creative commons

COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

공학석사 학위논문

협업 태깅 시스템의 문제를 고려한 사용자 인식 기반 음악 지도

User Perception Based Music Map:

Resolving Issues

in Collaborative Tagging System

2015년 2월

서울대학교 대학원

산업공학과 데이터마이닝 전공

한 송 원

협업 태깅 시스템의 문제를 고려한 사용자 인식 기반 음악 지도

User Perception Based Music Map:

Resolving Issues

in Collaborative Tagging System

지도 교수 조 성 준

이 논문을 공학석사 학위논문으로 제출함

2015년 2월

서울대학교 대학원

산업공학과 데이터마이닝 전공

한 송 원

한송원의 공학석사 학위논문을 인준함

2014년 12월

위 원 장 _____ 박 종 헌 _____ (인)

부위원장 _____ 조 성 준 _____ (인)

위 원 _____ 이 재 욱 _____ (인)

초 록

사람들은 다양한 상황과 분위기에서 음악을 듣는다. 각각의 상황 및 분위기에서 사람들이 선호하는 음악은 달라지는데, 이는 사람들로 하여금 재생 목록을 구성하게끔 하였다. 그러나 접할 수 있는 음악의 수가 크게 증가하면서, 사람들은 재생 목록을 구성하는 것에도 어려움을 겪기 시작했다.

본 논문에서는 사람들의 인식이 태그로 공유되는 협업 태깅 시스템을 활용하여, 음악을 상황 및 분위기에 따라 지도 형태로 나타내는 음악 시각화 프레임워크를 제안한다. 태그는 유용한 정보를 담고 있지만 사용 행태의 일관성 부족 등의 문제로 활용에 어려움이 따른다. 제안하는 프레임워크는 태그의 속성 분류 및 군집화에 기반하여 태그의 비효과성과 모호성, 그리고 중복성을 해결하고 Self-Organizing Map을 통해 많은 수의 음악을 2차원에 나타낸다. 논문의 후반부에서는 프레임워크가 음악 시각화에 효과적임을 보이기 위해, 웹 상의 실제 데이터에 프레임워크를 적용하고 정성적 및 정량적 평가를 수행한다.

본 논문의 결과물로 사람들은 지도에서 영역을 선택함으로써 상황과 분위기에 알맞은 음악을 들을 수 있다. 이는 선택지가 많아짐에도 만족도는 떨어지는 이른바 '선택의 역설'을 극복하는 하나의 대안이 될 것으로 기대한다.

주요어 : 음악 시각화, 비효과적 태그 제거, 태그 의미 명확화, 태그 군집화, Self-Organizing Map
학 번 : 2013-21089

목 차

제 1 장 서 론 1
제 2 장 관련 연구 5
제 3 장 제안 프레임워크 8
제 1 절 비효과적 태그 제거 13
제 2 절 태그 의미 명확화 14
제 3 절 태그 군집화 22
제 4 절 시각화 25
제 4 장 실험 결과 27
제 1 절 데이터 수집 27
제 2 절 제안 기법 적용 단계 31
제 3 절 성능 평가 36
제 5 장 결 론 44
참고 문헌 46
Abstract 52

표 목차

[표 1] 협업 태깅 시스템에서 태그가 갖는 문제 3
[표 2] 재생 목록에 나타난 주요 상황 및 분위기 29
[표 3] 정성적 평가를 위한 질문 구성 36
[표 4] 태그 군집화 결과 39
[표 5] QE, TE 결과 표 43

그림 목차

[그림 1] ‘Favorite’ 태그가 달려있는 음악만이 즐겨 찾는 음악인가? 9
[그림 2] ‘Catchy’ 음악은 어디에 위치해야 하는가? 10
[그림 3] ‘Christmas’ 분위기의 음악을 듣기 위해 세 군데 영역을 선택해야 하는 경우 11
[그림 4] 음악 시각화 프레임워크 12
[그림 5] 음악-태그 행렬 13
[그림 6] 음악 네트워크 15
[그림 7] 태그 의미 명확화 프로세스 20
[그림 8] 태그 분리 예시 21
[그림 9] 태그 네트워크 23
[그림 10] 텍스트 마이닝 전처리 과정 28
[그림 11] 재생 목록 탐색 예시 28
[그림 12] 음악 지도 34
[그림 13] 재생 목록 시각화 결과 35
[그림 14] 태그 ‘Chill’의 Component Plane 38
[그림 15] 의미적으로 유사한 재생 목록의 인접성 확인 41
[그림 16] QE, TE 결과 그래프 43

