. 39).

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

He, X., Dong, Y.-B., Wu, Y.-Y., Jiang, G.-R., & Zheng, Y. (2019). Factors affecting evolution of the interprovincial technology patent trade networks in China based on exponential random graph models. *Physica A: Statistical Mechanics and Its Applications*, *514*, 443–457.

Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association*, *97*(460), 1090–1098.

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

Holland, P. W., & Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*, *76*(373), 33–50.

Hummel, R. M., Hunter, D. R., & Handcock, M. S. (2012). Improving simulation-based algorithms for fitting ERGMs. *Journal of Computational and Graphical Statistics*, *21*(4), 920–939.

Hunter, D. R. (2007). Curved exponential family models for social networks. *Social Networks*, *29*(2), 216–230.

Hunter, D. R., Goodreau, S. M., & Handcock, M. S. (2008). Goodness of fit of social network models. *Journal of the American Statistical Association*, *103*(481), 248–258.

Jason, L. A., Light, J. M., Stevens, E. B., & Beers, K. (2014). Dynamic social networks in recovery homes. *American Journal of Community Psychology*, *53*(3-4), 324–334.

Kalbfleisch, J., & Prentice, R. (2002). *The statistical analysis of failure time data*. Hoboken, NJ: Wiley-Blackwell.

Kim, B., Lee, K. H., Xue, L., & Niu, X. (2018). A review of dynamic network models with latent variables. *Statistics Surveys*, *12*, 105–135.

Kinne, B. J. (2016). Agreeing to arm: Bilateral weapons agreements and the global arms trade. *Journal of Peace Research*, *53*(3), 359–377.

Kolaczyk, E. D. (2009). *Statistical analysis of network data: Methods and models*. New York, NY: Springer.

Koskinen, J., Caimo, A., & Lomi, A. (2015). Simultaneous modeling of initial conditions and time heterogeneity in dynamic networks: An application to Foreign Direct Investments. *Network Science*, *3*(1), 58–77.

Krivitsky, P. N., & Handcock, M. S. (2014). A separable model for dynamic networks. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *76*(1), 29–46.

Langholz, B., & Borgan, Ø. (1995). Counter-matching: A stratified nested case-control sampling method. *Biometrika*, *82*(1), 69–79.

Lebacher, M., Thurner, P. W., & Kauermann, G. (2019). A dynamic separable network model with actor heterogeneity: An application to global weapons transfers. arXiv e-prints arXiv:1803.02707v4.

Leifeld, P., Cranmer, S. J., & Desmarais, B. A. (2018). Temporal exponential random graph models with btergm: estimation and bootstrap confidence intervals. *Journal of Statistical Software*, *83*(6).

Lerner, J., Bussmann, M., Snijders, T., & Brandes, U. (2013). Modeling frequency and type of interaction in event networks. *Corvinus Journal of Sociology and Social Policy*, *4*(1), 3–32.

Lerner, J., & Lomi, A. (2019). Reliability of relational event model estimates under sampling: How to fit a relational event model to 360 million dyadic events. arXiv preprint arXiv:1905.00630.

Leskovec, J., Lang, K. J., Dasgupta, A., & Mahoney, M. W. (2009). Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, *6*(1), 29–123.

Liben-Nowell, D., & Kleinberg, J. (2007). The link-prediction problem for social networks. *Journal of the Association for Information Science and Technology*, *58*(7), 1019–1031.

Lusher, D., Koskinen, J., & Robins, G. (2012). *Exponential random graph models for social networks: Theory, methods, and applications*. Cambridge, UK: Cambridge University Press.

Marcum, C., & Butts, C. (2015). Constructing and modifying sequence statistics for relevent Using informR in R. *Journal of Statistical Software*, *64*(5), 1–36.

Morris, M., Handcock, M. S., & Hunter, D. R. (2008). Specification of exponential-family random graph models: Terms and computational aspects. *Journal of Statistical Software*, *24*(4), 1548.

de Nooy, W. (2002). The dynamics of artistic prestige. *Poetics*, *30*(3), 147–167.

Paranjape, A., Benson, A. R., & Leskovec, J. (2017). Motifs in temporal networks. *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, pp. 601–610.

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

Peto, R. (1972). Contribution to the discussion of the paper by Dr. Cox. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *34*(2), 187–220.

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

Raabe, I. J., Boda, Z., & Stadtfeld, C. (2019). The social pipeline: How friend influence and peer exposure widen the STEM gender gap. *Sociology of Education*, *92*(2), 105–123.

Ripley, R., Boitmanis, K., & Snijders, T. (2013). RSiena: Siena–simulation investigation for empirical network analysis. R package version 1.2-12. https://CRAN.R-project.org/package=RSiena

Robins, G., & Pattison, P. (2001). Random graph models for temporal processes in social networks. *Journal of Mathematical Sociology*, *25*(1), 5–41.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks*, *29*(2), 173–191.

