Music Recommendation Based on
Multiple Contextual Similarity Information

Chih-Ming Chen, Ming-Feng Tsai
Department of Computer Science & Program in Digital Content and Technology
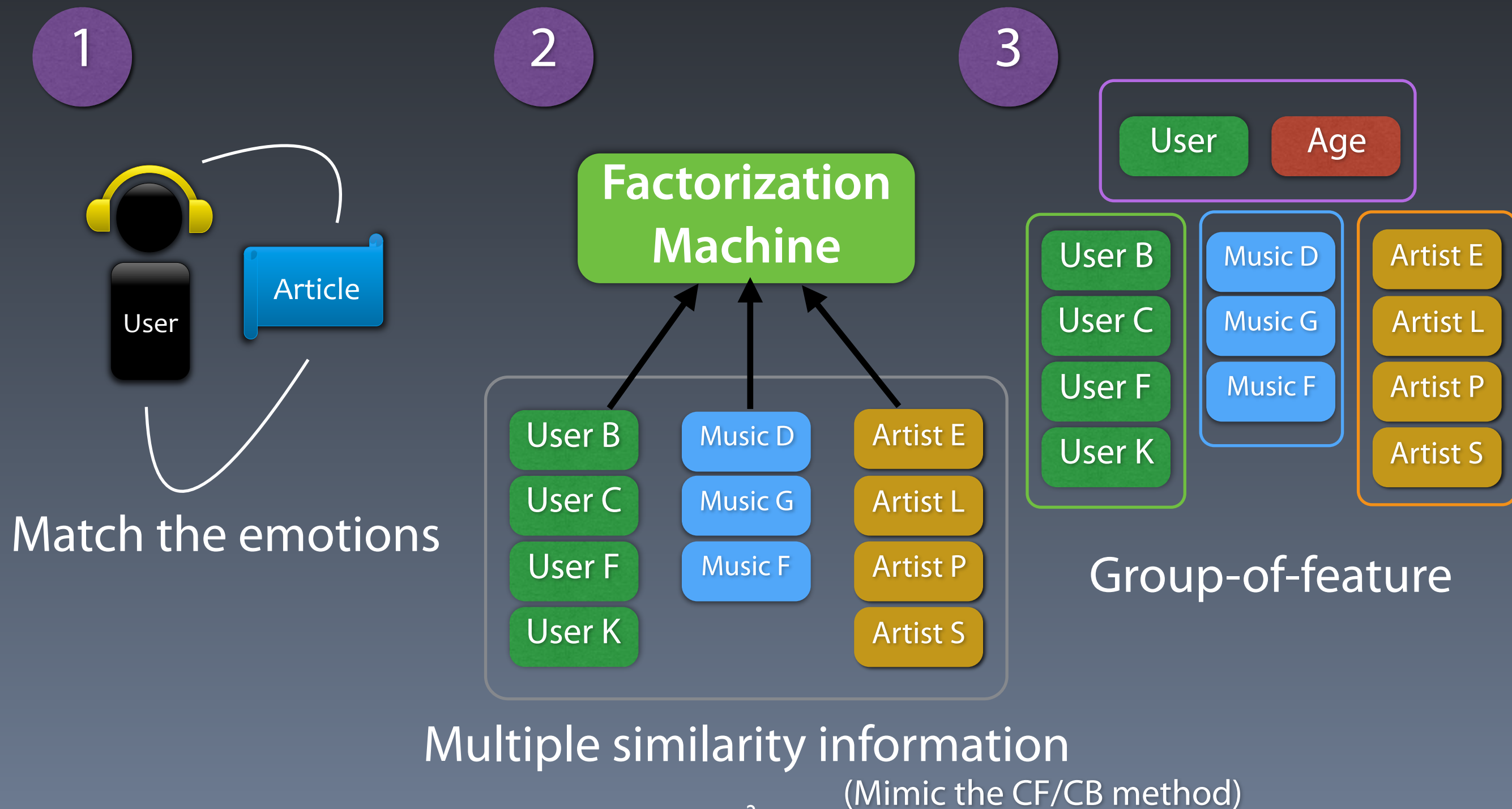National Chengchi University

Jen-Yu Liu, Yi-Hsuan Yang
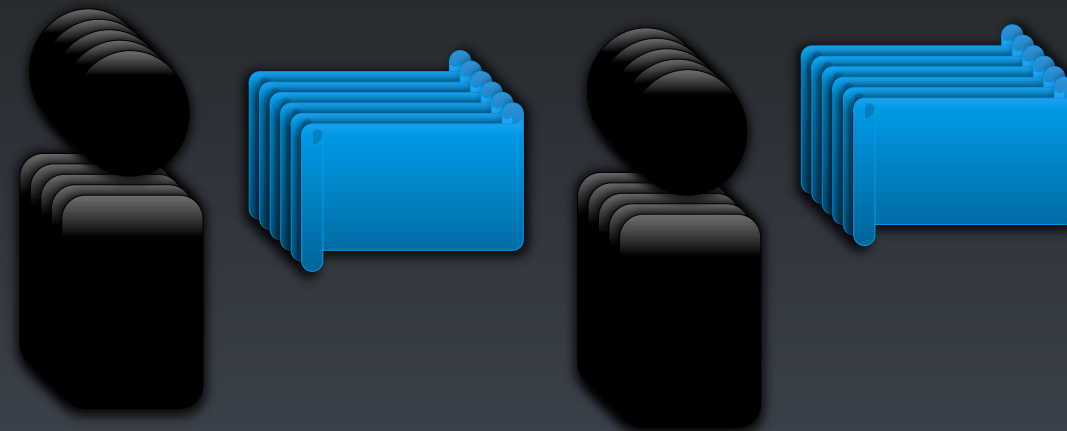Research Center for Information Technology Innovation
Academia Sinica

Taipei, Taiwan

WI-IAT-2013

Our Studies

1

User

Article

User

Match the emotions

2

Factorization Machine

User B	Music D	Artist E
User C	Music G	Artist L
User F	Music F	Artist P
User K		Artist S

Multiple similarity information

(Mimic the CF/CB method)

3

| User | Age |

User B	Music D	Artist E
User C	Music G	Artist L
User F	Music F	Artist P
User K		Artist S

