Supplement to
Friends with Motives: Using Text to Infer Influence on SCOTUS

Yanchuan Sim
Language Technologies Institute
Carnegie Mellon University
Pittsburgh, PA 15213, USA
ysim@cs.cmu.edu

Bryan Routledge
Tepper School of Business
Carnegie Mellon University
Pittsburgh, PA 15213, USA
routledge@cmu.edu

Noah A. Smith
Computer Science & Engineering
University of Washington
Seattle, WA 98195, USA
nasmith@cs.washington.edu

A Modeling details

Votes model. The joint likelihood of our votes model is

$$
\begin{aligned}
&\mathcal{L}^{\text{vote}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\psi}, \boldsymbol{\theta}, \boldsymbol{\Delta}, \boldsymbol{\chi}, \boldsymbol{\pi}, \boldsymbol{z}, \boldsymbol{v}, \boldsymbol{w}) \\
&\quad \propto \prod_{t=1}^{T} p(\boldsymbol{\phi}_t \mid \beta) \\
&\quad \times \prod_{j \in \mathcal{J}} p(\boldsymbol{\psi}_j \mid \sigma_J^2, \rho) p(\chi_j \mid \sigma_I^2) \\
&\quad \times \prod_{e \in \mathcal{E}} p(\pi_e \mid \sigma_P^2) \\
&\quad \times \prod_{i \in \mathcal{C}} p(a_i, b_i \mid \sigma_C^2) p(\boldsymbol{\theta}_i \mid \alpha) \prod_{k \in \mathcal{A}_i} p(\boldsymbol{\Delta}_{i,k} \mid \alpha) \\
&\quad \times \prod_{i \in \mathcal{C}} \prod_{n} p(z_{i,n}^{(m)} \mid \boldsymbol{\theta}_i) p(w_{i,n}^{(m)} \mid \phi_{z_{i,n}^{(m)}}) \\
&\quad \times \prod_{i \in \mathcal{C}} \prod_{k \in \mathcal{A}_i} \prod_{n} p(z_{i,k,n}^{(a)} \mid \boldsymbol{\theta}_i) p(w_{i,k,n}^{(a)} \mid \phi_{z_{i,k,n}^{(a)}}) \\
&\quad \times \prod_{i \in \mathcal{C}} \prod_{k \in \mathcal{A}_i} \exp(\eta^{\text{vote}}\, U_{i,k}^{\text{vote}}(\boldsymbol{\Delta}_{i,k})) \\
&\quad \times \prod_{i \in \mathcal{C}} \prod_{j \in \mathcal{J}_i} V_{i,j}
\end{aligned}
$$

where

$$V_{i,j} = p(v_{i,j} \mid a_i, b_i, \boldsymbol{\psi}_j, \boldsymbol{\theta}_i, \boldsymbol{\Delta}_i, \chi_j, \boldsymbol{\pi})$$

and

$$U_{i,k}^{\text{vote}}(\boldsymbol{\Delta}_{i,k}) = \frac{1}{|\mathcal{J}_i|} \sum_{j \in \mathcal{J}_i} V_{i,j} - H^2(\boldsymbol{\Delta}_{i,k}, \boldsymbol{\theta}_i)$$

We first initialize the topic mixtures $(\boldsymbol{\theta}, \boldsymbol{\Delta})$ and topic-word distributions $(\boldsymbol{\phi})$ of the model using latent Dirichlet allocation.[1] In each iteration, we sampled the latent variables $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\psi}, \boldsymbol{\theta}, \boldsymbol{\Delta}, \boldsymbol{\chi}$, and $\boldsymbol{\pi}$ in turn using the Metropolis-Hastings algorithm (Hastings, 1970). We discarded samples from the first 1,500 iterations (burn-in) and keeping every 10th subsequent sample to compute the posterior mean. In total, we performed 3,000 iterations over the data.

Opinions model. The joint likelihood of the opinion model is

$$
\begin{aligned}
&\mathcal{L}^{\text{opinion}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\psi}, \boldsymbol{\theta}, \boldsymbol{\Delta}, \boldsymbol{\chi}, \boldsymbol{\pi}, \boldsymbol{z}, \boldsymbol{\Gamma}, \boldsymbol{\tau}, \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{w}) \\
&\quad \propto \mathcal{L}^{\text{vote}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\psi}, \boldsymbol{\theta}, \boldsymbol{\Delta}, \boldsymbol{\chi}, \boldsymbol{\pi}, \boldsymbol{z}, \boldsymbol{v}, \boldsymbol{w}) \\
&\quad \times \prod_{j \in \mathcal{J}} p(\boldsymbol{\Gamma}_j \mid \alpha) \\
&\quad \times \prod_{i \in \mathcal{C}} \prod_{s \in \{\text{p,r}\}} p(\tau_i^s \mid \gamma^s(\boldsymbol{v}_i)) \\
&\quad \times \prod_{i \in \mathcal{C}} \prod_{s \in \{\text{p,r}\}} \prod_{n} p(x_{i,s,n} \mid \tau_i^s) \\
&\quad \times \prod_{i \in \mathcal{C}} \prod_{s \in \{\text{p,r}\}} \prod_{n} p(z_{i,s,n}^{(o)} \mid x_{i,s,n}, \Gamma, \boldsymbol{\theta}_i, \boldsymbol{\Delta}_i^s) \\
&\quad \times \prod_{i \in \mathcal{C}} \prod_{s \in \{\text{p,r}\}} \prod_{n} p(w_{i,s,n}^{(o)} \mid \phi_{z_{i,s,n}^{(o)}}) \\
&\quad \times \prod_{i \in \mathcal{C}} \prod_{k \in \mathcal{A}_i} \exp(\eta^{\text{opinion}}\, U_{i,k}^{\text{opinion}}(\boldsymbol{\Delta}_{i,k}))
\end{aligned}
$$