제 1 장 서 론

오늘날 음악은 사람들이 가장 쉽게 접할 수 있는 컨텐츠 중의 하나이다. 사람들은 출근 길에 지루함을 달래려 음악을 듣고, 업무로 쌓인 스트레스를 해소하기 위해 음악을 듣기도 하며, 자기 전에 피로를 풀기 위해 음악을 듣기도 한다. 여기서 주목할 점은 음악을 듣는 상황이 다양하고 상황에 따라 선호하는 음악이 달라진다는 것이다. 예를 들어, 고단한 하루를 마감하기 위해 음악을 듣는 경우라면 일반적으로 시끄러운 헤비메탈 음악보다는 잔잔한 어쿠스틱 음악 혹은 클래식 음악을 선호할 것이고, 운동할 때 듣는 경우라면 활기 넘치고 신나는 음악을 선호할 것이다.

그러나 사람들에게 있어서 상황에 맞는 음악을 선택하는 것은 매우 어려운 문제이다. 역설적이게도 이는 가용한 음악의 수가 매우 많아졌기 때문이다. 최근 인터넷 음원 사이트를 통한 음원 공급 확대로 사람들이 들을 수 있는 음악의 수가 급격하게 증가했는데, 사람들은 선택 가능한 대상이 많아짐에 따라 포기해야 할 대상에 대한 후회도 커지는 이른바 '선택의 역설'에 빠지게 되었다.

사람들은 이러한 문제를 재생 목록을 구성함으로써 해결하고자 하였다. 재생 목록이란 사용자가 장르나 분위기 등의 기준에 따라 몇 개의 음악을 선택해놓은 것으로, 상황에 맞는 음악들을 추후에 빠르게 선택하기 위해 만들어낸 개념이다. 구체적으로 말하자면, 'Mariah Carey'의 'Santa Claus is comin to town'이나 'Lynden David Hall'의 'All you need is love'와 같은 음악들로 'Songs for Christmas season'이라는 제목이 달린 재생 목록을 구성하는

방식이다. 일단 재생 목록을 구성해 놓으면 이후에는 상황에 적합한 재생 목록을 통째로 선택하면 된다는 장점이 있다. 하지만 현재로서는 수많은 음악으로부터 재생 목록을 구성하는 것 자체가 사람들에게 어려운 문제이다. 선택 가능한 음악의 수가 매우 많기 때문이다. 각종 기술의 발달로 사람들은 언제든 원하는 음악을 들을 수 있게 되었지만, 정작 자신이 어떤 음악을 듣고 싶어하는지를 알기 힘들게 된 것이다. 이로 인해, 사람들은 음악들을 사용자의 인식에 맞게 분류해주는 시스템의 필요성을 느끼게 되었다.

다양한 연구에서 재생 목록 구성 과정을 보다 쉽게 만들고자 데이터 기반의 방법을 제안하였는데, 그 중 대표적인 방법은 많은 수의 음악을 2차원 평면에 시각화하여 음악 지도를 만드는 방식이다. 이 경우 비슷한 분위기의 음악은 지도에서 가까이에 위치하게 되고, 따라서 지도에서 특정 영역에 모여있는 음악들을 선택함으로써 분위기가 일관된 재생 목록을 구성할 수 있게 된다. 재생 목록 구성을 위한 음악 시각화 기법은 어떠한 변수를 사용하느냐에 따라 내용 기반 기법(Content-based approach)과 협업적 기법(Collaborative approach) 두 가지로 나눌 수 있다.