Salathé, M., Vu, D., Khandelwal, S., & Hunter, D. (2013). The dynamics of health behavior sentiments on a large online social network. *EPJ Data Science*, *2*(1), 4.

Sarkar, P., & Moore, A. W. (2005). Dynamic social network analysis using latent space models. *ACM SIGKDD Explorations Newsletter*, *7*(2), 31–40.

Schweinberger, M. (2011). Instability, sensitivity, and degeneracy of discrete exponential families. *Journal of the American Statistical Association*, *106*(496), 1361–1370.

SIPRI (2019). SIPRI arms transfers database. https://www.sipri.org/databases/armstransfers

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

Snijders, T. (2002). The statistical evaluation of social network dynamics. *Sociological Methodology*, *31*(1), 361–395.

Snijders, T. (2005). Models for longitudinal network data. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 215–247). Cambridge, UK: Cambridge University Press.

Snijders, T. (2017). Comment: Modeling of coordination, rate functions, and missing ordering information. *Sociological Methodology*, *47*(1), 41–47. Available at https://doi.org/10.1177/0081175017723346

Snijders, T., & Koskinen, J. (2013). *Longitudinal models*. In D. Lusher, J. Koskinen, & G. Robins (Eds.), *Exponential random graph models for social networks* (pp. 130–140). Cambridge, UK: Cambridge University Press.

Snijders, T., Koskinen, J., & Schweinberger, M. (2010). Maximum likelihood estimation for social network dynamics. *The Annals of Applied Statistics*, *4*(2), 567.

Snijders, T., Pattison, P., Robins, G., & Handcock, M. (2006). New specifications for exponential random graph models. *Sociological Methodology*, *36*(1), 99–153.

Stadtfeld, C. (2012). Events in social networks: a stochastic actor-oriented framework for dynamic event processes in social networks. Karlsruhe, Germany: KIT Scientific Publishing.

Stadtfeld, C., & Block, P. (2017). Interactions, actors, and time: Dynamic network actor models for relational events. *Sociological Science*, *4*(14), 318–352.

Stadtfeld, C., & Hollway, J. (2018). goldfish: Goldfish – Statistical network models for dynamic network data. Retrieved from R package version 1.2. www.social-networks.ethz.ch/research/goldfish.html

Stadtfeld, C., Hollway, J., & Block, P. (2017). Dynamic network actor models: Investigating coordination ties through time. *Sociological Methodology*, *47*(1), 1–40.

Stansfield, S. E., Mittler, J. E., Gottlieb, G. S., Murphy, J. T., Hamilton, D. T., Detels, R., … Goodreau, S. M. (2019). Sexual role and hiv-1 set point viral load among men who have sex with men. *Epidemics*, *26*, 68–76.

Strauss, D., & Ikeda, M. (1990). Pseudolikelihood estimation for social networks. *Journal of the American Statistical Association*, *85*(409), 204–212.

Thurner, P. W., Schmid, C. S., Cranmer, S. J., & Kauermann, G. (2019). Network interdependencies and the evolution of the international arms trade. *Journal of Conflict Resolution*, *63*(7), 1736–1764.

Train, K. (2009). *Discrete choice methods with simulation*. Cambridge, UK: Cambridge University Press.

Tranmer, M., Marcum, C. S., Morton, B., Croft, D., & de Kort, S. (2015). Using the relational event model (REM) to investigate the temporal dynamics of animal social networks. *Animal Behaviour*, *101*, 99–105.

Vu, D., Asuncion, A., Hunter, D., & Smyth, P. (2011). Dynamic egocentric models for citation networks. *Proceedings of the 28th International Conference on Machine Learning*, pp. 857–864.

Vu, D., Hunter, D., Smyth, P., & Asuncion, A. (2011). Continuous-time regression models for longitudinal networks. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, & K. Q. Weinberger (Eds.), *Advances in neural information processing systems 24* (pp. 2492–2500). Red Hook, NY: Curran Associates.

Vu, D., Lomi, A., Mascia, D., & Pallotti, F. (2017). Relational event models for longitudinal network data with an application to interhospital patient transfers. *Statistics in Medicine*, *36*(14), 2265–2287.

Vu, D., Pattison, P., & Robins, G. (2015). Relational event models for social learning in MOOCs. *Social Networks*, *43*, 121–135.

Wang, P., Robins, G., & Pattison, P. (2006). PNet: A program for the simulation and estimation of exponential random graph models. Available from http://sna.unimelb.edu.au/PNet

Ward, M., Ahlquist, J. S., & Rozenas, A. (2013). Gravity's rainbow: A dynamic latent space model for the world trade network. *Network Science*, *1*(1), 95–118.

