Spotify Artist Feature Collaboration Network

A COMPREHENSIVE ANALYSIS

Daniele Borghesi

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

UNIVERSITÀ DI PISA

Introduction

WHAT IS THE SPOTIFY ARTIST FEATURE COLLABORATION NETWORK?

What is the Spotify artist feature collaboration network?

Data on approximately 150,000 artists

Over 300,000 collaboration edges, illustrating the interconnectedness of artists in the music industry

Enables in-depth analysis of artist collaborations and network dynamics within the music industry

Data preparation

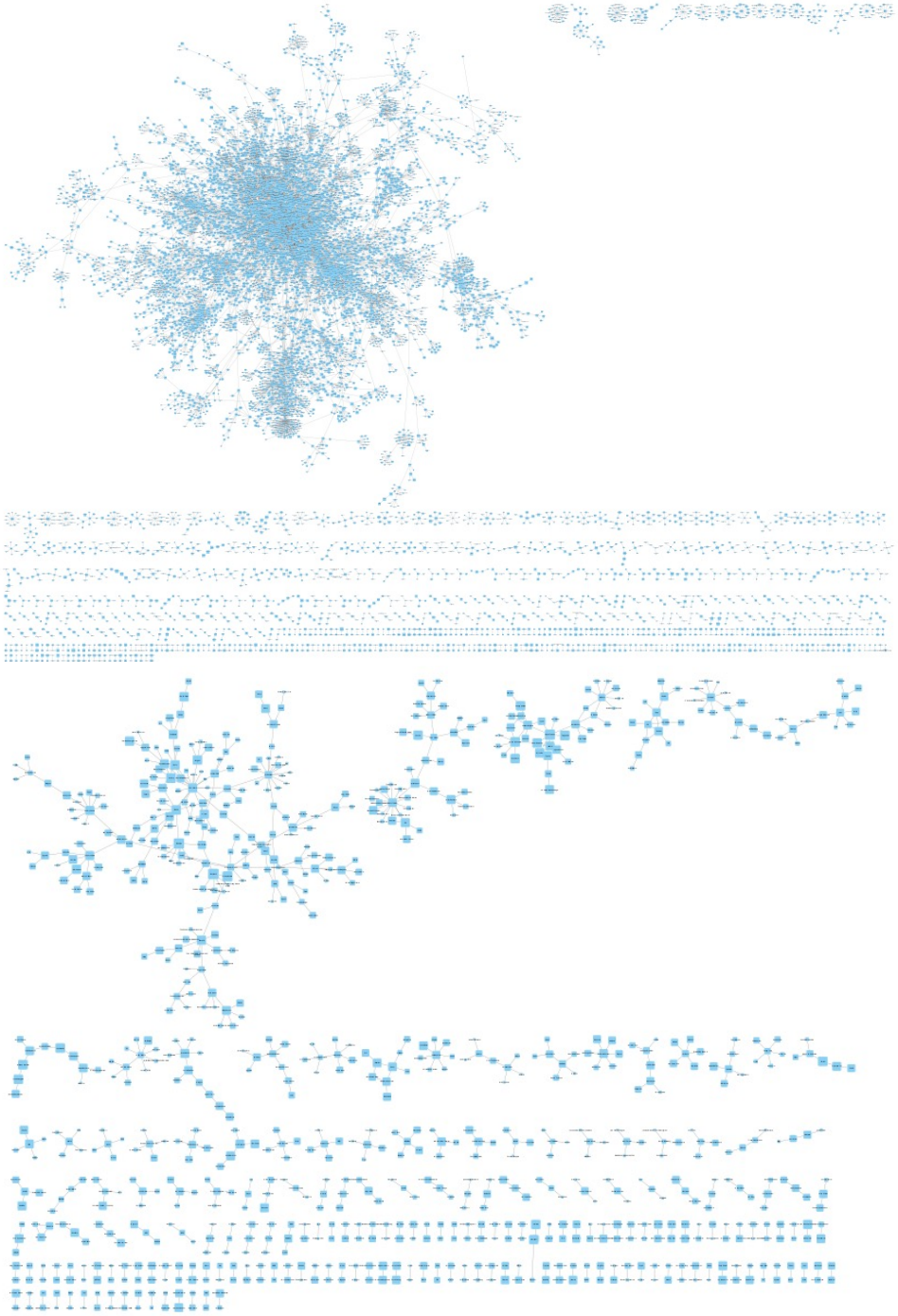
THE PROBLEM OF NETWORK SIZE

The problem of network size

With almost 150,000 nodes and over 300,000 edges, performing the analysis on the original network is extremely slow and difficult on a common hardware