Group-of-feature
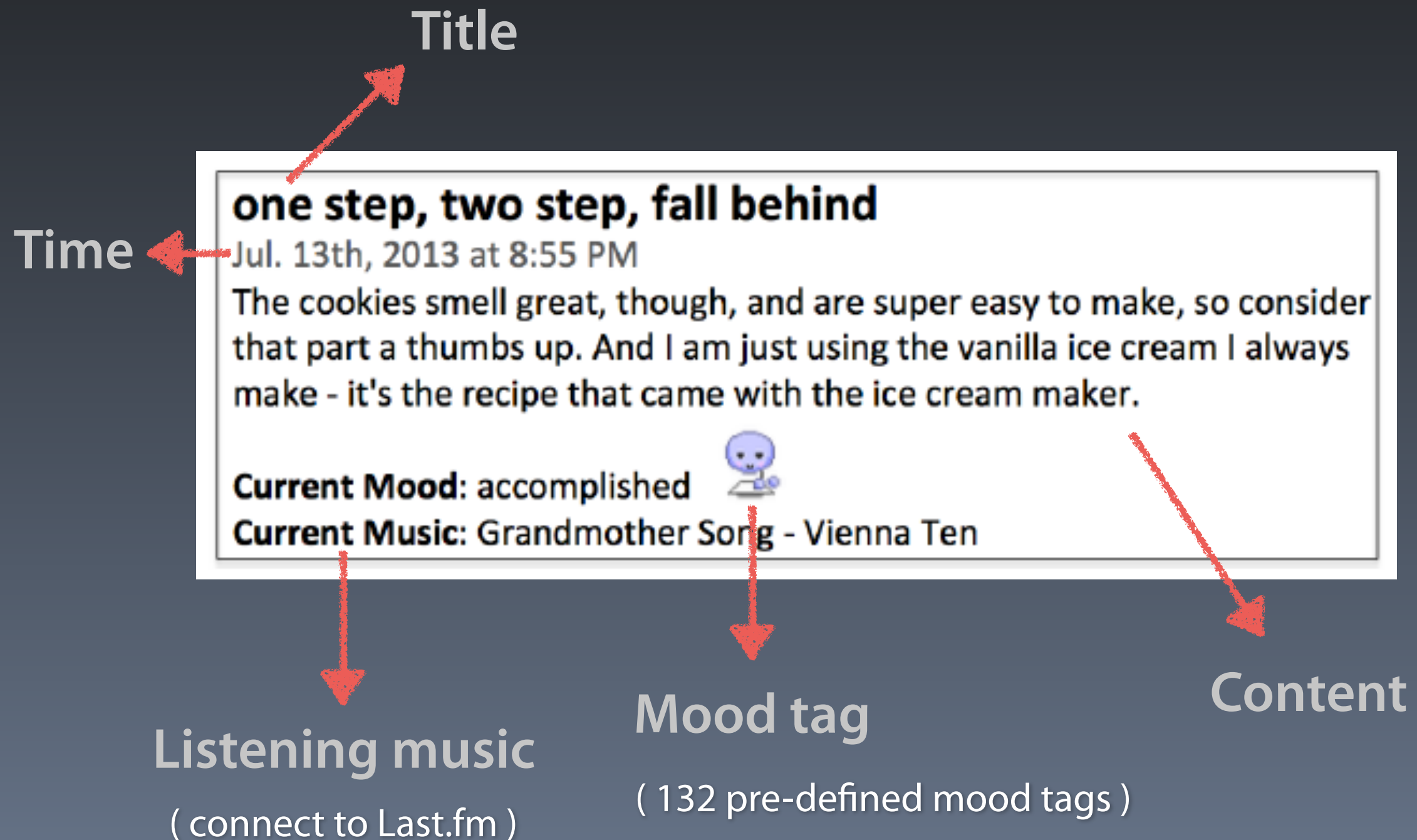
Music-related Dataset

Initial dataset

- 1,928,868 Listening records
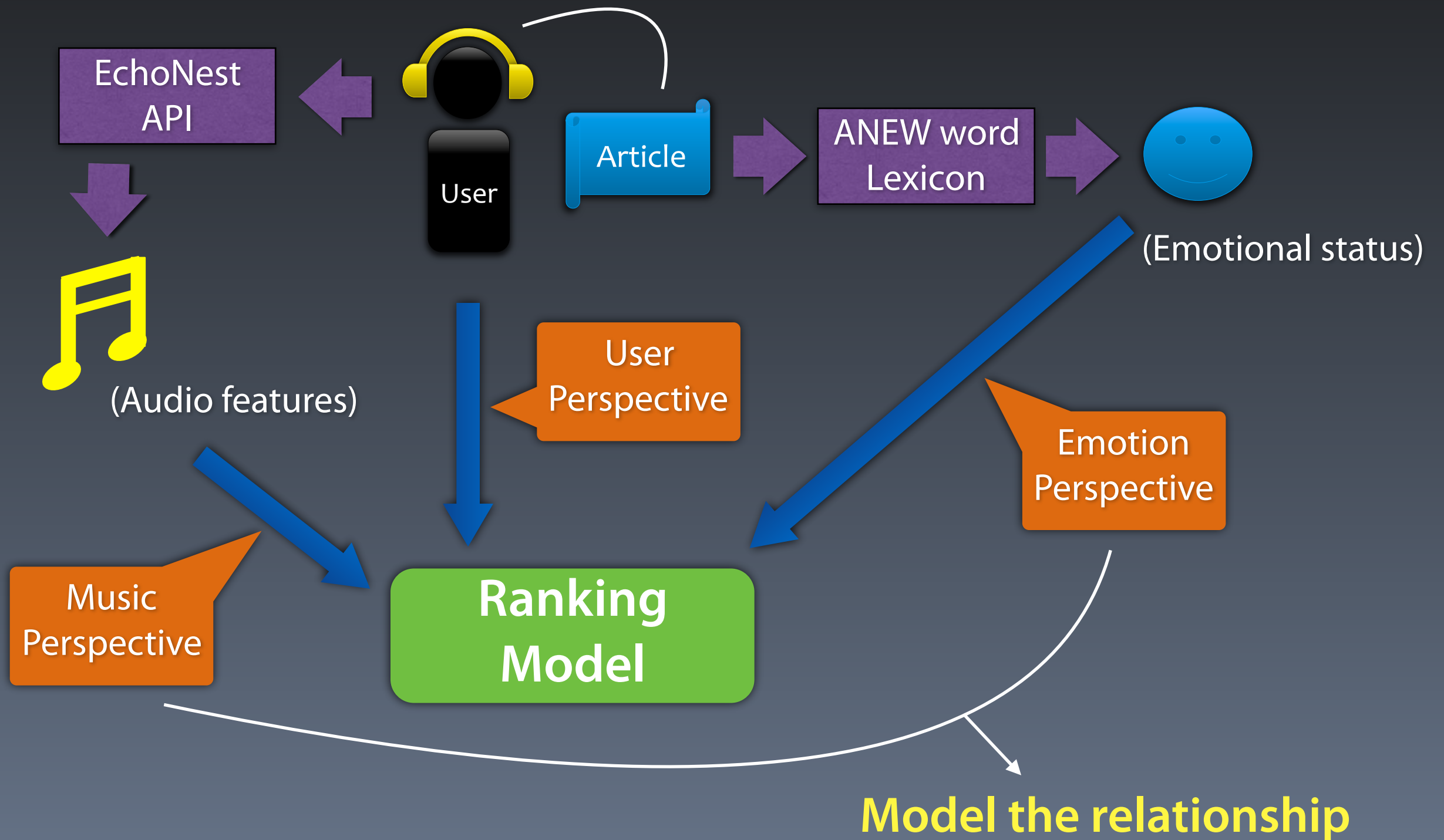- 674,932 Users
- 72,913 Songs

