docs_t^m	documents at t-th time point in sequence m
$w_{t,i}^m$	i-th word in docs_t^m
$z_{t,i}^m$	topic assigned to $w_{t,i}^m$
s_t^m, $\bar{\mathbf{s}}_t^m$	state at t-th time point in sequence m
a_t^m	social status at t-th time point in sequence m
$\mathbf{w}_t^m$, $\bar{\mathbf{w}}_t^m$	all words in docs_t^m, $\mathbf{w} - \{\mathbf{w}_t^m\}$
$\mathbf{z}_t^m$, $\bar{\mathbf{z}}_t^m$	topics assigned to $\mathbf{w}_t^m$, $\mathbf{z} - \{\mathbf{z}_t^m\}$
$\mathbf{d}_t^m$	document types of docs_t^m
$N_{m,t,j}^{MTZ}$	number of words assigned topic j in docs_t^m
$N_{j,w}^{ZW}$	# words w assigned topic j
$N_{c,k}^{SD}$	# occurrences of document type d in state c
$N_{c,j}^{SZ}$	# words assigned topic j in state c
$N_{c,a,c'}^{SAS}$	# transitions from state c to c' given status a

Table 1: Description of notation.

We use Gibbs sampling for inference. Each iteration samples $z_{t,i}^m$ and s_t^m according to the following probabilities.

$$\begin{aligned}
p(z_{t,i}^m = j | \bar{\mathbf{z}}_{t,i}^m, \mathbf{w}, s_t^m) &= \frac{p(z_{t,i}^m = j, w_{t,i}^m | \bar{\mathbf{z}}_{t,i}^m, \bar{\mathbf{w}}_{t,i}^m, s_t^m)}{w_{t,i}^m | \bar{\mathbf{z}}_{t,i}^m, \bar{\mathbf{w}}_{t,i}^m, s_t^m)} \\
&\propto= p(z_{t,i}^m = j, w_{t,i}^m | \bar{\mathbf{z}}_{t,i}^m, \bar{\mathbf{w}}_{t,i}^m, s_t^m) \\
&= p(z_{t,i}^m = j | \bar{\mathbf{z}}_{t,i}^m, s_t^m) p(w_{t,i}^m | z_{t,i}^m = j, \bar{\mathbf{z}}_{t,i}^m, \bar{\mathbf{w}}_{t,i}^m) \\
&= \int_{\boldsymbol{\theta}} p(z_{t,i}^m = j | \boldsymbol{\theta}) p(\boldsymbol{\theta} | \bar{\mathbf{z}}_{t,i}^m, s_t^m) d\boldsymbol{\theta} \int_{\phi} p(w_{t,i}^m | \phi) p(\phi | \bar{\mathbf{z}}_{t,i}^m, \bar{\mathbf{w}}_{t,i}^m) d\phi \\
&= \frac{N_{m,t,j}^{MTZ} + \alpha}{\sum_{j'} \left(N_{m,t,j'}^{MTZ} + \alpha \right)} \frac{N_{j,w_{t,i}^m}^{ZW} + \beta}{\sum_{w'} \left(N_{j,w'}^{ZW} + \beta \right)} \\
&\propto \left(N_{m,t,j}^{MTZ} + \alpha \right) \frac{N_{j,w_{t,i}^m}^{ZW} + \beta}{\sum_{w'} \left(N_{j,w'}^{ZW} + \beta \right)}.
\end{aligned}$$

$$
\begin{aligned}
p(s_t^m = c|\bar{\mathbf{s}}_t^m, \mathbf{z}, \mathbf{a}, \mathbf{d}) \propto & \left(\prod_{k=1}^{D} \frac{\Gamma\left(N_{c,k}^{SD} + \nu + N_{m,t,k}^{MTD}\right)}{\Gamma\left(N_{c,k}^{SD} + \nu\right)} \frac{\Gamma\left(\sum_{k'} (N_{c,k'}^{SD} + \nu)\right)}{\Gamma\left(\sum_{k'} (N_{c,k'}^{SD} + \nu) + |\mathbf{d}_t^m|\right)} \right)^{(1)} \\
& \times \left(\prod_{j=1}^{Z} \frac{\Gamma\left(N_{c,j}^{SZ} + \alpha + N_{m,t,j}^{MTZ}\right)}{\Gamma\left(N_{c,j}^{SZ} + \alpha\right)} \frac{\Gamma\left(\sum_{j'} (N_{c,j'}^{SZ} + \alpha)\right)}{\Gamma\left(\sum_{j'} (N_{c,j'}^{SZ} + \alpha) + |\mathbf{z}_t^m|\right)} \right)^{(2)} \\
& \times \left(\frac{N_{s_{t-1}^m, a_{t-1}^m, c}^{SAS} + \gamma}{\sum_{c'} \left(N_{s_{t-1}^m, a_{t-1}^m, c'}^{SAS} + \gamma\right)} \right)^{(3)} \left(\frac{N_{c, a_t^m, s_{t+1}^m}^{SAS} + \mathbf{1}(s_{t-1}^m = c = s_{t+1}^m) + \gamma}{\sum_{c'} \left(N_{c', a_t^m, s_{t+1}^m}^{SAS} + \mathbf{1}(s_{t-1}^m = c' = s_{t+1}^m) + \gamma\right)} \right)^{(4)}.
\end{aligned}
$$

