Interaction-Partitioned Topic Model (IPTM):
A Network Model for Dynamic Textual Communications[1]

Bomin Kim
Department of Statistics
Pennsylvania State University

Aaron Schein
College of Information and Computer Sciences
University of Massachusetts Amherst

Bruce Desmarais
Department of Political Science
Pennsylvania State University

Hanna Wallach
Microsoft Research NYC
University of Massachusetts Amherst

Abstract

We introduce the interaction-partitioned topic model (IPTM)—a probabilistic model for who communicates with whom about what, and when. Broadly speaking, the IPTM partitions time-stamped textual communications, according to both the network dynamics that they reflect and their content. To define the IPTM, we integrate a dynamic version of the exponential random graph model and latent Dirichlet allocation. The IPTM assigns each topic to an "interaction pattern"—a generative process for ties that is governed by a set of dynamic network features. Each communication is then modeled as a mixture of topics and their corresponding interaction patterns. We use the IPTM to analyze emails sent between department managers in Dare county government in North Carolina, and demonstrate that the model is effective at predicting and explaining continuous-time textual communications.

1 Introduction

In recent decades, real-time digitized textual communication has developed into a ubiquitous form of social and professional interaction (see, e.g., Kanungo and Jain, 2008; Szóstek, 2011; Burgess et al., 2004; Pew, 2016). From the perspective of the computational so-

cial scientist, this has lead to a growing need for methods of modeling interactions that manifest as text exchanged in continuous time (e.g., e-mails). A number of models that build upon topic modeling through Latent Dirichlet Allocation (Blei et al., 2003) to incorporate link data as well as textual content have been developed recently (McCallum et al., 2005; Lim et al., 2013; Krafft et al., 2012). These models are innovative in their extensions that incorporate network tie information. However, none of the models that are currently available in the literature integrate the rich random-graph structure offered by state of the art models for network structure—in particular, the exponential random graph model (ERGM) (Robins et al., 2007; Chatterjee et al., 2013; Hunter et al., 2008). The ERGM is the canonical model for network structure, as it is flexible enough to specify a generative model that accounts for nearly any pattern of tie formation (e.g., reciprocity, clustering, popularity effects) (Desmarais and Cranmer, 2017). We build upon recent extensions of ERGM that model time-stamped ties (Perry and Wolfe, 2013; Butts, 2008), and develop the interaction-partitioned topic model (IPTM) to simultaneously model the network structural patterns that govern tie formation, and the content in the communications.

ERGM, and models based on ERGM, provide a framework for explaining or predicting ties between nodes using the network sub-structures in which the two nodes are embedded (e.g., an ERGM specification may predict ties between two nodes that have many shared partners). ERGM-style models have been used for many applications in which the ties between nodes are annotated with text. The text, despite providing rich information regarding the strength, scope, and character of the ties, has been largely excluded from these analyses, due to the inability of ERGM-

style models to incorporate textual attributes of ties. These application domains include, among other applicaitons, the study of legislative networks in which networks reflect legislators' co-support of bills, but exclude bill text (Bratton and Rouse, 2011; Alemán and Calvo, 2013); the study of alliance networks in which networks reflect countries' co-signing of treaties, but exclude treaty text (Camber Warren, 2010; Cranmer et al., 2012b,a; Kinne, 2016); the study of scientific co-authorship networks that exclude the text of the co-authored papers (Kronegger et al., 2011; Liang, 2015; Fahmy and Young, 2016); and the study of text-based interaction on social media (e.g., users tied via 'mentions' on twitter) (Yoon and Park, 2014; Peng et al., 2016; Lai et al., 2017).

In defining and testing the IPTM we embed two core conceptual properties, in addition to modeling both text and network structure. First, we link the content component of the model, and network component of the model such that knowing who is communicating with whom at what time (i.e., the network component) provides information about the content of communication, and vice versa. **Second, we provide fixible parameterization in modeling the time of documents using generalized linear model approach**. In what follows we (1) present the generative process for the IPTM, describing how it meets our theoretical criteria, (2) derive the sampling equations for Bayesian inference, and (3) illustrate the IPTM through application to email corpora of internal communications by government officials in Dare County, NC.

