

Graph Centrality를 이용한 음원차트에 따른 프로듀서와 아티스트의 영향력 분석 및 신곡 추천

김건희(2017103958)

요약

Graph Centrality를 이용해 음원차트에 따른 국내 음악시장에서의 프로듀서와 아티스트 간의 사회연결망을 확인하고, 프로듀서와 아티스트의 각 관계가 어떤 영향을 끼치며, 어떠한 상호 관계가 있는지 분석한다. 이것을 토대로 신곡들 중 높은 순위를 차지할 가능성이 있는 음악들을 선별하는 것을 목표로 한다.

1. 서론

1.1 연구배경

최근 음원차트를 봤을 때 발견할 수 있는 점은 유명한 아티스트의 노래라고 하더라도 무조건 높은 순위를 차지하지는 않는다는 것이다. 많은 사람들이 알고 있는 발매 후 한달이 지난 '트와이스'의 신곡 'Talk that Talk'은 9월 20일 기준 인기차트에서 32위를 차지하고 있다. 그에 반하여 상대적으로 인지도가 적은 '김승민'의 '내 기쁨은 너가 벤틀리를 끄는 거야'는 발매된 지 1년이 넘었지만 인기차트 27위를 차지하고 있다. 이외에도 곡의 발매 시기와 관계없이 'sokodomo', '시운' 등 유명세가 낮은 아티스트의 곡들을 인기차트에서 꽤나 찾아볼 수 있다.

일반적으로 우리는 아티스트가 곡의 흥행을 좌우한다고 생각하는 경향이 있다. 그러나 위에서도 알 수 있듯이 꼭 그런 것만은 아니다. 프로듀서도 아티스트에 못지 않은 영향력을 끼친다. 음악은 기본적으로 아티스트와 프로듀서의 합작으로 이루어진다. 아티스트는 노래나 랩을 하고, 프로듀서는 비트를 제작하며 음원 제작을 총괄하는 역할을 맡는다.

이에 따라 본 프로젝트에서는 국내 음악의 다양한 분야에서 아티스트뿐만 아니라 프로듀서의 영향력이 음원 성적과 얼마나 긴밀한 연관이 있는지 Graph Centrality를 이용해 분석하고, 아티스트와 프로듀서 간의 상호 관계에 대해서도 연구한다. 이를 기반으로 신곡을 추천하는 것까지 나아갈 예정이다.

1.2 연구목표

국내에서 가장 이용자수가 많은 멜론 차트를 이용하여 댄스, 힙합, 인디음악 장르 차트에서 지난 3년간의 월별 데이터를 수집하고 정제한다. 그 과정에서 Graph Centrality를 이용하여 가장 적합한 중심성을 찾는다. 그렇게 나온 중심성 수치를 통해 각 분야별 아티스트와 프로듀서의 영향력을 확인해보고, 그것이 실제 차트와 연관성이 있는지 확인한다. 이후 분석을 토대로 영향력이 높은 아티스트와 프로듀서를 소개하고, 신곡들 중 높은 음원 순위를 차지할 가능성이 있는 곡들을 추천하는 웹 서비스를 제공할 계획이다.

2. 관련연구

2.1 Graph Centrality

그래프에서 중심성이란 그래프 혹은 사회 연결망에서 꼭짓점 혹은 노드의 상대적 중요성을 나타내는 척도이다. 이 중심성은 지수로 계산되는데 기준에 따라서 다양하다. 노드들 간의 최단 경로를 가지고 계산하는 매개 중심성(Betweenness centrality), 그래프의 노드와 다른 모든 노드들 사이의 최단 경로 길이의 합의 역수를 이용해서 구하는 근접 중심성(Closeness Centrality), 한 노드에 연결된 모든 엣지의 개수로 중심성을 평가하는 연결 중심성(Degree Centrality) 등 이 밖에도 다른 계산 방법들이 많이 존재한다[1].

2.2 Graph Centrality를 이용한 국내 힙합 프로듀서와 아티스트 간 영향력과 음원 성적 분석

지난 학기 캡스톤 디자인1에서 진행하였던 해당 연구에서는 Graph Centrality 중 가중연결 중심성 (Weighted Degree Centrality)을 이용해 힙합 장르에서 각 아티스트와 프로듀서의 영향력을 파악했고, 이를 토대로 음원 성적 사이의 연관성을 분석했다.

각 엣지에는 음원성적을 기반으로 가중치를 부여해 중심성을 도출해 시각화하였다. 그 결과, 해당 연도 장르에서 어떤 아티스트와 프로듀서가 높은 영향을 끼쳤는지 파악할 수 있었다. 아티스트는 프로듀싱까지 직접 하는 사람이 높은 영향력을 보였고, 프로듀서는 같은 회사 소속의 아티스트와 작업할 때 높은 영향력을 보였다[2].

