Response Aware Model-Based Collaborative Filtering

Guang Ling Haiqin Yang Michael R. Lyu Irwin King

{gling,hqyang,lyu,king}@cse.cuhk.edu.hk

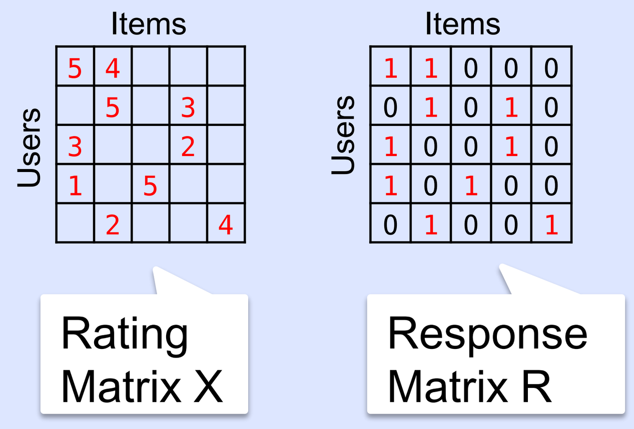
香港中文大學 The Chinese University of Hong Kong

at&t

Introduction

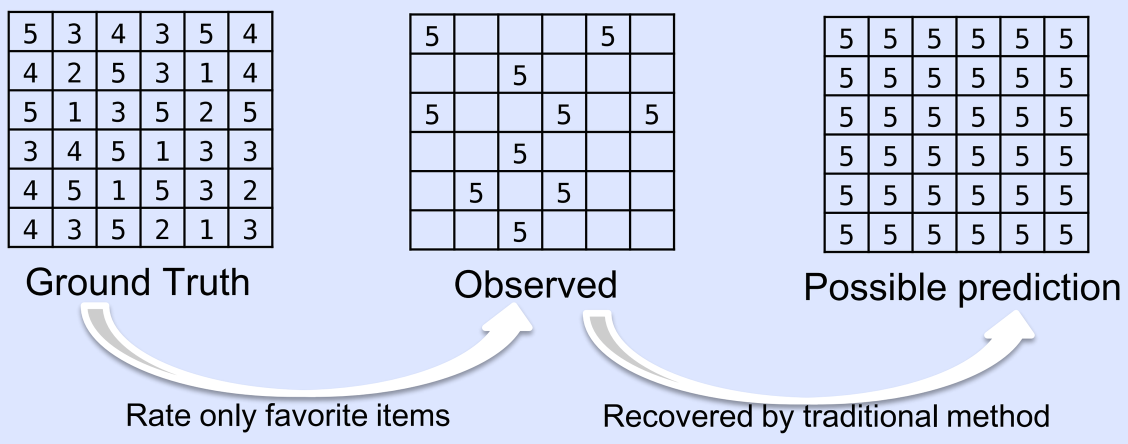
Two-fold information in CF rating data:

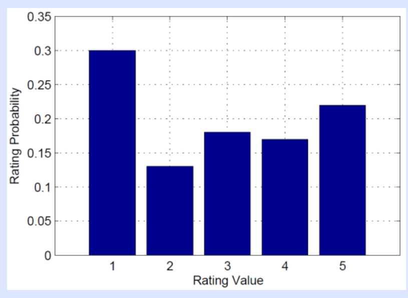
- Rating values
- Response patterns



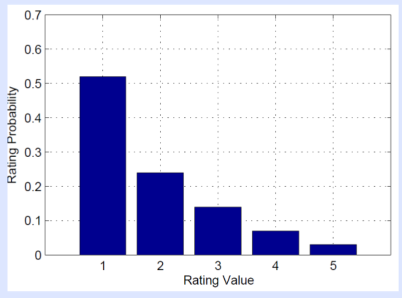
Response patterns are ignored in most previous work:

- Assumption: rate all inspected items or randomly selected items
- Is the assumption true? Unlikely
- Impact on prediction results?
  - Biased or even incorrect prediction
    - Example: assume only rate favourable items





