CS4347

Sound and Music Computing

L11a: Introduction to Music Information Retrieval (MIR) & Music Recommender Systems

Wang Ye

www.comp.nus.edu.sg/~wangye

wangye@comp.nus.edu.sg

Office: AS6-04-08

Topics to Cover (*selective approach*)

Part A: The Core

- Introduction
- Review of DFT, Audio Representation, and Machine Learning
- Music Representation, Analysis and Transcription
- Automatic Music Transcription (AMT)
- Automatic Speech Recognition (ASR)
- Generative Models for Text-to-Speech (TTS) & Singing Voice Synthesis (SVS)

Midterm break

Part B: The Breadth

- Spoken Language Assessment
- Singing Voice Processing
- Nonnegative Autoencoders with Applications to Music Audio Decomposing
- Automatic Music Generation
- Music Recommender Systems/Music Production Audio Effects
- Synthesis of Sound & Music – a DSP Approach
- Project presentations/demo/concert

Outline

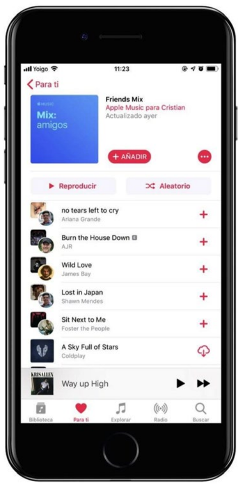
- Motivation
- Music Information Retrieval (MIR): an introduction
  - Collaborative filtering (CF)
  - Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  - Relevance Feedback
  - User Modeling
- Newer approach
- Conclusion

Why is music **recommendation** & search important? (1)

- The number of songs fit in the pocket changes largely as time goes on

10 songs
1981

1,000 songs
2001



countless songs
2022

Reference: RecSys tutorial in 2011 on music recommendation by Celma & Lamera

Why is music search & **recommendation** important? (2)

- Recommendation solves a matching problem
- The technologies can help music listeners find the next favorite song

Music Profile
- MID
- Audio
- Category
- Lyric
- Album
- Artist

……

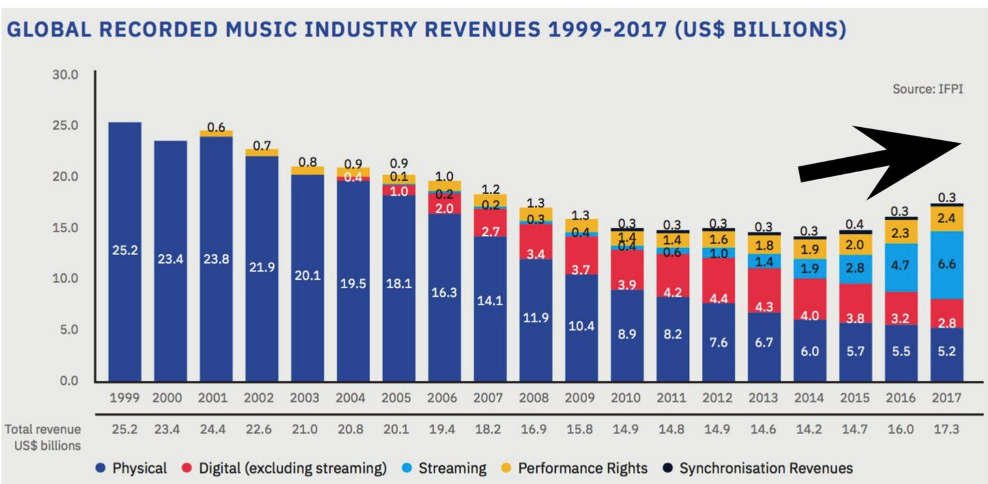
Music Recommendation System

User Profile
- UID
- Listening habits
- Rating history
- Age, Gender
- Personality

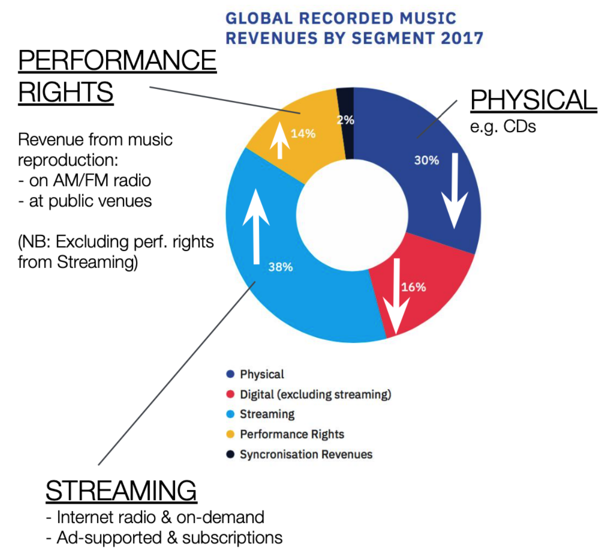
……

Why is music search & **recommendation** important? (3)

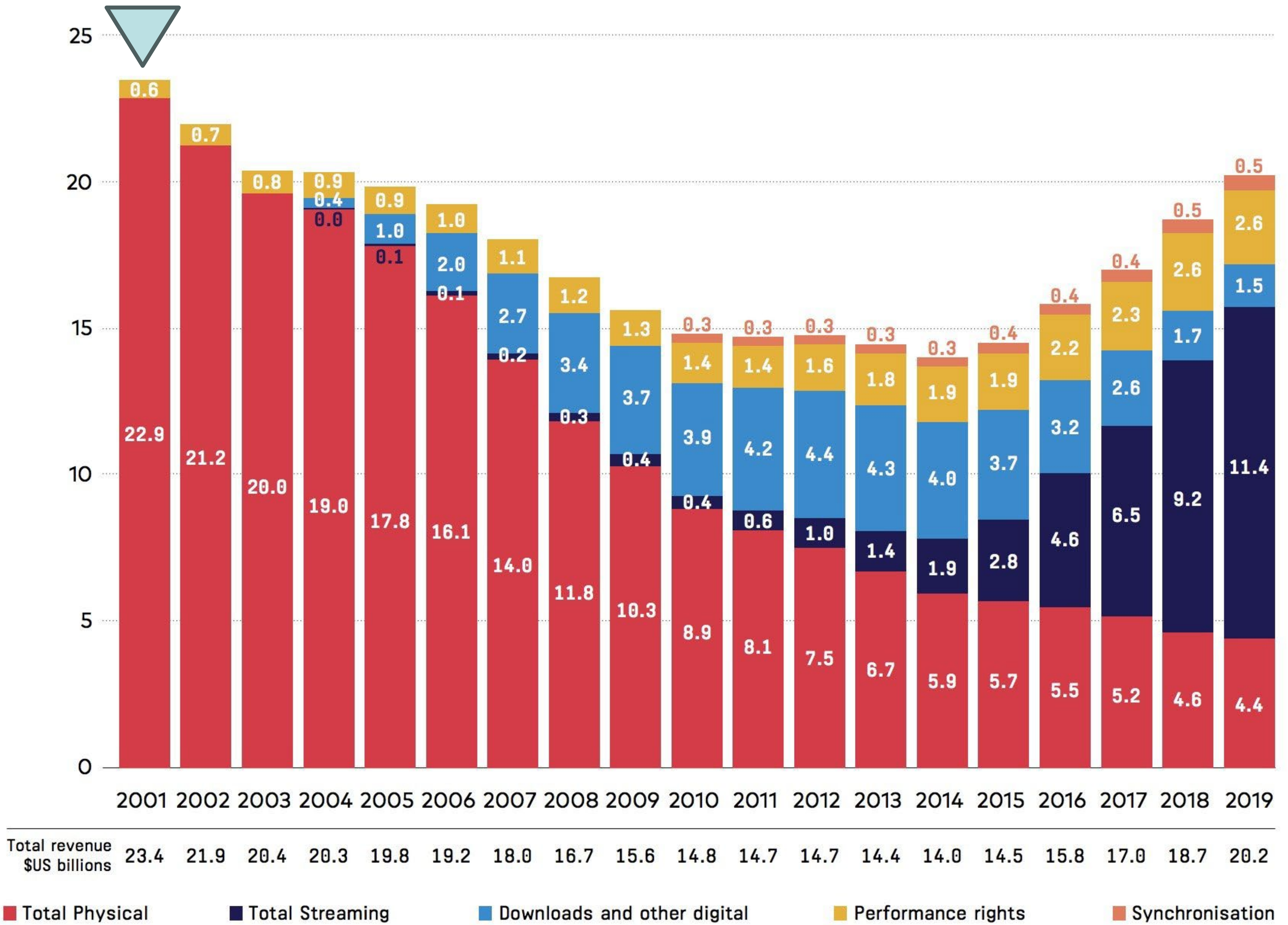
- strong impact on music industry (Schedl, Knees, Gouyon, 2018)
  - billions of global music industry revenue
  - accelerating transition: physical → streaming
  - change of paradigm: a product/item → recommending an experience



Figures adopted from: ISMIR Tutorials in 2018 on music recommendation by Schedl, Knees, Gouyon



Global Recorded Music Industry Revenues 2001-2019 (US$ Billions)



Where/how do you find your music?

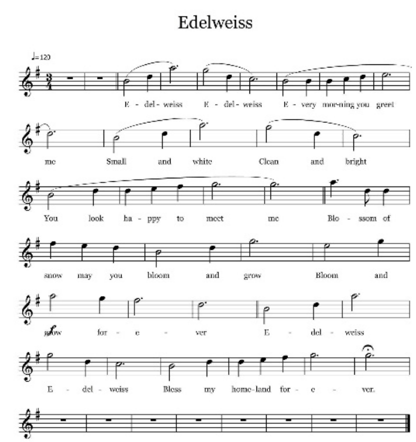
- Music stores (do any of you still purchase CDs in a physical store?)
- Online music retrieval systems

Outline

- Motivation
- Music Information Retrieval (MIR): an introduction
  - Collaborative filtering (CF)
  - Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  - Relevance Feedback
  - User Modeling
- Newer approach
- Conclusion

Multimodal Representation of Music

- **Music documents:** **Recordings (audio/image/video, MIDI), scores, lyrics, etc.**
- **Audio features**: rhythm, melody, tempo, genre, etc.
- **Metadata:** artists, album, year, descriptions, etc.



