

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

김건희¹ 송근영⁰¹ 한치근*

¹*경희대학교 컴퓨터공학과

tahm_kench@khu.ac.kr, rmsdud1626@naver.com, cghan@khu.ac.kr

Analysis of between Korean hip-hop producer and artists' influence and music chart using graph centrality

Geonhui Kim¹ Geunyoung Song⁰¹ Chigeun Han*

¹*Department of Computer Science and Engineering, KyungHee University

요 약

최근 세계적인 추세에 따라 우리 나라에도 힙합 장르의 영향력이 점점 커지고 있다. 이러한 힙합 분야는 아티스트와 프로듀서의 역할이 모두 중요하다. 따라서, Graph Centrality를 이용해 국내 랩/힙합 분야의 프로듀서와 아티스트의 영향력을 분석한다. 그 중, 가중연결 중심성(Weighted Degree Centrality)을 이용하며, 2차원 함수 가중치를 사용한다. 분석 결과 아티스트의 경우 셀프 프로듀싱을 하는 아티스트들이 높은 영향력을 보였고, 프로듀서의 경우 서로 비슷한 음악적 특색을 지닌, 같은 회사 소속의 아티스트와 작업을 하는 경우 높은 영향력을 보였다.

1. 서 론

최근 음원차트를 보면 눈에 띄는 장르가 있다. 세계적인 음원차트 빌보드에서는 힙합 장르가 주를 이룬 지 오래 되었고, 국내 음원차트 역시 몇 년 사이 힙합 음악의 영향력이 커졌다. 국내 최대 음원 스트리밍 서비스 '멜론'에서 집계한 통계를 보면 2020년도 히트곡 순위 1, 2위를 힙합 장르가 차지했으며, 2021년 또한 4위를 차지하였다.

일반적으로 아티스트가 곡의 흥행을 좌우한다고 생각하는 경우가 많지만, 프로듀서 역시 그에 못지않은 영향력을 끼치는 경우가 많다. 힙합 음악은 기본적으로 아티스트와 프로듀서의 합작으로 이루어진다. 아티스트는 랩을 하며, 프로듀서는 비트를 제작하고 믹싱, 마스터링 등 음원 제작을 총괄하는 역할을 한다.

이에 따라 본 연구에서는 국내 힙합 장르에서 아티스트와 프로듀서가 음원 성적에 끼치는 영향력을 Graph Centrality를 이용하여 분석하고, 아티스트와 프로듀서 사이의 관계에 대해서도 연구한다.

2. 관련 연구

2.1 Graph Centrality

Graph Centrality이란 그래프 혹은 사회 연결망에서 꼭짓점 혹은 노드의 상대적 중요성을 나타내는 척도이다. 이 중심성은 지수로 계산되는 데 중점으로 두는 기준에 따라 출력되는 결과값이 다르다.

노드들 사이의 최단 경로를 가지고 계산하는 매개 중심성(Betweenness Centrality), 그래프의 노드와 다른 모든

노드 사이의 최단 경로 길이 합 역수를 이용해서 구하는 근접 중심성(Closeness Centrality), 한 노드에 연결된 모든 엣지의 개수로 중심성을 평가하는 연결 중심성(Degree Centrality) 등 이 밖에도 다른 계산방법들이 다양하게 존재한다[1].

2.2 R 기반의 빅데이터 기술을 활용한 뉴스기사와 음원차트의 상관관계 분석

기존의 연구에선 뉴스 기사와 음원차트의 상관관계를 분석했다[2]. 해당 논문에선 홍보 기사의 수와 음원차트 순위의 연관성을 분석했고, '두 요인 사이에 연관성은 존재하지만, 높은 수준은 아니다'라는 결론을 도출했다. 본 연구에선 아티스트와 프로듀서의 영향력이 음원 차트에 끼치는 영향에 대해서 분석해본다.

3. 연구 목적 및 연구 방법

3.1 연구 목적

연구를 통해서 얻고자 하는 결론은 크게 두 가지다. 첫 번째, Graph Centrality를 이용해 힙합 장르에서 프로듀서와 아티스트의 영향력을 파악한다. 두 번째, 영향력 높은 아티스트와 프로듀서들이 어떤 특색을 지니며, 어떠한 연관성을 보이는지 파악한다.