Rating value distribution of user selected items



Rating value distribution of randomly selected items

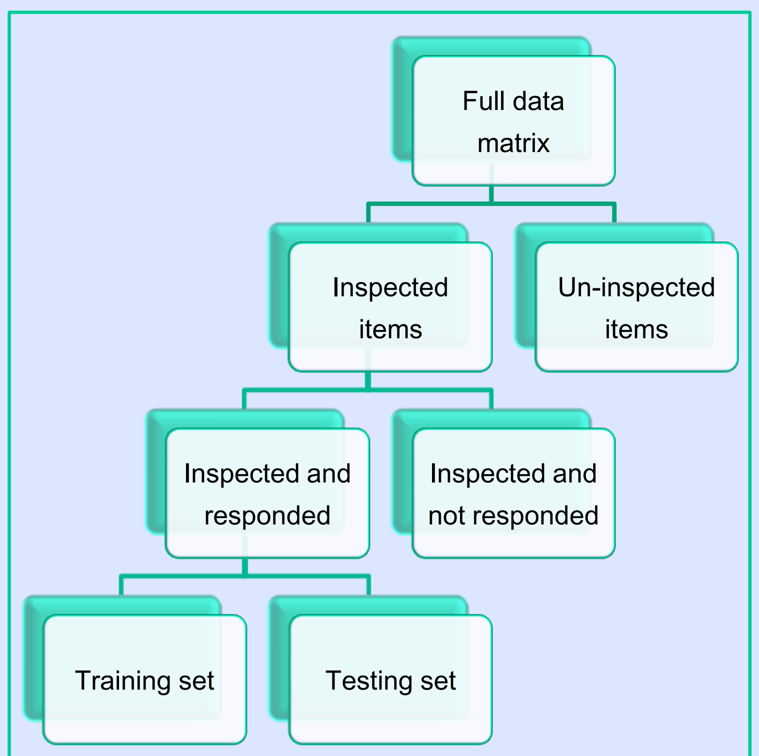
Our proposal:

- Model rating values as well as response patterns
- Response Aware Probabilistic Matrix Factorization (RAPMF)

Experiments and Results

Datasets

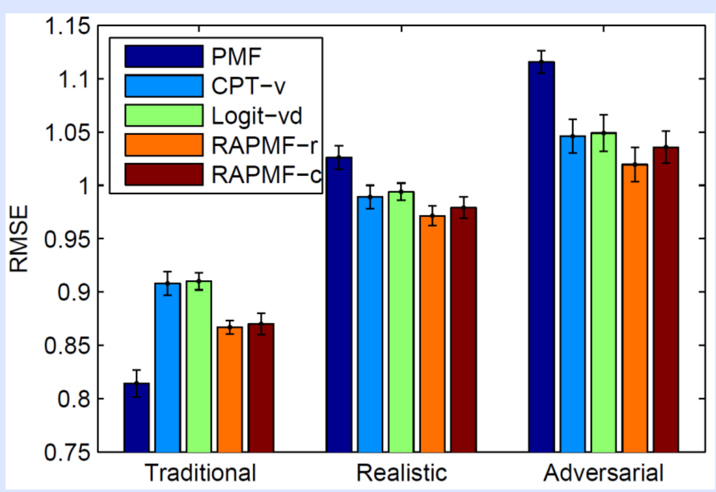
- Yahoo! music ratings for user selected and randomly selected songs
  - Normal interaction ratings (311704 ratings from 15400 users)
  - Survey ratings (54000 ratings on randomly selected songs from 5400 users)
- Synthetic dataset