Applications

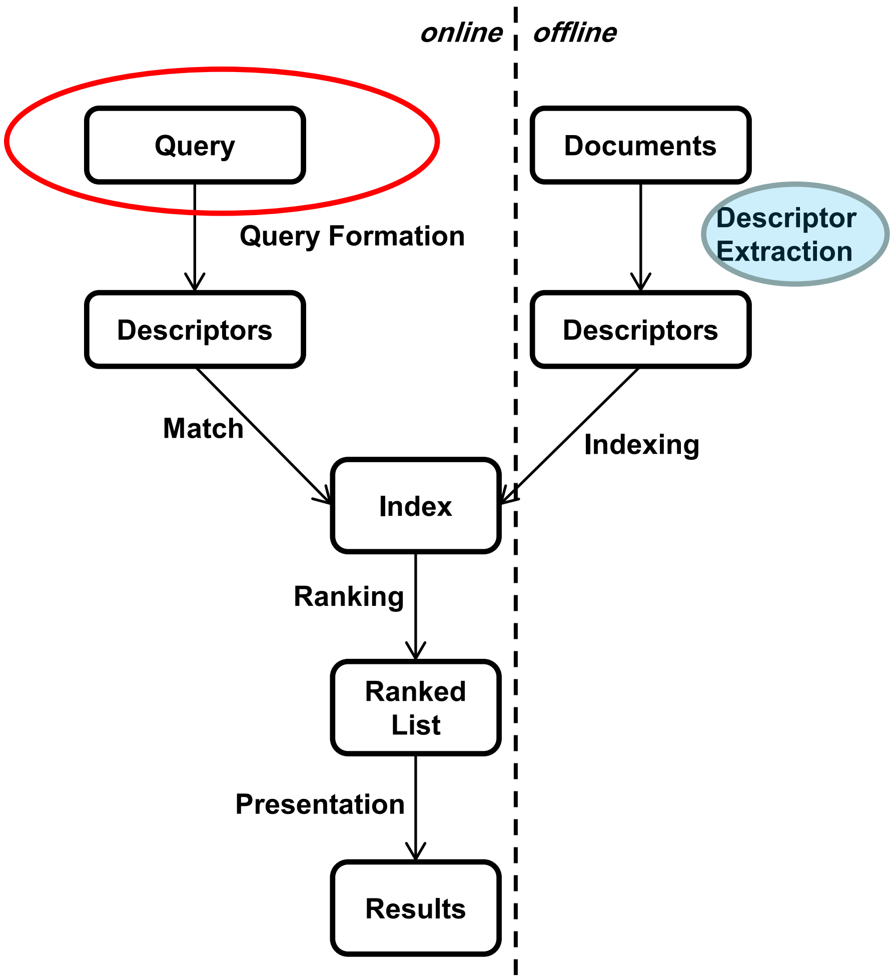
- Music Information Retrieval
  - Applications
    - ***Search*: YouTube, Shazam**
    - ***Recommendation*: Spotify**
    - *Browsing*: iTunes Store

Tasks

- Music Information Retrieval
  - Tasks
    - Genre Classification
    - Optical music recognition (OMR)
    - Content-based Similarity
    - Tempo Estimation
    - MIREX tasks…

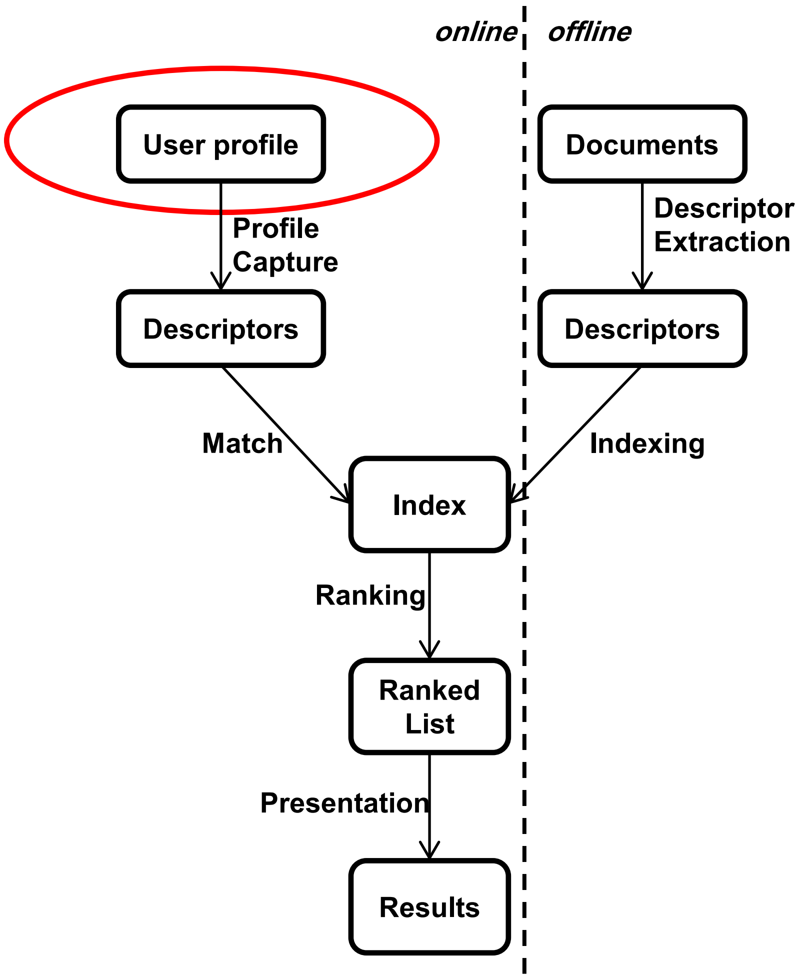
2019:Audio Classification (Train/Test) Tasks, incorporating:
2019:Audio Beat Tracking
2019:Audio Chord Estimation
2019:Audio Cover Song Identification
2019:Audio Downbeat Estimation
2019:Audio Key Detection
2019:Audio Onset Detection
2019:Audio Tempo Estimation
2019:Automatic Lyrics-to-Audio Alignment
2019:Drum Transcription
2019:Multiple Fundamental Frequency Estimation & Tracking
2019:Real-time Audio to Score Alignment (a.k.a Score Following)
2019:Structural Segmentation
2019:Discovery of Repeated Themes & Sections
2019:Audio Fingerprinting
2019:Set List Identification
2019:Query by Singing/Humming
2019:Singing Voice Separation
2019:Audio Tag Classification
2019:Audio Music Similarity and Retrieval
2019:Symbolic Melodic Similarity
2019:Audio Melody Extraction
2019:Query by Tapping

Applications -- Search



- User's information need: represented by explicit query: text, audio, etc.
- Retrieval: compare a query with documents in the database (similarity/distance measure).
- Ranking: using domain specific criteria (relevance, no. of hits, …)

Applications -- Recommendation



- Differs from music search: instead of an explicit Query, user's information need is represented <u>implicitly</u> by <u>user profiles</u>, e.g., ratings, listening history, etc.

What's unique for <u>music</u> recommendation? (1)

- large scale of item space
- low cost per item
- listeners with high passions
- importance of user intent and contextual usage
- re-recommendation is also appreciated
    - e.g., music vs. movie

What's unique for <u>music</u> recommendation? (2)

- a variety of modalities
  - audio, lyrics, user feedback, metadata
- lots of data sources

user-generated
(e.g., tags,
reviews)

interaction data
(e.g., listening
histories)

curated collections
(e.g., CD album
compilations, playlists,
radio channels)

Outline

- Motivation
- Music Information Retrieval (MIR): an introduction
  - Collaborative filtering (CF)
  - Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  - Relevance Feedback
  - User Modeling
- Newer approach
- Conclusion

Collaborative Filtering (CF)

Main idea: **user-item matrix (user clustering)**

Similar users

Similar items

	S1	S2	S3	S4	S5	S6
U1	5	3	2	1	?	4
U2	1	2	5	4	3	?
U3	3	2	4	4	5	1
U4	5	3	3	4	1	5

A new user:

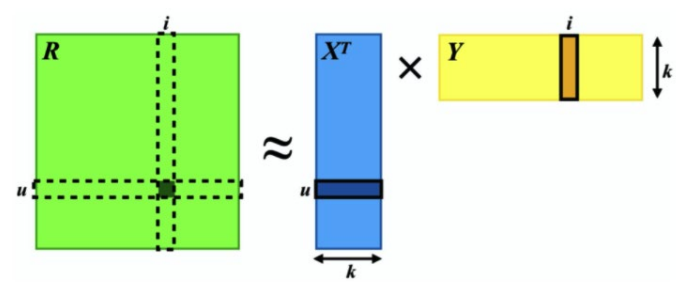
U5	**3**	**2**	**4**	**?**	**4**	**?**

Solving the Matrix Completion Problem

- Widely-employed technique
  - Matrix Factorization (cf. Singular Value Decomposition, SVD which can be considered as data-driven generalization of DFT)
- Rationale:
  - similar users or similar items should have similar embeddings
  - can train the model based on all interaction data by sharing user/item embeddings → using <u>only k dimensions</u> (lower rank) to estimate the Rating Matrix

		Music				
		U2	WP	AF	LT	...
User	J	117	49	?	?	?
	M	68	56	7	?	?
	P	?	55	?	6	?
	...	?	?	?	?	?

Rating Matrix (Interaction Matrix)



Reference: ISMIR Tutorials in 2018 on music recommendation by Schedl, Knees, Gouyon

User-User Collaborative Filtering (CF)

m users
n songs
rating matrix: $\mathbf{R} = \{r_{ij}\}_{m \times n}$

	S1	S2	S3	S4	S5
U1	5	?	?	1	?
U2	4	5	3	2	1
U3	3	2	?	?	5
U4	2	?	4	5	5

Centered cosine distance
(Pearson Correlation)

	S1	S2	S3	S4	S5
U1	2	?	?	-2	?
U2	1	2	0	-1	-2
U3	0	-1	?	?	2
U4	-1	?	1	2	2

	U1	U2	U3	U4
U1	8	4	**0**	-6
U2		10	-6	-7
U3			5	4
U4				10

Item-Item Collaborative Filtering (CF)

	S1	S2	S3	S4	S5
U1	5	?	?	1	?
U2	4	5	3	2	1
U3	3	2	?	?	5
U4	2	?	4	5	5

The Singular Value Decomposition (SVD)

Latent Factor Analysis

Matrix Factorization

SVD

	S1	S2	S3	S4	S5
F1	.2	.1	.7	.8	.2
F2	.7	.8	.3	.5	.1
F3	.3	.2	.3	.2	.9

	S1	S2	S3	S4	S5
S1		.64	.44	.57	.38
S2			.37	.52	.28
S3				.77	.44
S4					.39
S5					

Tools for CF Recommendation

Toolbox	**Coding Language**
MyMediaLite	C#
scikit-surprise	Python
Apache Mahout Recommenders	Spark
Spotlight	Python
Rival	Java

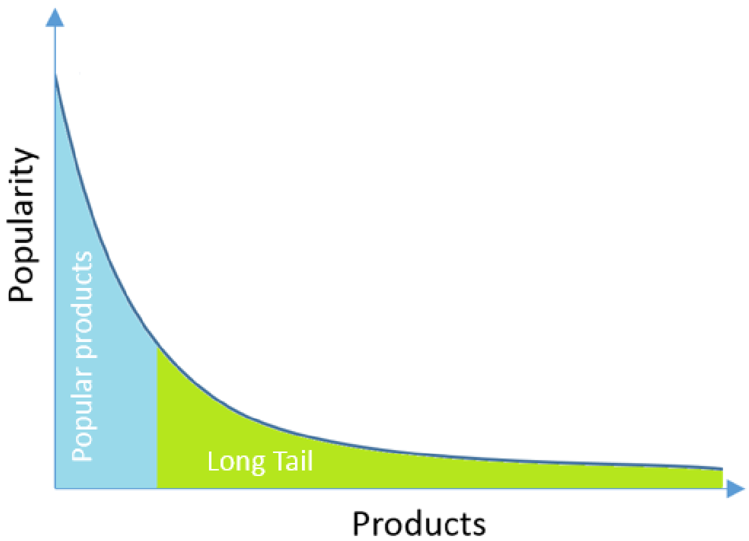
Toolboxes organized in ISMIR Tutorials in 2018 on music recommendation by Schedl, Knees, Gouyon

CF

- Pros
  - Industry standard
  - Very accurate when lots of data is available for users and songs
- Cons
  - Cold start problem for new users, new songs
  - Data is often not available to the public
  - Recommendation Transparency
    - We don't know *why* an item is being recommended