2 Model Definition

To define the IPTM, we begin by describing a probabilistic process by which documents are generated, where documents include author, recipients, contents, and timing. We provide a fully parametric definition of each component of the generative process, which enables the model to be used to simulate distributions of who communicates with whom about what, and when.

The data generated under the IPTM consists of D unique documents. A single document, indexed by $d \in \{1, \ldots, D\}$, is represented by the four components $(a_d, \boldsymbol{r}_d, t_d, \boldsymbol{w}_d)$. The first two are the author and recipients of the document: an integer $a_d \in \{1, \ldots, A\}$ indicates the identity of the author and a binary vector $\boldsymbol{r}_d = \{u_{di}\}_{i=1}^A$, which indicates the identity of the recipients. Next, t_d is the timestamp of the document d. For simplicity, we assume that documents are ordered by time such that $t_d < t_{d+1}$ for $d = 1, \ldots, D$. Lastly, $\boldsymbol{w}_d = \{w_{dn}\}_{n=1}^{N_d}$ is a set of tokens that comprise the text of the document, where N_d denotes the total number of tokens in a document.

2.1 Content Generating Process

In this section, we illustrate how the words $\boldsymbol{w}_d$ are generated according to latent Dirichlet allocation (Blei et al., 2003). First, we generate the corpus-wide global variables that describe the content via topics.

For each topic $k = 1, \ldots, K$:

- Choose a discrete distribution over V word types

$$\boldsymbol{\phi}^{(k)} \sim \text{Dirichlet}\Big(\beta, (\frac{1}{V}, \ldots, \frac{1}{V})\Big).$$

Next, given that the number of words N_d is known, we generate each token by drawing a topic from the document-topic distribution and then drawing a word from the chosen topic.

For each document $d = 1, \ldots, D$:

- Choose a discrete distribution over K topics

$$\boldsymbol{\theta}_d \sim \text{Dirichlet}\Big(\alpha, (m_1, \ldots, m_K)\Big).$$

- For $n = 1, \ldots, N_d$:
  - Choose a topic $z_{dn} \sim \text{Multinomial}(\boldsymbol{\theta}_d)$.
  - Choose a word $w_{dn} \sim \text{Multinomial}(\phi_{z_{dn}})$.

2.2 Interaction Patterns

They key idea that combines the IPTM component modeling "what" with the component modeling "who," "whom," and "when" is that different topics are associated with different interaction patterns. Each interaction pattern is characterized by a set of dynamic network features—such as the number of messages sent from i to j in some time interval— and corresponding coefficients. We associate each topic with the interaction pattern that best describes how people interact when talking about that topic.

For each topic $k = 1, \ldots, K$:

- Assign topic k to an interaction pattern

$$l_k \sim \text{Uniform}(1, C).$$

Then, we summarize each document's content as a distribution over interaction patterns:

$$\pi_{dc} = \frac{\sum_{k:l_k=c} N_{dk}}{N_d},$$

where N_{dk} is the number of times topic k appears in the document d. In other words, for each document and interaction pattern, we compute the fraction of tokens that were generated using a topic that was assigned to that interaction pattern. We then use this to generate the tie components, which are discussed in the next section.

2.3 Tie Generating Process

The IPTM generates ties and timestamps using a continuous-time process that depends on the interaction patterns' various features and corresponding coefficients. Conditioned on the content generated from Section 2.1, we assume the following four steps of tie generating process for each document d in a corpus of D documents.

2.3.1 Stochastic Intensity

We start by computing a stochastic intensity for every possible author–recipient pair by combining information about content and network structures.

For $i = 1, \ldots, A$, $j = 1, \ldots, A$ $(i \neq j)$ and $c = 1, \ldots, C$, we define the interaction-pattern-specific intensity:

$$\nu_{idjc} = \exp\Big(\boldsymbol{b}_c^\top \boldsymbol{x}_{idjc}\Big),$$

where $\boldsymbol{b}_c$ is the interaction-pattern-specific coefficients with prior $\boldsymbol{b}_c \sim N(\boldsymbol{\mu}_b, \Sigma_b)$, and $\boldsymbol{x}_{idjc}$ is the interaction patterns' dynamic network features which vary depending on the hypotheses regarding canonical processes relevant to network theory such as popularity, reciprocity, and transitivity.

