Exploiting Music Play Sequence for Music Recommendation

CSE/ECE471 | SPRING 2019

BY AYUSH KUMAR DWIVEDI, AMAN KRISHNA, GYANSHU AZAD SINGH, SHIVANI DIXIT
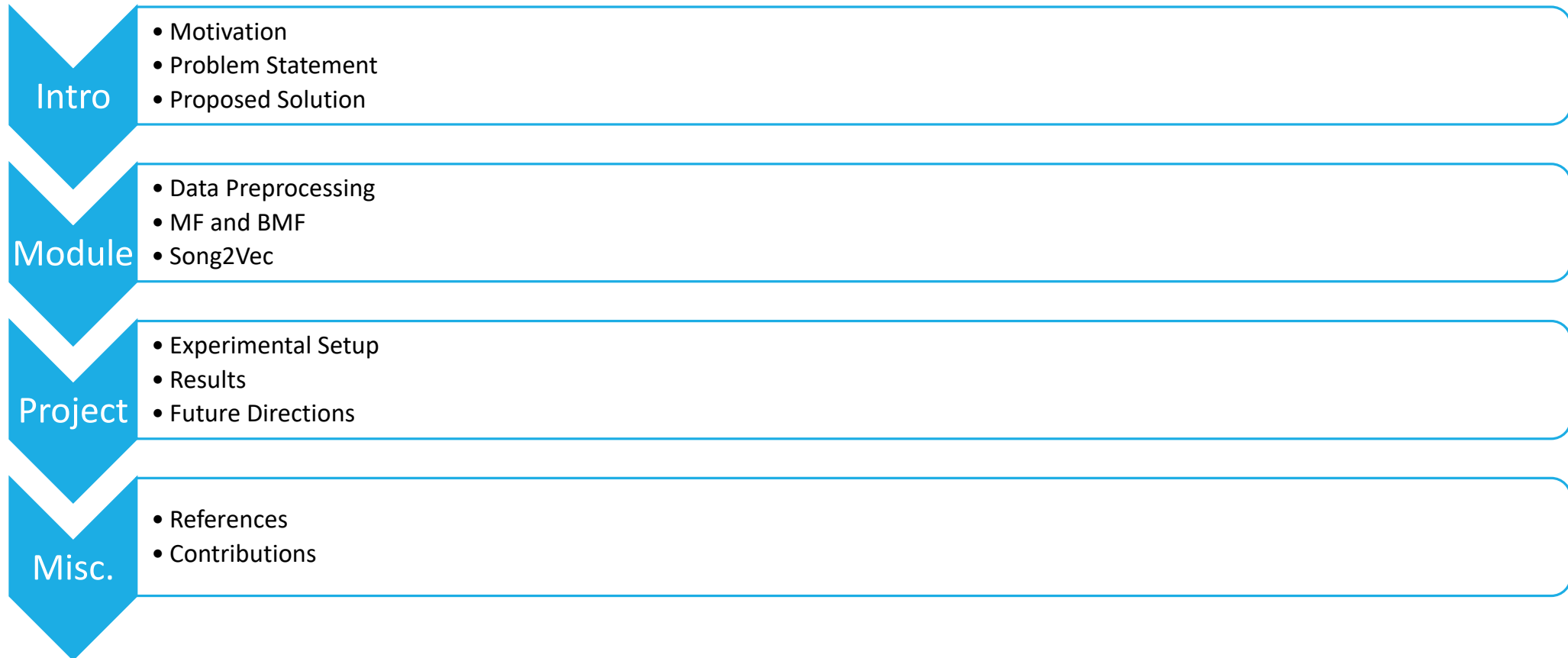
IIIT HYDERABAD

Base Paper

This project is an implementation of the following base paper:

- **Title:** Exploiting Music Play Sequence for Music Recommendation

- **Author:** Zhiyong Cheng and Jialie Shen, Lei Zhu, Mohan Kankanhalli, Liqiang Nie

- **Publication:** Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17)

Flow of presentation

Intro
- Motivation
- Problem Statement
- Proposed Solution

Module
- Data Preprocessing
- MF and BMF
- Song2Vec

Project
- Experimental Setup
- Results
- Future Directions

Misc.
- References
- Contributions

Introduction

- Users leave digital footprints when interacting with various music streaming services.

- Music play sequence (MPS), contains rich information about personal music preference and song similarity.

- MPS has been largely ignored in previously developed music recommender systems.