CF Problems: “Cold Start” and “Long Tail”

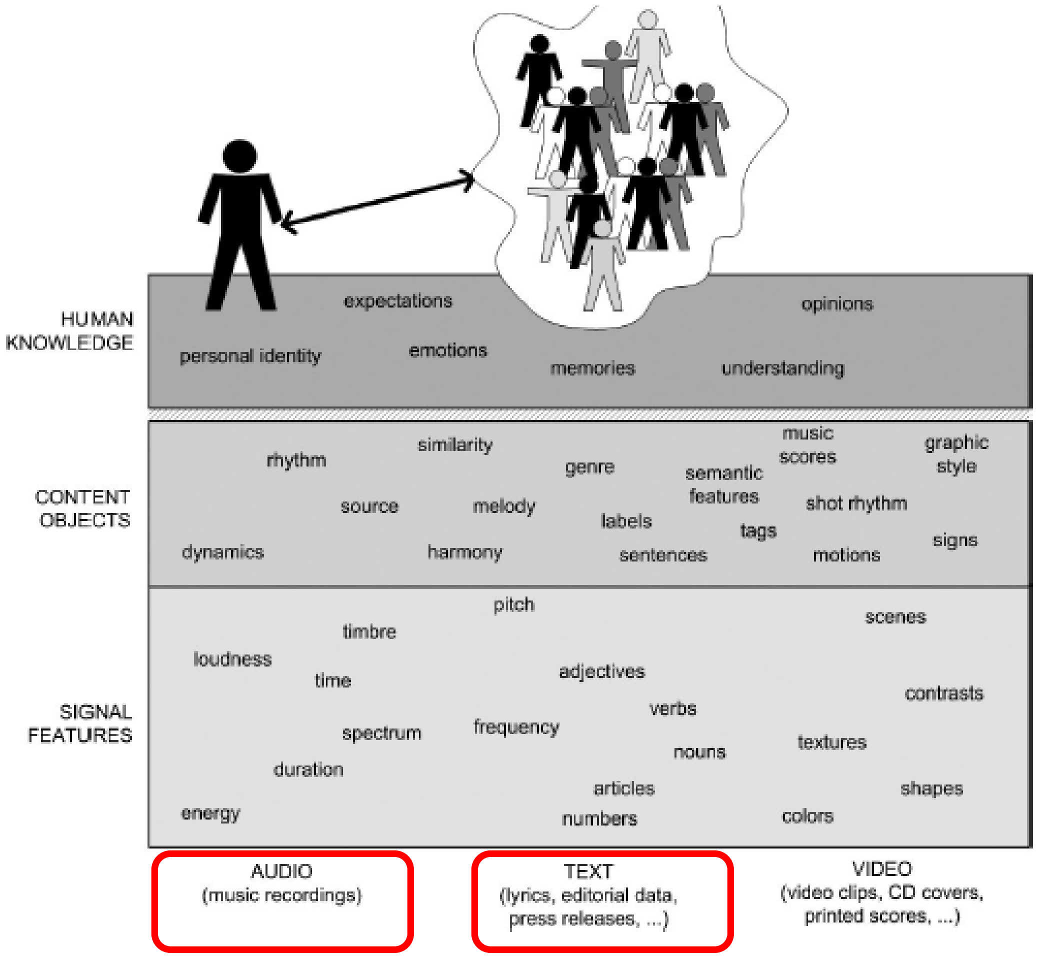
- When new users join or new music is added
  - No ratings available
- Less popular items get little chance
  - “Winner gets all”
  - Novelty and diversity suffer



Outline

- Motivation
- Music Information Retrieval (MIR): an introduction
  - Collaborative filtering (CF)
  - Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  - Relevance Feedback
  - User Modeling
- Newer approach
- Conclusion

Music Information Plane



Music

Similarity can be measured from different aspects.

Song1: New favorite – Alison Krauss

Song2: She is Beautiful - Andrew W.K.

Dissimilar

Song1	**Song2**
Female	Male
Gentle	Aggressive
Slow	fast

Similar

Guitar
Tempo: ~162 BPM (Beat Per Minute)

* O. C. Herrada. Music recommendation and discovery in the long tail. PhD thsis. 2008.

Outline

- Motivation
- Music Information Retrieval (MIR): an introduction
  – Collaborative filtering (CF)
  – Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  – Relevance Feedback
  – User Modeling
- Newer approach
- Conclusion

Text-based Method

- Basic principle: TF-IDF (Term Frequency - Inverse Document Frequency)
  – TF(t) = (Number of times term t appears in a document) / (Total number of terms in the document)
  – IDF(t) = log_e(Total number of documents / Number of documents with term t in it)

Text-based Methods: CS3245 Information Retrieval

Feature Extraction from Text Content

- Processing lyrics and user-generated content (e.g., reviews, blogs)
- Similar to basic information retrieval task
  - represent every document as a word vector
    - each word can be a feature
  - Methods: Bag-of-words model (simplest), TF-IDF, word2vec, Topic models
- Widely-used tools
  - NLTK (python), Gensim (python), StanfordNLP (python)
  - GATE (Java), ApacheNLP (Java), jMIR (Java)
  - MeTA (C++)

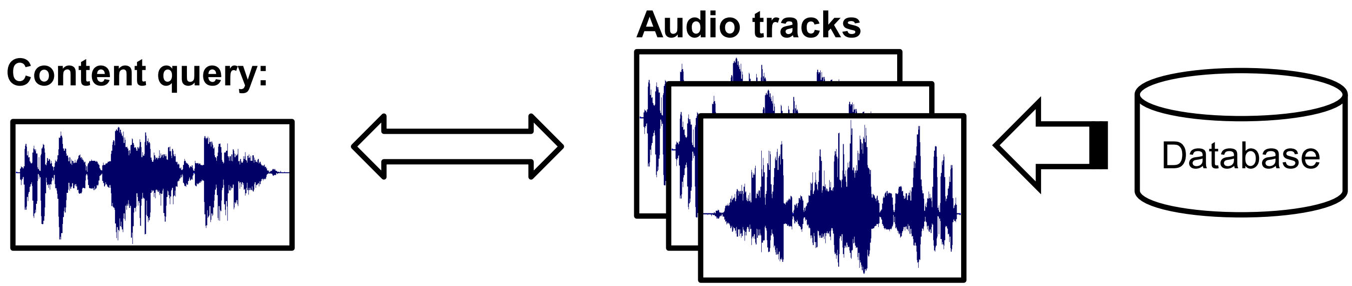
Text-based Method

- Pros
  - Simple & efficient
- Cons
  - Affected by noisy/wrong texts
  - Songs with no text cannot be retrieved
  - Require high-level domain knowledge to create good metadata
  - “Text retrieval on audio metadata” not pure music retrieval

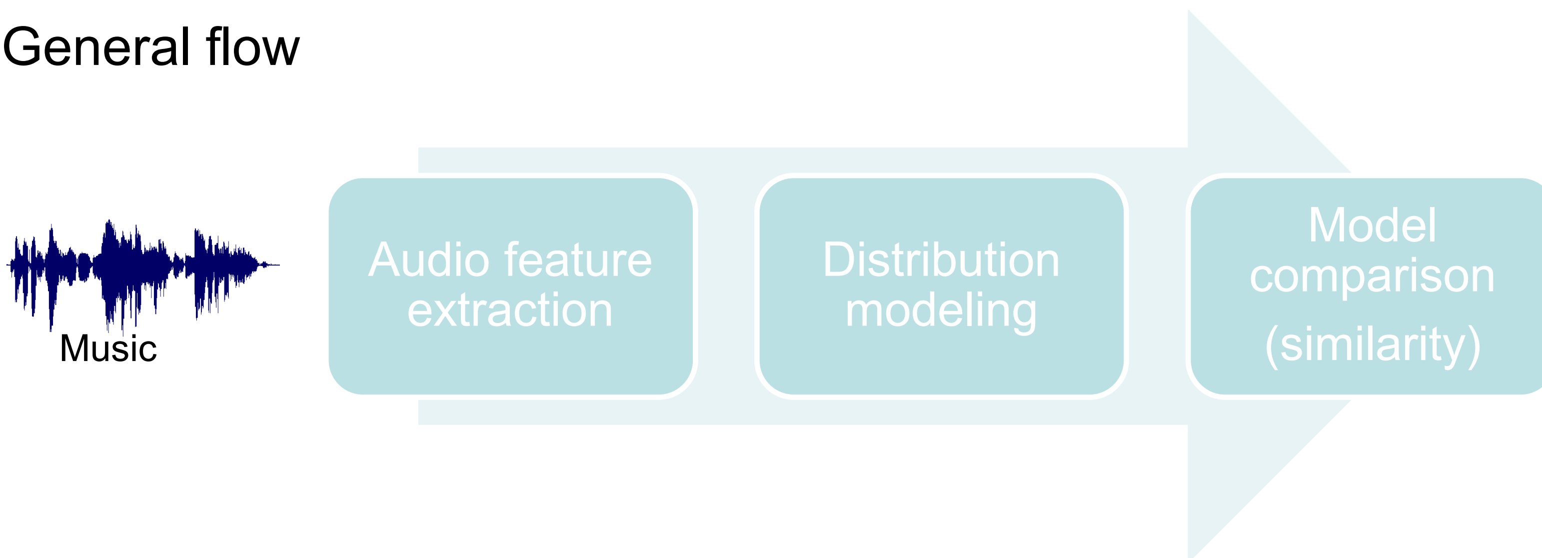
Outline

- Motivation
- Music Information Retrieval (MIR): an introduction
  - Collaborative filtering (CF)
  - Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  - Relevance Feedback
  - User Modeling
- Newer approach
- Conclusion

Audio Content-based Method



- General flow



Features from Audio Content (1)

- Low-level acoustic features extracted from the audio wave
  - tempo
  - mode
  - timbre
  - pitch
  - loudness
  - structure
  - harmony
  - ……
- Widely-used tools
  - Librosa (Python), Madmom (Python)
  - Essentia (C++, Python)
  - jMIR (Java)

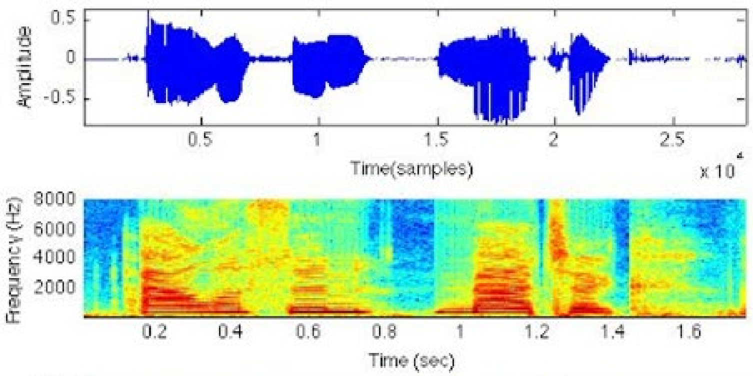


Fig. Audio signals in time and frequency domain

Features from Audio Content (2)

