UNC CHARLOTTE

Personalized TV Recommendation with Mixture Probabilistic Matrix Factorization

Huayu Li[*], Hengshu Zhu[#], Yong Ge[*], Yanjie Fu[+], Yuan Ge[−]

[*]Computer Science Department, UNC Charlotte
[#]Baidu Research-Big Data Lab
[+]Rutgers University
[−]Anhui Polytechnic University

Outline

UNC CHARLOTTE

- Introduction
- Challenges of TV Recommendation
- Data
- Methods
- Experiments
- Conclusion

Introduction

UNC CHARLOTTE

Nowadays, smart TV is very prevalent…

Introduction

UNC CHARLOTTE

TV Recommender System is very important!

However, which TV program should we watch?

Outline

UNC CHARLOTTE

- Introduction
- Challenges of TV Recommendation
- Data
- Methods
- Experiments
- Conclusion

UNC CHARLOTTE

Watching Groups

TV Program

Watching group refers to users who have similar preferences for TV programs in front of a television.

Challenges of TV Recommendation

UNC CHARLOTTE

1. How to infer the preference for different watching group from such a large number of individual watching records?

2. How to handle the implicit feedbacks of users, e.g. watching time ?

Outline

UNC CHARLOTTE

- Introduction
- Challenges of TV Recommendation
- Data
- Methods
- Experiments
- Conclusion

Data

1. Each watching record includes:
   - Television ID
   - Program ID
   - Time Information

   For example :

TV ID	Program ID	Watching Duration	Start Time	Total Time
2	ba000000000018817163	740	2014-03-12T00:00:00.000Z1	800

2. Each TV program includes:
   - Title
   - Two types of genres: first level genre and sub-genre

Data

UNC CHARLOTTE

1. Each watching record includes:
   - Television ID
   - Program ID
   - Time Information

# Televisions	# TV Programs	# Watching Records
230,196	4,289	14,159,678

2. Each TV program includes:
   - Title
   - Two types of genres: first level genre and sub-genre

Outline

- Introduction
- Challenges of TV Recommendation
- Data
- Methods
- Experiments
- Conclusion

Methods

UNC CHARLOTTE

Basic Framework

Step 1: Discover Watching Groups

Step 2: Learn Preference of Television

Methods – Discovery of Watching Groups

UNC CHARLOTTE

Television Clustering
(K-means)

→

Estimating Watching Groups
(Markov Clustering)

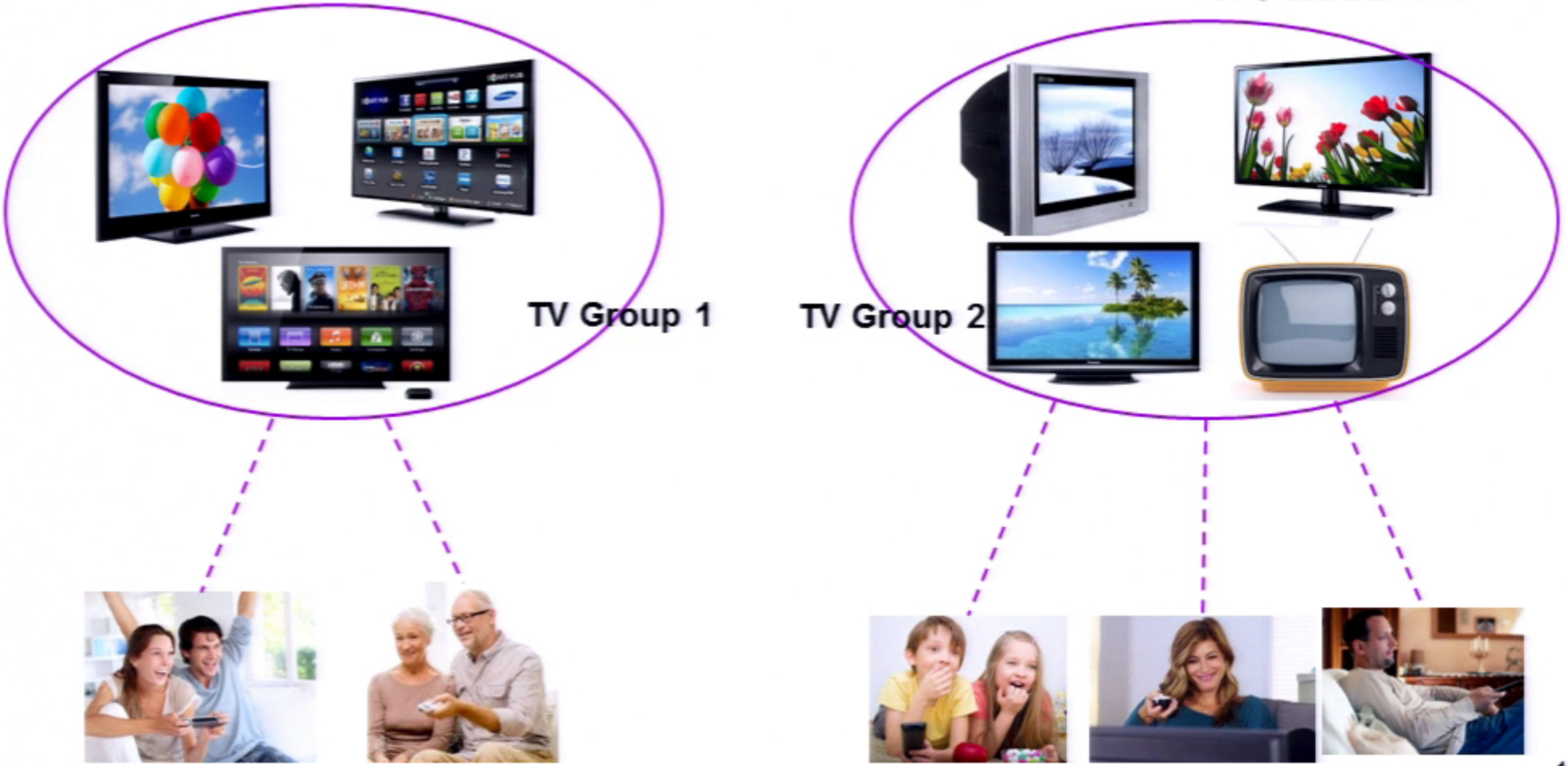
Feature:

- Watching frequency of TV program

Feature:

- First-level genre
- Sub-genre
- Watching time in a day
- Week day or weekend

Methods – Discovery of Watching Groups



Methods – Discovery of Watching Groups

Methods – Discovery of Watching Groups

UNC CHARLOTTE

Television Clustering (K-means) → **Estimating Watching Groups (Markov Clustering)**

Feature:

- Watching frequency of TV program

Feature:

- First-level genre
- Sub-genre
- Watching time in a day
- Week day or weekend

Methods – Discovery of Watching Groups

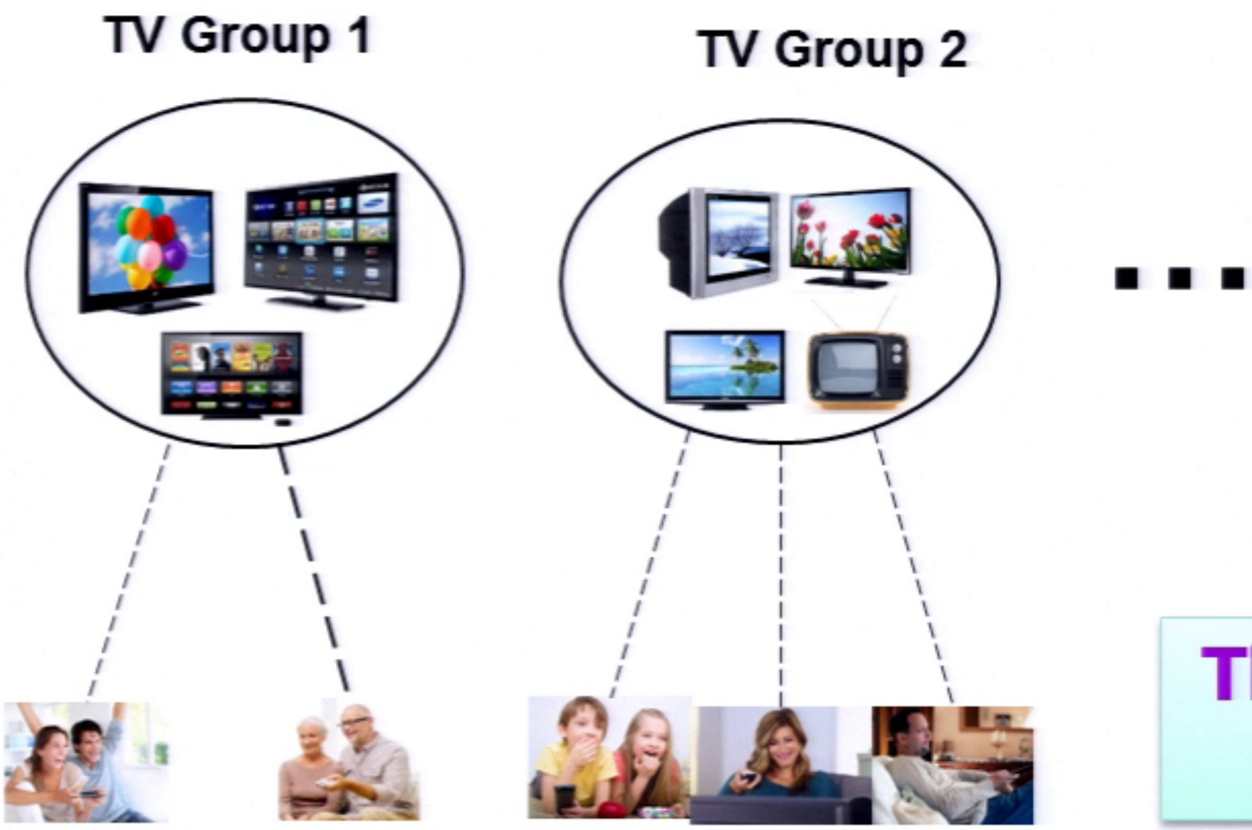
UNC CHARLOTTE

Television Clustering (K-means)	→	Estimating Watching Groups (Markov Clustering)

Feature:
