Artist Recommendations Based on Event Data

Amy Ruskin
MUMT 621

Motivation

- Festival line-ups and shared bills provide relationships between diverse sets of artists
- Aggregating event data → different pool of artists than recommendations based on user listening behavior?
- Get experience using APIs, different Python libraries

Image source

Example of diverse line-up at Bol' Festival 2019

Background: Other Uses of Event Data

- “VenueRank: Identifying Venues that Contribute to Artist Popularity” (Krasanakis et al. 2018)
  - Event data from Facebook, artist information from Spotify
- “Mining and Forecasting Career Trajectories of Music Artists” (Arakelyan et al. 2018)
  - Event data from Songkick
  - Three tasks: forecasting artist success, predicting the venues at which an artist will perform, and joint discovery of influential artists and venues
- “Evaluating Recommender System Algorithms for Generating Local Music Playlists” (Akimchuk et al. 2019)
  - Event data from Ticketfly and Facebook

Background: MusicLynx

- Graph-based artist discovery system
- Artists linked by relationships mined from sources like MusicBrainz, DBpedia, AcousticBrainz, Last.fm, and Wikidata
- Color-coded graph for the user to explore—for example:
  - red = “Other Musical Groups Disestablished in 2003”
  - pink = “Other English Alternative Rock Groups”’
  - orange = “Similar Artists by Timbre”

(Allik et al. 2018)

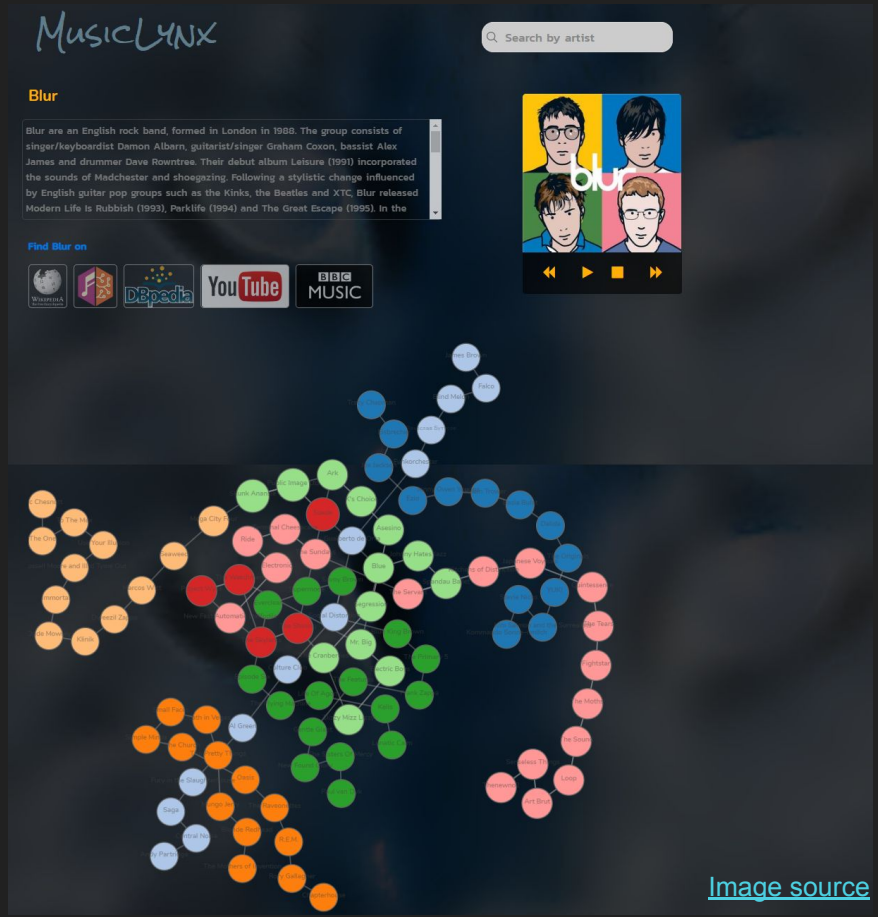

Image source

Graph of artists colored by relationship to Blur

Main Goal: Recommendation System

Deliverable

Interface where user inputs artist name, receives list of recommended artists

Basic approach

- Pull events for query artist
- Pull events occurring at venues of query artist's events
- Rank artists in resulting dataset by number of shared venues

Challenges

- Querying APIs is slow, but creating local copy of databases out-of-scope
- Combining data from multiple sources
  - Handling different metadata schemas
  - Matching same entities represented with different IDs
  - Using only existing API methods

Sources of Event Data

	MusicBrainz	Setlist.fm
Coverage and scope	Founded 2000 Initially collected CD information, Event entities added in 2015 As of 6 April 2020, 41,437 events[1]	Founded 2008[2] “wiki-like service to collect and share setlists”[3] As of 6 April 2020, 4,936,340 events[4]
API	No key required Rate limits unclear—50 requests per second?[5]	Need to register for key Default limit of 2 requests per second and 1440 max per day, with option to upgrade
Event representation	1 event object with multiple artists Few required fields (e.g., can add event without associated venue)	Separate event object per artist Stricter schema

[1] https://musicbrainz.org/statistics
[2] https://www.crunchbase.com/organization/setlist-fm#section-overview
[3] https://www.setlist.fm/about
[4] https://www.setlist.fm/
[5] https://musicbrainz.org/doc/XML_Web_Service/Rate_Limiting

Subgoals

Visualization

- Artist-venue graph?
- Map?
- Something like MusicLynx?

Data Analysis

- Explore difference in coverage, distribution of artists/venues/events by source
- For example:
  - Valhalla (200-capacity venue in New Zealand): 247 events in MusicBrainz vs. 148 events in Setlist.fm
  - Madison Square Garden: 127 events in MusicBrainz vs. 4058 events in Setlist.fm

References

Akimchuk, Daniel, Timothy Clerico, and Douglas Turnbull. 2019. “Evaluating Recommender System Algorithms for Generating Local Music Playlists.” https://arxiv.org/abs/1907.08687

Allik, Alo, Florian Thalmann, and Mark Sandler. 2018. “MusicLynx: Exploring Music Through Artist Similarity Graphs.” In *Companion Proceedings of the The Web Conference 2018 (WWW ’18)*, 167–70. https://doi.org/10.1145/3184558.3186970

Arakelyan, Shushan, Fred Morstatter, Margaret Martin, Emilio Ferrara, and Aram Galstyan. 2018. “Mining and Forecasting Career Trajectories of Music Artists.” In *HT ’18: 29th ACM Conference on Hypertext and Social Media*. https://doi.org/10.1145/3209542.3209554

Krasanakis, Emmanouil, Emmanouil Schinas, Symeon Papadopoulos, Yiannis Kompatsiaris, and Pericles Mitkas. 2018. “VenueRank: Identifying Venues that Contribute to Artist Popularity.” In *Proceedings of the 19th International Society for Music Information Retrieval Conference (ISMIR 2018)*, 702–8. https://doi.org/10.5281/zenodo.1492513

Swartz, Aaron. 2002. “Musicbrainz: A Semantic Web Service.” *IEEE Intelligent Systems 17* (1): 76–7.