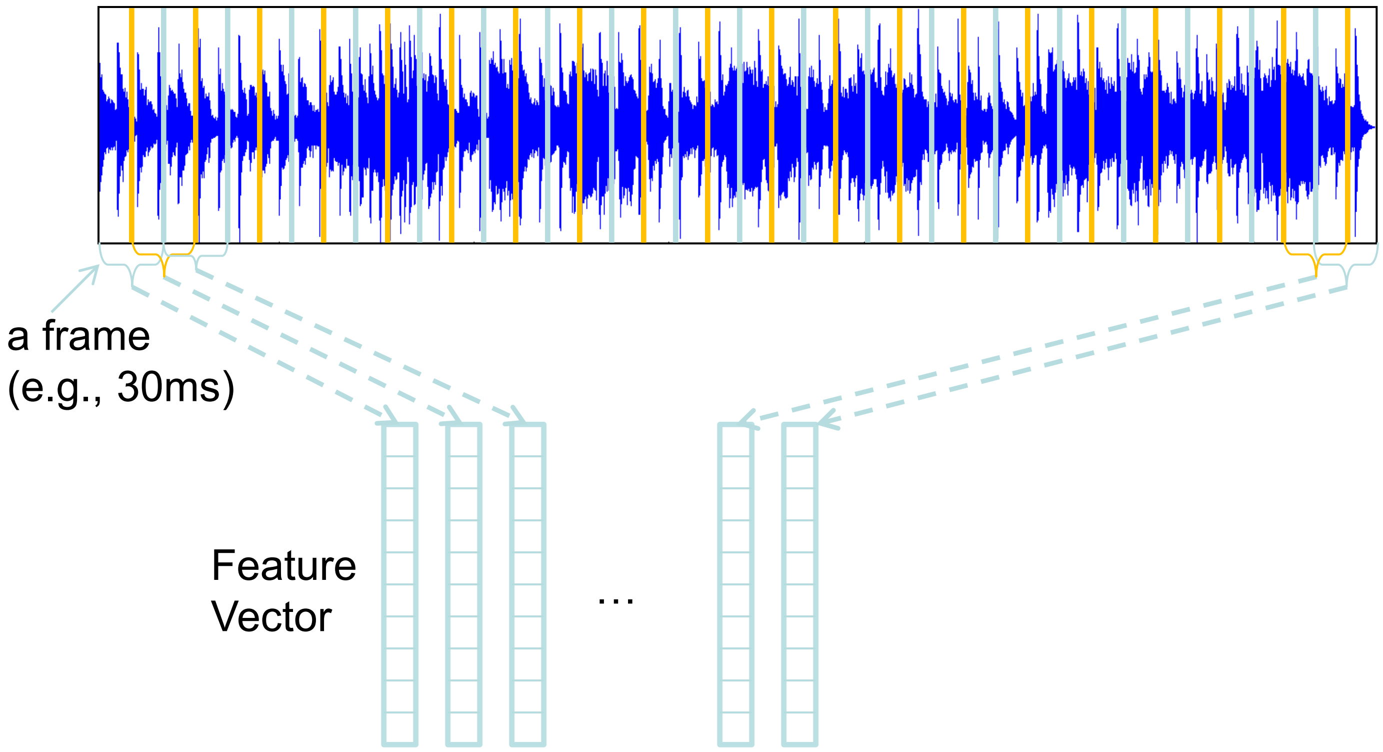
- High-level semantic features
  - e.g., mood, genre
  - can be predicted based on low-level acoustic features
  - with classification algorithms: SVM, RNN (Hu & Downie, 2010; Dong et al., 2019)
- Need exterior tagged data for training model

Music Mood

Genre

Audio Content-based Method

- Feature extraction → Distribution modeling → Model comparison
  - Frame-based features (capture short-term audio characteristics)



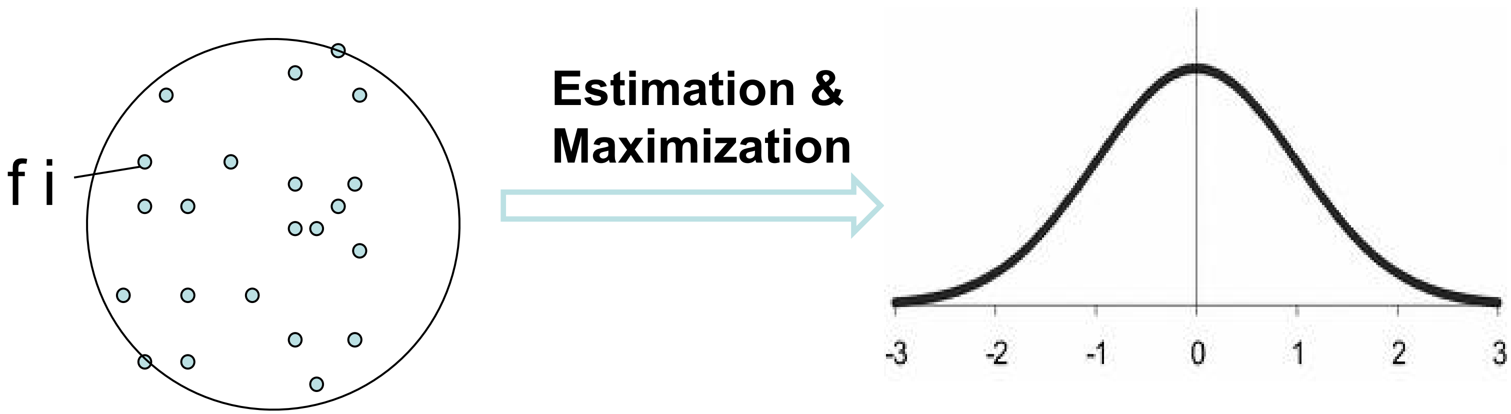
Error robustness of features (e.g., Shazam)

Audio Content-based Method

- Feature extraction → **Distribution modeling** → Model Comparison
  - **Directly combine low-level features** → single vector
    - Pitch, loudness, MFCC (Blum et al., 1999)
    - Histogram of MFCC (Foote, 1997)
    - Spectrum, rhythm, chord change (Tzanetakis, 2002)
  - **Low-level features → higher-level features**
    - Cluster MFCC → model comparison (Aucouturier, 2002)
    - MFCC → Gaussian Mixture Models → model comparison
    - MFCC →“anchor space”, compare probability models (Berenzweig et al., 2003)

Audio Content-based Method

- Feature extraction → **<u>Distribution modeling</u>** → Model Comparison
  - A single Gaussian model



Parameters:
mean, variance: *μ, Λ*
if feature components are assumed to be independent.
mean, co-variance matrix: *μ, Σ*
if feature components are assumed to be dependent.

Audio Content-based Method

- Feature extraction → **<u>Distribution modeling</u>** → Model Comparison
  - Gaussian Mixture models (GMMs)

Parameters:
weights, means, co-variance matrices: *w_i, μ_i, Σ_i*

Audio Content-based Method

- Feature extraction → Distribution modeling → **Model Comparison**
  - Basic distance/similarity measures
  - *Uniform-length feature vectors*: $A=(A_1, A_2, \ldots, A_n)$, $B=(B_1, B_2, \ldots, B_n)$
    - Euclidean distance

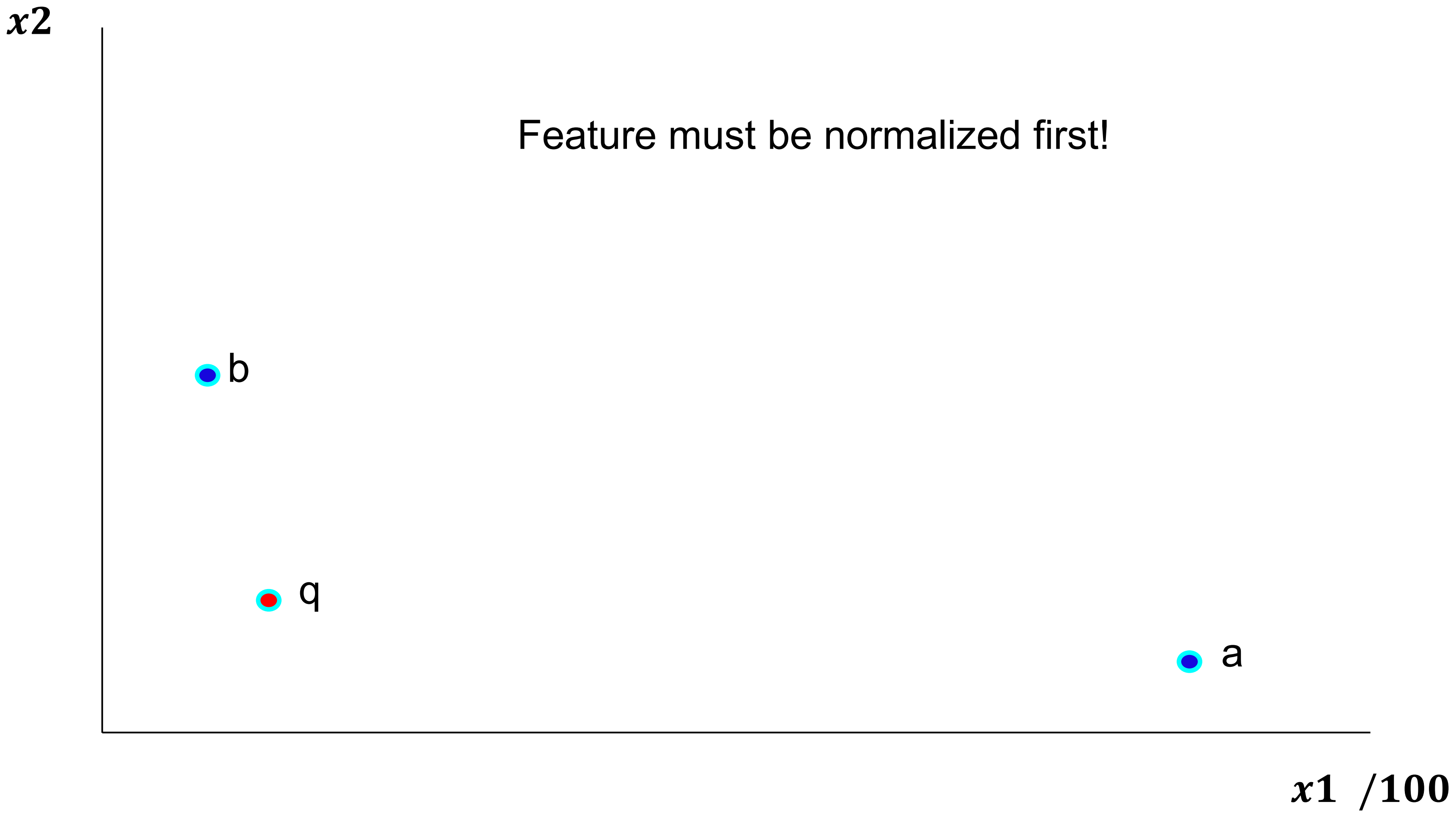
$$d(A,B)=\sqrt{\sum_{i=1}^{n}(A_i-B_i)^2}$$