Experimental dataset

- 225,652 Listening records
- 19,596 Users
- 30,260 Songs

* keep only the users who have more than 10 listening records

LiveJournal Example Post

Title

Time

one step, two step, fall behind

Jul. 13th, 2013 at 8:55 PM

The cookies smell great, though, and are super easy to make, so consider that part a thumbs up. And I am just using the vanilla ice cream I always make - it's the recipe that came with the ice cream maker.

Current Mood: accomplished

Current Music: Grandmother Song - Vienna Ten

Content

Mood tag

(132 pre-defined mood tags)

Listening music

(connect to Last.fm)

Recommendation Strategy

EchoNest API

Article

ANEW word Lexicon

User

(Emotional status)

(Audio features)

User Perspective

Emotion Perspective

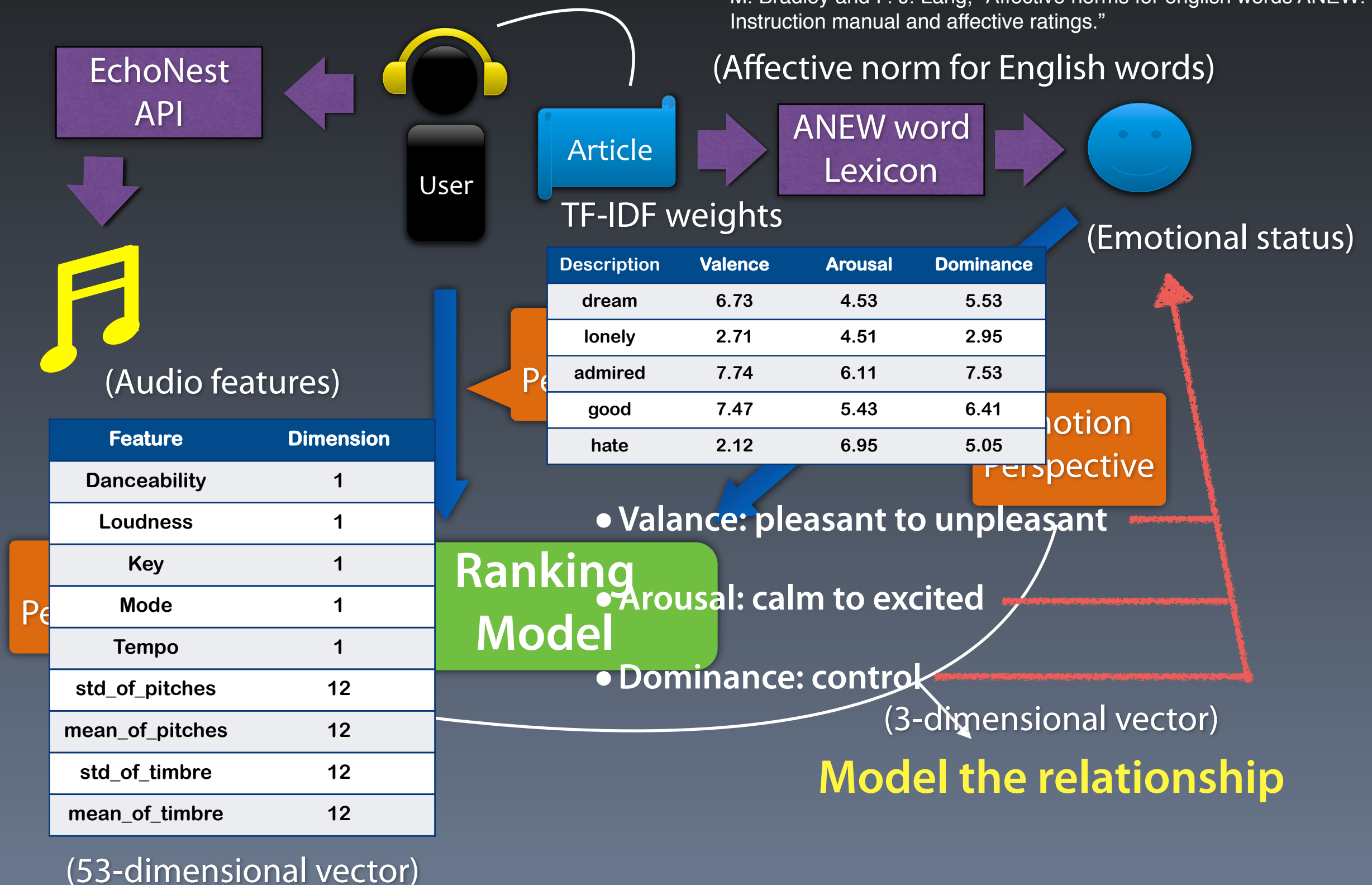Music Perspective

Ranking Model

Model the relationship

Match the Emotions

EchoNest API

User

M. Bradley and P. J. Lang, "Affective norms for english words ANEW: Instruction manual and affective ratings."

(Affective norm for English words)

