Context Aware Music Recommendation

Ethan Hamer

hamer073@morris.umn.edu
April 18, 2020
University of Minnesota Morris Computer Science Senior Seminar
Division of Science and Mathematics
University of Minnesota, Morris
Morris, Minnesota

Motivation

With increased mobile device and music streaming usage more people are listening to music in more diverse spaces

With this increased diversity comes the challenge of recommending music that fits a users current context

Outline

- Background
- ContextPlay System
  - Jin et al
- Venue Music System
  - Cheng and Shen
- Conclusion

Outline

- **Background**
  - Context Awareness
  - Bag of Audio Words
  - Topic Models
- ContextPlay System
- Venue Music System
- Conclusion

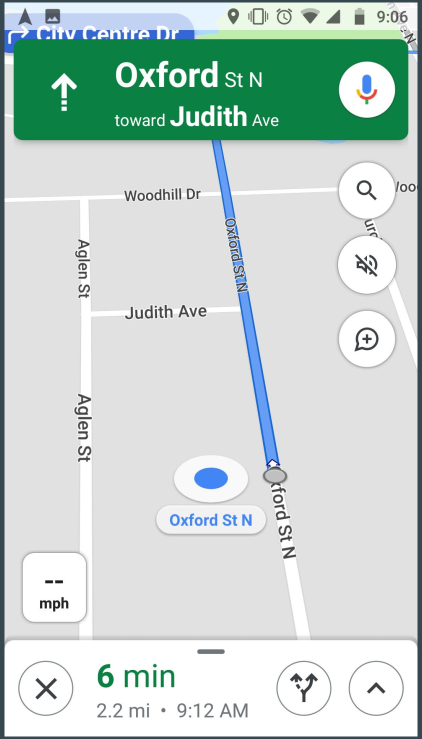
Background - Context Awareness

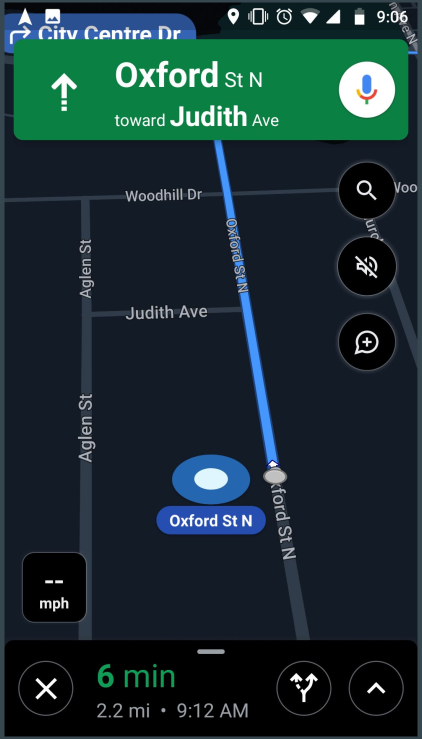
Obtaining a current context

- Time
- Location
- Etc

Using that context to perform different actions

- Turning on a dark theme at night
- Choosing specific music based on current location
- Etc





Background - Bag of Audio Words

Representation of the contents of an audio stream

1. Find all occurrences of chosen audio features
2. Keep a count of each feature
3. Most common features likely best representation of audio stream

Background - Topic Models

Representation of the *topics* in an audio stream

1. Uses a fixed vocabulary
   a. Consists of potential *topics* and *concepts* related to those topics
2. Look at the *concepts* in the audio stream
3. Link *concepts* to their topics
4. Gives a probabilistic model of topics

Outline

- Background
- **ContextPlay System**
  - ContextPlay User Interface
  - Experiment Setup and Results
- Venue Music System
- Conclusion

ContextPlay - Motivation

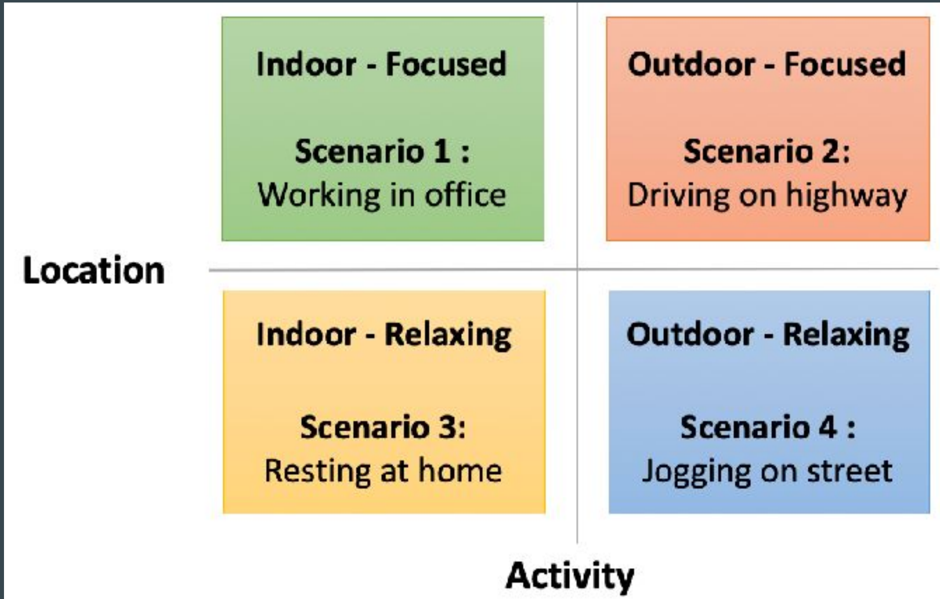
Goal: To study how different aspects of context control affect user perceptions of a system