Motivation

- Include the much neglected Music Play Sequence (MPS) to make the song recommendation more relevant

Problem Statement:

- Incorporate the effect of Music Play Sequence (MPS) along with Matrix Factorization (MF) methods to generate more relevant music recommendations.
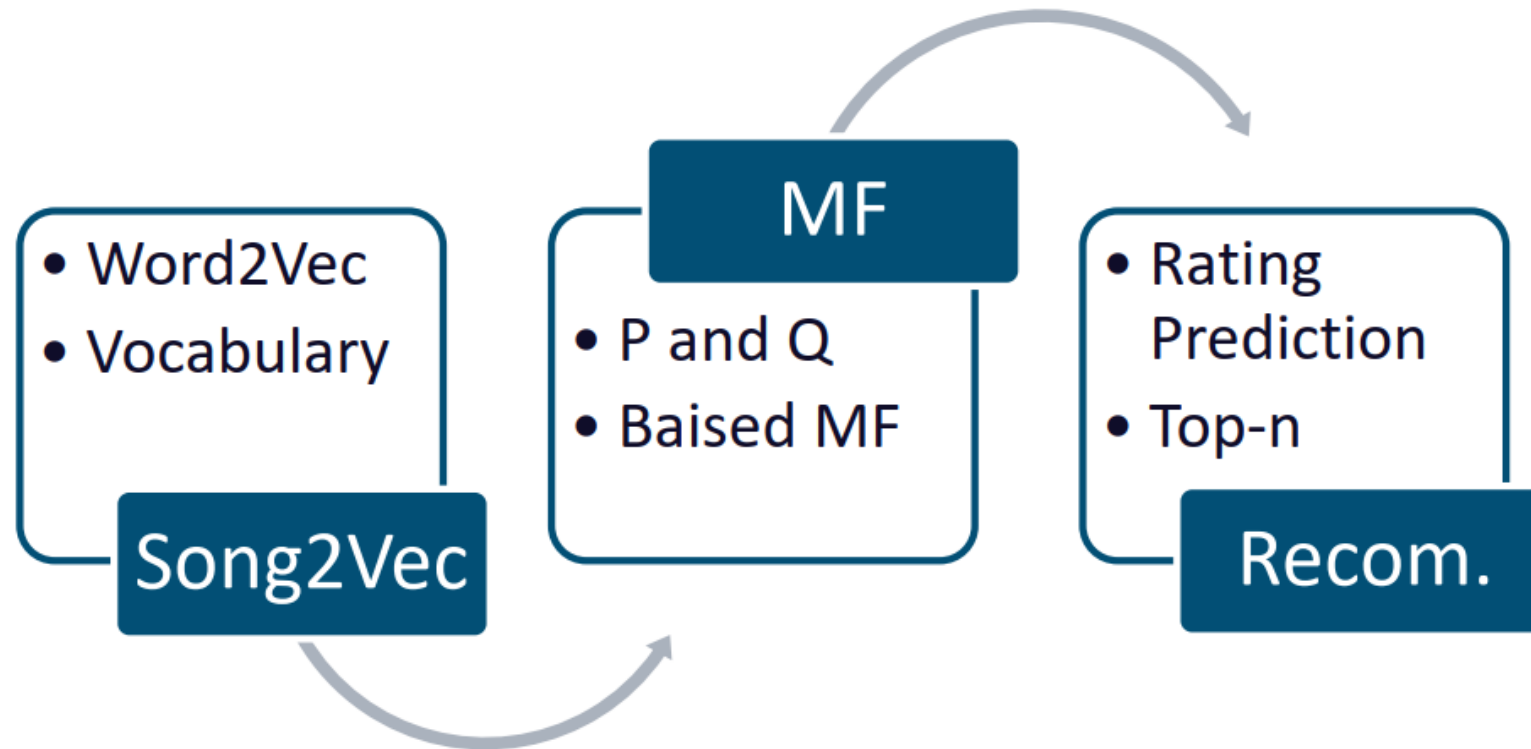
Proposed Solution:

- Use word embedding techniques in MPS to estimate the similarity between songs.

