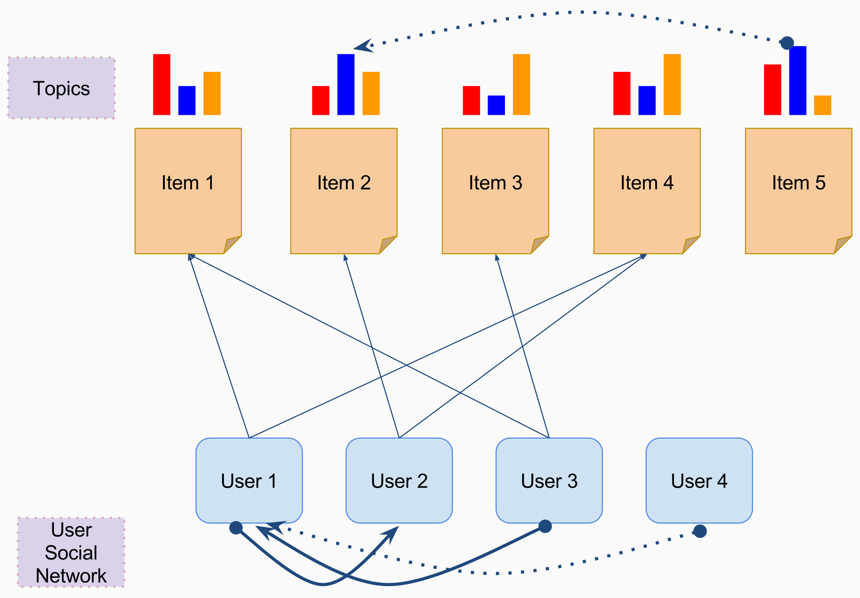
Content-Based Social Recommendation with Poisson Matrix Factorization

Eliezer de Souza da Silva (eliezer.souza.silva@ntnu.no)
Helge Langseth (helge.langseth@ntnu.no)
Heri Ramampiaro (heri@ntnu.no)
https://github.com/zehsilva/poissonmf_cs
Department of Computer Science, Norwegian University of Science and Technology

Introduction

- Classic MF recommendations models:
  - Matrix factorization on user-item explicit feedback matrix
  - Gaussian likelihood
- Poisson Factorization:
  - Principled way to model count/binary data and implicit feedback
  - Fast inference by taking advantage of sparsity of observations
  - Sparse latent representations induced by Gamma priors
- Our proposal - Poisson matrix factorization with joint user social network and item topic content side information:
  - Item topic model is also a Poisson-Gamma model, where the observations correspond to word-document count data
  - Include user social influence on the user-item rating conditional distribution