- Watching frequency of TV program

Feature:
- First-level genre
- Sub-genre
- Watching time in a day
- Week day or weekend

Methods – Discovery of Watching Groups



TV Groups

The hidden watching group number

Methods – mPMF

UNC CHARLOTTE

Basic frame work

Step 1: Discover Watching Groups

Step 2: Learn Preference of Television

Mixture Probabilistic Matrix Factorization (mPMF)

Methods – mPMF

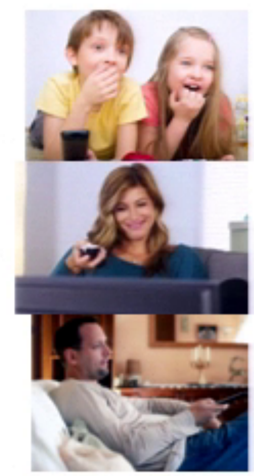
UNC CHARLOTTE

Assumption: The preferences of a television for TV programs could be decomposed into a mixture preference of the hidden watching groups.

Preference of TV

= Mixture

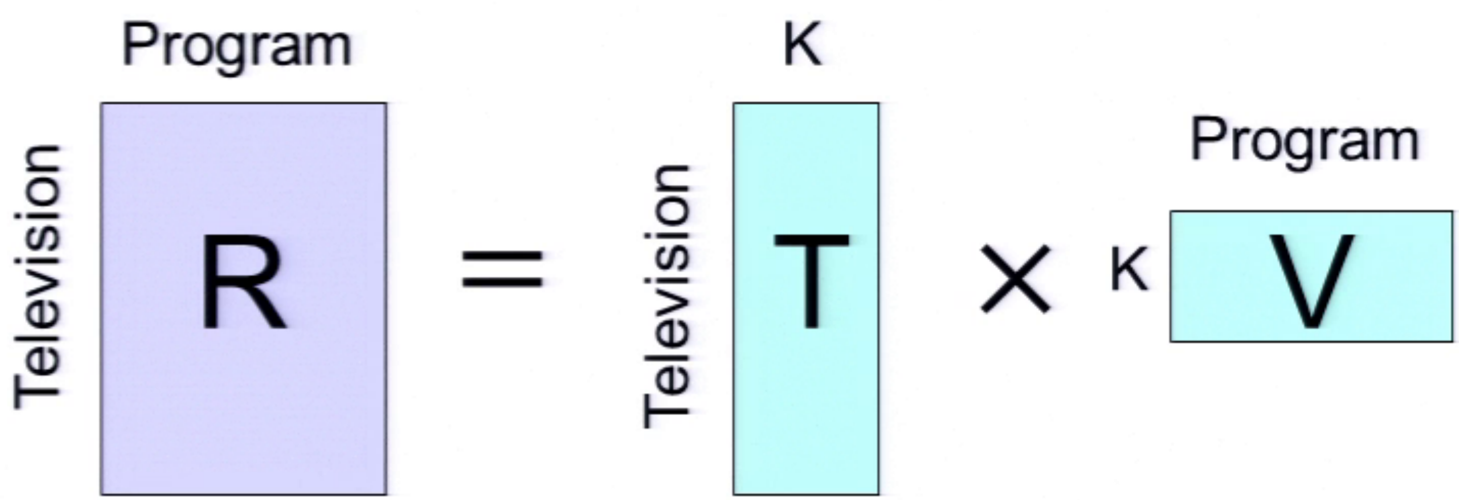
Preferences of Watching Groups



...