ContextPlay - Experiment Setup

Created 4 Scenarios

- Based on two factors
  - Location
  - Activity Level

Gathered participants through Amazon Mechanical Turk



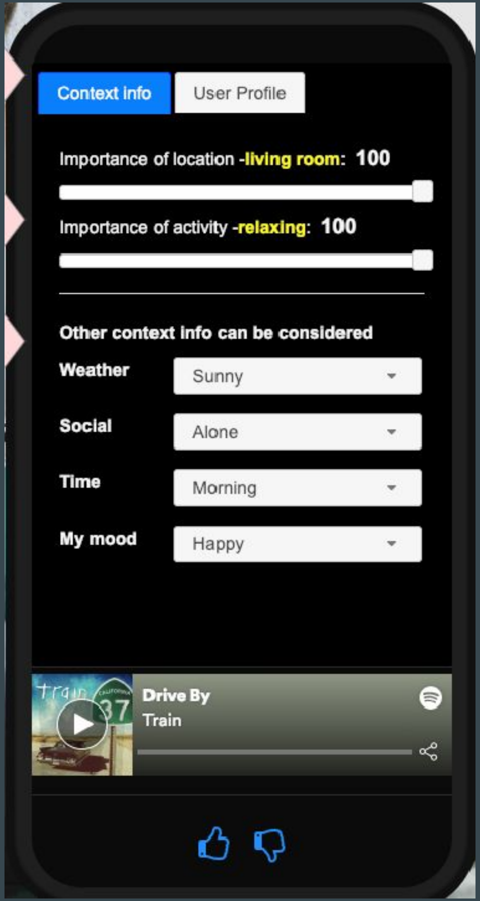
Y. Jin, N. N. Htun, N. Tintarev, and K. Verbert

ContextPlay - User Interface

Simulated Mobile Phone

Two Screens

1. Context Info Section
    a. Change weighting of location and activity
    b. Change other context characteristics
        i. Mood, weather, time, and social
2. User Profile Screen
    a. Personal contextual information
        i. Favorite artists, tracks, etc



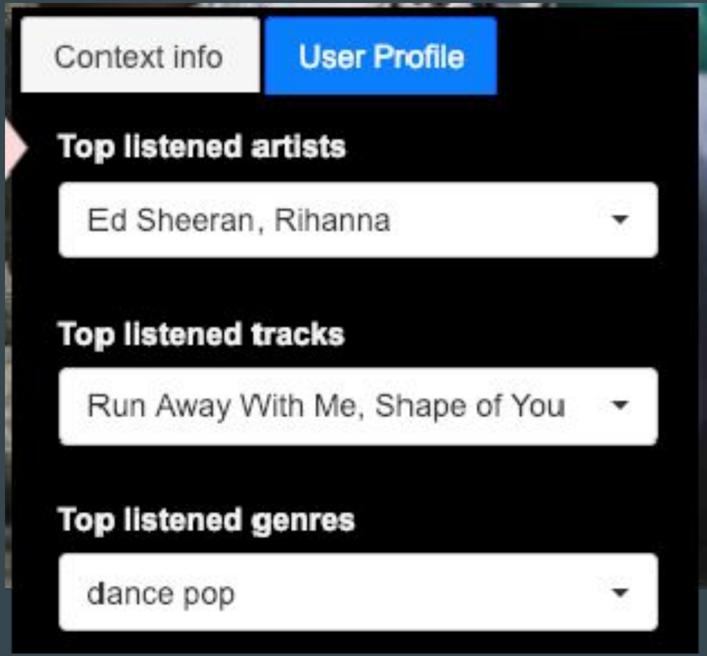
Y. Jin, N. N. Htun, N. Tintarev, and K. Verbert

ContextPlay - User Interface

Simulated Mobile Phone

Two Screens

1. Context Info Section
   a. Change weighting of location and activity
   b. Change other context characteristics
      i. Mood, weather, time, and social
2. User Profile Screen
   a. Personal contextual information
      i. Favorite artists, tracks, etc