- Embed learned similarity into matrix factorization

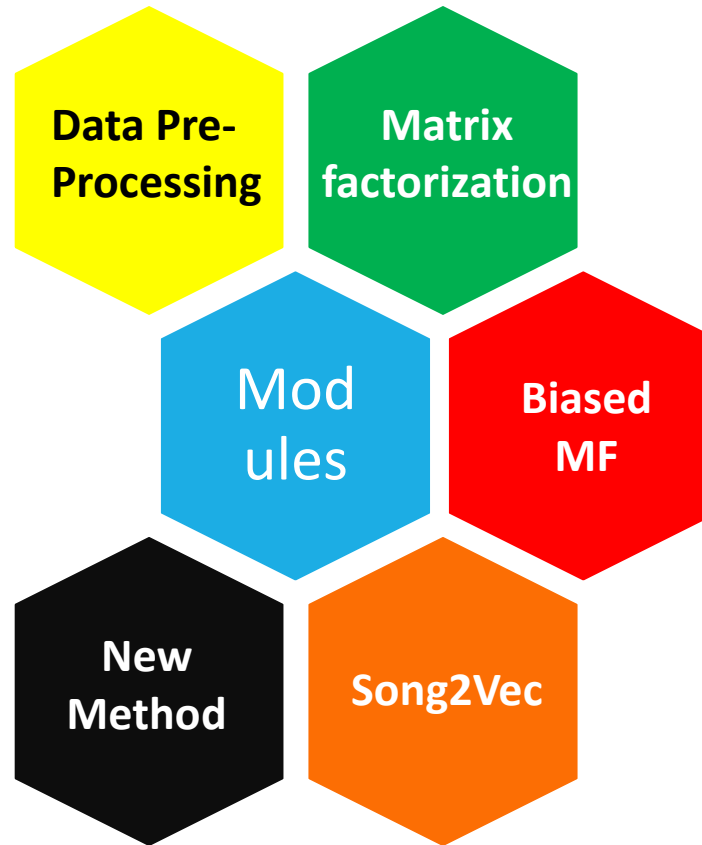- k-nearest songs in the learning process to avoid the increase in time complexity

Overview:

Modules

Data Processing

We are using Last.fm dataset, which has the following characteristics:

- No. of data points : 4752899

- No. of unique songs : 72678

- No. of unique users : 249

Snapshot of Raw Data

```
user_000001    2009-05-03T15:10:18Z    463a94f1-2713-40b1-9c88-dcc9c0170cae    Minus 8 4e78efc4-e545-47af-9617-05ff816
d86e2    Elysian Fields

user_000001    2009-05-03T15:04:31Z    ad0811ea-e213-451d-b22f-fa1a7f9e0226    Beanfield    fb51d2c4-cc69-4128-92f5
-77ec38d66859    Planetary Deadlock

user_000001    2009-05-03T14:56:25Z    309e2dfc-678e-4d09-a7a4-8eab9525b669    Dj Linus    4277434f-e3c2-41ae-9ce3
-23fd157f9347    Good Morning Love Coffee Is Ready

user_000001    2009-05-03T14:50:51Z    6f3d4a7b-45b2-4c08-9306-8d271e92cb4f    Alif Tree    1151b040-8022-4965-96d2
-8a4605ce456c    Deadly Species

user_000001    2009-05-03T14:46:29Z    463a94f1-2713-40b1-9c88-dcc9c0170cae    Minus 8 f78c95a8-9256-4757-9a9f-213df5c
6854e    Cold Fusion

user_000001    2009-05-03T14:39:20Z    45bdb5be-ec03-484f-b58d-d22afc944b24    Wei-Chi c4fc8802-d186-4c4d-85cd-d5d063b
935c2    Clouds
```

Data Processing

The following are the preprocessing steps used before using the data to train the model.

1. Exclude the users only listened to less than 10 songs
2. Exclude the songs which have been played by less than 10 users.
3. Exclude the irrelevant data columns i.e. song id, artist id, artist name

	UserID	TimeStamp	Song
0	1	2009-05-04T23:08:57Z	Fuck Me Im Famous (Pacha Ibiza)-09-28-2007
1	1	2009-05-04T13:54:10Z	Composition 0919 (Live_2009_4_15)
2	1	2009-05-04T13:52:04Z	Mc2 (Live_2009_4_15)
3	1	2009-05-04T13:42:52Z	Hibari (Live_2009_4_15)
4	1	2009-05-04T13:42:11Z	Mc1 (Live_2009_4_15)

Music Play Sequence (MPS)

- Play event to be part of a session if it occurs no later than 800 seconds

- We only keep the listening sessions with no less than 10 songs.

- Number of sessions 125181 (with each session having 10 or more songs)

```
{1: [['Lust', 'The Essence', 'Idioteque', 'Change Of Seasons',
idal Reprise', 'Landing', 'Detchibe', 'Watching Windows', 'Rid
less', 'Id', 'Zazen Bo', 'What You Gonna Do?', 'Rusty Gears Lo
'], ['Ozma', 'Hint Oyaji', 'Cow', 'Cow', 'Extra Ignored', 'Hib
n The Forest (Interlude Mix)', 'Waltz For Jason (Full Nine Yar
', 'Gum', 'Clap & Whistle & Walking', 'The Star Spangle-Gayo']
Tune', 'More Than Ever People', 'Appreciation (Radio Mix)', 'P
re Is The Line', 'Vökuró', 'Öll Birtan', 'Who Is It', 'Submari
ss', 'Where Is The Line', 'Vökuró', 'Öll Birtan', 'Who Is It',
e Down', 'Surrender', 'Misunderstanding', '花狂い', 'Element Wa
de Mix)', 'Waltz For Jason (Full Nine Yards Re-Edit) (2 Banks
gle-Gayo', 'Music', 'Gum', 'Clap & Whistle & Walking', 'The St
```