Methods – mPMF

Given: The learned number of watching groups for each television group



Methods – mPMF

UNC CHARLOTTE

Given: The learned number of watching groups for each television group

1. Draw television-specific latent factor from a mixture of Gaussian distribution
2. The mixture number is the number of watching groups

Methods – mPMF

UNC CHARLOTTE

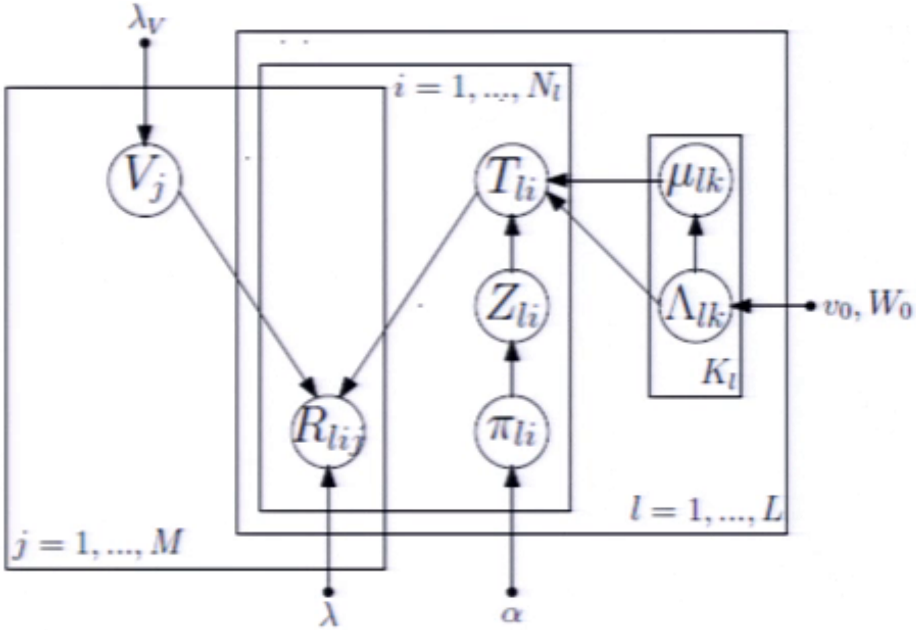


Figure 1: The Graphical Model of mPMF.

Table 1: The generative process.

1. For each program j,
   a. Draw $V_j \sim \mathcal{N}(V_j \mid 0, \lambda_V^{-1}\mathbf{I})$.
2. For each group l,
   a. For each television i,
      - Draw $\pi_{li} \sim Dirichlet(\alpha)$.
      - Pick a Gaussian $Z_{li} \sim discrete(\pi_{li})$.
      - Draw $T_{li} \sim \mathcal{N}(T_{li} \mid \mu_{lz_{li}}, \Lambda_{lz_{li}}^{-1})$.
   b. For each typical user k,
      - Draw $u_{lk} \sim \mathcal{N}(u_{lk} \mid 0, (\beta_0\Lambda_{lk})^{-1})$.
      - Draw $\Lambda_{lk} \sim \mathcal{W}(\Lambda_{lk} \mid W_0, v_0)$.
3. For each non-missing entry (l, i, j),
   a. Draw $R_{lij} \sim \mathcal{N}(R_{lij} \mid T_{li}^T V_j, \lambda^{-1})$.

Methods – mPMF

UNC CHARLOTTE

$$\mathcal{L} = -\frac{\lambda}{2}\sum_{l,i,j} I_{lij}(R_{lij} - T_{li}^T V_j)^2 - \frac{\lambda_V}{2}\sum_j V_j^T V_j$$
$$+ \sum_{l,i} \ln \sum_k \pi_{lik}\mathcal{N}(T_{li} \mid \mu_{lk}, \Lambda_{lk}^{-1}) + \sum_{l,i,k}(\alpha_k - 1)\ln \pi_{lik}$$
$$+ \frac{1}{2}\sum_{l,k}\{(v_0 - D)\ln|\Lambda_{lk}| - \beta_0\mu_{lk}\Lambda_{lk}\mu_{lk}^T + Tr(W_0^{-1}\Lambda_{lk})\}.$$

Alternating Least Square for the parameter estimation.