    - Cosine distance

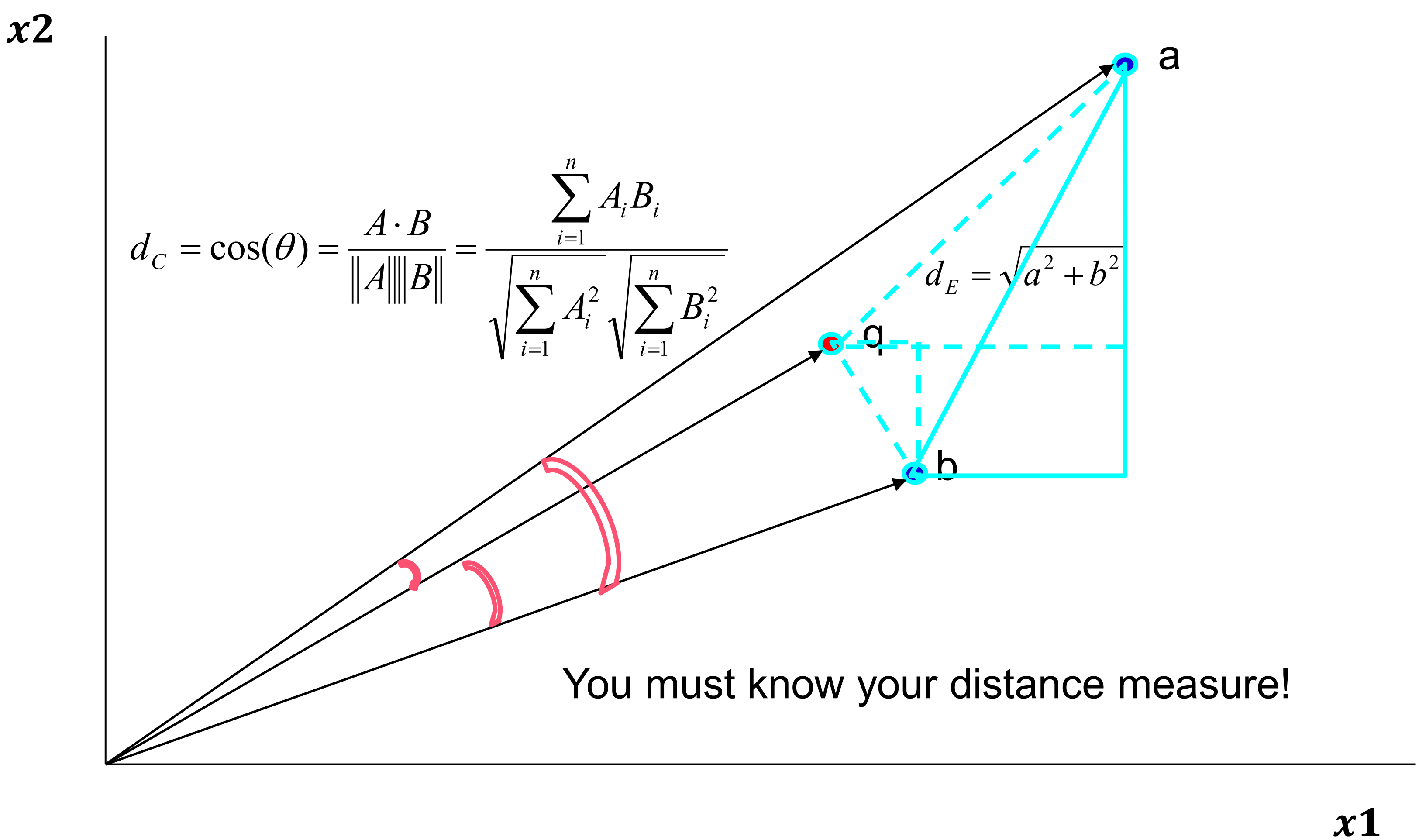
$$s(A,B)=\cos(\theta)=\frac{A\cdot B}{\|A\|\|B\|}=\frac{\sum_{i=1}^{n}A_i\times B_i}{\sqrt{\sum_{i=1}^{n}(A_i)^2}\sqrt{\sum_{i=1}^{n}(B_i)^2}}$$

Berenzweig, A., Logan, B., Ellis, D. P., and Whitman, B. P., *A Large-Scale Evaluation of Acoustic and Subjective Music-Similarity Measures*. Computer Music Journal. 28, 2004,

Which song is closer to my query (1)?



Which song is closer to my query (2)?



$$d_C = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}}$$

$$d_E = \sqrt{a^2 + b^2}$$

You must know your distance measure!

Audio Content-based Method

- Pros
  - Can deal with new songs with no or few texts.
  - Save human labors from annotating each song manually
- Cons
  - Complexity is relatively high.
  - The average performance is reaching the glass ceiling of around 65% in accuracy.
  - Features ≠ semantics:
    - Two songs with very similar features may **sound** very differently.
    - Two songs that sound very similar may have entirely different features
    - One song can even be processed in such a way that humans cannot tell the difference but the features are completely changed.

Outline

- Motivation
- Music Information Retrieval (MIR): an introduction
  - Collaborative filtering (CF)
  - Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  - Relevance Feedback
  - User Modeling
- Newer approach
- Conclusion

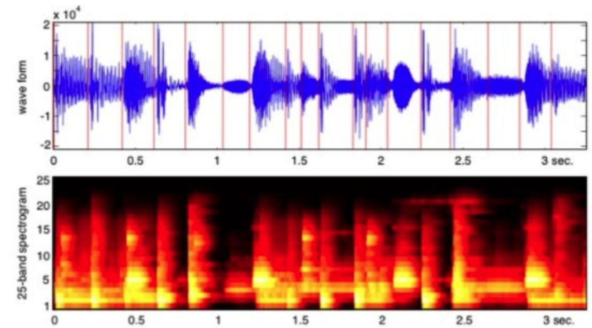
Hybrid Filtering

- Weighted (linear combination)
  - e.g., CF * 0.4 + CB * 0.6
- Cascade
  - e.g., First apply CF, then reorder by CB

Music

		U2	WP	AF	LT	...
	J	117	49	?	?	?
User	M	68	56	7	?	?
	P	?	55	?	6	?
	...	?	?	?	?	?

collaborative filtering

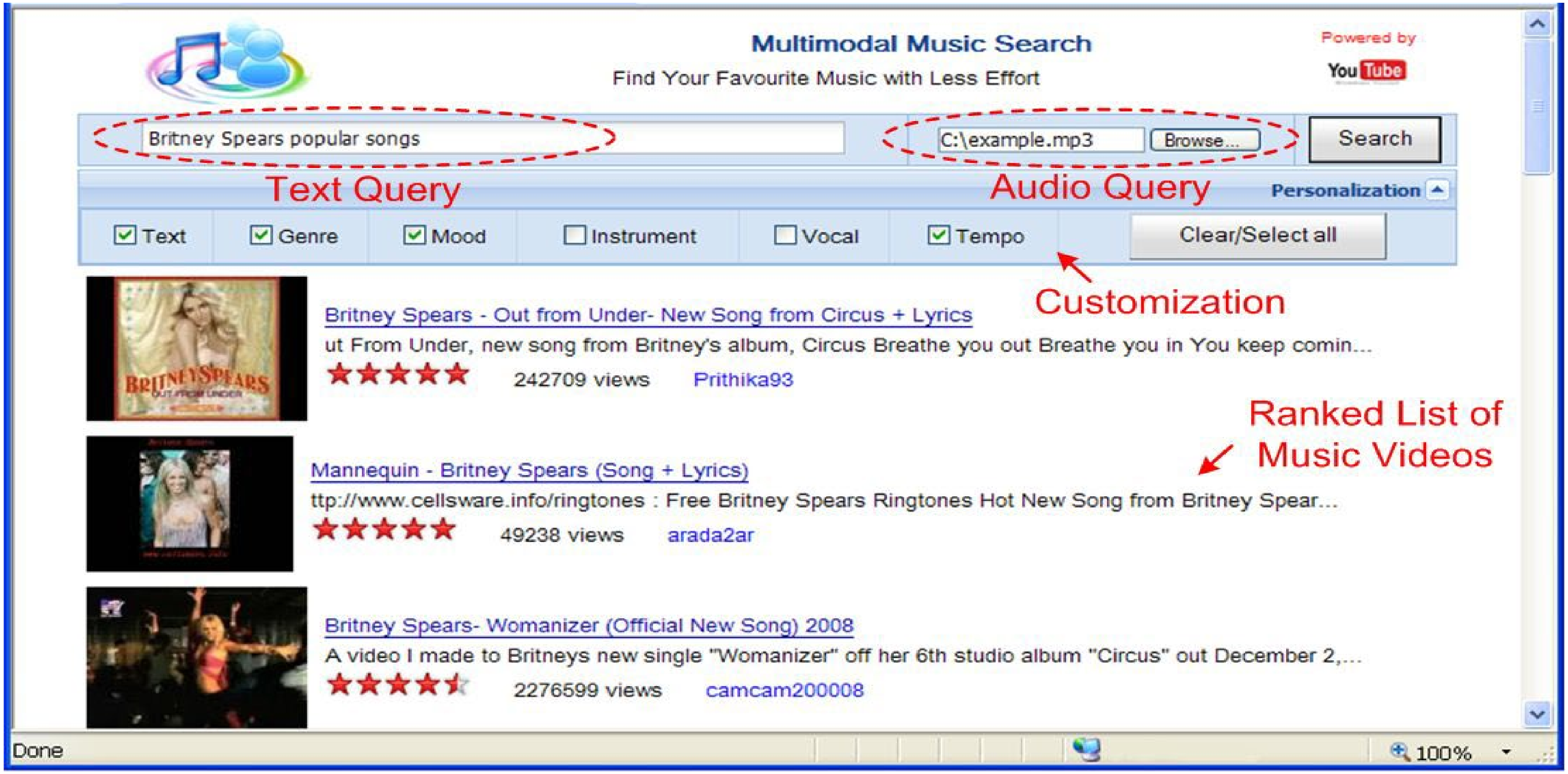


content-based

Hybrid

Multimodal Fusion Method

- An example



B. Zhang, J. Shen, Q. Xiang, and Y. Wang. *CompositMap: a novel framework for music similarity measure*. ACM Multimedia, 2009.

Multimodal Fusion Method

- Training

Music

Metedata
Semantic tags

Document Vector (DV)
…

Audio

Fuzzy music semantic vector (FMSV)
…

- Testing

Textual query — match — DV

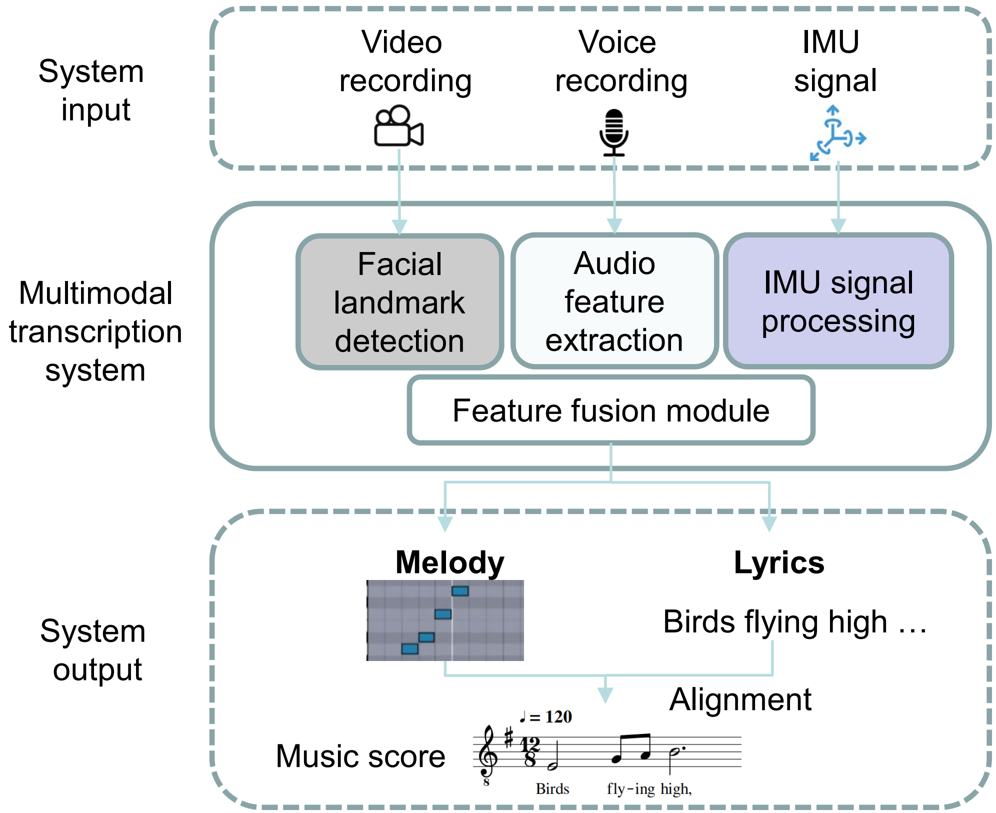
Audio query — Query FMSV — match — FMSVs of songs in database