2.3 기존 연구의 문제점과 해결 방안

기존 연구는 2020년 1월부터 12월까지의 데이터만 이용하였으므로 신뢰성이 상대적으로 떨어진다고 볼 수 있다. 그러므로 2018, 2019년도의 데이터도 이용할 것이다. 또한 기존 연구는 힙합 장르에만 한정하여 진행했기 때문에 이번 연구에서는 댄스 및 인디음악 분야로 확장하여 다른 분야에서는 아티스트와 프로듀서가 어떠한 특징을 보이는지 확인할 것이다.

3. 프로젝트

3.1 프로젝트 내용

본 프로젝트는 지난 학기 진행하였던 'Graph Centrality를 이용한 음원차트에 따른 국내 힙합 프로듀서와 아티스트의 영향력 분석'을 기반으로 먼저 힙합 장르에서는 기존 2020년 데이터만 이용한 것을 2019, 2018년 데이터까지 분석을 진행하여 신뢰성을 늘릴 것이다. 그리고 댄스 및 인디음악까지 분야를 확장하여 각 분야에서의 아티스트와 프로듀서의 영향력과 서로의 상관관계를 분석해볼 것이다. 다양한 중심성을 사용해보고 가장 적합한 중심성을 선택할 것이다. 이후 장르별 분석을 기반으로 신곡들 중 차후 높은 순위를 차지할 가능성이 있는 곡들을 선별하는 것까지 나아갈 계획이다. 해당 선별 방식은 예전의 월별 신곡 데이터를 수집한 후, 해당 신곡들 중 아티스트와 프로듀서의 영향력을 기반으로 높은 음원순위를 기록할 가능성이 있는 곡들을 선별한 후 실제로 선별된 곡들이 다음 달 음원 순위에서 높은 성적을 기록하였는지 확인하여 해당 예측 방식에 대해 검증할 예정이다. 이후 해당 정보들을 웹을 통하여 제공할 것이다.

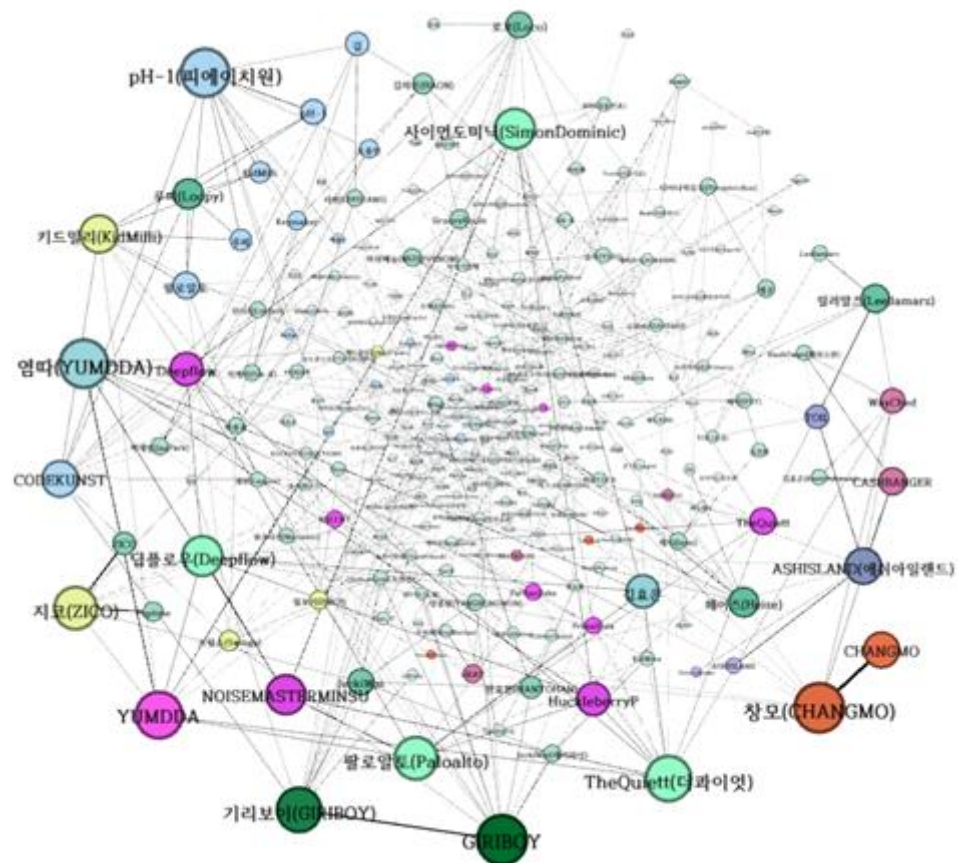


그림 1 [시각화된 그래프]

3.2 요구사항

3.2.1 정제된 데이터

위와 같은 프로젝트를 위해 데이터 수집 및 정제가 필요하다. 본 프로젝트는 두 가지 경로로 데이터를 수집한다. 첫째, 음원 서비스 멜론에서 제공하는 음원차트를 사용한다. 멜론은 현재 음원 스트리밍 서비스 중 국내에서 최대 사용자를 보유하고 있다. 둘째, 한국음악저작권협회에서 제공하는 음원 저작권자 데이터를 이용한다. 이를 통해 곡의 정확한 아티스트와 프로듀서를 확인한다. 이후 시각화 도구 'Gephi'와 호환될 수 있도록 데이터를 교정 및 통합해 정제한다.