Outline

- Introduction
- Challenges of TV Recommendation
- Data
- Methods
- Experiments
- Conclusion

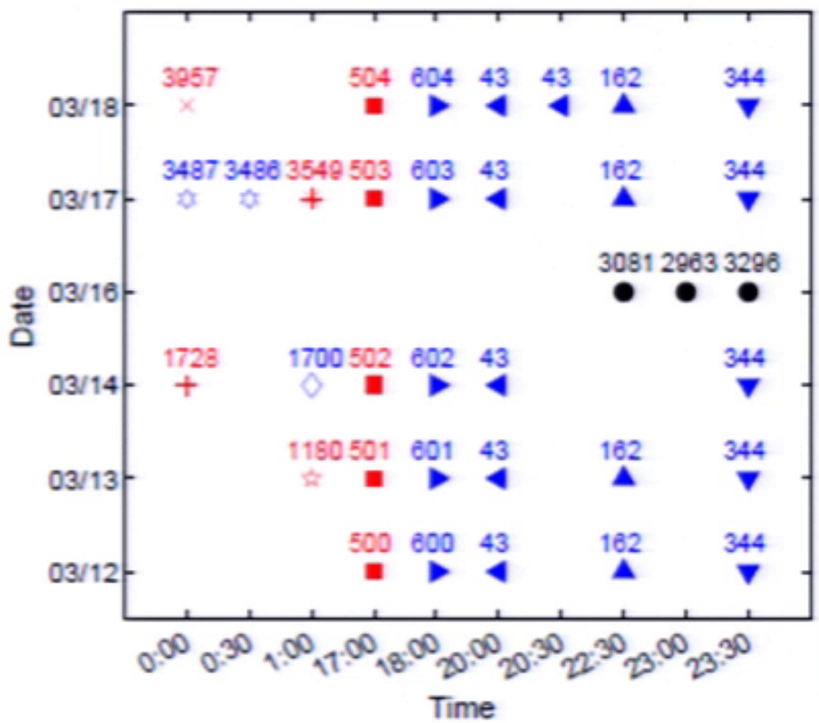
Experiments

UNC CHARLOTTE

- Show an example of clustering
- Evaluate the proposed method's performance
  - Prediction Accuracy
  - Ranking Accuracy
  - Top-K Recommendation
- Compare different data conversion methods

Experiments – An Example of clustering

UNC CHARLOTTE



ID	Name
500-504	The Chew
600-604	General Hospital
43	Jeopardy!
162	ABC World News With Diane Sawyer
344	Wheel of Fortune
3081	House Hunters International
2936	House Hunters International
3296	House Hunters

ID	Genre
500-504	Talk-Show
600-604	Drama
43	Game-Show
162	News
344	Game-Show
3081	Reality-TV
2936	Reality-TV
3296	Documentary

An example of clustering: Left is the clustering result, and Right is the corresponding program names and main genres.

Experiments – Prediction Accuracy

UNC CHARLOTTE

Rating Conversion

- Cumulative ratio of watching time to the total time of a program played

Baselines

- PMF
- mPMF
  - ✓ Random # watching group
  - ✓ # watching group as 1
  - ✓ # watching group as 3