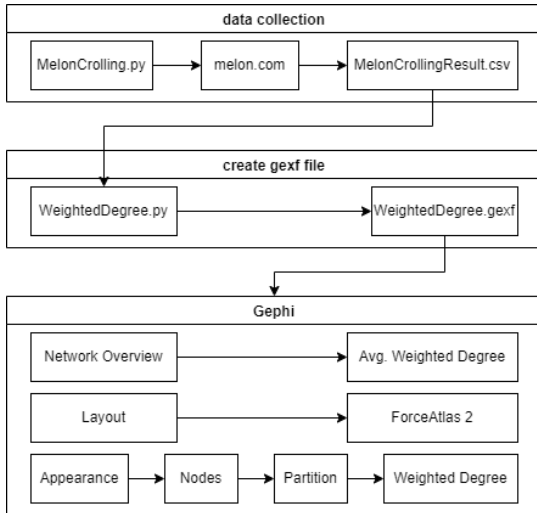
3.2 연구 방법

프로듀서와 아티스트의 영향력을 분석하고, 음원 차트와의 상관관계를 분석하기 위해 중심성을 선택할 필요가 있었다. 그중

¹ "본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학 사업의 연구결과로 수행되었음"(2017-0-00093)

하나인 매개 중심성(Betweenness Centrality)은 셀프 프로듀싱하는 경우, 소수의 사람과 작업하는 경우에 대하여 제대로 영향력을 반영하지 못한다. 근접 중심성(Closeness Centrality) 역시 마찬가지다. 그렇지만 가중연결 중심성(Weighted Degree Centrality)은 이러한 사항들을 적절하게 반영하기 때문에 가중연결 중심성(Weighted Degree Centrality)을 채택했다. 가중 연결 중심성(Weighted Degree Centrality)은 엣지의 가중치에 따라 노드의 크기와 엣지의 굵기가 변하는 특징을 가진다.

본 연구에서 결론을 도출하기 위해선 음원차트 성적 요인이 굉장히 중요하다. 따라서 그래프의 엣지에 음원 순위에 따른 가중치를 부여했으며, 가중치 식은 2차원 함수를 사용하였다. 가중치 식으로 1차원 함수를 사용하게 된다면 단편적인 예시로 1위를 한 번 한 사람과 50위를 2번 한 사람의 가중치가 동일하다는 문제가 생긴다. 그렇기 때문에 높은 순위 곡의 영향력을 늘리기 위하여 가중치 식으로 2차원 함수를 사용하였다.



[그림 1] 프로젝트 모듈 구조

본 연구의 프로젝트 모듈 구조는 [그림 1]과 같다. 모듈의 처리 순서는 데이터 수집 단계, 데이터 정제 단계, 그래프 생성 및 시각화 단계 순으로 진행된다. 웹 크롤링을 통해 음원 서비스에서 데이터를 수집한다. 이후, 수집된 데이터를 시각화 프로그램 Gephi에 적합한 형태로 정제 및 가공하는 과정을 거친다. 그렇게 그래프 정보를 포함해 생성된 파일은 Gephi에서 가중연결 중심성(Weighted Degree Centrality)을 이용해 분석하고 원하는 결과를 도출한다.

3.3 데이터 추출 및 정제

음원 차트 제공 서비스로는 국내 최대 음원 사이트 멜론을 사용했다. 현재 음원 스트리밍 서비스 중 최대 사용자를 보유하고 있으며, 집계하는 방식을 명백히 밝히고 있기 때문에 신뢰성이 있다. 멜론에서는 1개월 단위로 장르별 차트를 제공한다. 본 연구에서 그래프의 데이터 대규모 수집과 정확도 향상을 위하여, 2020년 1월부터 12월까지 총 12개월의 데이터를 수집하여 정제 작업을 실행했다.