Semantic query
- match — DV
- Query FMSV — match — FMSVs of songs in database

Fusion — Results

Multimodal Fusion Method

Questions related to a default group project

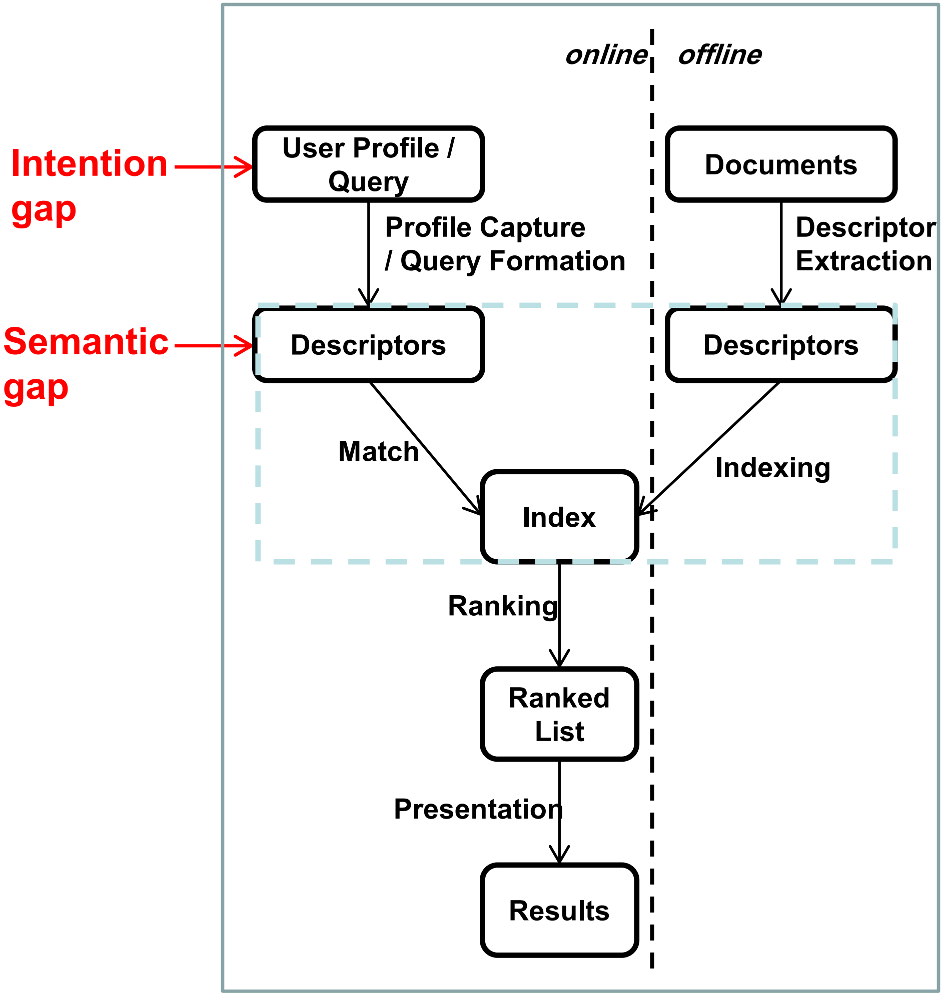


word war world

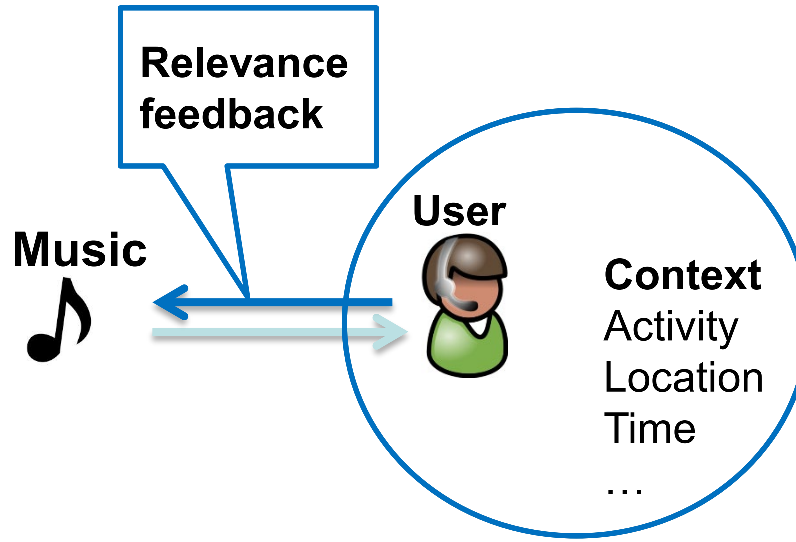
Outline

- Motivation
- Music Information Retrieval (MIR): an introduction
  - Collaborative filtering (CF)
  - Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  - Relevance Feedback
  - User Modeling
- Newer approach
- Conclusion

Relevant Issues

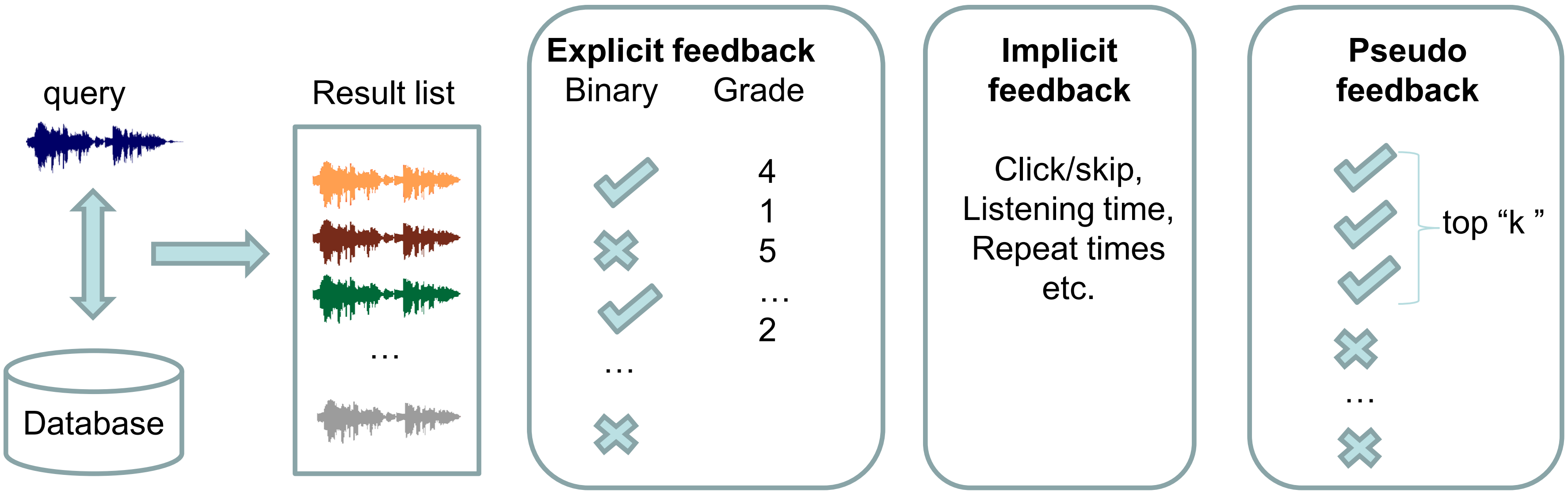


Intention gap: Intention ←→ query
Semantic gap:
descriptors ←→ semantic meaning of music



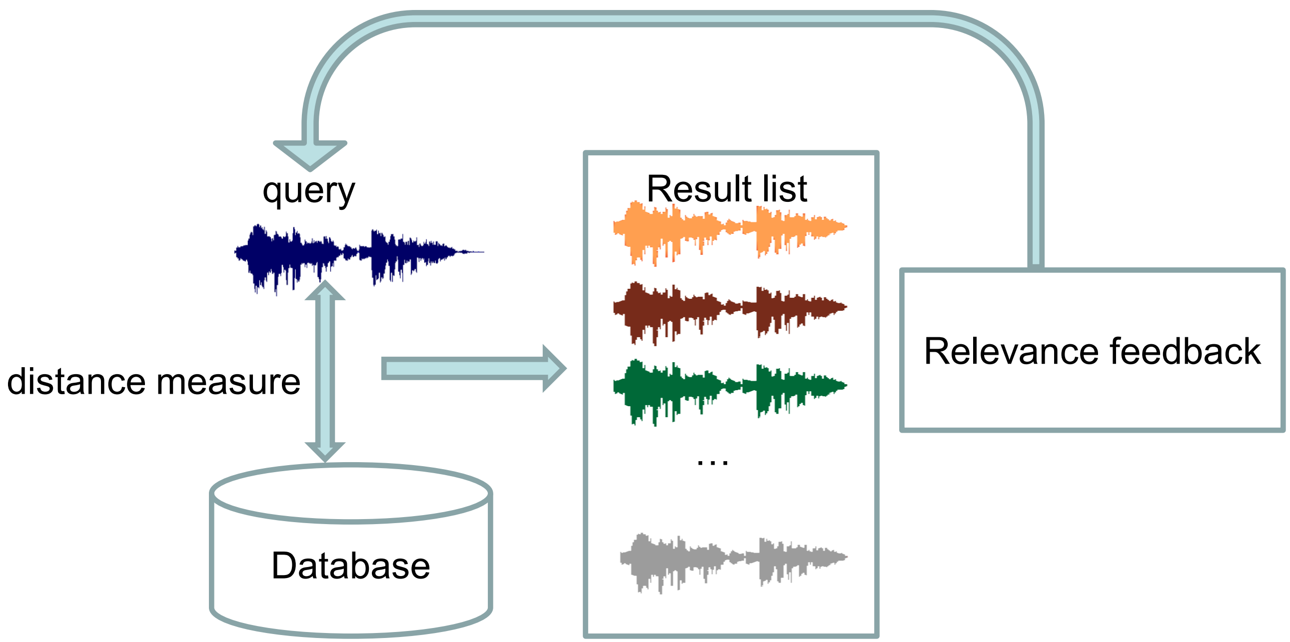
Relevance Feedback

- Explicit feedback: explicitly label relevance.
- Implicit feedback: inferred from user behavior.
- Pseudo feedback: an automatic method without user interaction
  - Assume the top "k" ranked songs are relevant.