Article

ANEW word Lexicon

TF-IDF weights

(Emotional status)

Description	Valence	Arousal	Dominance
dream	6.73	4.53	5.53
lonely	2.71	4.51	2.95
admired	7.74	6.11	7.53
good	7.47	5.43	6.41
hate	2.12	6.95	5.05

(Audio features)

Feature	Dimension
Danceability	1
Loudness	1
Key	1
Mode	1
Tempo	1
std_of_pitches	12
mean_of_pitches	12
std_of_timbre	12
mean_of_timbre	12

(53-dimensional vector)

Ranking Model

● **Valance: pleasant to unpleasant**

●**Arousal: calm to excited**

●**Dominance: control**

notion Perspective

(3-dimensional vector)

Model the relationship

6

Our Ranking Approach

$$\hat{y}(x) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{j=1}^{n} \sum_{j'=j+1}^{n} x_j x_{j'} \sum_{f=1}^{\kappa} v_{jf} v_{j'f}$$

Global Bias   Feature Weights   Weights of pair of features   [Rendle, ICDM 2010]

- **Factorization Machine（FM）**

  - A **competitive model** for ranking problem.

  - **Easy to embed** various kinds of feature in the data.

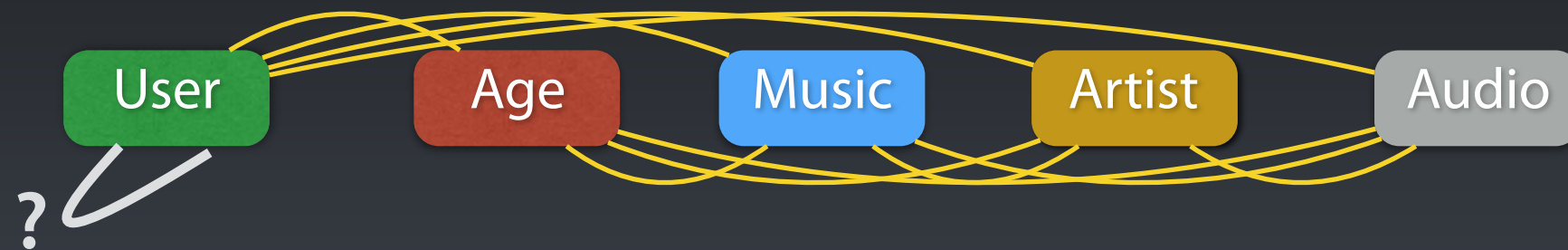 - Capable of learning the interactions from **pair of features**.

7

The Data Format

$$\hat{y}(x) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{j=1}^{n} \sum_{j'=j+1}^{n} x_j x_{j'} \sum_{f=1}^{\kappa} v_{jf} v_{j'f}$$

Global Bias   Feature Weights   Weights of pair of features

Rating	User				Age	Music				Artist			Audio			...
4	1	0	0	0	12	1	0	0	0	1	0	0	0.2	0.7	0.1	...
0	1	0	0	0	12	0	1	0	0	0	1	0	0.8	0.1	0.1	...
2	0	1	0	0	18	0	1	0	0	0	1	0	0.8	0.1	0.1	...
1	0	1	0	0	18	0	0	1	0	0	0	1	0.3	0.6	0.1	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Y	A	B	C	D	G	A	B	C	D	A	B	C	A	B	C	

Features Interaction

Similarity Information

- Enable the **missing connections** between the features

Simulate the user-based KNN method.