We then compute the weighted average of $\{\nu_{idjc}\}_{c=1}^C$ and obtain the stochastic intensity—the likelihood of document d being sent from i to j— using the the document's distribution over interaction patterns as mixture weights:

$$\lambda_{idj} = \sum_{c=1}^C \pi_{dc}\, \nu_{idjc}.$$

2.3.2 Hypothetical Recipients

Next, we generate a set of latent recpients for each possible author. In other words, we hypothesize "If i were the author of document d, who would its recipients be?" To do this, we draw each author's set of recipients from a non-empty Gibbs measure (Fellows and Handcock, 2017), which is a probability measure we defined in order to prevent from obtaining zero recipient as well as intractible normalizing constants. Derivation of the normalizing constant is given in Appendix A.

For each author $i = 1, \ldots, A$:

- Choose a binary vector $\boldsymbol{u}_{id} = (u_{id1}, \ldots, u_{idA})$

$$\boldsymbol{u}_{id} \sim \text{Gibbs}(\delta, \boldsymbol{\lambda}_{id}),$$

where δ is a real-valued parameter with prior $\delta \sim N(\mu_\delta, \sigma_\delta^2)$ that controls the mean number of recipients,

and $\boldsymbol{\lambda}_{id} = (\lambda_{id1}, \ldots \lambda_{idA})$ are the stochoastic intensities for the author i. In particular, the log-likelihood of $\text{Gibbs}(\delta, \boldsymbol{\lambda}_{id})$ is defined as

$$l_{gibbs}(\boldsymbol{u}_{id}) \propto \log\Big(\mathrm{I}(\|\boldsymbol{u}_{id}\|_1 > 0)\Big) + \sum_{j \neq i} \Big(\delta + \log(\lambda_{idj})\Big) u_{idj},$$

where $\|\boldsymbol{u}_{id}\|_1$ is the l_1-norm of the binary vector $\boldsymbol{u}_{id}$.

2.3.3 Hypothetical Timestamps

We then generate a hypothetical timestamp for each author–recipient-set pair by saying, "If i were the author of document d and its recipients were $\boldsymbol{u}_{id}$, when would it be sent?"

Similar to Section 2.3.1, for each author $i = 1, \ldots, A$, we define the interaction-pattern-specific rate:

$$\xi_{idc} = \exp\Big(\boldsymbol{\eta}_c^\top \text{GeomMean}(\{\boldsymbol{y}_{idjc}\}_{j:u_{idj}=1})\Big),$$

where $\boldsymbol{\eta}_c$ is the interaction-pattern-specific coefficients with prior $\boldsymbol{\eta}_c \sim N(\boldsymbol{\mu}_\eta, \Sigma_\eta)$, and $\boldsymbol{y}_{idjc}$ is the interaction patterns' time-related linear predictors, which are not limited to dynamic network features. We take the geometric mean of the predictors over the chosen recipients in order to compute the averaged vector of predictors.

We then calculate the expected value of timestamp as

$$\mu_{id} = \sum_{c=1}^C \pi_{dc}\xi_{idc},$$

which is again the weighted average of $\{\xi_{idc}\}_{c=1}^C$ that combines information about content (via $\{\pi_{dc}\}_{c=1}^C$) and the linear predictors.

In modeling the timestamps, we do not assume specific distribution; instead, we provide huge flexibility and follow the generalized linear model approach as below.

$$E(\tau_{id}) = \mu_{id},$$
$$V(\tau_{id}) = V(\mu_{id}),$$

where τ_{id} is assumed to be generated from a particular distribution in the exponential family with positive support (i.e. $\tau_{id} \in (0, \infty)$) with the mean μ_{id}. Possible choice of distributions include Exponential, Gamma, lognormal, and loglogistic distributions, which are commonly used in time-to-event modeling.

2.3.4 Actual Data

Finally, we choose the document's actual author, recipients, and timestamp by selecting the author–recipient-set pair with the earliest timestamp:

$$a_d = \operatorname{argmin}_i(\tau_{id}),$$
$$\boldsymbol{r}_d = \boldsymbol{u}_{a_d d},$$
$$t_d = t_{d-1} + t_{a_d d}.$$

Therefore, it is an author-driven process in that the author of a document determines its recipients and its timestamp, based on the author's urgency to send the document to chosen recipients.