첫 번째 방식인 내용 기반 기법은 전통적인 음악 정보 검색 분야의 연구를 따른 것으로, 리듬이나 멜로디 등과 같이 음악을 잘 요약한다고 알려진 다양한 특성을 수치화하여 변수로 활용하는 방식이다. 반면 두 번째 방식인 협업적 기법은 대중의 인식을 변수로 활용한다. 사람들은 자신이 즐겨 듣는 음악을 상황이나 분위기를 의미하는 태그로 묘사하는데, 오늘날에는 수많은 사람들이 기록한 태그가 협업 태깅 시스템이라는 거대한 시스템 상에서 공유된다. 즉, 협업 태깅 시스템에는 대중의 인식이 나타나 있으며,

이를 변수로 사용했을 때 시각화 결과에서 사람들이 비슷한 분위기로 인식하는 음악들이 가까이 위치할 것을 기대할 수 있다. 또한 전자의 경우에 결과 해석을 위해 평면의 각 영역에 위치한 음악을 듣고, 어떤 음악인지 묘사하는 별도의 과정이 필요하지만, 후자의 경우 영역에 위치한 음악들이 많이 포함하는 태그가 레이블 역할을 한다는 점을 주목할 필요가 있다. 다만 협업 태깅 시스템에서 태그는 다수의 대중에 의해 기록되어 일관성 부족 등의 문제를 갖기 때문에 분석에 제대로 활용되고 있지 못하다. 따라서 태그가 갖는 문제를 해결할 수 있다면, 협업적 기법은 음악을 인식에 따라 시각화하고 재생 목록을 구성하는 데에 있어서 더욱 적합한 방법이 될 것이다.

그렇다면 협업 태깅 시스템에서 태그가 갖는 문제란 무엇인가? 다양한 연구에서 태그의 비효과성과 모호성, 그리고 중복성을 기법의 정확도와 해석력을 떨어뜨리는 요소로 지적하였다(Gemmell et al., 2009; Gupta et al., 2010). 각각의 항목에서 구체적으로 다루는 내용은 [표 1]과 같다.

[표 1] 협업 태깅 시스템에서 태그가 갖는 문제

비효과성: 태그가 효과적인 의미를 전달하지 못함
모호성: 태그가 여러 맥락에서 사용되어 혼동을 줌
중복성: 여러 태그가 동일한 의미로 사용됨

본 논문에서는 태그의 비효과성과 모호성, 그리고 중복성을 해결함으로써 보다 정확하고 해석이 용이한 음악 지도를 만드는

프레임워크를 제안한다. 제안하는 프레임워크에서 비효과적 태그 제거와 태그 의미 명확화, 그리고 태그 군집화를 통해 데이터 셋의 모든 태그를 효과적이고 명확한 의미를 가지며, 분별력 있는 태그로 변환한 후 Self-Organizing Map(Kohonen, 1990)을 이용하여 2차원 평면상의 지도로 시각화한다. 프레임워크의 결과물인 음악 지도에서 각 영역은 상황과 분위기를 포함한 사용자의 인식을 의미한다. 이 때 Self-Organizing Map의 결과물인 가중치 벡터를 통해 지도의 영역을 구분하고 재생 목록을 정의함으로써 사용자들로 하여금 재생 목록을 보다 직관적이고 간편하게 선택할 수 있도록 한다.

본 논문의 구성은 다음과 같다. 2장에서는 재생 목록 구성을 위한 음악 시각화, 그리고 협업 태깅 시스템의 태그가 갖는 문제들과 관련된 연구를 정리한다. 3장에서는 태그의 비효과성과 모호성, 그리고 중복성을 해결하며 음악을 시각화하는 프레임워크를 기술한다. 4장에서는 웹 상의 데이터의 3장의 프레임워크를 적용하고 결과를 제시한다. 이 과정에서 재생 목록 공유 사이트로부터 사람들이 재생 목록을 구성할 때 고려하는 주요 상황과 분위기를 추출한 후, 이에 관련된 음악에 대하여 음악 협업 태깅 시스템에 나타난 사용자 인식을 분석한다. 마지막으로 5장에서는 본 논문의 결론과 추후 연구 방향을 정리한다.