Relevance Feedback

- The goal is to improve the retrieval accuracy.
  - **Query shifting ➔** better capture user's intention
  - **Feature re-weighting ➔** better distance measure (q🡐➔ songs)



Outline

- Motivation
- Music Information Retrieval (MIR): an introduction
  - Collaborative filtering (CF)
  - Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  - Relevance Feedback
  - User Modeling
- Newer approach
- Conclusion

User Profile

- A collection of data associated to a specific user

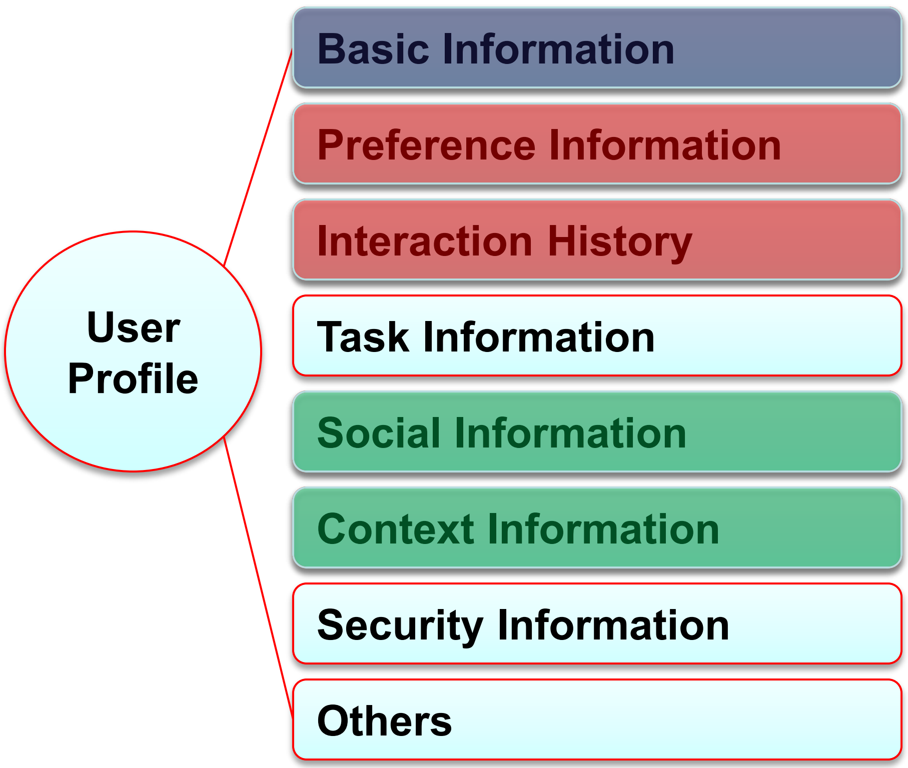
User Profile

Basic Information	•User's name, gender, age, …
Preference Information	•Content: high/low level features, …
Interaction History	•Log: query, ratings, …
Task Information	•Tasks or schedules of upcoming events, …
Social Information	•Friends, neighbors, … virtually
Context Information	•Location, environment , … in real life.
Security Information	•Security control of the user data.
Others	

- **Increasingly used in recommendations, adaptive hypermedia, web personalization and, etc.**

User Profile

- Currently Used Information



Relatively harder to obtain

Most frequently used

Getting more and more popular

Outline

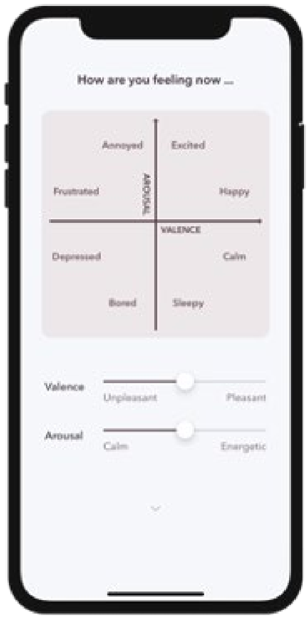
- Motivation
- Music Information Retrieval (MIR): an introduction
  - Collaborative filtering (CF)
  - Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  - Relevance Feedback
  - User Modeling
- Newer approach
- Conclusion

Newer Tech: Context-aware Method (1)

- Context categories
  - Environment-related context: irrelevant to users
    - e.g., location, time, weather, noise...
  - User-related context: relevant to listeners
    - psychology & sociology research driven data
    - e.g., demographics, personality traits, mood, cultural background, social context, music listening preference, music experience, activity

Reference: Schedl, M., Knees, P., McFee, B., Bogdanov, D., & Kaminskas, M. (2015). Music recommender systems. In *Recommender systems handbook* (pp. 453-492). Springer, Boston, MA.

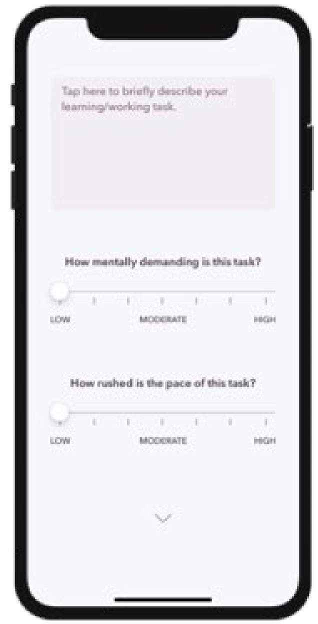
Newer Tech: Context-aware Method (2)



Emotion

Semantic space for emotion (Scherer, 2005)

Circumplex model of emotion (Russel, 1989)



Scenario

Task description

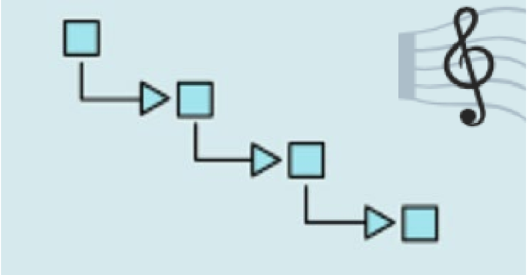
NASA task load index (Hart, 1988)

Xinxi Wang, David Rosenblum, Ye Wang, (2012) Context-Aware Mobile Music Recommendation for Daily Activities, ACM Multimedia 2012, Nara, Japan

Reference: Li, F., Hu, X., & Que, Y. (2020). Learning with background music: a field experiment. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 224-229).

Newer Tech: Sequence-aware Method (1)

- Motivation: Identify optimal sequence for recommendation
  - considerations of ordering constraints
- How to obtain ordering information?
  - Deal with domain knowledge
    - learn from music rotation rules
  - Mining user interactions with RecSys
    - e.g., Jack plays rock and roll on weekends for party, following classical music for study on weekdays.



sequence of music played

Newer Tech: Sequence-aware Method (2)

Often used approaches:

Frequent pattern mining

- rationale: identify patterns (e.g., subsequences) that occur frequently in a dataset
- e.g., learn from sequential co-occurrence of music tracks in playlists (Baccigalupo et al., 2008)
  → reduce the weights of tracks that are not adjacent

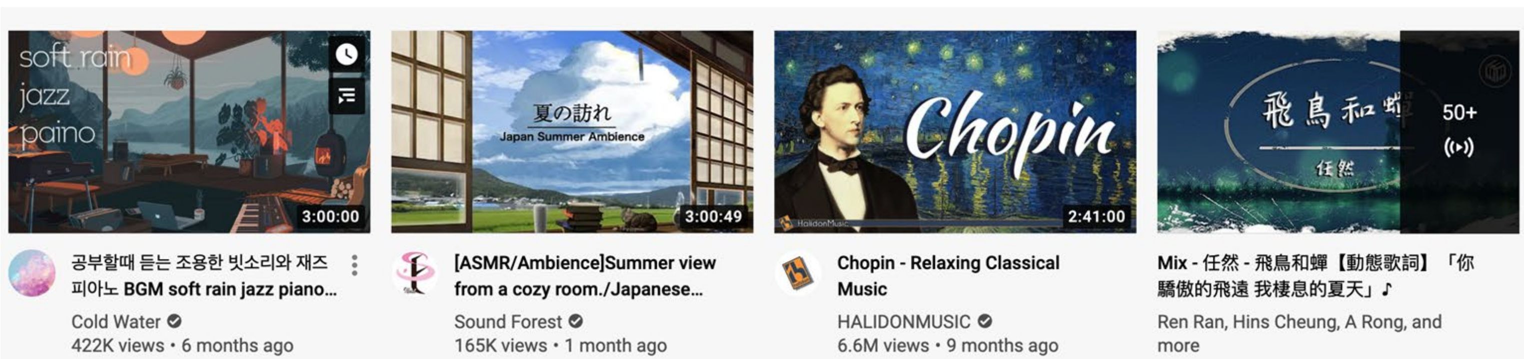
Markov model

- rationale: model transitions among sequential items
- e.g., model coherent music playlists as paths (Chen et al., 2012)
  → learn the positions of songs
  → represent transition probabilities by measuring distance between songs
  → predict next item based on transition probabilities

Transparency/Interpretability

- Transparent AI: Explainable AI: RecSys is a kind of AI
- explain how the system works
- increase the interpretability of the recommendation mechanism
- enhance users’ trust in RecSys

Potential question for Youtube: why am I recommended this?



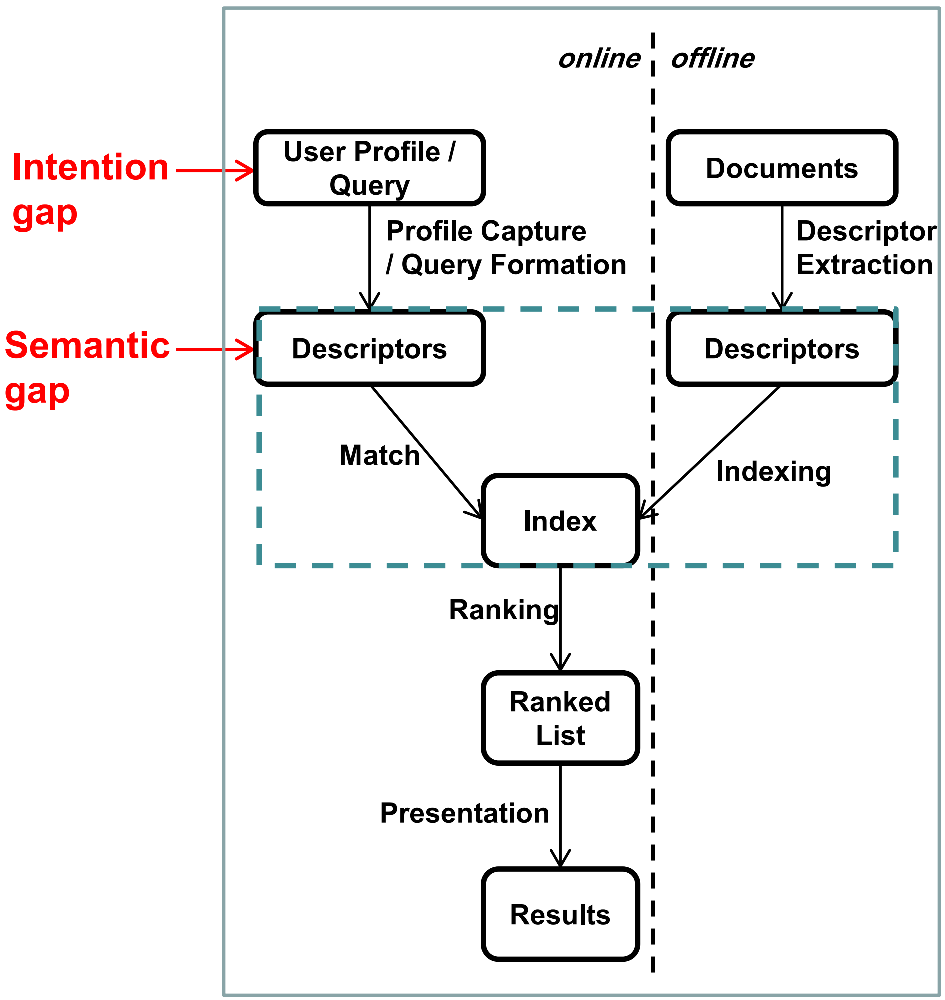
Outline

- Motivation
- Music Information Retrieval (MIR): an introduction
  - Collaborative filtering (CF)
  - Content Based filtering (CB)
    - Text-based Method
    - Audio Content-based Method
    - Multimodal Fusion Method
- Relevance Issues
  - Relevance Feedback
  - User Modeling
- Newer approach
- Conclusion