Obtaining usable representations of the network can also be extremely challenging

To solve the problem, we need to simplify the network reducin its size



Sampling the network

Sample 1 (project part 1)

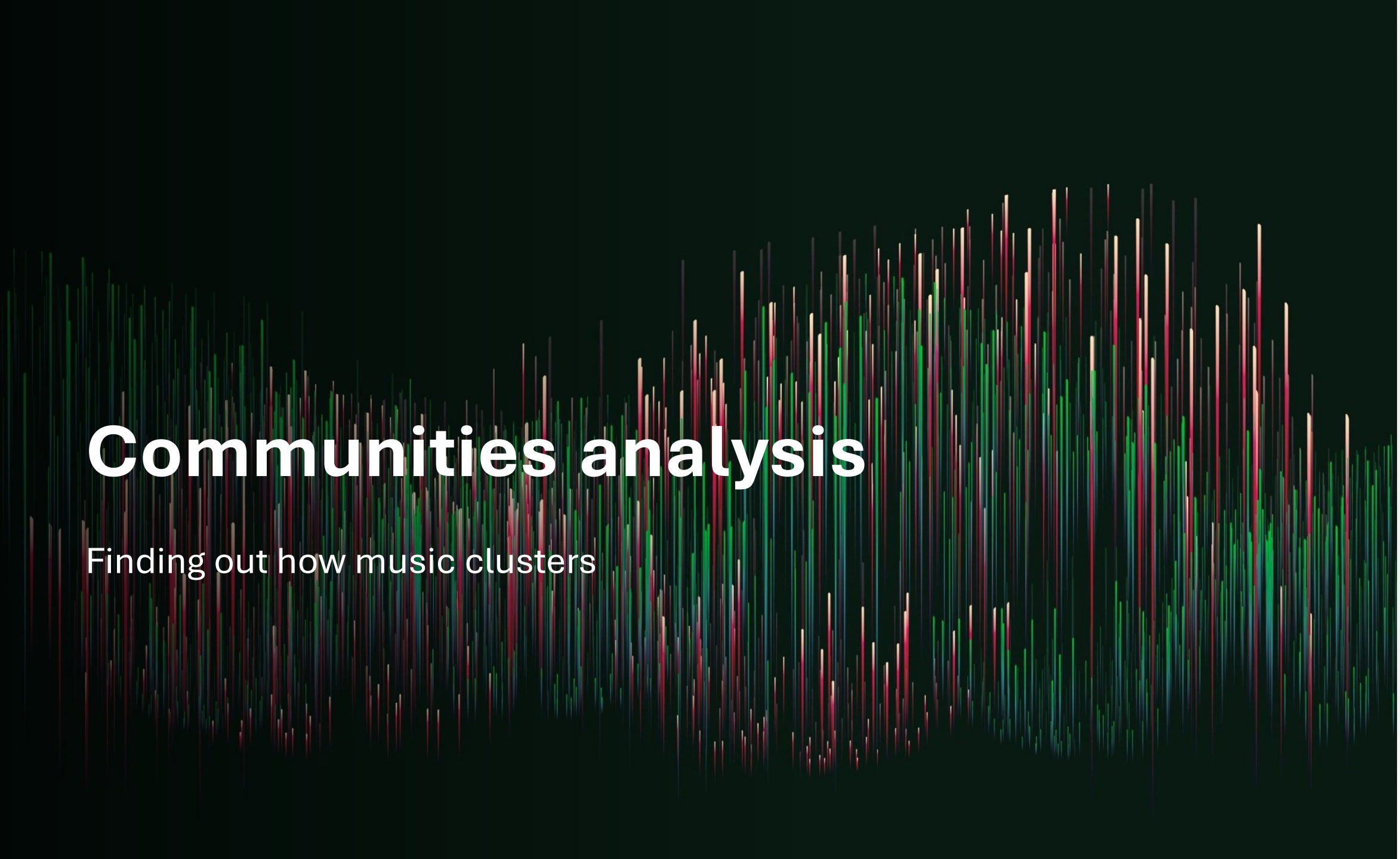
About *10,000* nodes and *12,000* edges. Used for preliminary analyses.