3 Inference

In this section we consider model inference and parameter estimation based on the IPTM. We take a Bayesian approach to infer the latent variables including the topics, the topic–interaction pattern assignments, and the interaction patterns' coefficients, given the content, authors, recipients, and timestamps of the observed real data. Our inference goal is to draw samples from the joint posterior distribution

$$p(\boldsymbol{z}, \boldsymbol{l}, \boldsymbol{b}, \boldsymbol{\eta}, \boldsymbol{u}, \delta | \boldsymbol{w}, \boldsymbol{a}, \boldsymbol{r}, \boldsymbol{\tau}, \alpha, \beta, \boldsymbol{m}, \boldsymbol{\mu}_b, \Sigma_b, \boldsymbol{\mu}_\eta \Sigma_\eta, \mu_\delta, \sigma_\delta^2)$$
$$\propto p(\boldsymbol{z}|\alpha, \boldsymbol{m})p(\boldsymbol{w}|\boldsymbol{z}, \beta)p(\boldsymbol{l})p(\boldsymbol{b}|\boldsymbol{\mu}_b, \Sigma_b)p(\boldsymbol{\eta}|\boldsymbol{\mu}_\eta \Sigma_\eta)$$
$$\times p(\delta|\mu_\delta, \sigma_\delta^2)p(\boldsymbol{u}, \boldsymbol{a}, \boldsymbol{r}, \boldsymbol{\tau}|\boldsymbol{z}, \boldsymbol{l}, \boldsymbol{b}, \boldsymbol{\eta}, \delta),$$

where $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ are integrated out using Dirichlet-multinomial conjugacy (Griffiths and Steyvers, 2004). The remaining unobserved variables are sequentially sampled from their joint posterior distribution using Markov chain Monte Carlo (MCMC) methods. Note that we draw the hypothetical recipients $\boldsymbol{u}$ and impute the data by employing data augmentation schemes (Tanner and Wong, 1987). A straightforward Gibbs sampling method are applied for categorical variables $(\boldsymbol{z}, \boldsymbol{l}, \boldsymbol{u})$, while we rely on Metropolis-Hasting for the rest of latent variables that do not have exact conditional posterior distributions. We omit large part of the sampling equations for the sake of brevity, however, here we illustrate the derivation of the joint posterior distribution of tie variables for dth document, $p(\boldsymbol{u}_{\cdot d}, a_d, \boldsymbol{r}_d, t_d | \boldsymbol{z}_d, \boldsymbol{l}, \boldsymbol{b}, \boldsymbol{\eta}, \delta)$, which is the term involved in most of the sampling equations.

$$p(\boldsymbol{u}_{\cdot d}, a_d, \boldsymbol{r}_d, t_d | \boldsymbol{z}_d, \boldsymbol{l}, \boldsymbol{b}, \boldsymbol{\eta}, \delta)$$
$$= \prod_{i=1}^A p(\boldsymbol{u}_{id} | \boldsymbol{z}_d, \boldsymbol{l}, \boldsymbol{b}, \delta) \times p(\tau_{a_d d} | \boldsymbol{z}_d, \boldsymbol{l}, \boldsymbol{\eta})$$
$$\times \prod_{i \neq a_d} p(\tau_{id} > \tau_{a_d d} | \boldsymbol{z}_d, \boldsymbol{l}, \boldsymbol{\eta})$$
$$= \prod_{i=1}^A L_{gibbs}(\boldsymbol{u}_{id}; \delta, \boldsymbol{\lambda}_{id}) \times \varphi_\tau\big(\tau_{a_d d}; \mu_{a_d d}, V(\mu_{a_d d})\big)$$
$$\times \prod_{i \neq a_d} \Big(1 - \Phi_\tau\big(\tau_{a_d d}; \mu_{id}, V(\mu_{id})\big)\Big),$$

where L_{gibbs} is the likelihood of non-empty Gibbs measure in Section 2.3.2, φ_τ and Φ_τ are the probability density function (pdf) and cumulative distribution function (cdf) of the specified distribution of timestamps in Section 2.3.3, respecitvely. According to the

tie generative process, this joint distribution can be interpreted as 'probability of the observed and latent recipients come from Gibbs measure $\times$ probability of the observed timestamp generated from the specified distribution of timestamps $\times$ probability of all hypothetical timestamps greater than the observed time given the specified distribution of timestamps.' For further details, see Appendix B.