제 2 장 관련 연구

음원 공급의 폭발적 증가로 재생 목록 구성에 대한 어려움이 커지면서, 여러 연구에서 음악을 상황에 따라 시각화함으로써 문제를 해결하고자 하였다. 음악을 시각화하는 방법은 어떠한 요소를 변수로 활용하느냐에 따라서 내용 기반 기법과 협업적 기법 두 가지로 구분할 수 있다. 시각화 과정에서 리듬이나 멜로디 등 음악 자체에서 추출한 값을 변수로 사용하면 내용 기반 기법으로, 태그와 같이 음악에 대한 대중의 의견을 변수로 사용하면 협업적 기법으로 구분된다.

여러 연구에서 위의 두 접근 방식을 사용하여 음악을 시각화하고자 하였다. 아직까지 대부분의 연구에서는 전통적인 음악 정보 검색 분야의 연구 흐름을 따라 내용 기반 기법을 사용하였으나, 다양한 웹 서비스를 통해 사용자의 의견이 활발히 공유되면서 이를 활용하는 협업적 기법에 기반한 연구도 진행된 바 있다. 여기에서는 두 접근 방식을 비교하고, 본 논문에서 협업적 기법을 사용하는 이유를 명확히 한다.

재생 목록 구성을 위한 음악 시각화에서 협업적 기법의 가장 큰 장점은 지도의 각 영역에 레이블을 할당하기 쉽다는 것이다. 시각화 결과에서 비슷한 음악들이 가까이에 위치하게 되는데, 각 영역에 밀집한 음악들이 재생 목록으로 취급되기 위해서는 영역 내의 음악들이 사용자에게 어떻게 인식 되는지를 레이블로 나타낼 수 있어야 한다. 기존의 내용 기반 기법(Mörchen et al., 2006; Rauber et al., 2003; Vembu and Baumann, 2005)에서는 음악을

나타내는 객관적인 지표에 의해 음악을 시각화할 수 있다는 장점이 있지만, 시각화된 결과에서 각 영역에 레이블을 할당하기 위해서는 사람이 직접 듣고 판단해야 한다는 치명적인 단점이 있다. 따라서 데이터 셋의 음악에 사용자 인식에 관련된 레이블이 달려있지 않은 경우 지도의 영역에 의미를 부여하기 어렵다. 위의 연구들에서는 주로 메타 정보를 활용하여 지도의 영역에 장르를 할당하였다. 이 경우 사람들이 재생 목록을 구성할 때 음악이 어울리는 상황이나 분위기를 고려한다는 점을 반영하지 못한다. 반면 Lehwark et al.(2008)의 연구와 같은 협업적 기법에서 변수로 사용되는 태그는 그 자체가 음악에 대한 대중의 인식이고 레이블이다. 따라서 내용 기반 기법의 문제를 극복할 수 있는 대안이 된다. 또한 태그가 음악 시각화를 위한 충분한 정보를 가지고 있는지도 중요한 문제인데, Laurier et al.(2009)은 음악 분류 문제에서 태그를 사용한 음악 분류가 음악 전문가들의 음악 분류와 매우 유사한 결과를 도출함을 보였다. 이는 태그를 사용하여 음악을 시각화 했을 때, 사람이 직접 듣고 시각화 한 것과 유사한 성능을 보일 수 있음을 간접적으로 보여주는 결과이다.

그러나 아직까지 협업 태깅 시스템의 태그를 활용하기에는 많은 제약이 따른다. 태그는 한 사람의 전문가가 명확한 기준에 입각하여 생성하는 것이 아니고 대중의 각자의 기준에 따라, 자신의 언어로 생성하기 때문에 일관성 부족 등의 문제를 갖기 때문이다. 이와 관련하여 여러 연구에서 문제점을 지적한 바 있다. Gupta et al.(2010), Lamere(2008)의 논문에서 태그의 사용 목적과 태그가 담고 있는 정보의 속성을 분류하였으며, 이 중에는 분석 목적에 따라 유효하지 않은 태그들이 있음을 언급하였다. 또한 Gemmell et

al.(2009)은 태그가 언어적 특성을 가지고 있기 때문에 발생하는 문제인 모호성과 중복성이 협업 태깅 시스템에서의 대표적인 예측 문제인 태그 추천 문제에서 성능 손실을 야기함을 실험을 통해 증명하였다. 또한 여러 연구에서 다양한 기법을 통해 모호성과 중복성을 해결하고자 하였다(Au Yeung et al., 2009; Begelman et al., 2006; Dattolo et al., 2011; Weinberger et al., 2008).