Sample 2 (project part 2)

About *900* nodes and *700* edges. Used for communities analysis and fault tolerance

Most central nodes (artists)

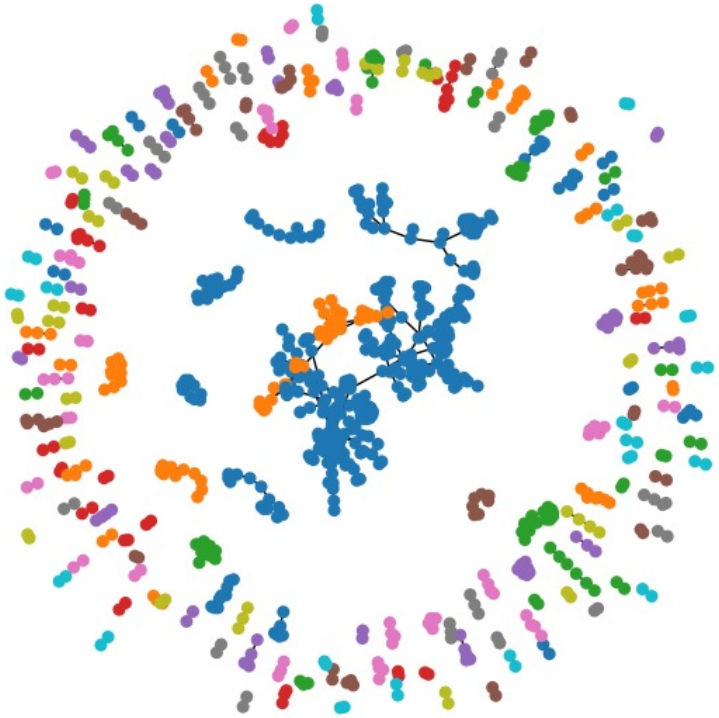
	Degree Centrality	Closeness Centrality	Betweenness Centrality	PageRank	Hubs	Authorities	Eigenvector Centrality
Mc Gw	0.0186	-	-	0.0056	0.037	0.037	0.5929
Diplo	0.0108	0.2278	0.2075	0.0028	-	-	-
Snoop Dogg	0.0091	0.2164	0.1218	0.0026	-	-	-
Jean Sibelius	0.0125	-	-	0.0057	-	-	-
Anitta	-	0.2156	0.0702	-	-	-	-