$$s_{ij} = \frac{O(T_i) \cap O(T_j)}{|O(T_i)|^{1-\alpha}|O(T_j)|^{\alpha}}$$

Modified version of Cosine Similarity

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}.$$

L2 Distance Function
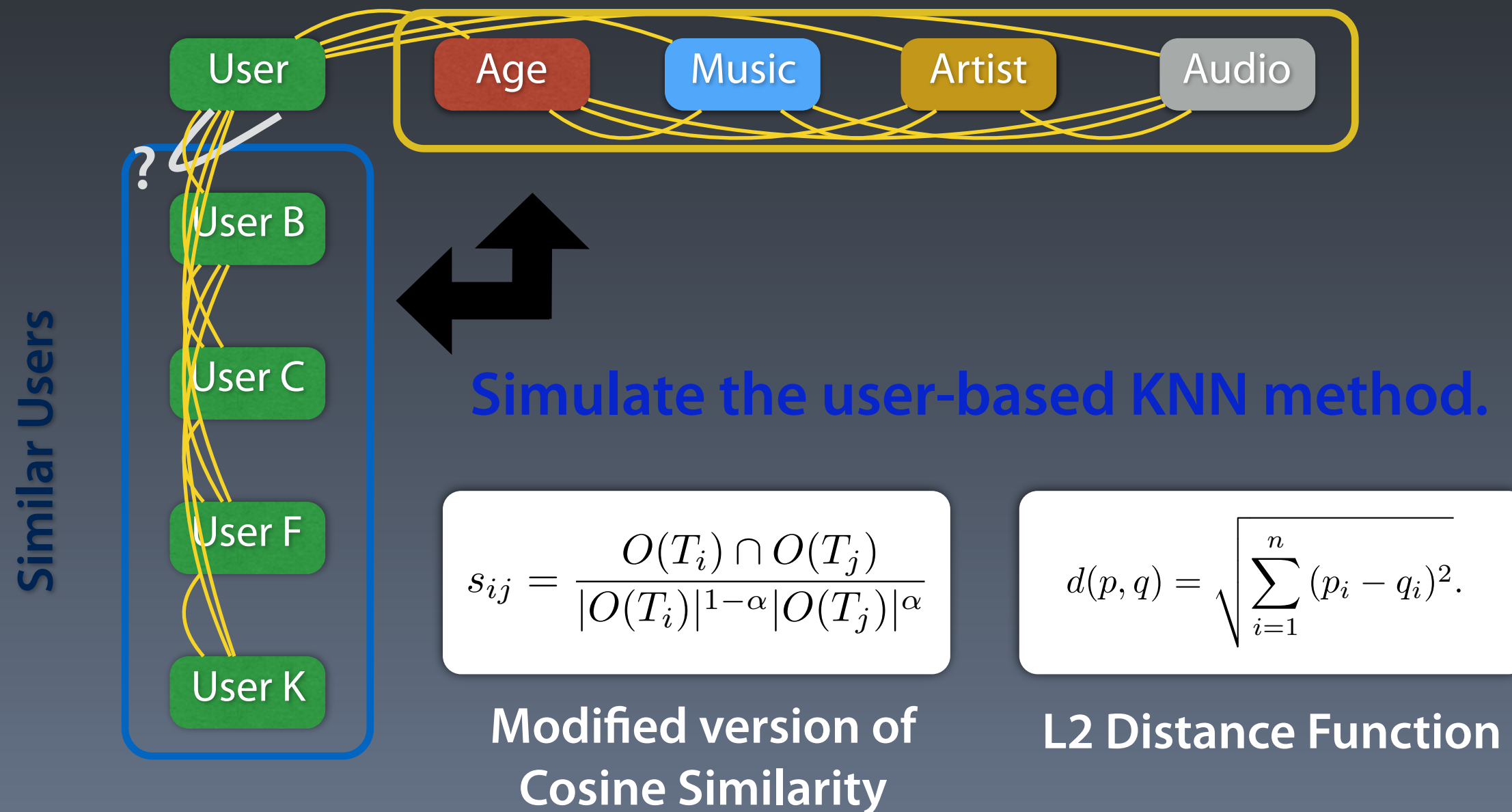
Example

User

Rating	A	B	C	D	E
3	1	0	0	0	0
2	1	0	0	0	0
0	0	1	0	0	0
1	0	1	0	0	0
…	…	…	…	…	…

User similarity

A	B	C	D	E
0	0	0.8	0	0.9
0	0	0.8	0	0.9
0	0	0	0.85	0
0	0	0	0.85	0
…	…	…	…	…

Music

A	B	C	D	E
1	0	0	0	0
0	1	0	0	0
0	1	0	0	0
0	0	1	0	0
…	…	…	…	…

Music similarity

A	B	C	D	E
0	0.8	0	0.7	0
0.85	0	0	0	0
0.85	0	0	0	0
0.7	0	0	0	0.9
…	…	…	…	…

User & Music Similarity

- U: User

- US: User similarity

- S: Song

- SS: Song similarity

(Mean Average Precision)

Features	MAP@10	Recall