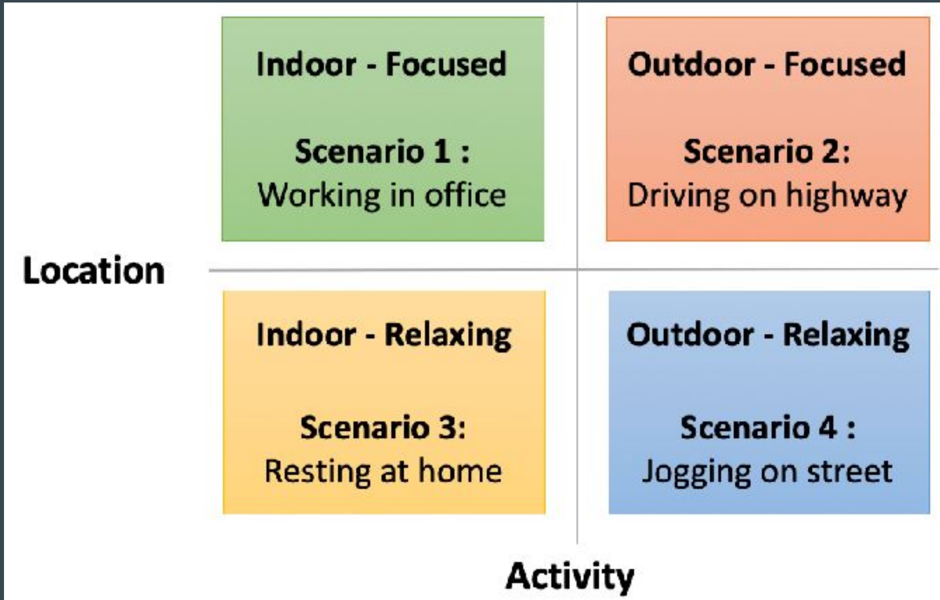
Y. Jin, N. N. Htun, N. Tintarev, and K. Verbert

ContextPlay - Experiment Setup

Mixed Design User Study with Two Designs

1. User has control over context controls
2. User doesn't control the contexts

Take a demographic survey before experiment

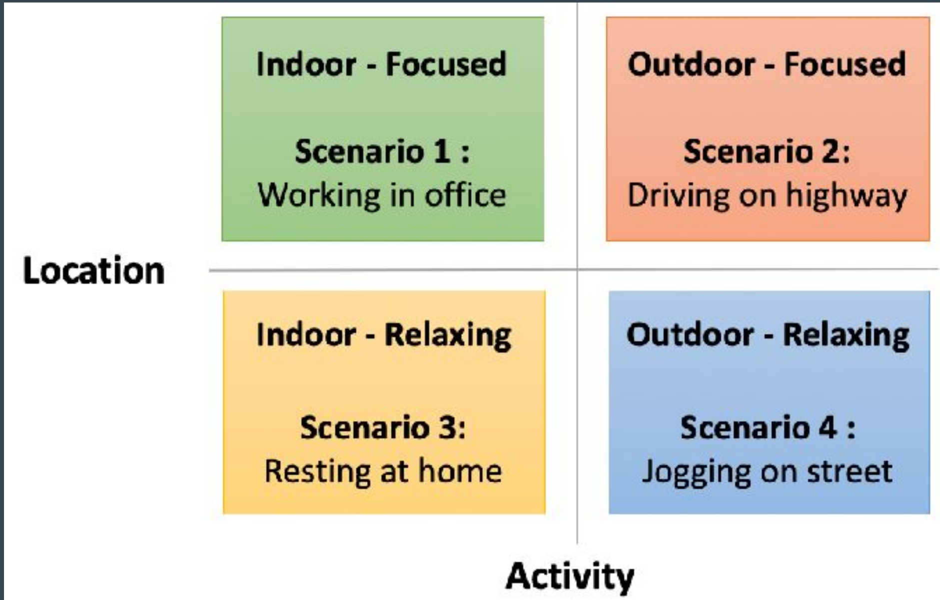


Y. Jin, N. N. Htun, N. Tintarev, and K. Verbert

ContextPlay - Experiment

Participants asked to find 5 good matches for the current scenario

1. If available, given time to adjust context controls
2. Listen to 30 seconds of a song
3. After 20 seconds, given option to rate
4. Once 5 songs are rated well, system resets and moves on to next scenario



Y. Jin, N. N. Htun, N. Tintarev, and K. Verbert

ContextPlay - Post Experiment

Participants are given a post-study questionnaire

Looks at:

- Perceived Music Quality
- Perceived Musical Diversity
- System Effectiveness
- Cognitive Load

ContextPlay - Results

Control Settings have a direct correlation with perceived quality

This in turn increases perceived effectiveness, allowing users to finish the task quicker

Control Settings have no significant effect on perceived diversity and cognitive load

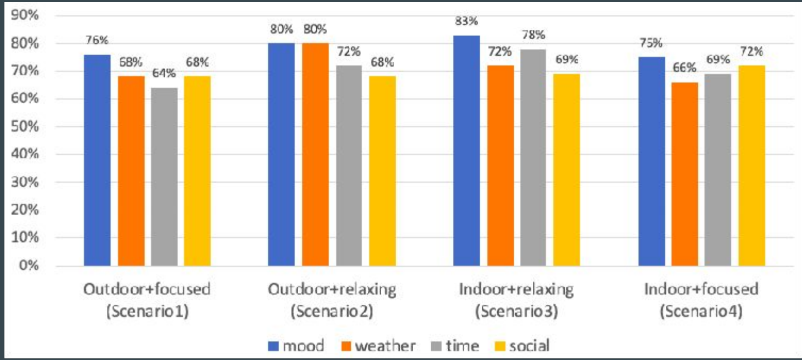


Figure 5: The percentage of users who controlled each context information in four scenarios.