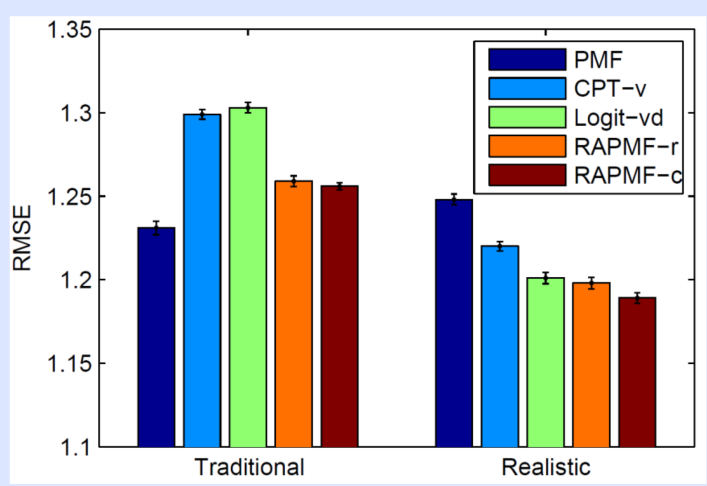
Protocols:

- Traditional
  - Train on training set
  - Test on testing set
- Realistic
  - Train on training set
  - Test on un-inspected items
- Adversarial
  - Train on training set
  - Test on Inspected and not responded