U + S **(baseline)**	0.3817	0.5216
U + S + US	0.4310	0.5712
U + S + SS	0.4635	0.6194
U + S + US + SS	**0.4712**	**0.6251**

Feature Similarity

- It is also applicable to other kinds of feature (under FM)

| User | Age | Music | Artist | Audio |

| User B | Age 15 | Music D | Artist E | Audio K |

- **Some users prefer the songs similar in melody,**
- **Some users prefer the songs similar in lyrics.**

Audio-based KNN

| User C | Age 18 | Music G | Artist L |

- **…**

Age-based KNN   **Artist-based KNN**

| User F | | Music F |

Music-based KNN

| User K |

User-based KNN   **(Mimic many state-of-the-art CF/CB algorithms)**
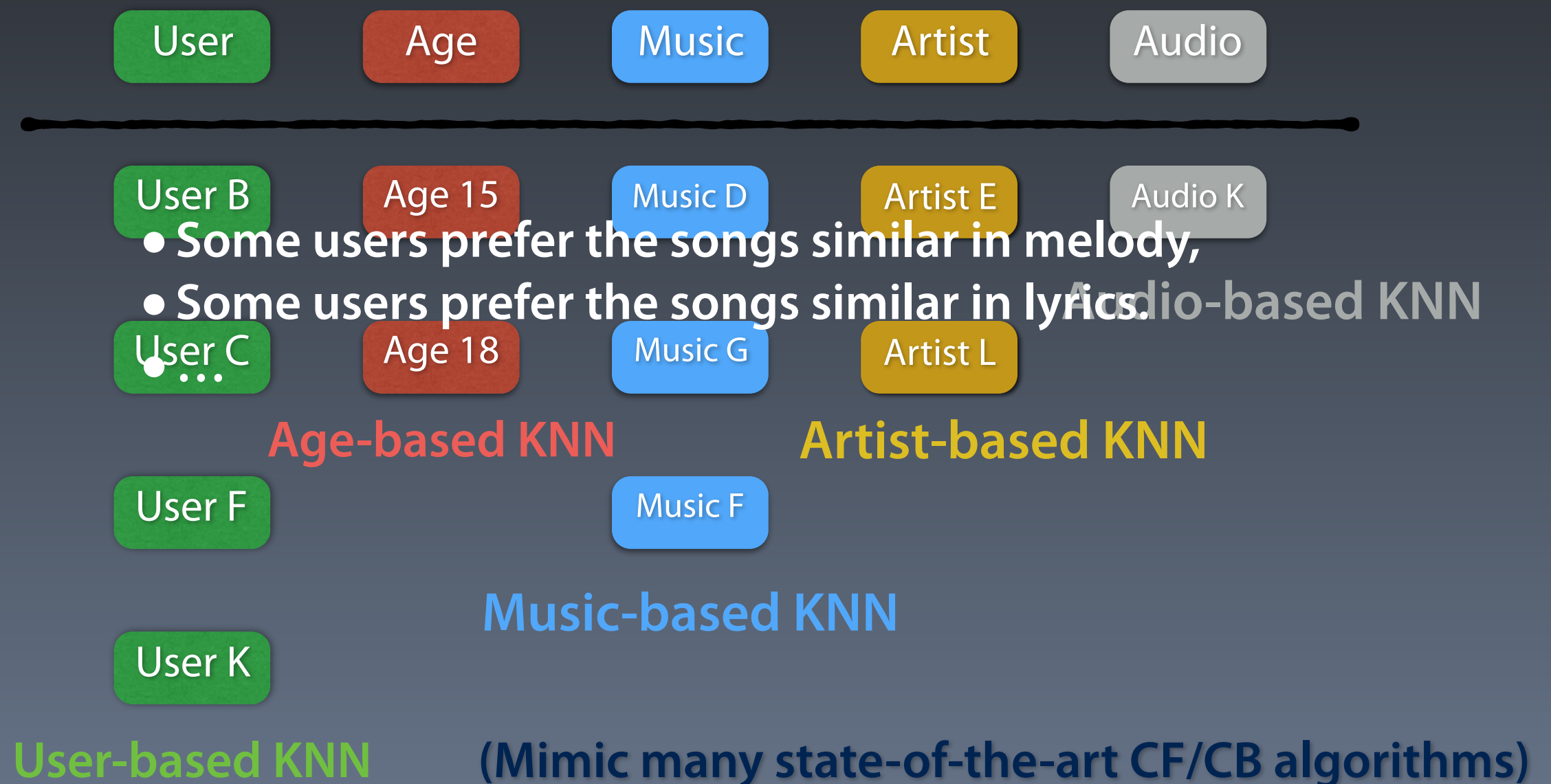
Results for Feature Similarity