Wasserman, S., & Faust, K. (1994). *Structural Analysis in the Social Sciences. Social network analysis: Methods and applications*. Cambridge, UK: Cambridge University Press.

White, L. A., Forester, J. D., & Craft, M. E. (2018). Covariation between the physiological and behavioral components of pathogen transmission: Host heterogeneity determines epidemic outcomes. *Oikos*, *127*(4), 538–552.

World Bank (2019). World Bank Open Data, Real GDP.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Fritz C, Lebacher M, Kauermann G. Tempus volat, hora fugit: A survey of tie-oriented dynamic network models in discrete and continuous time. *Statistica Neerlandica*. 2020;74:275–299. https://doi.org/10.1111/stan.12198

APPENDIX

ADDITIONAL DESCRIPTIVES

Figures A1 and A2 depict the distributions of in- and out-degrees in the two networks. Building on the in- and out-degrees of all nodes, these distributions represent the relative frequency of all possible in- and out-degrees in the observed networks, which is calculated with the `igraph` package in `R` (Csardi & Nepusz, 2006).

In the arms trade network, a strongly asymmetric relation is revealed, indicating that about 70% of the countries do not export any weapons, whereas a small percentage of countries account

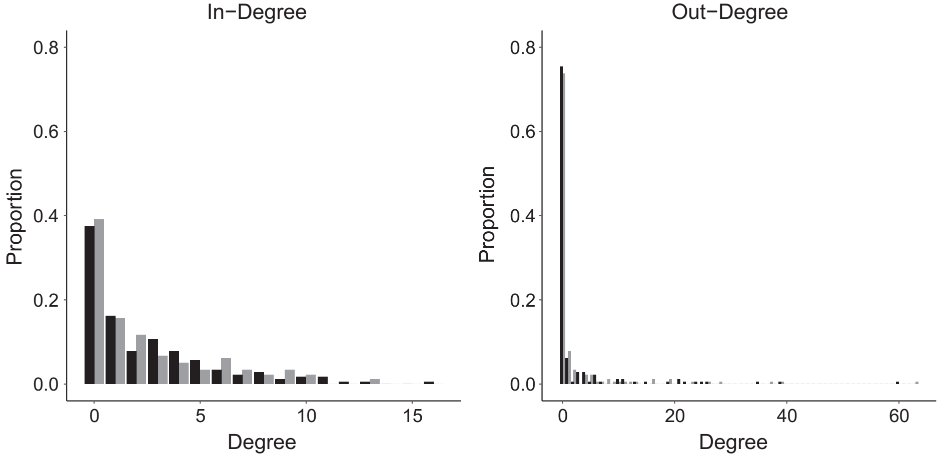

FIGURE A1 Arms trade network: Bar plots indicating the distribution of the in- and out-degrees. Black bars indicate the values of year 2016 and the grey bars, year 2017

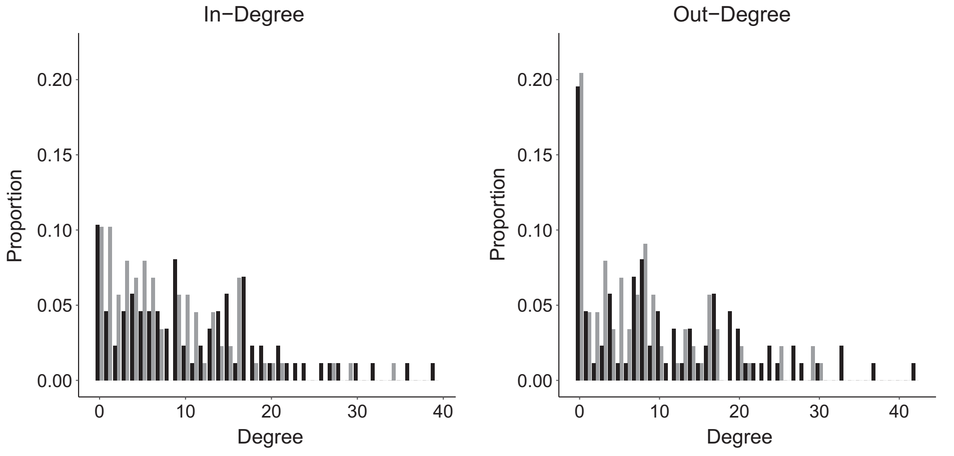

FIGURE A2 Email exchange network: Bar plots indicating the distribution of the in- and out-degrees. Black bars indicate the values of Period 1 and the grey bars, Period 2

for the major share of trade relations. The distribution of the in-degree is not that extreme, but still, we have roughly one third of all countries not importing at all.

The email exchange network shows a different structure. Here, many medium-sized in-degrees can be found, and only roughly 10% of all nodes have received no emails. For the out-degree, this number doubles (roughly 20% have not sent emails). Furthermore, the distribution of the out-degree is more skewed than the one for the in-degree.