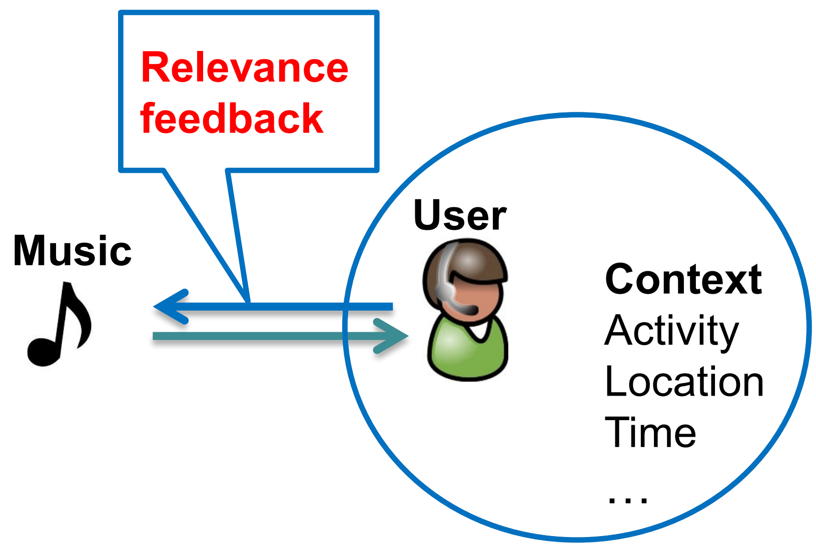
Conclusion

- What makes MIR so challenging?
  - Diverse users requirements with short & long term preferences
    - Context-aware
  - Complex music information:
    - Multimodal: audio, MIDI, score, metadata, social , …
    - Multicultural: e.g., modern art, Indian ragas, …
    - Multifaceted: melody, tempo, beat, …
    - …
- Better understanding of the users and music content
  - User modeling is a key
  - Bridging the semantic gap

Relevance Feedback (for MIR)



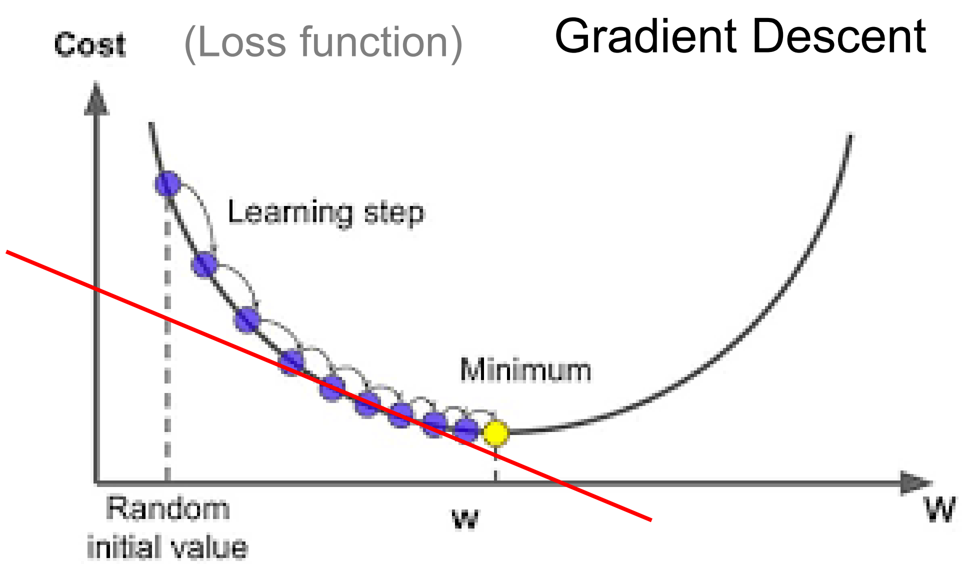
Intention gap: Intention ←→ query
Semantic gap:
descriptors ←→ semantic meaning of music



The rationales for the initial/midterm/final surveys after a major revision of the syllabus

✓ Fine-tune the content, delivery and assessment

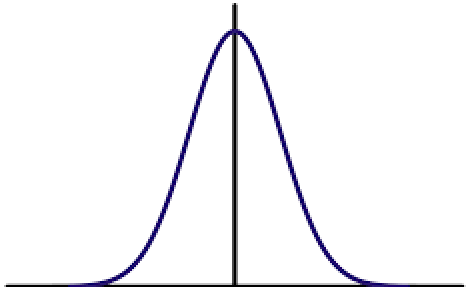
- Close the intention gap
- Give students tools for life
- Learn in ways that make sense
- Enable students to be successful in whatever they do



CS4347/CS5647

Please remember to give feedback to the course.

Comments are much more important than the score!



Department means: 4.2

Your feedbacks are important for the management to evaluate our teaching.

Your feedbacks are crucial for the teaching team to improve our teaching!

References from our lab

- Bingjun Zhang, Jialie Shen, Qiaoliang Xiang and Ye Wang, “CompositeMap: A Novel Framework for Music Similarity Measure,” ACM SIGIR2009, July 19-23, 2009, Boston, Massachusetts, USA
- Bingjun Zhang, Qiaoliang Xiang, Huanhuan Lu, Jialie Shen and Ye Wang, “Comprehensive Query-dependent Fusion using Regression-on-folksonomies: a Case Study of Multimodal Music Search, ” ACM Multimedia 2009, Beijing, China
- Zhonghua Li, Qiaoliang Xiang, Jason Hockman, Jianqing Yang, Yu Yi, Ichiro Fujinaga and Ye Wang, “A Music Search Engine for Therapeutic Gait Training,” ACM Multimedia 2010, Florence, Italy
- Zhendong Zhao, Xinxi Wang, Qiolaing Xiang, Andy Sarroff, Zhonghua Li and Ye Wang, “Large-scale Music Tag Recommendation with Explicit Multiple Attributes,” ACM Multimedia 2010, Florence, Italy
- Xinxi Wang, David Rosenblum, Ye Wang, Context-Aware Mobile Music Recommendation for Daily Activities, ACM Multimedia 2012, Nara, Japan
- Wang, Xinxi, and Ye Wang. “Improving Content-based and Hybrid Music Recommendation using Deep Learning.” ACM Multimedia 2014, Florida, USA
- Xinxi Wang, Yi Wang, David Hsu, and Ye Wang. “Exploration in Interactive Personalized Music Recommendation: A Reinforcement Learning Approach” ACM Trans. Multimedia Comput. Commun. Appl., Vol. 11, No. 1, Article 7. August 2014
- Zhe Xing, Xinxi Wang and Ye Wang. “Enhancing Collaborative Filtering Music Recommendation by Balancing Exploration and Exploitation” ISMIR 2014, Taipei, Taiwan

Other References (1)

ISMIR Tutorials in 2007 on music recommendation by Celma & Lamere

RecSys Tutorials in 2011 on music recommendation by Celma & Lamere

ISMIR Tutorials in 2018 on music recommendation by Schedl, Knees, Gouyon

CIKM Tutorials in 2019 on graph-based recommendation by Wang, He, Chua

Oramas, S., et al. (2017, August). A deep multimodal approach for cold-start music recommendation. In Proceedings of the 2nd workshop on deep learning for recommender systems (pp. 32-37).

Chen, S., et al. (2012, August). Playlist prediction via metric embedding. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 714-722).

Baccigalupo, C., et al. (2008). Uncovering Affinity of Artists to Multiple Genres from Social Behaviour Data. In ISMIR (pp. 275-280).

Fan, W., et al. (2019, May). Graph neural networks for social recommendation. In *The World Wide Web Conference* (pp. 417-426).

Hu, X., & Downie, J. S. (2010). Improving mood classification in music digital libraries by combining lyrics and audio. In Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, Gold Coast, QLD, Australia (pp.159-168).

Dong, Y., Yang, X., Zhao, X., & Li, J. (2019). Bidirectional Convolutional Recurrent Sparse Network (BCRSN): An Efficient Model for Music Emotion Recognition. IEEE Transactions on Multimedia, 21(12), 3150-3163.

Hu, X., Li, F., & Ng, J. (2018). On the relationships between music-induced emotion and physiological signals. In Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR), Sept. 23 - 27, 2018

Hu, X., Y. Que, N. Kando, W. Lian (2019). Analyzing user interactions with music information retrieval system: An eye tracking approach. In Proceedings of the 20th International Society for Music Information Retrieval Conference (ISMIR),

Que, Y., et al. (2021, April). Reading with and without Background Music: An Exploration with EEG, Eye Movement and Heart Rate. International Conference on Learning Analytics & Knowledge (LAK21), CrossMMLA Workshop.

Other References (2)

Hu, X., et al. (2019, March). Can Background Music Facilitate Learning? Preliminary Results on Reading Comprehension. In Proceedings of the 9th International Conference on Learning Analytics & Knowledge (pp. 101-105).

Li, F., et al. (2020, March). Learning with background music: a field experiment. In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (pp. 224-229).

Baccigalupo, C., Plaza, E., & Donaldson, J. (2008). Uncovering Affinity of Artists to Multiple Genres from Social Behaviour Data. In ISMIR (pp. 275-280).

Chen, S., Moore, J. L., Turnbull, D., & Joachims, T. (2012, August). Playlist prediction via metric embedding. In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 714-722).book recommendation (Hu, et al., 2020)

Hu, X., Ng, J. T. D., Yang, C. & Chu, S. K. W. (2020). Personalized book recommendation to young readers: Two online prototypes and a preliminary user evaluation, Accepted to Proceedings of the IEEE/ACM Joint Conference on Digital Libraries

Pu, P., Chen, L., & Hu, R. (2011, October). A user-centric evaluation framework for recommender systems. In Proceedings of the fifth ACM conference on Recommender systems (pp. 157-164).

Hu, X. (2019). Evaluating mobile music services in China: An exploration in user experience. Journal of Information Science, 45(1), 16-28.

Hu, X., et al. (2017). The MIREX Grand Challenge: A Framework of Holistic User Experience Evaluation in Music Information Retrieval, Journal of the Association for Information Science and Technology.

Lee, J.H., Hu, X., Choi, K., & Downie, J.S. (2015). MIREX grand challenge 2014 on user experience: Qualitative analysis of user feedback. In Proceedings of the 16th International Conference on Music Information Retrieval (ISMIR) (pp. 1–7). Malaga, Spain: ISMIR.