3.2.2 Graph Centrality 요구사항

본 프로젝트를 수행하기 위한 핵심 알고리즘이다. 노드와 엣지로 표현된 그래프에서 노드들의 상대적 중요성을 계산하는 기법이다. 계산하는 방법에 따라 크게 매개 중심성(Betweenness Centrality), 연결 중심성(Degree Centrality), 근접 중심성(Closeness centrality), 고유벡터 중심성(Eigenvector Centrality)로 분류된다. 현재 프로젝트에서는 지난 프로젝트에서 셀프 프로듀싱이 반영되었고, 소수의 순위 높은 가진 곡들을 가진 경우를 반영하였던 가중연결 중심성(Weighted Degree Centrality)를 기준으로 삼되 다른 중심성들과 비교하여 가장 적합한 중심성을 선택할 것이다.

3.2.3 Gephi

대규모 그래프를 시각화하고 분석하기 위한 오픈 소프트웨어이다. 3D 렌더링 엔진을 사용해 실시간으로 그래프를 표시하고 탐색 속도를 향상해준다. 이를 사용해 음원차트 기반으로 제작한 아티스트, 프로듀서 그래프를 탐색, 분석, 조작이 가능하다.

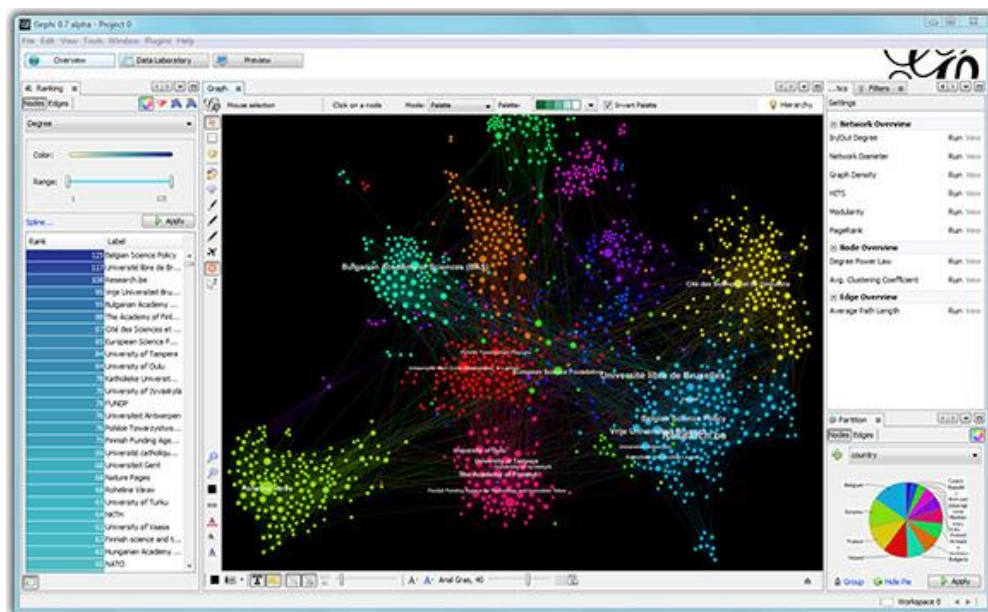


그림 2 [Gephi 화면 예시]

4. 향후 일정

진행 일정	내용
~ 10/07	데이터 수집 및 정제
10/8 ~ 10/14	그래프 중심성 측정 및 분석
10/15 ~ 10/21	신곡 추천에 필요한 데이터 수집 및 신곡 추천 기준 설정
10/22 ~ 10/28	중간 보고서 작성
10/29 ~ 11/18	서비스 제공을 위한 웹 개발
11/19 ~ 11/25	최종 보고서 작성
11/26 ~ 12/05	발표 준비

5. 결론 및 기대효과

최근 음원 차트를 보면 댄스, 힙합, 인디음악 등 다양한 분야의 노래들이 섞여 있다. 다양한 장르가 모두 인기있는 만큼 차트에서 높은 순위를 차지하기는 그만큼 더욱 어려워졌다. 그렇기 때문에 장르별로 높은 영향력을 보이는 아티스트와 프로듀서를 분석하여 어떠한 특징을 보이는지 살펴본다. 아티스트나 프로듀서들이 어떠한 방식으로 곡을 제작하면 높은 순위를 기록할 수 있을지 전략을 제공할 수 있으며, 이러한 정보들을 통하여 신곡들 중 높은 순위를 차지할 가능성이 있는 곡들을 선별할 수 있다. 위와 같은 정보들은 웹을 통하여 제공할 예정이다.

6. 참고문헌

[1] 조태수, 한치근, 이상훈, "그래프 중심성들을 이용한 그래프 유사도 측정", 한국컴퓨터정보학회논문지, 2018.

[2] 김건희, 송근영, 한치근, "Graph Centrality를 이용한 국내 힙합 프로듀서와 아티스트 간 영향력과 음원 성적 분석", 한국정보과학회 한국컴퓨터종합학술대회, 2022.