협업적 기법을 이용한 음악 시각화 연구인 Lehwark et al.(2008)의 논문에서는 다양한 태그들 중 장르를 나타내는 소수의 태그에 대해서만 데이터 셋을 정의하고 음악을 시각화하였다. 따라서 사람들이 재생 목록을 구성할 때 중요시하는 상황이나 분위기를 반영하지 못한다. 또한 장르에 대한 태그는 사실상 전문가에 의한 분류 체계를 따르기 때문에 태그 사용의 일관성 부족이라는 문제를 따로 고려하지 않았다는 점에서 한계를 갖는다.

Lehwark et al.(2008)은 음악을 이차원 평면에 시각화하기 위해 Self-Organizing Map을 사용하였다. 비교사 학습 알고리즘의 하나인 Self-Organizing Map은 원 데이터의 위상을 잘 보존하여 시각화에 유용한 것으로 알려져 있으며, 마케팅(Kiang et al., 2006), 생물학(Xiao et al., 2003), 금융(Kiviluoto, 1998), 제조(Lee et al., 2014) 등의 다양한 분야에서 데이터에 대한 직관을 얻기 위해 활용한 바 있다. 또한 하나의 문서가 여러 개의 단어로 표현된다는 점에서 본 문제와 유사성을 갖는 문서 분류(Lagus et al., 2004; Lee et al., 2014) 목적으로도 사용된 바 있어 본 문제를 푸는데 적합하다고 할 수 있다.

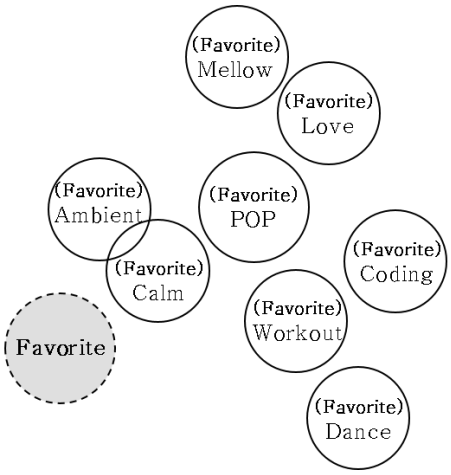
제 3 장 제안 프레임워크

본 논문에서는 협업 태깅 시스템에서의 태그가 갖는 문제인 태그의 비효과성과 모호성, 그리고 중복성을 해결하며 음악을 시각화하는 일련의 프레임워크를 제안한다. 먼저 태그가 갖는 세 가지 문제와 해결 방안에 대해 구체적으로 설명한 것은 아래와 같다.

첫 번째로, 비효과적 태그 제거를 통해 모든 태그가 효과적인 의미를 갖도록 한다. 협업 태깅 시스템에서 태그는 정보의 특성에 따라 여러 종류로 분류할 수 있는데, 그 중에는 효과적인 의미를 제공하지 못하는 태그들이 있다. 본 논문의 목적을 고려하여 태그를 나누었을 때, 효과적인 태그는 음악이 어울리는 상황이나 음악의 장르 등 음악의 특징을 나타내는 태그를 말한다. 이를테면 'Downtempo'나 'Reading' 혹은 'Coding' 등이 이에 속한다. 반면 위와 달리 음악의 특징을 나타내지 않는 태그는 효과적인 의미를 갖지 않는다. 'Favorite'과 "Albums I own'과 같이 주로 음악에 대한 호감도를 강조하기 위한 태그들이 이에 속한다. 이와 더불어 음악의 특징을 나타내더라도 출현 빈도가 매우 낮은 태그도 효과적인 의미를 전달할 수 없다.