Y. Jin, N. N. Htun, N. Tintarev, and K. Verbert

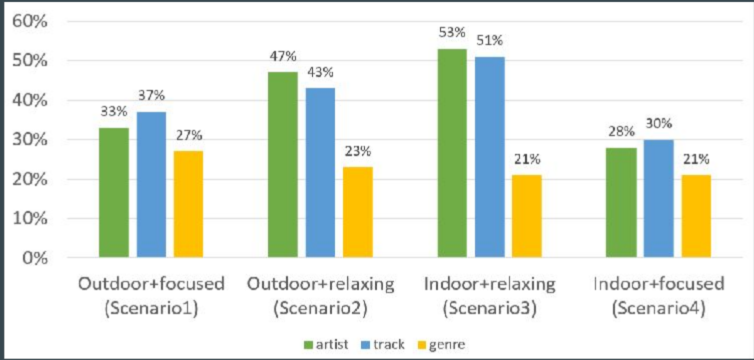


Figure 6: The percentage of users who controlled each user prof le element in four scenarios.

Y. Jin, N. N. Htun, N. Tintarev, and K. Verbert

Outline

- Background
- ContextPlay System
- **Venue Music System**
  - Music Concept Sequence Generation
  - Location Aware Topic Model
  - Experimental Results
- Conclusion

VenueMusic - Music Concept Sequence Generation

Overview

- Audio Partitioning
  - Fixed Length Windows
  - Music Segment Detection
- Audio Feature Extraction
- Feature Probability Estimation
- Feature Filtering

Music Concept Sequence Generation - Partitioning

Fixed length windows

- Choose an interval ahead of time e.g. 5 seconds
- Split the music into chunks of the chosen interval size e.g. 5 seconds chunks

This is the method used in the paper

Music Concept Sequence Generation - Partitioning

Music segment detection

- Detects sections of the music e.g. phrase endings, verses, etc
- Splits the music on these sections.

Music Concept Sequence Generation - Feature Extraction

Feature Extraction is done using publicly available toolboxes

Looking for 4 audio feature types:

1. Timbral Features
   a. What sounds are in the audio partition
2. Spectral Features
   a. Analyzing the actual waveform of the audio
   b. 70 total dimensions analyzed
3. Rhythmic Features
4. Temporal Features

The extracted features then become the *concepts* used throughout the rest of the process

Music Concept Sequence Generation - Probability Estimation

Goal is to estimate how likely a music *concept* is to be in a segment

First define n *music dimensions* to look at

- Dimension examples include genre, mood, instrumentation, etc.

Music Concept Sequence Generation - Probability Estimation

Each *dimension* is then made up of a number of *concepts*

As an example, the instrumentation dimension could be made up of 3 concepts

1. Trombone
2. Guitar
3. Tuba

Music Concept Sequence Generation - Probability Estimation

With the 3 *concepts* in the instrumentation *dimension* we can then build probabilities

Each *concept* is assigned a probability that the *concept* is in a given segment

1. Trombone - 65%
2. Guitar - 22%
3. Tuba - 80%

Music Concept Sequence Generation - Feature Filtering

Threshold filtering

1. Define a probability threshold value for each audio feature type
   a. Cheng and Shen determine these threshold values through experimentation
2. Remove any features whose probability falls below that threshold value

Continuing the instrumentation example from before

If we give a threshold of 30%, guitar would be removed from the instrumentation *dimension*, keeping trombone and tuba

Music Concept Sequence Generation - Feature Filtering

Frequency Filtering

1. Construct all possible *concept patterns*
   a. Take a concept from every music dimension in a segment
2. Determine the frequency of each concept pattern
3. Compare created concept patterns to pre-determined lists of frequent and infrequent concept patterns
4. Remove the concepts that most often in the infrequent patterns or least often in the frequent patterns