4 Applications

The IPTM is intended for any network with timestamped, text-valued ties, however, in our application of the model we focus on the analysis of email network, which is the canonical example of dynamic textual communication.

4.1 Data

We use a subset of the North Carolina county government email dataset collected by ben Aaron et al. (2017) that includes internal email corpora covering the inboxes and outboxes of managerial-level employees of North Carolina county governments. Out of over twenty counties, we chose Dare County, (1) in order to see whether and how communication networks surrounding a notable national emergency—Hurricane Sandy—differed from those surrounding other governmental functions, and (2) to limit the scope of this initial application. The Dare County email network contains around 2,210 emails, sent and received by 27 department managers over a period of 3 months in 2012.

4.2 Dynamic Network Statistics

In this Section, we introduce detailed specifications of the dynamic network statsitics in our model. We develop a suite of nine different effects to be used as the components of $\boldsymbol{x}_t^{(c)}(i, j)$, (intercept, outdegree, indegree, send, receive, 2-send, 2-receive, sibling, and cosibling). These statistics capture common network properties such as popularity, centrality, reciprocity, and transitivity. Each network statistic is calculated for each interaction pattern $c = 1, ..., C$, which means that each interaction pattern can be understood in terms of the ways that network dynamics shape tie formation within the interaction pattern. In addition to assigning interaction-pattern specific intercepts, we introduce the degree, dyadic, and triadic network statistics in this paper, and below are the detailed specifications.

We follow Perry and Wolfe (2013) and define each network feature to have potentially different effects within a number of intervals of recency in the formation of the

ties that contribute to the network feature. We partition the interval $[-\infty, t)$ into $L = 4$ sub-intervals with equal length in the log-scale, by setting $\Delta_l = $ (6 hours) $\times 4^l$ for $l = 1, ..., L-1$ such that Δ_l takes the values 24 hours (=1 day), 96 hours (=4 days), 384 hours (=16 days).

5 Experiments

In this section, we conduct a set of posterior predictive experiments using the Dare County email network and the Enron dataset, to showcase the IPTM's predictive performance as compared to alternative modeling approaches.

5.1 Tie Prediction

For a randomly chosen document $d \in \{M, M + 1, ..., D\}$, we fit the IPTM to the corpus consisting of the first $d = \{1, ..., D-1\}$ documents, then use the inferred posterior distributions to generate a distribution of predicted tie data for document D conditional on the content in document D, $\boldsymbol{w}^{(D)}$. A reasonable choice for M would be $D/2$, to assure a sufficient size training set for the first document in the test set. The variables that need to be sampled are the tie data $(i_o^{(D)}, J_o^{(D)}, t_o^{(D)})$, and we would compare the simulated ones to the observed data. We compared the IPTM's predictions to those made by some simple baseline regression models.

5.2 Topic Coherence

Topic coherence metrics Mimno et al. (2011) are often used to evaluate the semantic coherence in topic models. In order to test whether the IPTM's incorporation of network features improves the ability of modeling text, we compared the coherence of topics inferred using our model with the coherence of topics inferred using the latent dirichlet allocation (LDA). Instead of re-fit the data using standard LDA algorithms, we used the topic assignments from the IPTM with $C = 1$, which simply makes the IPTM reduced to LDA in terms of topic assignments by unlinking the text and networks. For each model, we varied the number of topics from 1 to 100 and draw five samples from the joint posterior distribution over the latent variable. We evaluated the topics resulting from each sample and averaged over the five samples, where the results are shown in Figure 1. Combined with the results in Section 5.1, this result demonstrates that the IPTM can achieve good predictive performance while producing coherent topics.
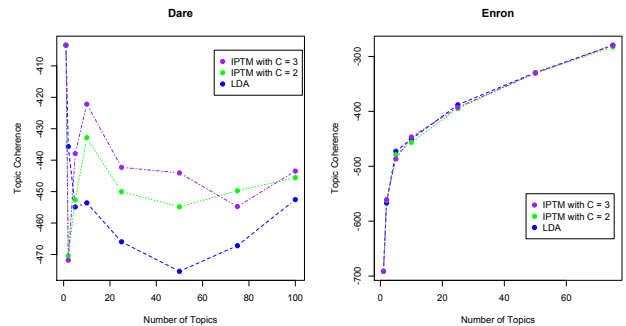