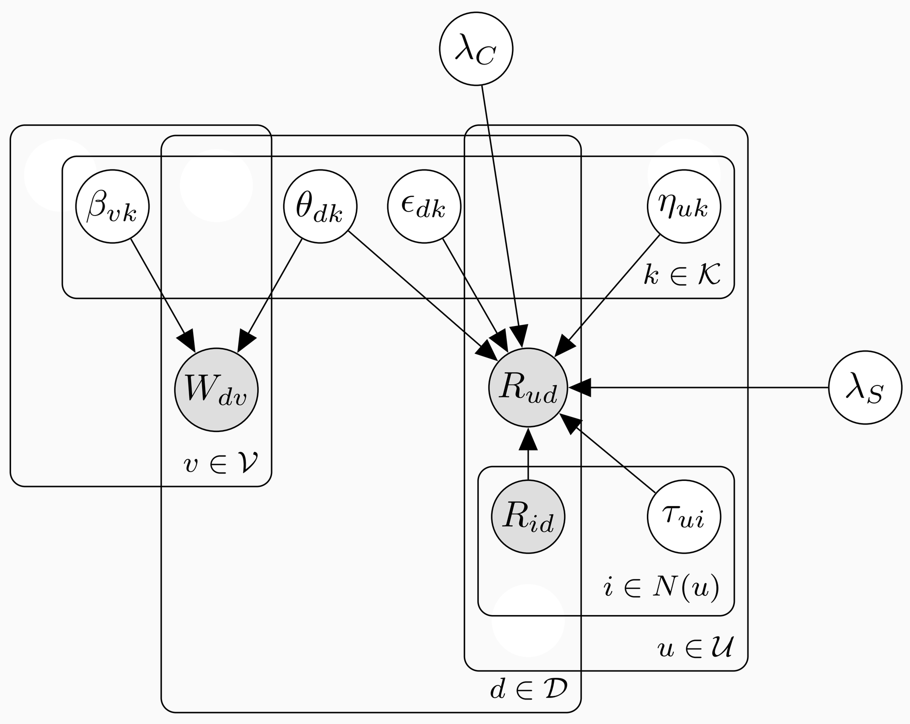
PoissonMF-CS model

- for all topics $k \in \mathcal{K}$:
  - for all words $v \in \mathcal{V}$: $\beta_{vk} \sim \text{Gamma}(a_\beta^0, b_\beta^0)$
  - for all documents $d \in \mathcal{D}$: $\theta_{dk} \sim \text{Gamma}(a_\theta^0, b_\theta^0)$ and $\epsilon_{dk} \sim \text{Gamma}(a_\epsilon^0, b_\epsilon^0)$
  - for all users $u \in \mathcal{U}$: $\eta_{uk} \sim \text{Gamma}(a_\eta^0, b_\eta^0)$
    - for all user $i \in N(u)$: $\tau_{ui} \sim \text{Gamma}(a_\tau^0, b_\tau^0)$
- Content weight: $\lambda_C \sim \text{Gamma}(a_C^0, b_C^0)$
- Social weight: $\lambda_S \sim \text{Gamma}(a_S^0, b_S^0)$
- for all observed document–word pairs dv :

$$W_{dv}|\boldsymbol{\beta}_v, \boldsymbol{\theta}_d \sim \text{Poisson}(\boldsymbol{\beta}_v^T \boldsymbol{\theta}_d)$$

- for all observed user–document pairs ud :

$$R_{ud}|\mathbf{R}_{\mathbf{N(u)},\mathbf{d}}, \boldsymbol{\eta}_\mathbf{u}, \boldsymbol{\epsilon}_\mathbf{d}, \boldsymbol{\theta}_\mathbf{d} \sim \text{Poisson}(\lambda_C \boldsymbol{\eta}_u^T \boldsymbol{\theta}_d + \boldsymbol{\eta}_u^T \boldsymbol{\epsilon}_d + \lambda_S \sum_{i \in N(u)} \tau_{ui} R_{id})$$



Full conditional distributions

Topic specific auxiliary latent counts variables (conditionally conjugate model)

$$\begin{aligned} Y_{dv,k}|\beta_{vk}, \theta_{dk} &\sim \text{Poisson}(\beta_{vk}\theta_{dk}) \\ Z^M_{ud,k}|\lambda_C, \eta_{uk}, \theta_{dk} &\sim \text{Poisson}(\lambda_C \eta_{uk}\theta_{dk}) \\ Z^N_{ud,k}|\eta_{uk}, \epsilon_{dk} &\sim \text{Poisson}(\eta_{uk}\epsilon_{dk}) \\ Z^S_{ud,i}|\lambda_S, \tau_{ui}, R_{id} &\sim \text{Poisson}(\lambda_S \tau_{ui} R_{id}) \end{aligned}$$

$$\text{with} \sum_k Y_{dv,k} = W_{dv}$$

$$\text{, and} \sum_k Z^M_{ud,k} + Z^N_{ud,k} + \sum_{i \in N(U)} Z^S_{ud,i} = R_{ud}.$$

- Gamma distribution latent factors:

$$\begin{aligned} \theta_{dk}|* &\sim \text{Gamma}(a_\theta^0 + \textstyle\sum_v Y_{dv,k} + \sum_u Z^M_{ud,k}, \\ &\qquad b_\theta^0 + \textstyle\sum_v \beta_{vk} + \lambda_C \sum_u \eta_{uk}) \\ \beta_{vk}|* &\sim \text{Gamma}(a_\beta^0 + \textstyle\sum_d Y_{dv,k}, b_\beta^0 + \sum_d \theta_{dk}) \\ \eta_{uk}|* &\sim \text{Gamma}(a_\eta^0 + \textstyle\sum_d Z^M_{ud,k} + Z^N_{ud,k}, \\ &\qquad b_\eta^0 + \lambda_C \textstyle\sum_d \theta_{dk} + \sum_d \epsilon_{dk}) \\ \epsilon_{dk}|* &\sim \text{Gamma}(a_\epsilon^0 + \textstyle\sum_u -Z^N_{ud,k}, b_\epsilon^0 + \sum_u \eta_{uk}) \\ \tau_{ui}|* &\sim \text{Gamma}(a_\tau^0 + \textstyle\sum_d Z^S_{ud,i}, b_\tau^0 + \lambda_S \sum_d R_{id}) \\ \lambda_C|* &\sim \text{Gamma}(a_C + \textstyle\sum_{u,d,k} Z^M_{ud,k}, b_C + \sum_{u,d,k} \eta_{uk}\theta_{dk} \\ \lambda_S|* &\sim \text{Gamma}(a_S + \textstyle\sum_{u,d,i} Z^S_{ud,i}, b_S + \sum_{u,d,i} \tau_{ui} R_{id}) \end{aligned}$$