VenueMusic - Location Aware Topic Model

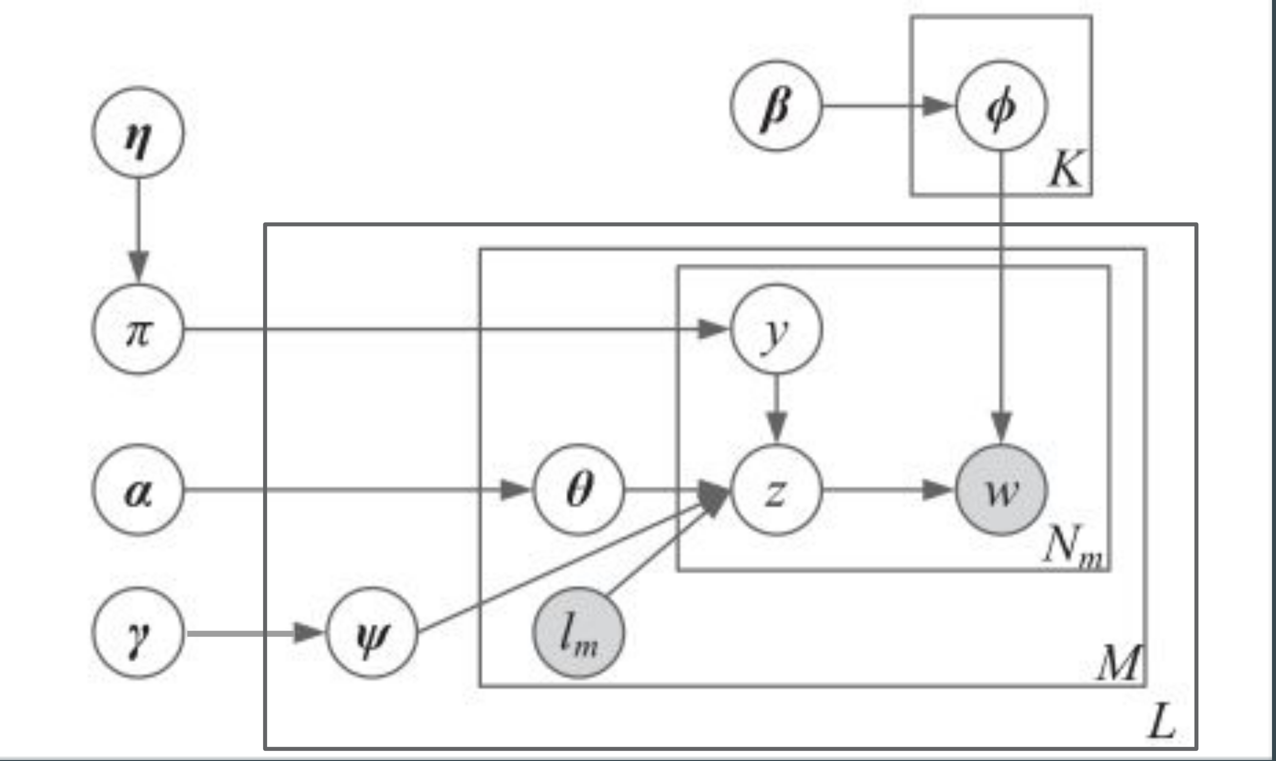
Goal

- Model the relation between songs and venues

Songs and venues are represented as a mixture of *topics*

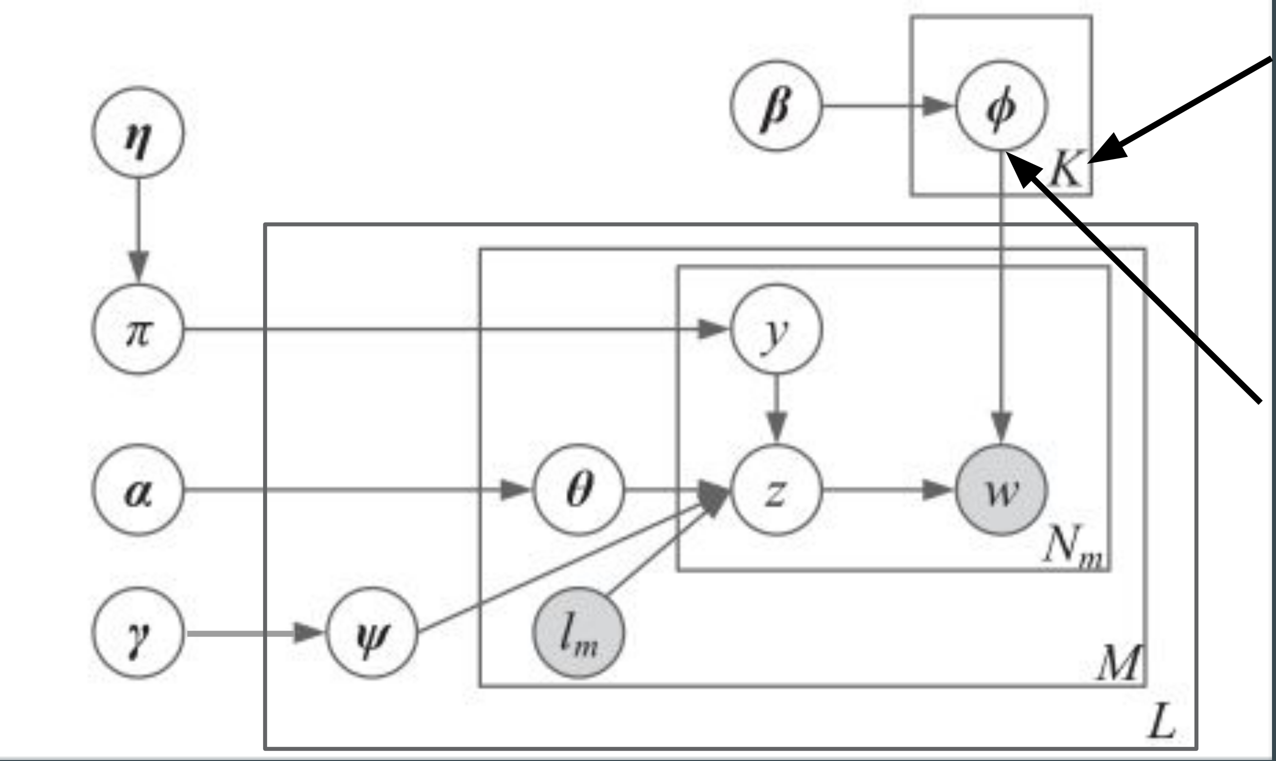
Topics are represented as a mixture of music *concepts*