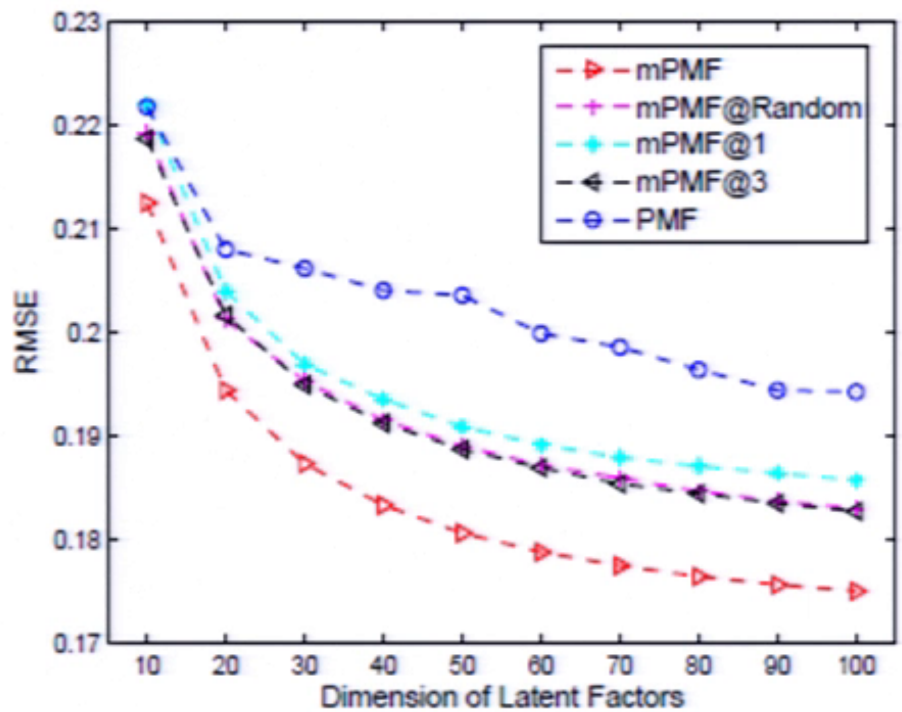


Figure 3: RMSE

Experiments – Ranking Accuracy

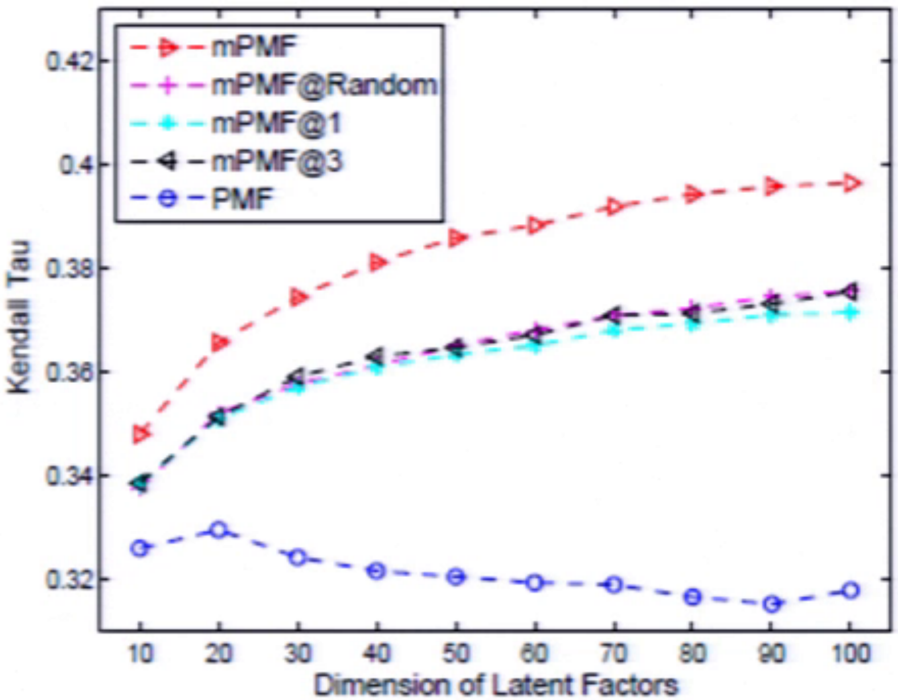


Figure 4: Kendall Tau

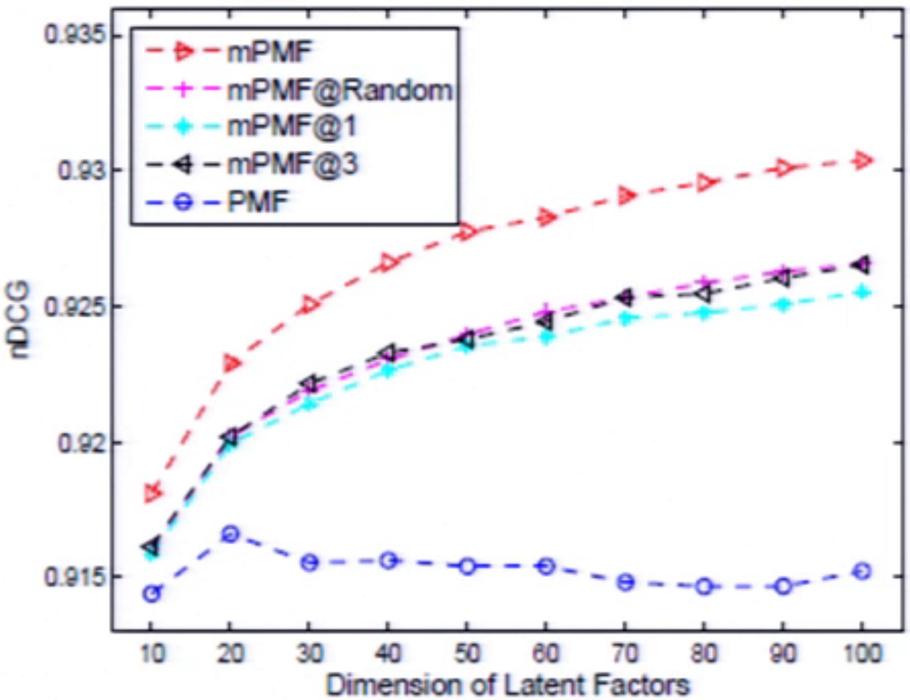