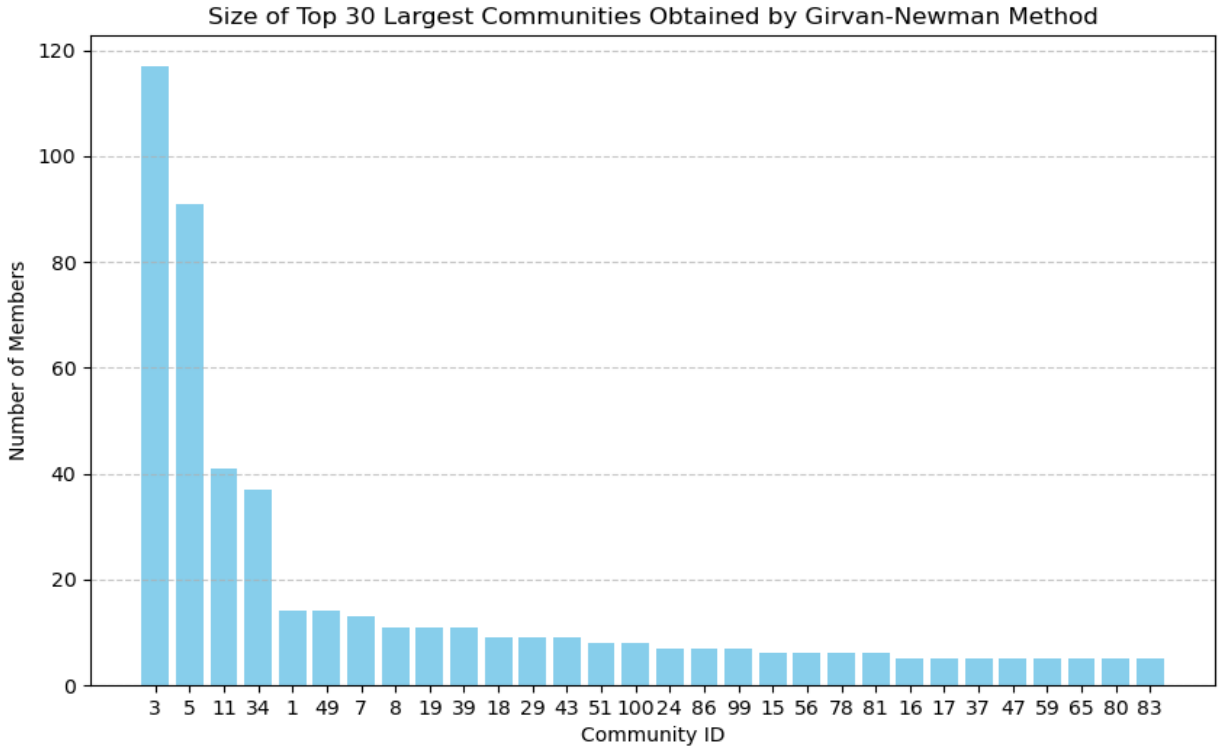


Communities analysis

Finding out how music clusters

Communities analysis: Girvan-Newman methodology





Communities analysis: Girvan-Newman methodology

Robustness and fault tolerance

What would happen to the music world in the case of an unexpected event?

What would happen to the music world in the case of an unexpected event?

Random failure

A group of various artists are taken hostage during a conference

Targeted attack

The world's most famous music artists are targeted and kidnapped at the same time for ransom

Random failure

A group of various artists are taken hostage during a conference

Original network

- Size: **890**
- Clustering Coefficient: **0.014**
- Average Degree: **1.62**
- Giant Component Size: **208**
- Giant Component Clustering Coefficient: **0.015**
- Giant Component Avg Degree: **2.12**
- Giant Component Diameter: **16**
- Giant Component Avg Shortest Path Length: **7.26**

After Random Failure

- Size: **846**
- Clustering Coefficient: **0.016**
- Average Degree: **1.53**
- Giant Component Size: **90**
- Giant Component Clustering Coefficient: **0.03**
- Giant Component Avg Degree: **2.2**
- Giant Component Diameter: **14**
- Giant Component Avg Shortest Path Length: **5.92**

Random failure

A group of various artists are taken hostage during a conference

- **Giant Component:** the size of the giant component decreases to 95 artists, signaling a fragmentation of the network and a loss of connectivity among artists.
- **Giant Component Clustering Coefficient:** Despite the decrease in size, the clustering coefficient of the giant component remains consistent, underscoring the persistence of collaboration within smaller clusters of artists.
- **Giant Component Average Shortest Path Length**: The decrease in average shortest path length to 5.92 suggests that, while paths between artists may be shorter within the giant component, the overall network experiences increased fragmentation and isolation.

Targeted attack

The world's most famous music artists are targeted and kidnapped at the same time for ransom

Original network

- Size: **890**
- Clustering Coefficient: **0.014**
- Average Degree: **1.62**
- Giant Component Size: **208**
- Giant Component Clustering Coefficient: **0.015**
- Giant Component Avg Degree: **2.12**
- Giant Component Diameter: **16**
- Giant Component Avg Shortest Path Length: **7.26**

After Targeted Attack

- Size: **840**
- Clustering Coefficient: **0.007**
- Average Degree: **1.32**
- Giant Component Size: **42**
- Giant Component Clustering Coefficient: **0.027**
- Giant Component Avg Degree: **2.095**
- Giant Component Diameter: **10**
- Giant Component Avg Shortest Path Length: **4.27**

Targeted attack

The world's most famous music artists are targeted and kidnapped at the same time for ransom

- **Clustering Coefficient**: The decrease in clustering coefficient to 0.007 signifies a disruption in collaboration patterns, as the removal of influential artists destabilizes local clusters within the network.
- **Average Degree**: The decrease in average degree to 1.32 indicates a significant reduction in collaboration opportunities, with fewer connections between remaining artists.
- **Giant Component**: The size of the giant component dwindles to 42 artists, highlighting the severe fragmentation and loss of connectivity within the network.
- **Giant Component Average Shortest Path Length**: The decrease in average shortest path length to 4.27 reflects the increased fragmentation and isolation within the network, as artists struggle to navigate pathways between collaborators

Conclusions

Random failure scenario: Moderate impact, reduction in collaboration opportunities, and connectivity. Network experiences fragmentation and decrease in size, but cohesive clusters of collaboration persist.

Targeted attack on popular artists: Pronounced and concentrated disruption, significant loss of connectivity and collaboration opportunities. Heightened fragmentation and isolation among remaining artists.

Effects of targeted attack: Profound destabilization of collaboration patterns and navigation pathways throughout the network.

Conclusion: Emphasizes the critical role of influential nodes in sustaining network resilience. Highlights network susceptibility to targeted disruptions aimed at prominent figures.

Thank you