VenueMusic - Location Aware Topic Model



Cheng and Shen

VenueMusic - Location Aware Topic Model

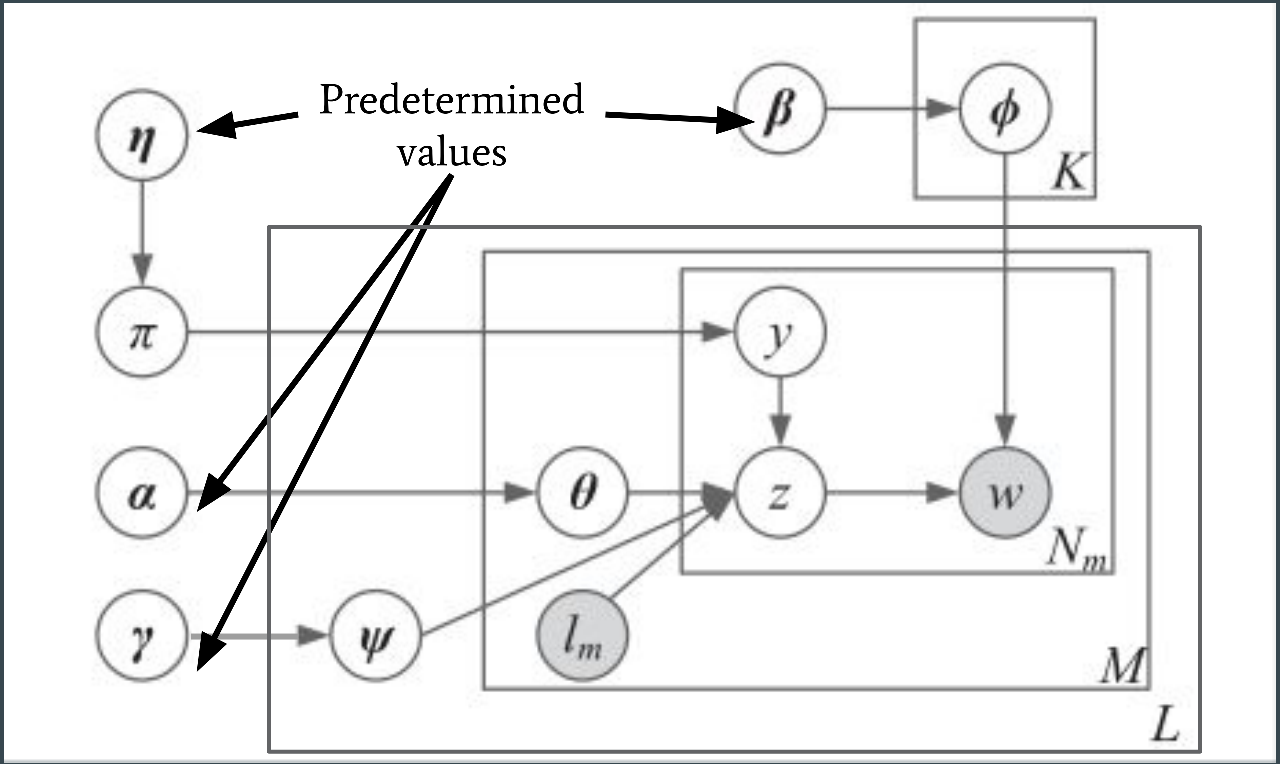


Total number of *topics*

Distribution of words specific to *topic k*

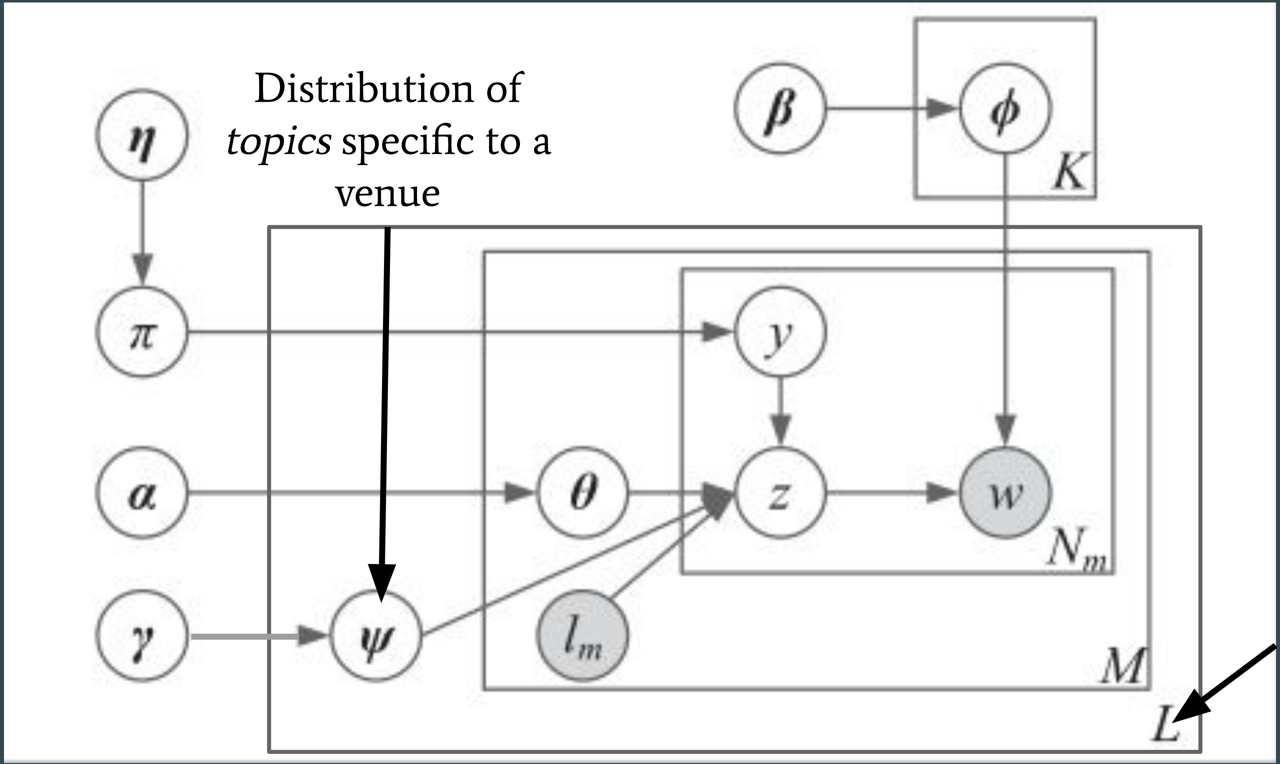
Cheng and Shen

VenueMusic - Location Aware Topic Model



Cheng and Shen

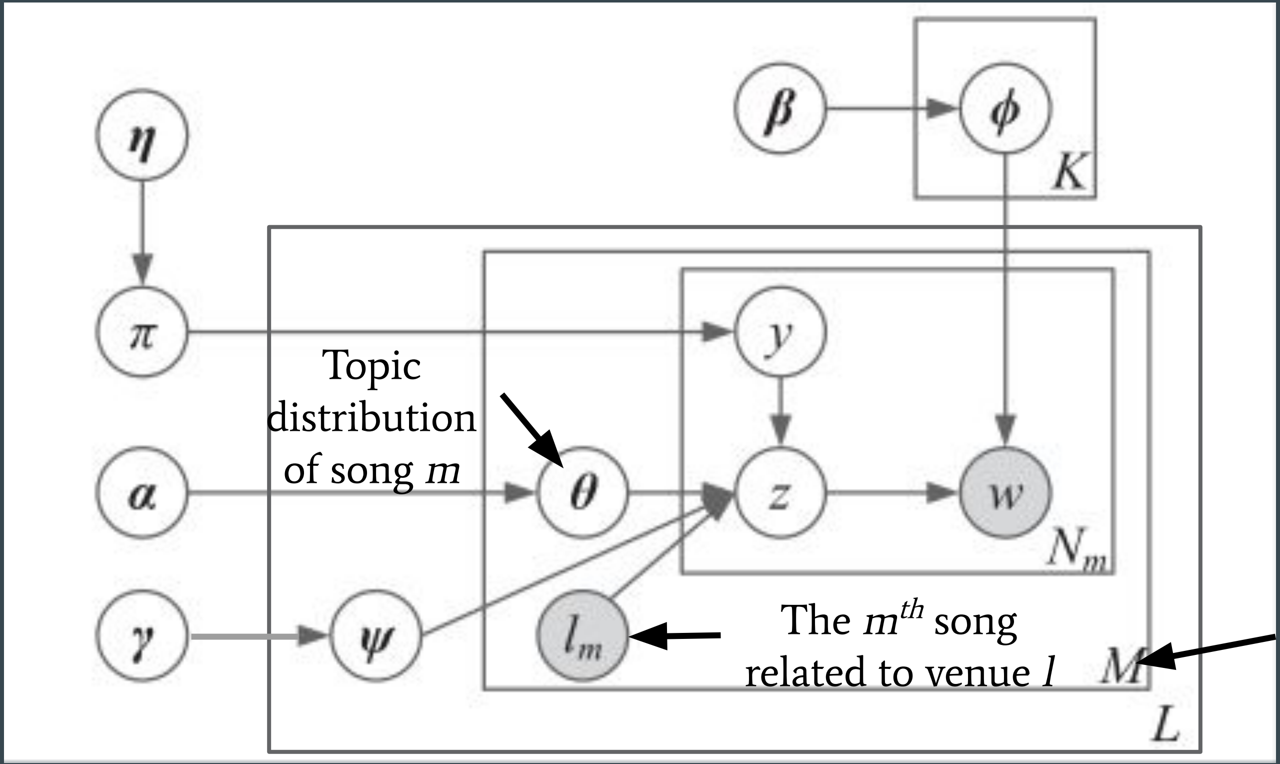
VenueMusic - Location Aware Topic Model



Total number of *venue* types

Cheng and Shen

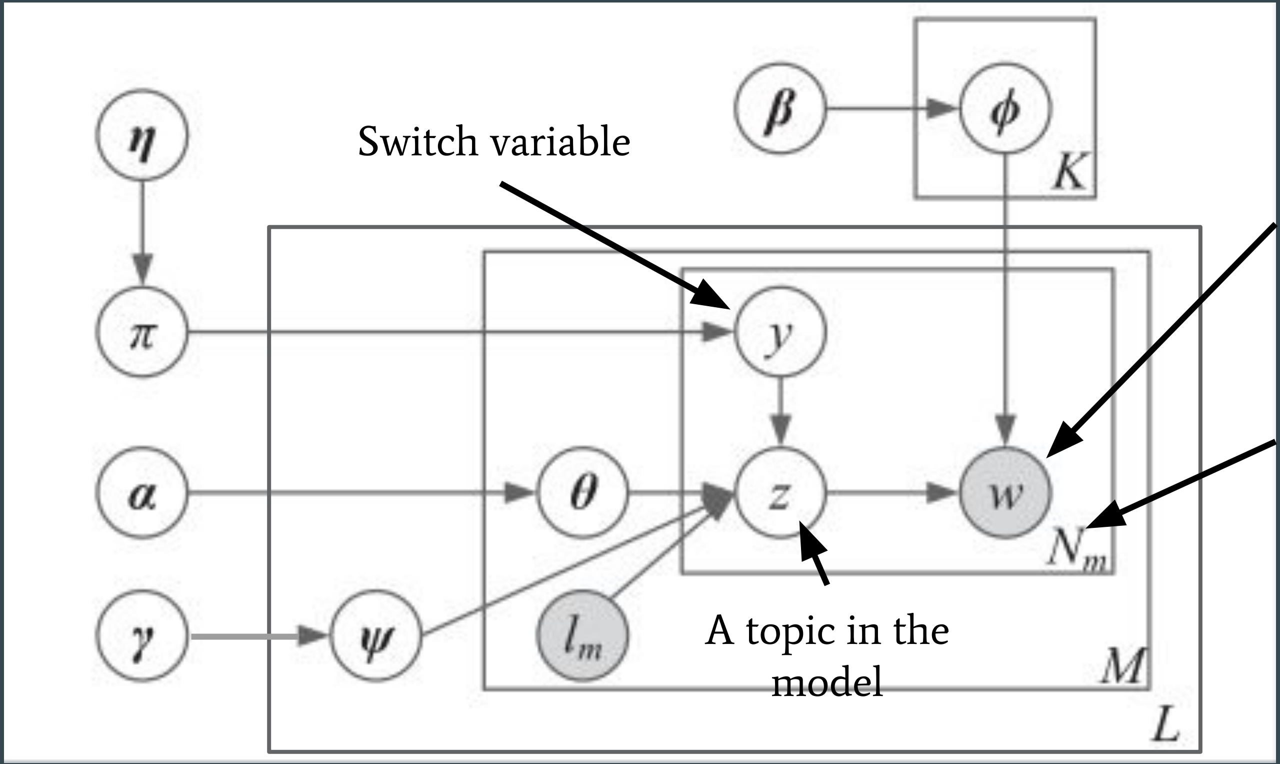
VenueMusic - Location Aware Topic Model



Total number of songs

Cheng and Shen

VenueMusic - Location Aware Topic Model