Figure 1: Average topic coherence scores: (*left*) Dare County email network. (*right*) Enron dataset.

5.3 Posterior Predictive Checks

We finally perform posterior predictive checks Rubin et al. (1984) in order to evaluate the appropriateness of the model specification for the Dare County email network. For the test of goodness-of-fit in terms of network dynamics, we defined multiple network statistics that summarize meaningful aspects of the Dare County email networks: indegree distribution for author activities, outdegree distribution for recipient activities, recipient size distribution, document time-increment distributions, the edgewise shared partner distribution, and the geodesic distance distribution. For content-wise goodness-of-fit, we employed mutual information (MI) in Mimno and Blei (2011), which is often used to evaluate "bag of words" model assumptions. We then generated 1,000 synthetic networks and texts from the posterior predictive distribution implied by the IPTM and Dare County email network. We applied each discrepancy function to each synthetic network to yield the distributions over the values of the six network statistics and MI. If the model is appropriate, the observed data should not be an outlier with respect to distributions of new data drawn from the posterior predictive distribution.

6 Conclusions

The IPTM is, to our knowledge, the first model to be capable of jointly modeling the author, recipients, timestamps and contents in time stamped text-valued networks. The IPTM incorporates innovative components, including the modeling of multicast tie formation and the conditioning of ERGM style network generative features on topic-based content. The application to North Carolina county government email data demonstrates, among other capabilities, the effectiveness at the IPTM in separating out both the content and relational structure underlying the normal day-to-day function of an organization and the man-
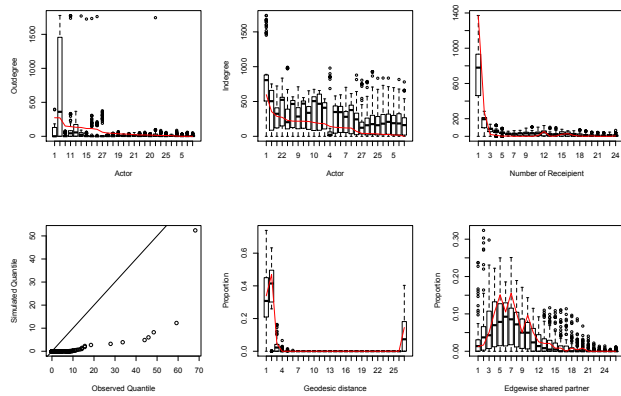
Figure 2: Posterior predictive checks for the Dare County email network: (a) outdegree, (b) indegree, (c) recipient size, (d) QQplot of time-increments, (e) geodesic distance, and (f) edgewise shared partners.

agement of a highly time-sensitive event—Hurricane Sandy. The IPTM is applicable to a variety of networks in which ties are attributed with textual documents. These include, for example, economic sanctions sent between countries and legislation attributed with sponsors and co-sponsors.

Acknowledgements

References

Alemán, E. and Calvo, E. (2013). Explaining policy ties in presidential congresses: A network analysis of bill initiation data. *Political Studies*, 61(2):356–377.

ben Aaron, J., Denny, M., Desmarais, B., and Wallach, H. (2017). Transparency by conformity: A field experiment evaluating openness in local governments. *Public Administration Review*, 77(1):68–77.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022.

Bratton, K. A. and Rouse, S. M. (2011). Networks in the legislative arena: How group dynamics affect cosponsorship. *Legislative Studies Quarterly*, 36(3):423–460.

Burgess, A., Jackson, T., and Edwards, J. (2004). Email overload: Tolerance levels of employees within the workplace. In *Innovations Through Information Technology: 2004 Information Resources Management Association International Conference, New Orleans, Louisiana, USA, May 23-26, 2004*, volume 1, page 205. IGI Global.

