


Beauty and Joy of Computing:

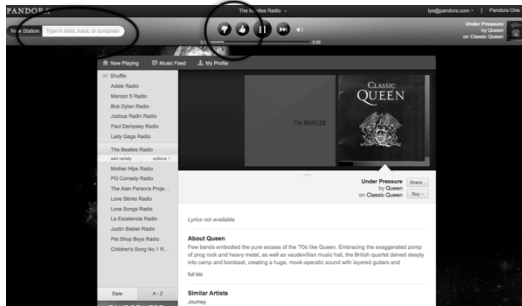
internet radio

Tao Ye
Pandora Playlist Engineering

PANDORA MEDIA, INC.



What is Pandora?



PANDORA MEDIA, INC. CS10 Fall 2011

Joy: Pandora's Mission

- ◆ **What:**
 - ◆ Enrich people's lives with the music they love (one station at a time)
 - ◆ Not song on demand
- ◆ **How:**
 - ◆ Curated tracks (900,000 '+' 15 million)
 - ◆ Analyzed by music analysts (Music Genome Project)
 - ◆ Streamed as Internet radio stations, everywhere
 - ◆ Personalized by users through thumbs
- ◆ **Business: legal 'freemium' music service**
 - ◆ Royalty payments to Sound Exchange
 - ◆ Advertising revenue
 - ◆ Subscription revenue


PANDORA MEDIA, INC. CS10 Fall 2011

Beauty: Technology

- ◆ Recommend relevant and awesome songs
- ◆ Scale streaming service to 100 million users
- ◆ Deal with devices

PANDORA MEDIA, INC. CS10 Fall 2011

Recommendation System



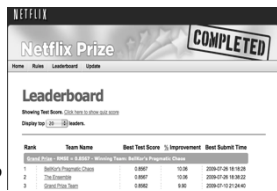
Why Rec Sys?

- ◆ Too much information, too little time
 - ◆ What do you really want?
 - ◆ Based on what you told us so far...
- ◆ They're important
 - ◆ Amazon (More than 1/3 of product sales result from recs)
 - ◆ Google News (Recommendations generate 38% more click-throughs)
 - ◆ Netflix (2/3 of movies rented are recommended)
 - ◆ Beer

PANDORA MEDIA, INC. CS10 Fall 2011

Clicker Question

- How much was the Netflix prize?
 - A. \$100,000
 - B. \$500,000
 - C. \$1,000,000
 - D. \$2,000,000
 - E. I have never heard of this.
- Goal: improves RMSE by 10%
- Simon Funk blogged about SVD early and it was used by every team
- Teams combined at the end



PANDORA MEDIA, INC.

CS10 Fall 2011

Problems Pandora is Solving

- Music recommendations is special
 - Long tail
 - ~3 min per song, low 'cost'/item
 - RMSE not the only goal!
 - Continuous, current, diverse
- Playlist
 - Different people
 - Different stations



PANDORA MEDIA, INC.

CS10 Fall 2011

Pandora's Recommender

- Content based: Music Genome Project
 - ~400 genes (attributes) rated by analysts on a 10 point scale
- Example: Adele's Someone Like You
 - Sad lyrics
 - It's about love
 - Voice has vibrato
 - Has alternative influence
 - Has soul influence
 - ...
- Distance between two songs: weighted sum of diff in each attribute
- Shortest distances win

PANDORA MEDIA, INC.

CS10 Fall 2011

Pandora's Recommender

- Crowd based: Collaborative Filtering
 - Crowd vote with collective thumbs
 - Stations evolve over time
 - Tenured listeners hear less popular songs
- Personalization: Real Time
 - Thumb up brings in songs similar to the thumbed one
 - Thumbed down bans a song on station

PANDORA MEDIA, INC.

CS10 Fall 2011

Let's Hear it



PANDORA MEDIA, INC.

CS10 Fall 2011

Should Play?



A

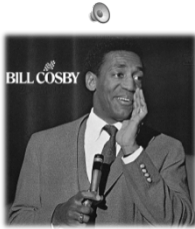


B

PANDORA MEDIA, INC.

CS10 Fall 2011

Content IS Important



- ◆ Requires clean comedy
- ◆ Content based filter: The explicit gene

PANDORA MEDIA, INC.

CS10 Fall 2011

What I Do

- ◆ Measure playlists
 - ◆ Relevance / Accuracy
 - ◆ Diversity
 - ◆ Novelty and Serendipity
 - ◆ Awesomeness
 - ◆ Currency
- ◆ Improve the genomic matches with crowd opinion



Image source: iso.org

PANDORA MEDIA, INC.

CS10 Fall 2011

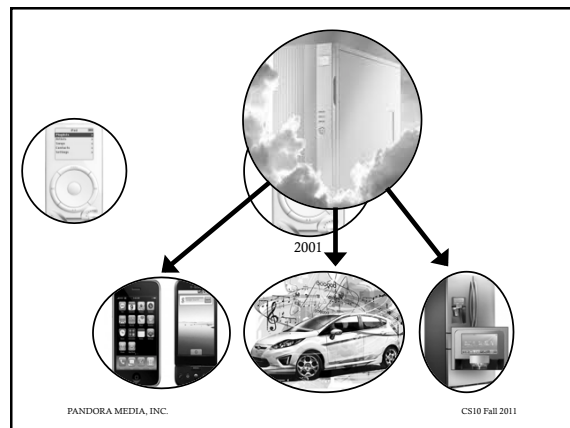
Scaling in the Cloud



Copyright: Matthew Sarnoff, msarnoff.org

PANDORA MEDIA, INC.

CS10 Fall 2011



PANDORA MEDIA, INC.

CS10 Fall 2011

Numbers

- ◆ 100,000,000 registered users
- ◆ 10,000,000 active daily users
- ◆ 900,000 songs, in multiple encoding formats
- ◆ 10,000,000,000 thumbs

PANDORA MEDIA, INC.

CS10 Fall 2011

Clicker Question

- ◆ How many media servers does Pandora have?
 - A. 100 - 500
 - B. 500 - 1000
 - C. 1000 - 2000
 - D. > 2000
 - E. I have no idea

PANDORA MEDIA, INC.

CS10 Fall 2011

Big Data: Streaming not Storage

- ◆ Take advantage of data locality
 - ◆ Small percentage of songs are played a large number of times
- ◆ Tiered media servers
 - ◆ Treat the cloud as one big computer
 - ◆ Most played songs in memory
 - ◆ Next tier on local hard disks
 - ◆ Next in the SAN

PANDORA MEDIA, INC.

CS10 Fall 2011

Future

- ◆ More listener personalization
- ◆ Better support for indie artists
- ◆ Broader music selection
- ◆ Everywhere

PANDORA MEDIA, INC.

CS 10 Fall 2011