Figure 5: nDCG

Experiments – Top-K Recommendation

UNC CHARLOTTE

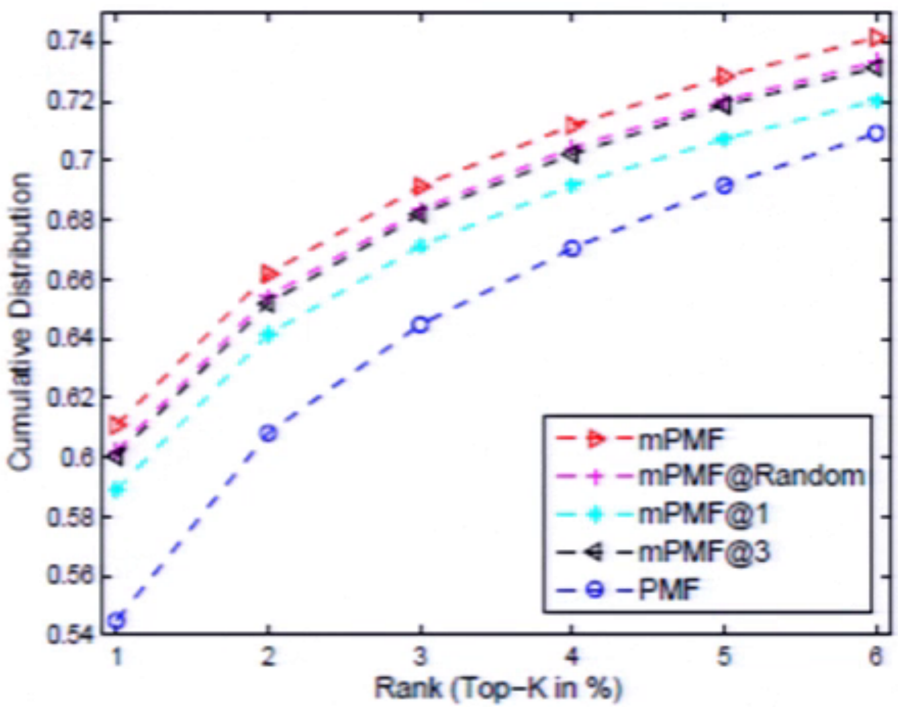


Figure 6: CD (D = 10)

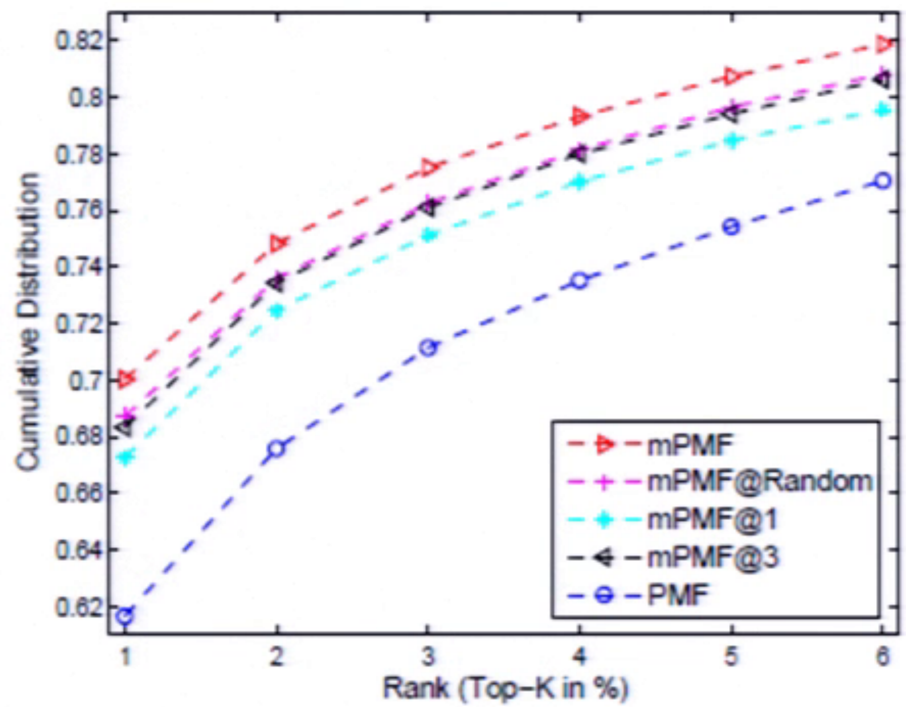


Figure 7: CD (D = 30)