[1] We used the online variational Bayes algorithm (Hoffman et al., 2010) implementation found in Python `scikit-learn` (Pedregosa et al., 2011).

where

$$\gamma^s(\boldsymbol{v}_i) = \begin{bmatrix} V_{i,1} \\ \vdots \\ V_{i,|\mathcal{J}_i|} \\ 1 \\ 1 \end{bmatrix},$$

$$U^{\text{opinion}}_{i,k}(\boldsymbol{\Delta}_{i,k}) \\ = 1 - H^2(\boldsymbol{\Delta}_{i,k}, \boldsymbol{\Omega}^{s_{i,k}}) - H^2(\boldsymbol{\Delta}_{i,k}, \boldsymbol{\theta}_i),$$

and

$$\boldsymbol{\Omega}^s = \left[\boldsymbol{\Gamma}_1, \ldots, \boldsymbol{\Gamma}_{|\mathcal{J}_i|}, \boldsymbol{\theta}_i, \boldsymbol{\Delta}^s_i\right] \frac{\gamma^s(\boldsymbol{v}_i)}{\|\gamma^s(\boldsymbol{v}_i)\|_1}$$

After estimating the parameters for the vote model, we held all the parameters of the vote model fixed and estimate the opinion model parameters ($\boldsymbol{\Gamma}, \boldsymbol{\tau}, \boldsymbol{x}$) using the Metropolis-Hastings algorithm. We discarded samples from the first 2,500 iterations (burn-in) and keeping every 10th subsequent sample to compute the posterior mean. In total, we performed 5,000 iterations over the opinions model.

B Hyperparameters

Table 1 presents the hyperparameters for our final models. We experimented with $\sigma^2_I \in \{0.25, 0.5, 1, 2\}$, $\sigma^2_P \in \{0.25, 0.5, 1, 2\}$, $\eta^{\text{vote}} \in \{0.125, 0.25, 0.5, 1, 2, 4\}$, and $\eta^{\text{opinion}} \in \{0.125, 0.25, 0.5, 1, 2, 4\}$, to find the best parameters using 5-fold cross validation on accuracy and perplexity.

C Justice opinion topics

Table 2 presents the top 3 topics associated with each justice's opinions topic mixture.

References

W. K. Hastings. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.

Matthew Hoffman, Francis R. Bach, and David M. Blei. 2010. Online learning for latent Dirichlet allocation. In *Advances in Neural Information Processing Systems 23*.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830. Available at `http://scikit-learn.org/`.

Description	Symbol	Value
No. of topics	T	128
Document-topic ($\boldsymbol{\theta}, \boldsymbol{\Delta}, \boldsymbol{\Gamma}$) prior	α	$\frac{1}{T}$
Topic-word ($\boldsymbol{\phi}$) prior	β	0.001
Justice IP ($\boldsymbol{\psi}$) diagonal variance	σ_J^2	1.0
Justice influenceability (χ) scale	σ_I^2	0.5
Amicus persuasiveness (π) scale	σ_P^2	1
Case parameters (a, b) variance	σ_C^2	4
Vote model utility function weight	η^{vote}	1
Opinion model utility function weight	η^{opinion}	2

Table 1: Final hyperparameter settings for our model.

John G. Roberts
32: speech, first amendment, free speech, message, expression 61: eeoc, title vii, discrimination, woman, civil rights act 52: sec, fraud, security, investor, section ##b
Ruth B. Ginsburg
61: eeoc, title vii, discrimination, woman, civil rights act 80: class, settlement, rule ##, class action, r civ 96: taxpayer, bank, corporation, fund, irs
Antonin Scalia
94: 42 USC 1983, qualified immunity, immunity, official, section #### 57: president, senate, executive, article, framer 80: class, settlement, rule ##, class action, r civ
Elena Kagan
34: candidate, buckley, 424 US 1, contribution, fec 96: taxpayer, bank, corporation, fund, irs 105: fda, drug, manufacturer, product, federal law
Stephen Breyer
96: taxpayer, bank, corporation, fund, irs 61: eeoc, title vii, discrimination, woman, civil rights act 15: plea, trial counsel, strickland, magistrate, guilty plea
Anthony M. Kennedy
57: president, senate, executive, article, framer 94: 42 USC 1983, qualified immunity, immunity, official, section #### 15: plea, trial counsel, strickland, magistrate, guilty plea
Sonia Sotomayor
22: sentence, offense, release, guidelines, guideline 23: legislature, voter, race, 42 USC 1973, minority voter 52: sec, fraud, security, investor, section ##b
Samuel A. Alito
32: speech, first amendment, free speech, message, expression 61: eeoc, title vii, discrimination, woman, civil rights act 52: sec, fraud, security, investor, section ##b
Clarence Thomas
5: federal government, framer, commerce, commerce clause, lopez 32: speech, first amendment, free speech, message, expression 72: due process, liberty, fourteenth amendment, hearing, forfeiture

Table 2: Top 3 topics contributed to Court opinions for recently active justices ($\boldsymbol{\Gamma}$).