- U: User

- S: Song

- A: Artist

- M: Mood tag

- Au: Audio information

- VAD: Emotional status

- R: Region

Features	MAP@10	Recall
U + S	0.3817	0.5216
U + S + A	0.5025	0.6538
U + S + A + AS	0.5125	0.6640
U + S + M	0.4635	0.6194
U + S + M + MS	0.4712	0.6251
U + S + Au	0.4254	0.5809
U + S + Au + AuS	0.4576	0.6114
U + S + VAD	0.4438	0.5905
U + S + VAD + VADS	0.4511	0.5935
U + S + R	0.4283	0.5723
U + S + R +RS	0.4382	0.5834

Some Issues

Non-informative connection

User

Age

Confused interactions

Age 15

Age 16

Age 18

Audio

Mode

Tempo

Loudness

Danceability

…

High dimensions

Music   Article

Music Features

Region   VAD

Mood   Emotion
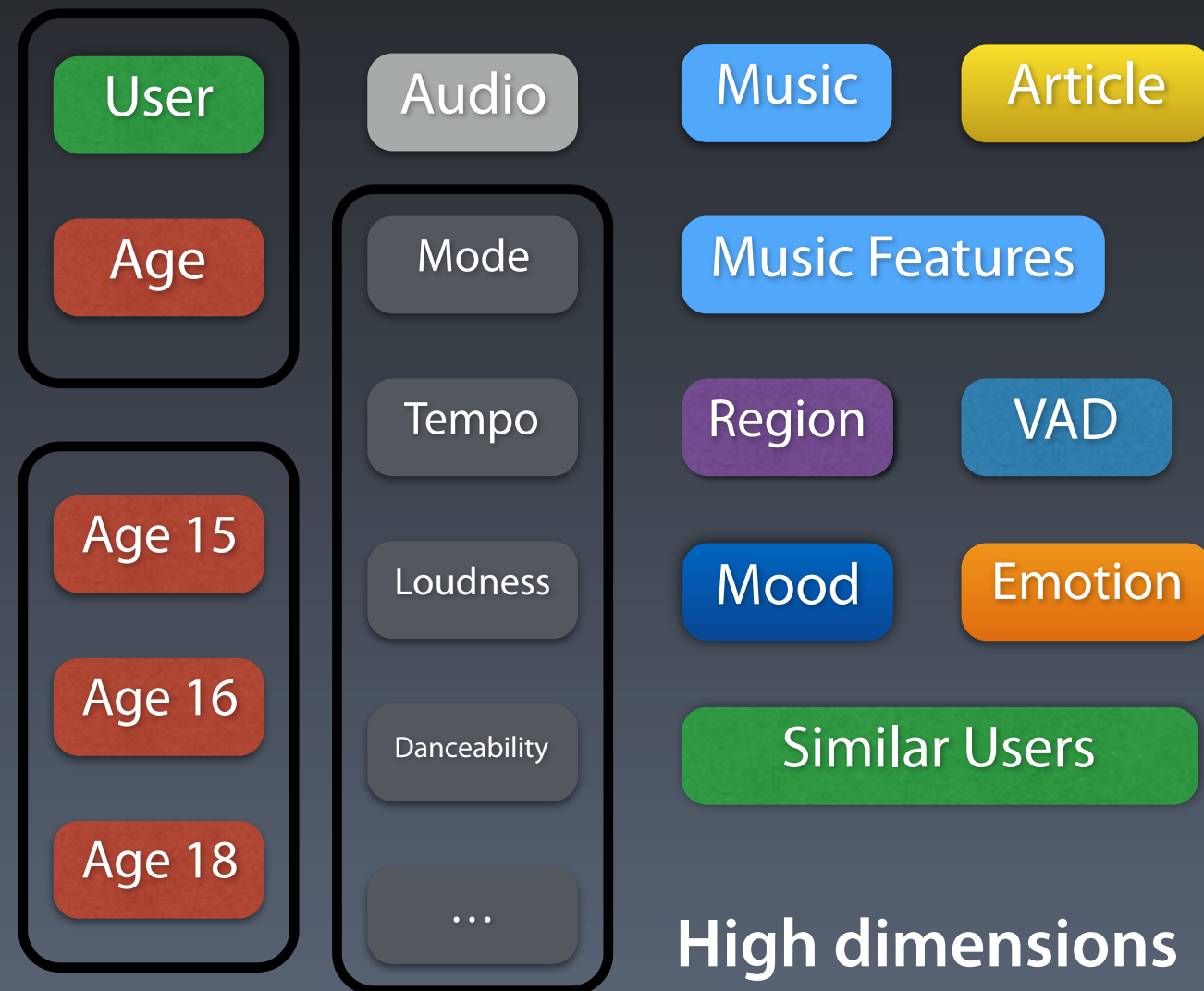
Similar Users

- **High computation cost**
- **High complexity**

Example for Group-of-Feature

Group A

User

Age

Group B

User B

User C

User F

User K

x

Music

Mode

Temp

Loudn

Dance

…
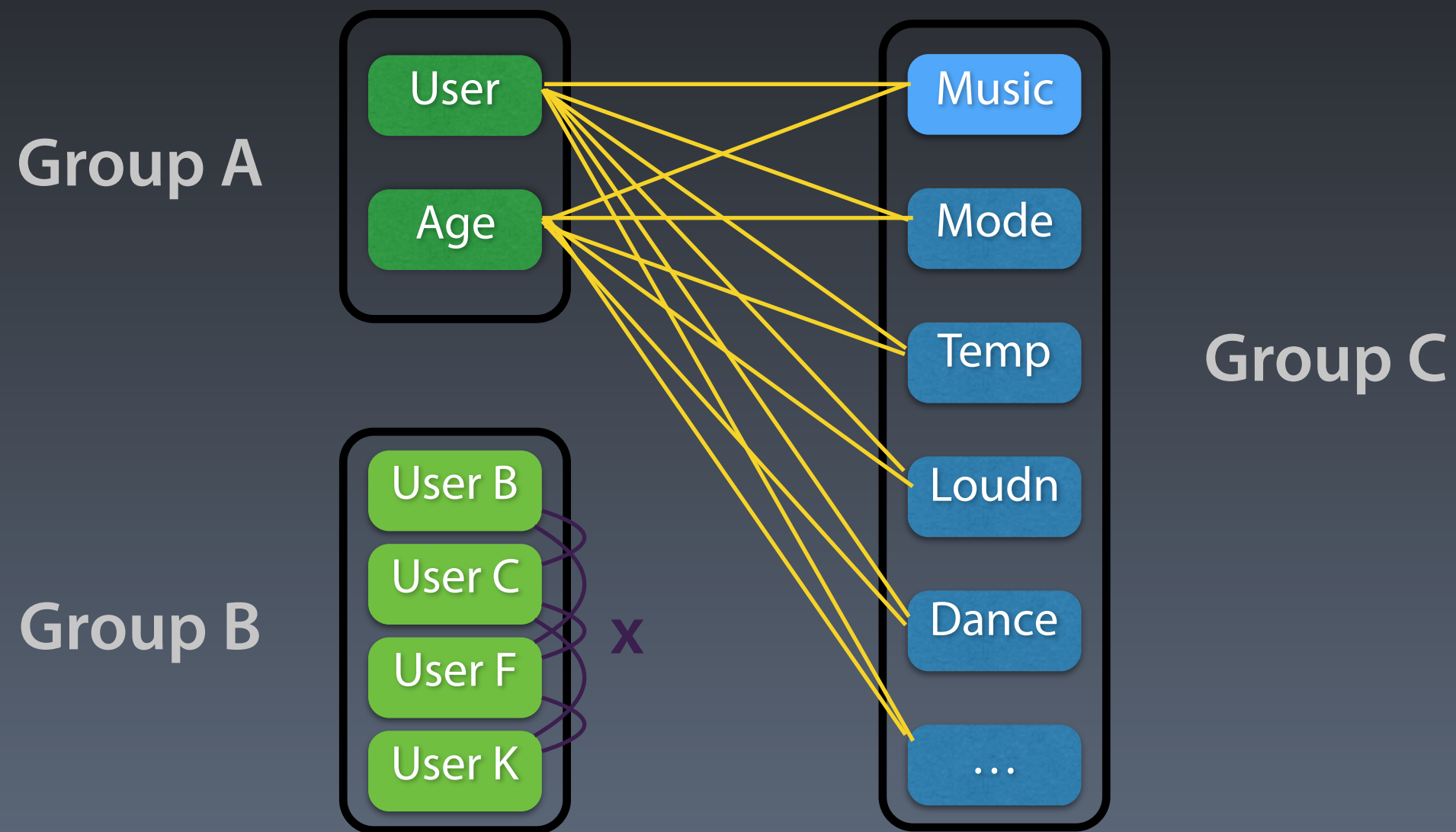
Group C

- **Reduce computation cost**
- **Reduce complexity**

Grouping Method

$$\hat{y}(x) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{j=1}^{n} \sum_{j'=j+1}^{n} x_j x_{j'} \sum_{f=1}^{\kappa} v_{jf} v_{j'f}$$

Interaction between each pair of features

$$\hat{y}(x) = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{j \in G(j)} \sum_{j' \notin G(j)} x_j x_{j'} \sum_{f=1}^{\kappa} v_{jf} v_{j'f}$$