A music *concept* in song *m*

Number of *concepts* in song *m*

Cheng and Shen

VenueMusic - Location Aware Topic Model

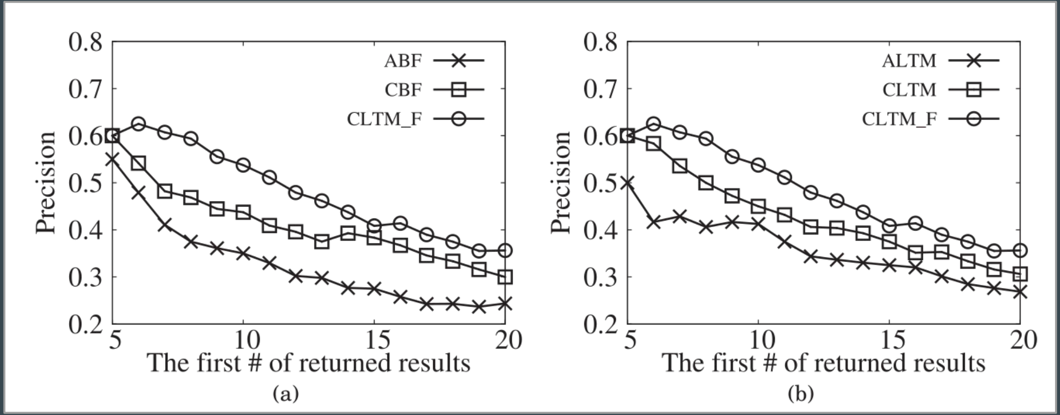
Once the model is built it can then be used to generalize to unknown songs

This is done through repeated sampling of potential *topics* in the new song, updating the *topics* fit to the new song on every pass

VenueMusic - Results

Found that compared to existing methods, VenueMusic offered higher precision when generating topics

The topics generated better captured associations between music and venus



Conclusion

As Jin et al show control over context and it's importance greatly improves user satisfaction with a system

With Sheng and Chen's paper we can see that there's a lot of work to still be done with recommender systems, especially music recommendation

Acknowledgements

Many thanks to both Nic McPhee and KK Lamberty for providing excellent guidance

Thanks to my family and girlfriend Clare for putting up with me through this process and encouraging me throughout

Questions?

Resources

Plate Notation: https://en.wikipedia.org/wiki/Plate_notation

ContextPlay: https://doi.org/10.1145/3320435.3320445

VenueMusic: https://doi.org/10.1145/2846092