	A	B	C	D	E
1	Month	Rank	Title	Artist	Producer
2		1	1 METEOR	창모 (CHA CHANGMO	
3		1	2 아무노래	지코 (ZICCO, Poptime	
4		1	3 아마두 (fe: 옴따	YUM NOISEMASTER	
5		1	4 빌었어	창모 (CHA CHANGMO	
6		1	6 BAND	창모 (CHA CASHBANGER	

[그림 2] 추출한 데이터의 일부

프로그래밍을 통해 웹사이트를 크롤링해 데이터를 수집했다. 멜론에서는 달, 음원차트, 제목, 아티스트 정보를 추출했고, 한국저작권협회에서 프로듀서의 데이터를 추출 후 정제하였다.

정제 후 노드의 개수는 276개, 엣지의 개수는 450개가 되었다. 추출된 데이터는 위와 같이 csv파일로 저장돼 있다.

3.4 그래프 생성 및 시각화

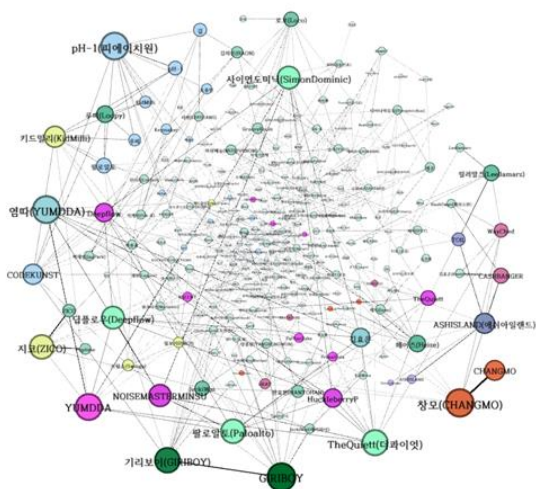
이러한 데이터를 시각화하고 보다 정확하게 분석하기 위해 가장 적합한 자료구조 ‘그래프’를 채택했다. 아티스트와 프로듀서를 각각 노드로 생성해주었고, 이들이 같은 곡을 작업했을 경우 두 노드를 엣지로 이어주었다. 또한, 이 엣지에는 음원차트를 기반으로 한 가중치를 부여해주었다. 가중연결 중심성(Weighted Degree Centrality) 가중치 공식은 먼저 $(가중치1) = [(1 - (해당 곡의 순위 / 101)^2) * 100]$ 로 설정하였다. 그리고 아티스트 및 프로듀서 각각 (가중치1)을 모두 합친다. 이를 바탕으로 작성된 전체 아티스트와 프로듀서의 엣지의 가중치 중 최대, 최소값을 구하고 나서 $(가중치2) = (가중치1 - 최소값) / (최대값 - 최소값)$ 의 식으로 가중치를 계산한다. 마지막으로 각 두 노드 간 협업된 곡들의 가중치2 값을 총합하여 최종 가중치를 구한다.

그래프를 쉽게 파악하기 위해 데이터 시각화 도구 ‘Gephi’ 프로그램을 사용했다. ‘Gephi’는 대규모 데이터를 시각화하고 분석하기 위한 오픈 소프트웨어이다[3]. 노드, 엣지, 가중치를 설정했다면 Gephi를 이용해 그래프 중심성 수치를 계산하고 쉽게 파악할 수 있다.

3.5 그래프 시각화

앞선 연구를 통해 프로듀서와 아티스트의 각 장르에서 영향력을 도출했다. 이제, 영향력과 음원 차트의 상관관계를 보다 정확하고 쉽게 분석하기 위해 그래프를 작성해 분석했다.

Gephi를 사용해 그래프를 시각화하였다. 노드의 개수는 총 276개, 엣지의 개수는 총 450개가 생성됐다. 제작된 그래프는 아래와 같은 형태를 띈다.

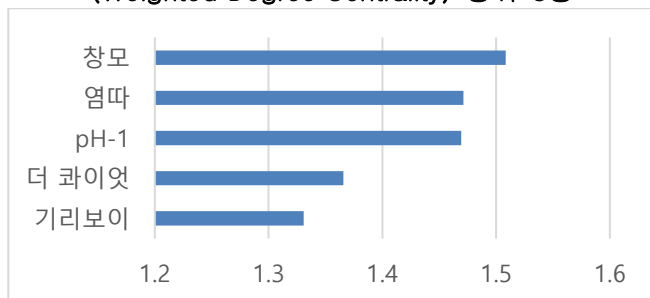


[그림 3] 'Gephi로 시각화한 그래프

앞서 분석했던 가중연결 중심성(Weighted Degree Centrality)을 보다 손쉽게 파악하기 위해 시각화하였다. 노드의 크기는 각 프로듀서의 영향력을, 엣지의 굵기는 음원 차트의 성적을 의미한다. 프로듀서의 영향력이 높을수록 노드의 크기가 커지며, 아티스트와 프로듀서 간 높은 성적을 거뒀을 경우 굵은 엣지가 나타난다.