**This way can eliminate the inner interaction
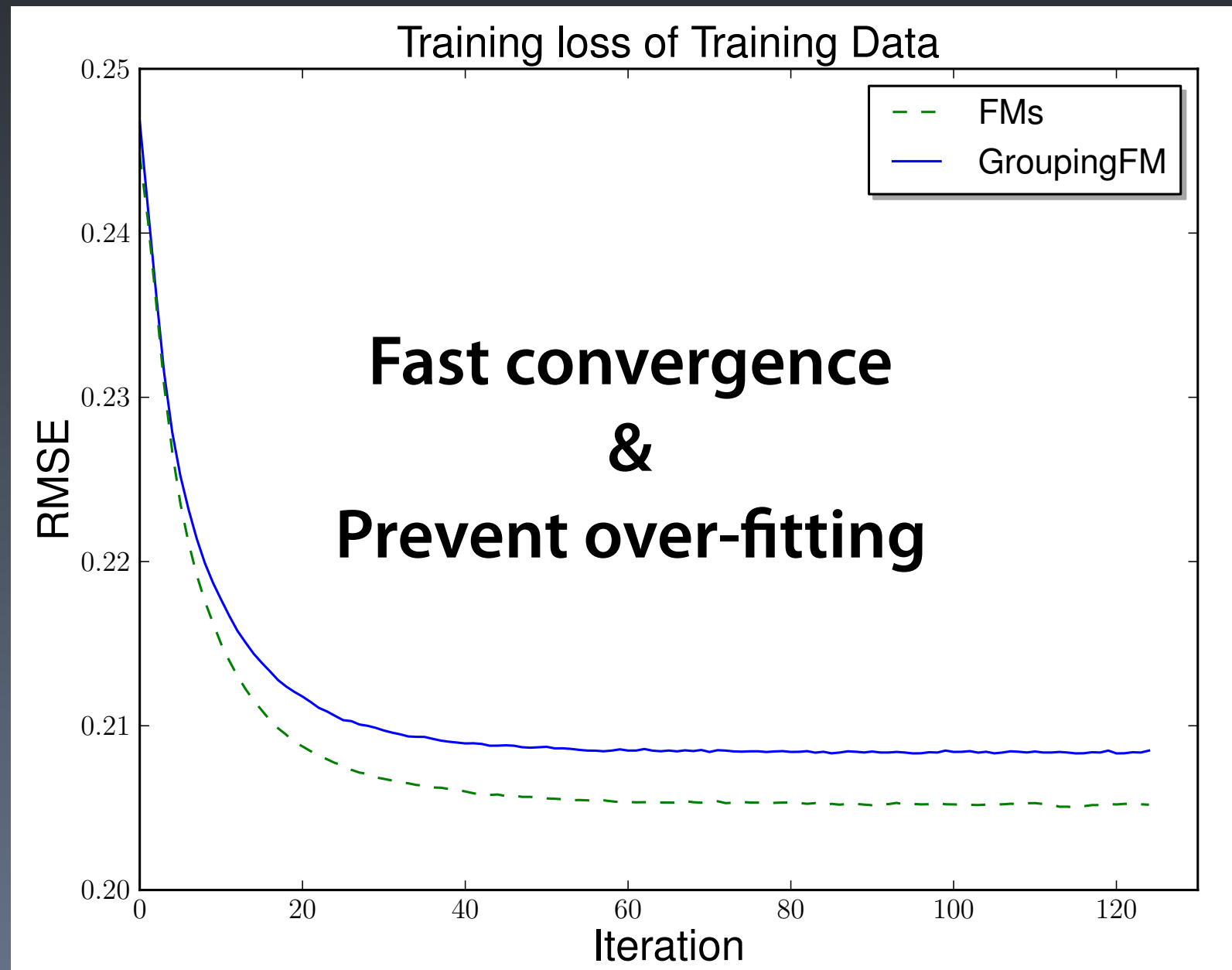(If the two features are in the same group)**

Mean Average Precision
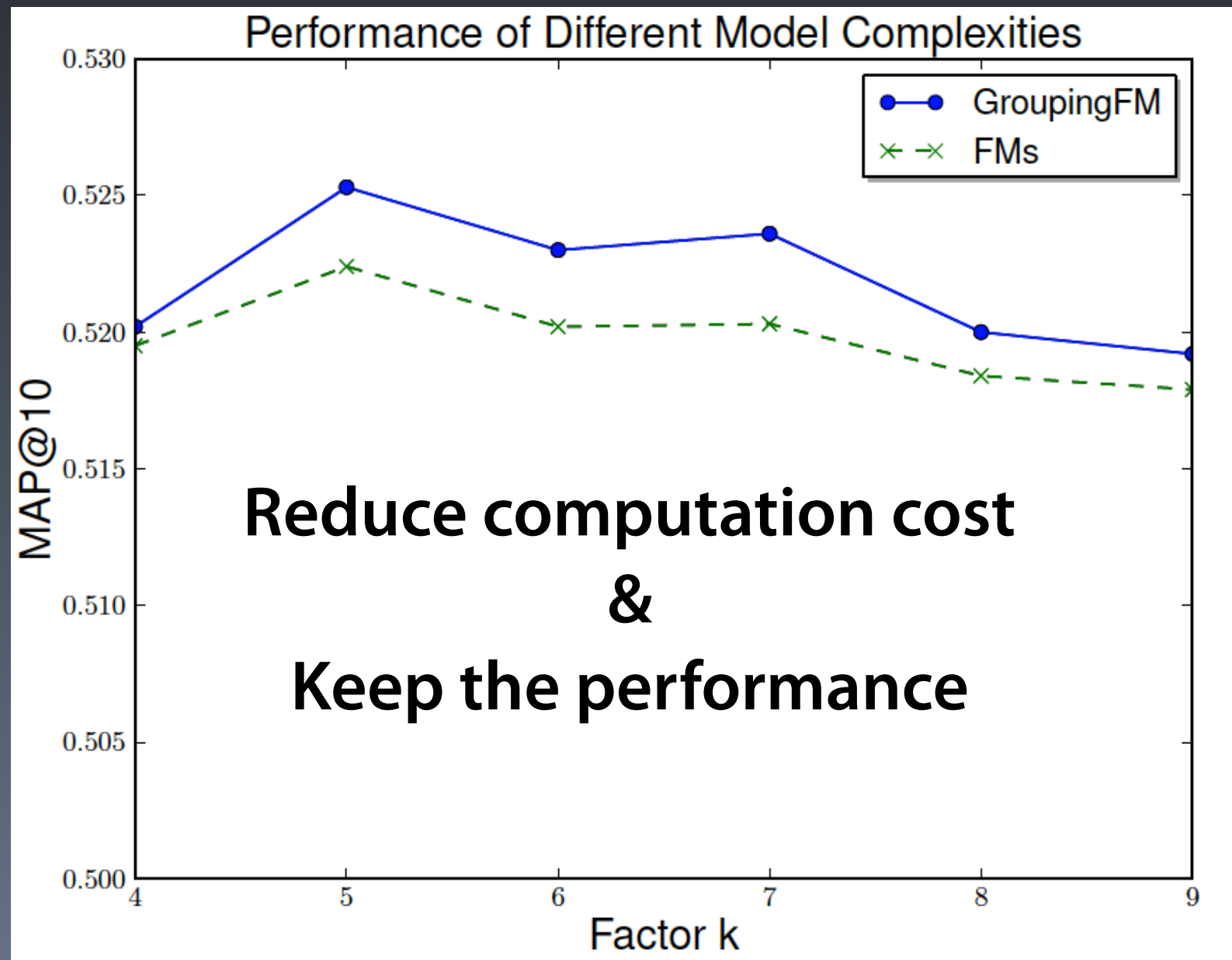
Training Loss

Conclusion

- **Music Recommendation**

  - Match the music by capturing the **emotions**

- **Recommendation Model**

  - **Factorization Machine** is used for ranking purpose

  - Integrate the **multiple similarity information**

  - Apply the **group-of-feature concept** to FM model

Thank You!

Chih-Ming Chen
changecandy@gmail.com