(2) can be seen in two ways. First, it is the expectation of $\mathbb{E}_{\boldsymbol{\theta} \sim \mathrm{Dir}(\boldsymbol{\alpha} + \bar{\mathbf{z}}_t^m)} [\mathbf{z}_t^m | \boldsymbol{\theta}]$, where $\mathrm{Dir}(\boldsymbol{\alpha} + \bar{\mathbf{z}}_t^m)$ is a Dirichlet distribution whose parameter is $\boldsymbol{\alpha}$ added by counter $\bar{\mathbf{z}}_t^m$. Second, instead of thinking about the probability of $\mathbf{z}_t^m$ being drawn all at once, we can think of sequential draws of each element of $\mathbf{z}_t^m$. That is, if $\mathbf{z}_t^m = [z_{t,1}^m, \cdots, z_{t,k}^m]$, the probability of sequentially drawing $\mathbf{z}_t^m$ is

$$
\frac{N_{c, z_{t,1}^m}^{SZ} + \alpha}{\sum_j' \left(N_{c,j'}^{SZ} + \alpha\right)} \frac{N_{c, z_{t,2}^m}^{SZ} + \alpha}{\sum_j' \left(N_{c,j'}^{SZ} + \alpha\right)} \cdots \frac{N_{c, z_{t,k}^m}^{SZ} + \alpha}{\sum_j' \left(N_{c,j'}^{SZ} + \alpha\right)}.
$$

This is a slight abuse of notation. N^{SZ} in the $(i+1)$th term is actually N^{SZ} in the ith term after $N_{c, z_{t,i}^m}^{SZ}$ is increased by 1. This is clear if we think about $p(\mathbf{z}_t^m | \bar{\mathbf{z}}_t^m) = p(z_{t,1}^m | \bar{\mathbf{z}}_t^m) p(z_{t,2}^m | z_{t,1}^m, \bar{\mathbf{z}}_t^m) \cdots p(z_{t,k}^m | z_{t,1}^m, \cdots, z_{t,k-1}^m, \bar{\mathbf{z}}_t^m)$. (1) can be derived similarly.

(4) can also be thought of as two consecutive draws of s_t^m and s_{t+1}^m and can be derived in the same way as above.

From the sampling results, we can estimate

$$
\begin{aligned}
\phi_{j,w} &= \frac{N_{j,w}^{ZW} + \beta}{\sum_{w'} \left(N_{j,w'}^{ZW} + \beta\right)}, \theta_{c,j} = \frac{N_{c,j}^{SZ} + \alpha}{\sum_{j'} \left(N_{c,j'}^{SZ} + \alpha\right)}, \\
\psi_{c,k} &= \frac{N_{c,k}^{SD} + \nu}{\sum k' \left(N_{c,k'}^{SD} + \nu\right)}, \theta_{t,j}^m = \frac{N_{m,t,j}^{MTZ} + \alpha}{\sum_{j'} \left(N_{m,t,j'}^{MTZ} + \alpha\right)}, \\
\pi_{c,b,c'} &= \frac{N_{c,b,c'}^{SAS} + \gamma}{\sum_{c'} \left(N_{c,a,c'}^{SAS} + \gamma\right)},
\end{aligned}
$$

where θ_t^m is the topic distribution of docs_t^m. After inference, we can infer the state of each document by the state assigned to the document during the sampling process. For unseen documents, we may infer a document's state using a Viterbi algorithm based on the document type distribution, topic distribution and state transition distribution of each state. Once states are finalized, maximum likelihood estimation can be applied to estimate the topic distribution of each document.