- Multinomial distributed auxiliary variables:

$$\mathbf{Y}_{dv}|* \sim \text{Mult}(W_{dv}; \boldsymbol{\phi}_{\mathbf{dv}})$$

$$\text{with } \phi_{dv,k} = \frac{\beta_{vk}\theta_{dk}}{\sum_k \beta_{vk}\theta_{dk}}$$

$$\mathbf{Z}_{ud}|* \sim \text{Mult}(R_{ud}; \boldsymbol{\xi}_{\mathbf{ud}})$$

$$\text{with } \xi_{ud,k} = \begin{cases} \xi^M_{ud,k} = \frac{\lambda_C \eta_{uk}\theta_{dk}}{\sum_k \eta_{uk}(\lambda_C\theta_{dk}+\epsilon_{dk})+\lambda_S \sum_{i\in N(u)} \tau_{ui}R_{id}} \\ \xi^N_{ud,k} = \frac{\eta_{uk}\epsilon_{dk}}{\sum_k \eta_{uk}(\lambda_C\theta_{dk}+\epsilon_{dk})+\lambda_S \sum_{i\in N(u)} \tau_{ui}R_{id}} \\ \xi^S_{ud,i} = \frac{\lambda_S \tau_{ui}R_{id}}{\sum_k \eta_{uk}(\lambda_C\theta_{dk}+\epsilon_{dk})+\lambda_S \sum_{i\in N(u)} \tau_{ui}R_{id}} \end{cases}$$

Variational Inference

Evidence Lower BOund (ELBO):

$$\operatorname*{argmin}_\Psi L(\Psi) = \text{E}_q[\log p(\mathbf{R}, \mathbf{W}, \Theta) - \log q(\Theta|\Psi)]$$

$$\begin{aligned} q(\Theta|\Psi) = {} & q(\lambda_C|a_{\lambda_C}, b_{\lambda_C}) q(\lambda_S|a_{\lambda_S}, b_{\lambda_S}) \textstyle\prod_{u,k,i} q(\tau_{ui}|a_{\tau_{ui}}, b_{\tau_u} \\ & \times \textstyle\prod_{d,v,k} q(\epsilon_{dk}|a_{\epsilon_{dk}}, b_{\epsilon_{dk}}) q(\theta_{dk}|a_{\theta_{dk}}, b_{\theta_{dk}}) q(\beta_{vk}|a_{\beta_{vk}}, \\ & \times \textstyle\prod_{d,v,u} q(\mathbf{Z}_{dv}|\phi^*_{dv}) q(\mathbf{Y}_{ud}|\boldsymbol{\xi}^{M*}_{ud}, \boldsymbol{\xi}^{N*}_{ud}, \boldsymbol{\xi}^{S*}_{ud}) \end{aligned}$$

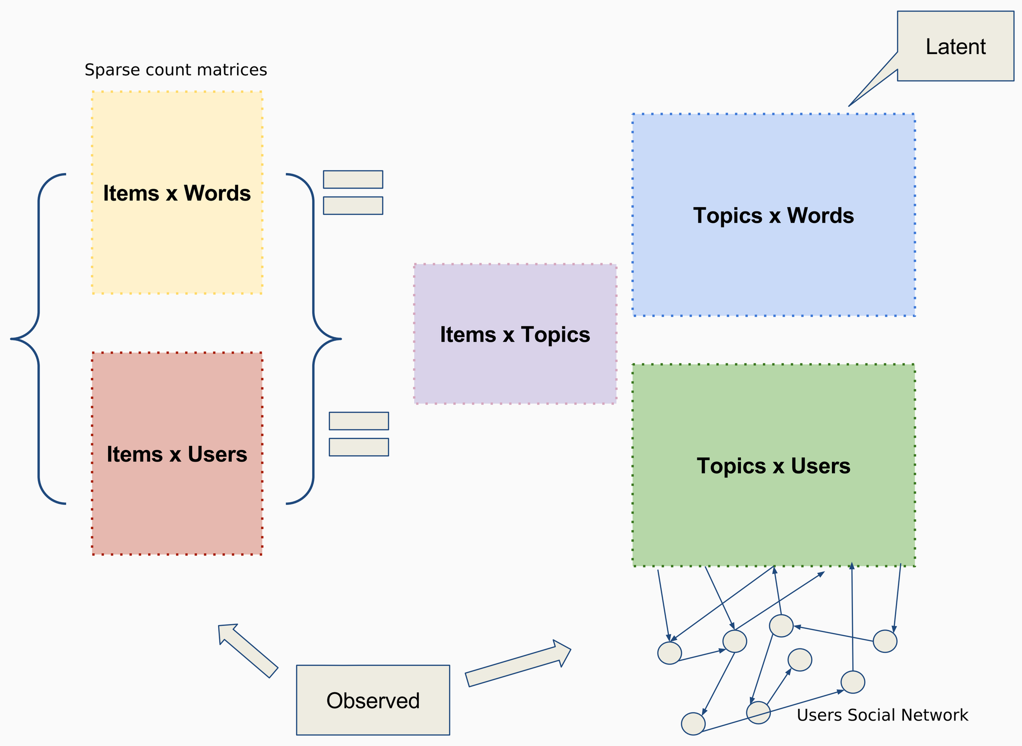
- **User u latent factor**:

$$q(\eta_{uk}) = \text{Gamma}(a_\eta^0 + \textstyle\sum_d R_{ud}(\xi^{M*}_{ud,k} + \xi^{N*}_{ud,k}), b_\eta^0 + \\ \textstyle\sum_d \text{E}_q[\lambda_C] \frac{a_{\theta_{dk}}}{b_{\theta_{dk}}} + \frac{a_{\epsilon_{dk}}}{b_{\epsilon_{dk}}})$$

Recommendations

Rank items for each users using the expected value:

$$\text{E}[R_{ud}] \approx \text{E}_q[\eta_u]^T (\text{E}_q[\lambda_C]\, \text{E}_q[\theta_d] + \text{E}_q[\epsilon_d]) \\ + \text{E}_q[\lambda_S] \sum_{i \in N(u)} \text{E}_q[\tau_{ui}] R_{id}$$



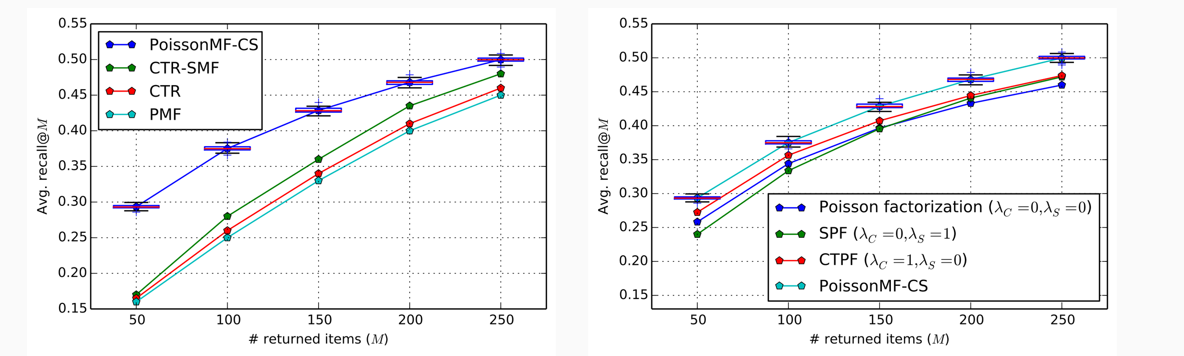
Experiments

- Task: artist recommendation using LastFM dataset consisting of: 1892 users, 17632 artists, 11946 tags, 25434 user–user connections, 92834 user–items interactions, and 186479 user–tag–items entries.
- Metric: recall at the top M items (recall@M) for a user, defined as:

$$\text{recall@}M = \frac{\text{number of items the user likes in Top } M}{\text{total number of relevant items for the user}}$$

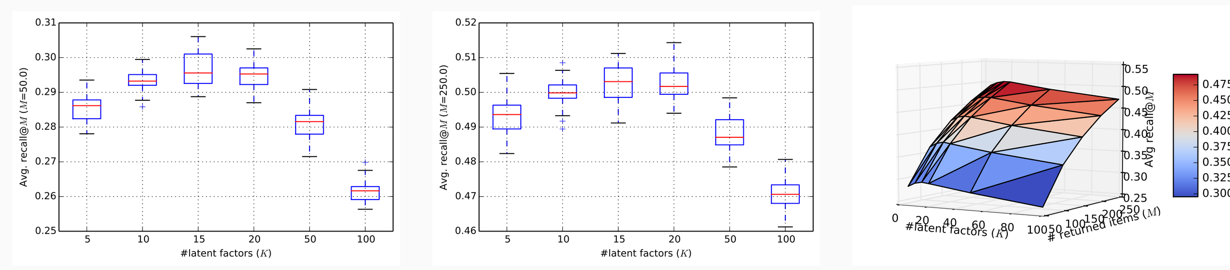
- Compare with: CTR [4], CTR-SM [3], SPF [1], CTPF [2]