Results

- Top 6 recommendation for the song "The Fox"

```
get_song("The Fox")
```

```
Top 6 Songs for: The Fox

Name                        Similarity Score

Start Toge                  0.8167651891708374
Milkshake                   0.796122312545776
Let'S Call                  0.79258641242981
Leave You                   0.792604029178619
The End Of                  0.786170005798339
One Song F                  0.785908639430998
```

Results

- Top 6 recommendation for the song "Bedragaren I Murmansk"

```
get_song("Bedragaren I Murmansk")

Top 6 Songs for: Bedragaren I Murmansk

Name                            Similarity Score

Ännu Mera                       0.9780382513999939
Bilder Av                       0.977543473243713
M Som I                         0.9768289923667908
Sån                     0.9757107496261597
Hur Många                       0.9739065170288086
Trösta Mig                      0.9728430509567261
```

Results

- Top 6 recommendation for the song "A little Bit Of Pain"

```
get_song("A Little Bit Of Pain")
```
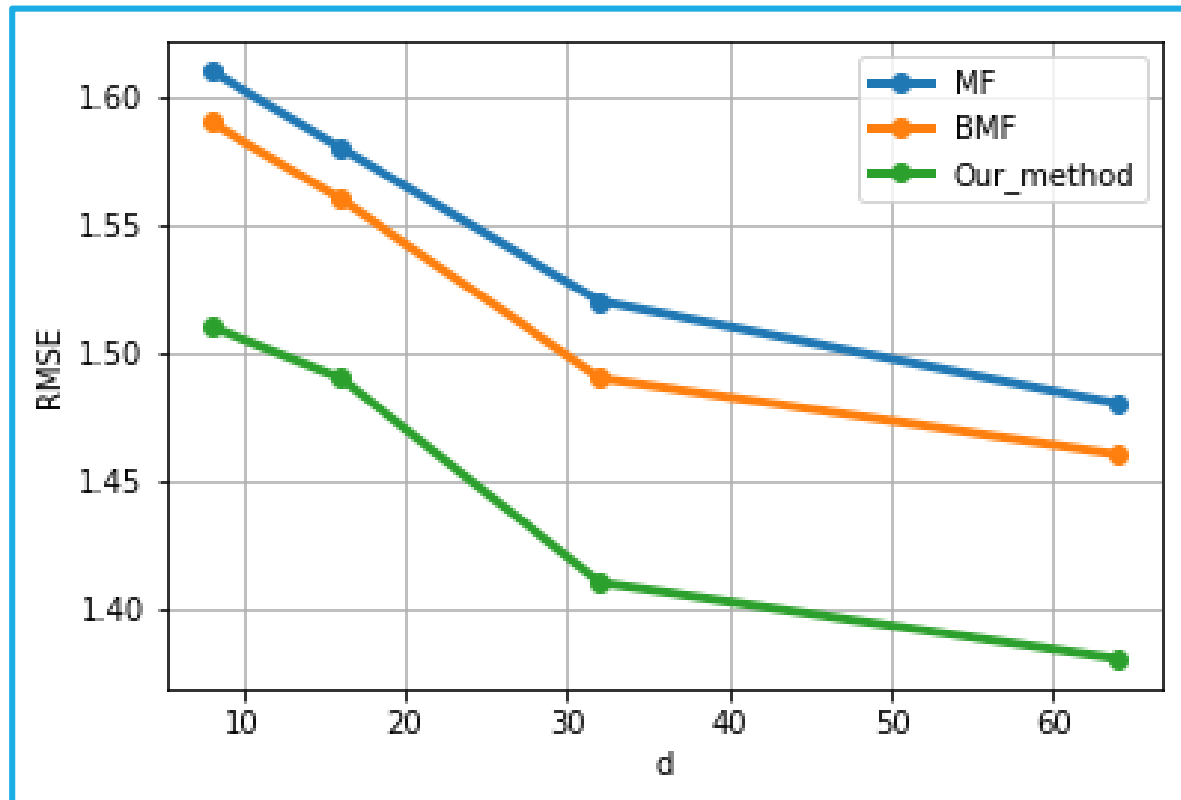
Top 6 Songs for: A Little Bit Of Pain

Name	Similarity Score
Make You L	0.9642292261123657
Ilfracombe	0.9522489309310913
Autumn'S H	0.9332128763198853
Silent Tra	0.9329584836959839
I Would Fo	0.9248721599578857
Aphrodite	0.9218252897262573

Results

- RMSE vs d (No. of dimensions in latent factor space)

Results

- Variation of RMSE vs k (No. of nearest neighbors)