Experiments – Top-K Recommendation

UNC CHARLOTTE

	Precision@5	Recall@5	Precision@10	Recall@10	MAP
10D Latent Features					
mPMF	**0.0432**	**0.0411**	0.0277	0.0527	**0.0377**
mPMF@Random	0.0430	**0.0411**	**0.0284**	**0.0540**	0.03619
mPMF@1	0.04105	0.0391	0.0259	0.0494	0.0342
mPMF@3	0.0429	0.0409	0.0281	0.0534	0.0354
PMF	0.0320	0.0304	0.0211	0.0402	0.0276
30D Latent Features					
mPMF	0.0517	0.0493	0.0322	0.0613	**0.0469**
mPMF@Random	**0.0529**	**0.0503**	**0.0326**	**0.0620**	0.0452
mPMF@1	0.0488	0.0464	0.0296	0.0564	0.0417
mPMF@3	0.0516	0.0491	0.0318	0.0606	0.0439
PMF	0.0392	0.0373	0.0242	0.0461	0.0359
60D Latent Features					
mPMF	**0.0584**	**0.0556**	**0.0356**	**0.0679**	**0.0534**
mPMF@Random	0.0581	0.0553	0.0352	0.0671	0.0497
mPMF@1	0.0534	0.0508	0.0316	0.0603	0.0457
mPMF@3	0.0562	0.0535	0.0338	0.0643	0.0479
PMF	0.0499	0.0475	0.0292	0.0556	0.0466

Table 4: Performance Comparisons(Precision, Recall, MAP).

Experiments – Data Conversion Methods

UNC CHARLOTTE

- Data Conversion Methods
  - Cumulative Ratios
  - Frequency
  - Binary
  - Confidence Level

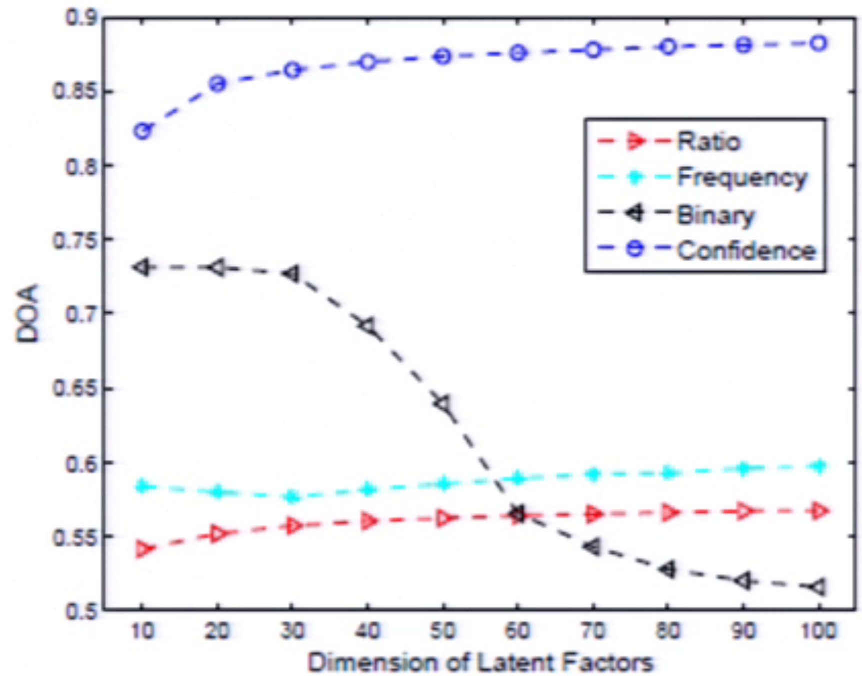


Figure 8: DOA on mPMF

Conclusion

UNC CHARLOTTE

- Design a two-stage framework
  - Employ clustering to discover the watching groups
  - Develop probabilistic model to learn the preference of television for TV program based on mixture Gaussian distributions
- Evaluate the proposed model in real-world data with various metrics