Spotify Artist Network Analysis: Community Detection and Centrality Insights

Abhir Iyer
abhiyer@iu.edu
03/16/2025

Abstract

This project analyzes personal music preferences by constructing a Spotify Playlist and Artist Network. The objective was to discover hidden patterns and communities among favorite artists based on their co-occurrence in user-created playlists. Data collection was automated using the Spotify API via Python and Spotipy, resulting in a dataset containing playlists, songs, and corresponding artists. Network analysis methods, including community detection (Louvain algorithm) and centrality analysis, were implemented to uncover artist clusters and influential nodes. Findings revealed distinct artist communities reflecting genre-based or stylistic similarities and identify key influential artists within the network, such as Guns N' Roses and Queen.

1. Introduction and Background

Music streaming platforms like Spotify allow users to curate personalized playlists, reflecting individual music tastes and preferences. Analyzing these playlists through network analysis methods can reveal insights into relationships among artists based on listening habits. This project explores these relationships by creating an artist co-occurrence network constructed from personal Spotify playlists. This network provides insights into underlying patterns, communities, and influential artists within personal playlists.

2. Data Collection

The dataset was collected programmatically using Spotify's API with the Spotipy Python library. Authentication was conducted through Spotify Developer credentials, and playlist data was retrieved, including playlist names, artist names, and song titles. The dataset comprised multiple playlists with a total of 232 unique artists and 15,191 connections.

The network construction clearly defined:

- Nodes: Unique artists appearing across playlists.
- Edges: Co-occurrences of artists within the same playlists, with weights indicating the frequency of these occurrences.

To enhance clarity, the network was filtered to include only edges representing artists appearing together in at least two playlists, resulting in a simplified network of 36 nodes and 216 edges.

3. Analysis

The analysis applied two primary network methodologies:

- Community Detection: The Louvain algorithm identified four distinct communities reflecting genre or stylistic similarities among artists. The visualization of these communities clearly demonstrated the network structure and meaningful artist clusters.

- Centrality Analysis: Degree centrality was calculated to identify influential artists within the network. Guns N' Roses and Queen emerged as the top influential artists, each with a degree centrality score of 0.971, followed by Nirvana at 0.657. These results indicate artists who frequently appear in playlists with numerous other artists, signifying their central role.

The interactive network visualization enhanced clarity and interpretability, highlighting distinct musical communities and pinpointing influential artists.

4. Conclusion and Discussion

This network analysis of Spotify playlist data revealed clear patterns in artist co-occurrences, artist communities, and influential figures within the user's music preferences. The identified communities offer insights into underlying genre and style preferences, and the centrality analysis provided clarity on artists who significantly influence the structure of the playlists. Future analyses could include temporal analysis to examine how playlist structure and artist popularity evolve over time.

5. References

[1] Spotify. (2024). Web API Documentation. Retrieved from https://developer.spotify.com/documentation/web-api/

[2] Spotipy Documentation. (2024). Spotipy: A lightweight Python library for the Spotify Web API. Retrieved from https://spotipy.readthedocs.io/

The complete dataset and the Jupyter notebook with code implementation are available on GitHub: https://github.com/abhiriyer/SMM-Assignment-2