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

Camber Warren, T. (2010). The geometry of security: Modeling interstate alliances as evolving networks. *Journal of Peace Research*, 47(6):697–709.

Chatterjee, S., Diaconis, P., et al. (2013). Estimating and understanding exponential random graph models. *The Annals of Statistics*, 41(5):2428–2461.

Cranmer, S. J., Desmarais, B. A., and Kirkland, J. H. (2012a). Toward a network theory of alliance formation. *International Interactions*, 38(3):295–324.

Cranmer, S. J., Desmarais, B. A., and Menninga, E. J. (2012b). Complex dependencies in the alliance network. *Conflict Management and Peace Science*, 29(3):279–313.

Desmarais, B. A. and Cranmer, S. J. (2017). Statistical inference in political networks research. In Victor, J. N., Montgomery, A. H., and Lubell, M., editors, *The Oxford Handbook of Political Networks*. Oxford University Press.

Fahmy, C. and Young, J. T. (2016). Gender inequality and knowledge production in criminology and criminal justice. *Journal of Criminal Justice Education*, pages 1–21.

Fellows, I. and Handcock, M. (2017). Removing phase transitions from gibbs measures. In *Artificial Intelligence and Statistics*, pages 289–297.

Griffiths, T. L. and Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, 101(suppl 1):5228–5235.

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., and Morris, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of statistical software*, 24(3):nihpa54860.

Kanungo, S. and Jain, V. (2008). Modeling email use: a case of email system transition. *System Dynamics Review*, 24(3):299–319.

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

Krafft, P., Moore, J., Desmarais, B., and Wallach, H. M. (2012). Topic-partitioned multinetwork embeddings. In Pereira, F., Burges, C., Bottou, L., and Weinberger, K., editors, *Advances in Neural Information Processing Systems 25*, pages 2807–2815. Curran Associates, Inc.

Kronegger, L., Mali, F., Ferligoj, A., and Doreian, P. (2011). Collaboration structures in slovenian scientific communities. *Scientometrics*, 90(2):631–647.

Lai, C.-H., She, B., and Tao, C.-C. (2017). Connecting the dots: A longitudinal observation of relief organizations' representational networks on social media. *Computers in Human Behavior*, 74:224–234.

Liang, X. (2015). The changing impact of geographic distance: A preliminary analysis on the co-author networks in scientometrics (1983-2013). In *System Sciences (HICSS), 2015 48th Hawaii International Conference on*, pages 722–731. IEEE.

Lim, K. W., Chen, C., and Buntine, W. (2013). Twitter-network topic model: A full bayesian treatment for social network and text modeling. In *NIPS2013 Topic Model workshop*, pages 1–5.

McCallum, A., Corrada-Emmanuel, A., and Wang, X. (2005). The author-recipient-topic model for topic and role discovery in social networks, with application to enron and academic email. In *Workshop on Link Analysis, Counterterrorism and Security*, page 33.

Mimno, D. and Blei, D. (2011). Bayesian checking for topic models. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 227–237. Association for Computational Linguistics.

Mimno, D., Wallach, H. M., Talley, E., Leenders, M., and McCallum, A. (2011). Optimizing semantic coherence in topic models. In *Proceedings of the conference on empirical methods in natural language processing*, pages 262–272. Association for Computational Linguistics.

Peng, T.-Q., Liu, M., Wu, Y., and Liu, S. (2016). Follower-followee network, communication networks, and vote agreement of the us members of congress. *Communication Research*, 43(7):996–1024.

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

Pew, R. C. (2016). Social media fact sheet. *Accessed on 03/07/17.*

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

Rubin, D. B. et al. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *The Annals of Statistics*, 12(4):1151–1172.

Szóstek, A. M. (2011). ?dealing with my emails?: Latent user needs in email management. *Computers in Human Behavior*, 27(2):723–729.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association*, 82(398):528–540.

Yoon, H. Y. and Park, H. W. (2014). Strategies affecting twitter-based networking pattern of south korean politicians: social network analysis and exponential random graph model. *Quality & Quantity*, pages 1–15.