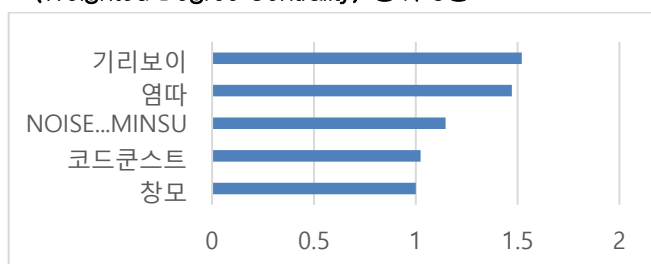
3.6 그래프 중심성 분석

[표 1] 아티스트 가중연결 중심성 (Weighted Degree Centrality) 상위 5명



가중연결 중심성(Weighted Degree Centrality)을 통해 분석한 결과, 해당 연도에서 아티스트 '창모'가 가장 높은 중심성을 지닌 것을 파악했다. 해당 아티스트의 'METEOR'라는 곡은 매달 국내 랭/힙합 부문 차트 10위 안에 들었다. 그 밖의 다른 많은 곡 역시 꾸준히 차트인하며, 음원 차트에서 좋은 모습을 보여주었다.

[표 2] 프로듀서 가중연결 중심성 (Weighted Degree Centrality) 상위 5명



프로듀서 부분에서는 기리보이가 가장 높은 중심성을 보여주었다. 기리보이가 프로듀싱한 곡 중 차트인 한 곡은

20곡이 넘으며 2위, 4위 등 좋은 순위를 거둔 곡들도 존재하기 때문에 높은 가중치를 보인 것으로 판단된다. 그 외에도 도출된 결과값을 통해 많은 프로듀서의 영향력을 파악했다.

이를 토대로 영향력이 높은 아티스트와 프로듀서의 특색과 연관성을 분석했다. 먼저 아티스트 측면으로 분석했다. 가중연결 중심성(Weighted Degree Centrality)이 높은 아티스트 상위 10명을 분석한 결과, 차트인한 그들의 곡 중 셀프 프로듀싱의 비중이 높은 것을 파악할 수 있었다. 스스로 본인의 특색을 잘 알기 때문에 직접 프로듀싱을 한 것이 순위에 도움이 된 것으로 해석된다.

다음으로 프로듀서 측면으로 분석했다. 가중연결 중심성(Weighted Degree Centrality)이 높은 프로듀서 상위 10명을 분석한 결과, 차트인 한 곡 대부분이 같은 소속사의 아티스트와 작업한 곡임을 확인할 수 있었다. 같은 회사일 경우, 서로 비슷한 특색을 지닌 경우가 많은데, 이것이 높은 음원 순위를 기록하는 데 긍정적인 영향을 끼쳤다고 해석된다.

4. 결론 및 향후연구

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.

향후, 다른 연도에도 본 연구를 적용하여 분석해 얻은 결론의 신뢰성을 높일 것이다. 또한, 더 나아가 힙합 장르에만 국한하는 것이 아니라 다른 대중 음악 차트에서의 아티스트와 프로듀서의 연관성을 분석할 예정이다.

참고문헌

- [1] 조태수, 한치근, 이상훈, “그래프 중심성들을 이용한 그래프 유사도 측정”, 한국컴퓨터정보학회논문지, 2018.
- [2] 하정철, 강동훈, 박재모, 장으뜸, 이은영, 임성현, 길준민, “뉴스 기사와 음원차트 간의 상관관계 추출을 위한 R언어 기반 분석기의 설계 및 구현”, 한국정보과학회 학술발표논문집, 122-124, 2016.
- [3] <https://gephi.org/>