Results



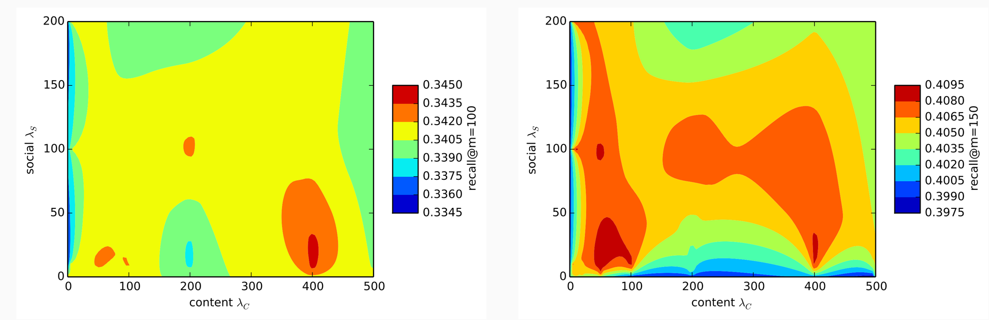
(a) PoissonMF-CS (K=10) and Gaussian based models (b) PoissonMF-CS (K=10) and other PF models

Figure : Each subplot is the result of running the PoissonMF-CS recommendation algorithm over 30 random splits of the *Hetrec2011-lastfm* dataset for a fixed number of latent features K (in this case, $K = \{10, 15\}$). The values for CTR-SMF, CTR, and PMF was taken from [3], and according to the reported results, they are the best values found after a grid search.



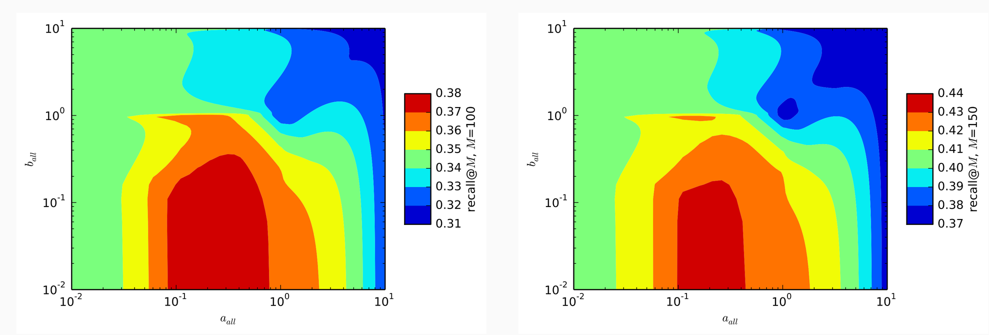
(a) M=50 (b) M=250 (c) 3D visualization

Figure : Impact of the number of latent variables (K) parameter on the Av. Recall@M metric for different number of returned items (M). Each subplot is the result of running the PoissonMF-CS recommendation algorithm over 30 random splits of the dataset with K varying in (5,10,15,20,50,100)



(a) M=100 (b) M=150

Figure : Evaluation of the impact of content and social weight parameters (in all experiments in this figure $K = 10$)



(a) M=100 (b) M=150

Figure : Evaluation of the impact of latent Gamma hyperpriors on the recall (in all experiments in this figure $K = 10$)

- A. J. Chaney, D. M. Blei, and T. Eliassi-Rad. A probabilistic model for using social networks in personalized item recommendation. In *RecSys 2015*, pages 43–50, 2015.
- P. Gopalan, L. Charlin, and D. M. Blei. Content-based recommendations with poisson factorization. In *NIPS 2014*, pages 3176–3184, 2014.
- S. Purushotham and Y. Liu. Collaborative topic regression with social matrix factorization for recommendation systems. In *ICML 2012*. icml.cc / Omnipress, 2012.
- C. Wang and D. M. Blei. Collaborative topic modeling for recommending scientific articles. In *KDD 2011*, pages 448–456, 2011.

NTNU
Norwegian University of Science and Technology