Results

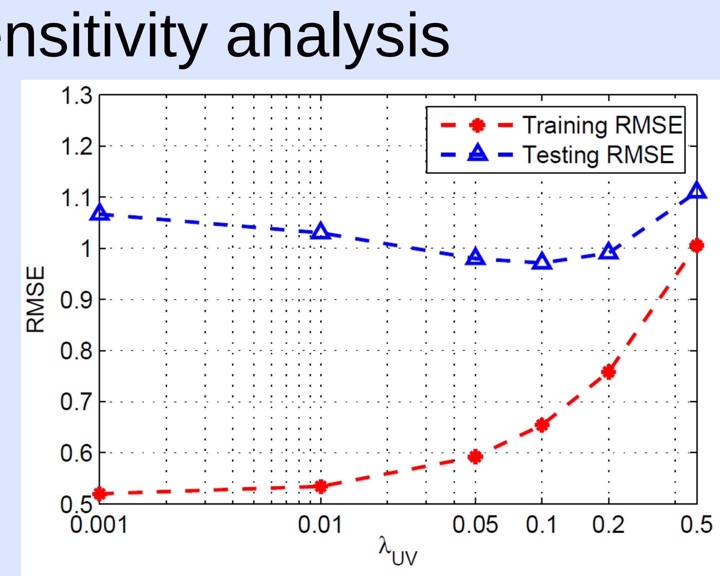


Performance on synthetic dataset

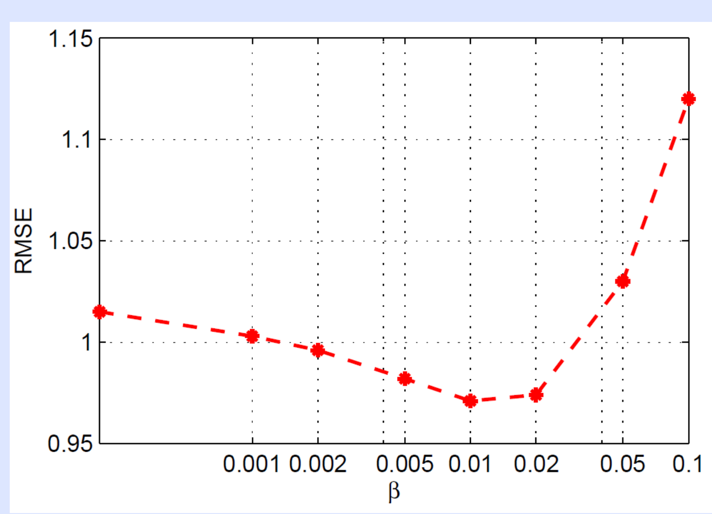


Performance on Yahoo! dataset

Sensitivity analysis



Impact of regularization parameter

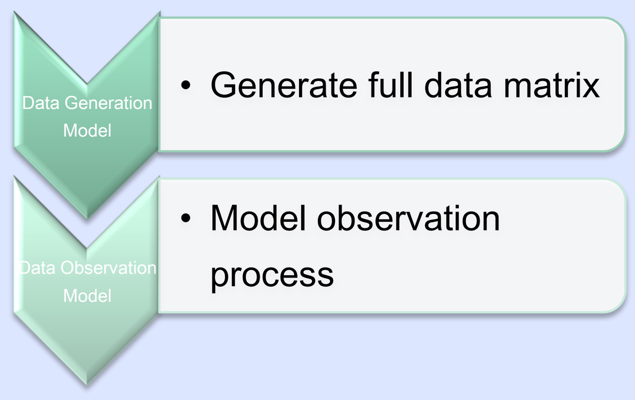


Impact of discount parameter

Model

Model the response patterns explicitly using two-step procedure:

- Data model: generate the full data
- Response model: model the response patterns



$$P(R,X|\mu,\theta) = P(R|X,\mu)P(X|\theta)$$

Response Model Data Model

Response aware probabilistic matrix factorization

- Data model: probabilistic matrix factorization (PMF)
- Response model?

$$P(R,X|U,V,\mu,\sigma^2) = P(R|X,U,V,\mu,\sigma^2)P(X|U,V,\sigma^2)$$

Rating dominant response model

- Observation: rating value affects response decision
- Model response prob. by a Bernoulli distribution $R_{ij} = \mathrm{Bernoulli}(\mu_{X_{ij}})$

$$P(R|X,U,V,\mu,\sigma^2) = P(R|U,V,\mu,\sigma^2)$$
$$= \prod_{i=1}^{N}\prod_{j=1}^{M}\sum_{k=1}^{D}(\mu_k^{[r_{ij}=1]}(1-\mu_k)^{[r_{ij}=0]})P(x_{ij}=k|U,V,\sigma^2)$$

Response prob. when rating is k

N users M items D rating levels Adopt a soft assignment

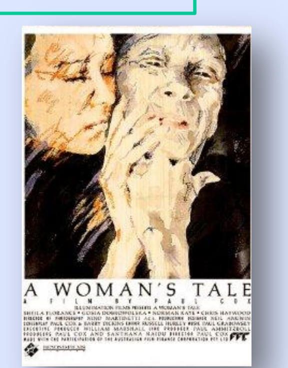
$$P(R|X,U,V,\mu,\sigma^2) = P(R|U,V,\mu,\sigma^2)$$
$$\propto \prod_{i=1}^{N}\prod_{j=1}^{M}\left(\sum_{k=1}^{D}(\mu_k^{[r_{ij}=1]}(1-\mu_k)^{[r_{ij}=0]})\right)P(x_{ij}=k|U,V,\sigma^2)^{\beta}$$

Discount parameter, control relative weight of response model

Context aware response model

- Context matters
  - Rating value
  - Heavy raters vs. Light raters
  - Hot items vs. Obscure items

Hot item



Obscure item

- Model response prob. by a Bernoulli distribution $R_{ij} = \mathrm{Bernoulli}(\mu_{ijk})$
- Adopt the same structure as rating dominant response model

$$\mu_{ijk} = \frac{1}{1+\exp(-(\delta_k + U_i^T\boldsymbol{\theta}_U + V_j^T\boldsymbol{\theta}_V))}$$

Rating value User feature Item feature

Model inference

- Alternating gradient ascent
- Let $\mathcal{L}$ be the log-likelihood of the full model $P(R,X|U,V,\mu,\sigma^2)$

$$U_i \leftarrow U_i + \eta\frac{\partial\mathcal{L}}{\partial U_i} \qquad \boldsymbol{\theta}_U \leftarrow \boldsymbol{\theta}_U + \eta\frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}_U}$$

$$\mu_k \leftarrow \mu_k + \eta\frac{\partial\mathcal{L}}{\partial\mu_k}$$

$$V_j \leftarrow V_j + \eta\frac{\partial\mathcal{L}}{\partial V_j} \qquad \boldsymbol{\theta}_V \leftarrow \boldsymbol{\theta}_V + \eta\frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}_V}$$

Summary of RAPMF

- Subsumes PMF as a special case
- Performs better on randomly selected items (a better way of evaluating a recommender system)