위와 같은 비효과적 태그는 정보의 가치가 없고 데이터의 차원과 희박성을 증가시킨다. 또한 시각화 결과에서 왜곡된 정보를 제공할 수 있으므로 제거해야 한다. 태그 'Favorite'을 예로 들어 왜곡 원인 중 한 가지를 설명할 수 있다. [그림 1]에서 나타나듯 'Favorite'이라는 태그는 사실상 모든 음악에 내포되어 있다.

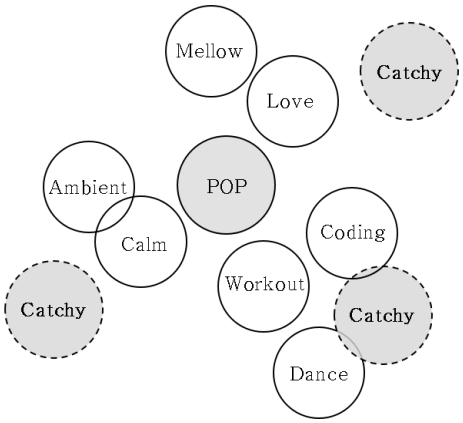
기본적으로 태그가 달린 음악들은 사람들이 선호하는 음악이기 때문이다. 그러나 태그 사용의 일관성 부족으로 모든 음악이 'Favorite'을 포함하지는 않는데, 이를 제거하지 않으면 일부 음악만 'Favorite'에 해당되는 듯한 결과를 얻을 수 있다.



[그림 1] 'Favorite' 태그가 달려있는 음악만이 즐겨 찾는 음악인가?

다음으로 태그 의미 명확화를 수행하여 모호한 태그를 맥락에 맞게 분리한다. 협업 태깅 시스템에서는 수많은 사람들이 태그로 음악을 묘사하기 때문에, 태그는 종종 여러 맥락에서 사용된다. 이러한 문제를 태그의 모호성이라고 하는데, 특정 장르가 여러 상황에 연관되어 있거나, 특정 상황에 어울리는 음악이 여러 장르로 구성되어 있는 경우에 주로 발생한다. 대표적인 예시로 'Pop'이라는 태그를 들 수 있다. 이 태그는 Pop 장르에 속하는 음악을 묘사하기 위해 사용되는데, 'Pop' 음악에는 'Mellow' 음악도, 자기 전에 듣기 좋은 'Calm' 음악도 있고, 운동할 때 주로 듣는 'Workout' 음악도 있어 여러 맥락에서 언급된다. 이렇게 모호한 태극들이 많아지면 2차원 평면에 나타내기 어려워지는데, [그림 2]에서 이를

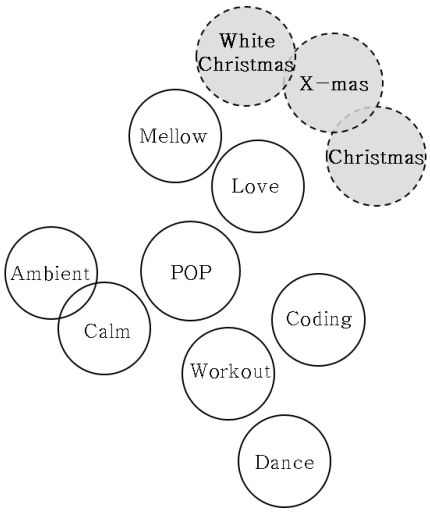
묘사하였다. [그림 2]에서 여러 가지 태그를 갖는 음악들을 시각화하기 위해서는, 같은 'Catchy' 음악이라도 세부 분위기에 따라 다르게 취급하는 것이 바람직함을 알 수 있다. 따라서 모호한 태그를 사용된 맥락에 따라 분리하여 사용할 필요가 있다. 본 논문에서는 모호한 태그를 포함한 음악들을 군집화하고 각 음악 군집을 하나의 세부 맥락으로 취급한다. 각 군집에서 원래의 모호한 태그 대신 분리된 태그를 사용함으로써 태그의 세부 맥락을 고려한다.



[그림 2] 'Catchy' 음악은 어디에 위치해야 하는가?

마지막으로, 태그 군집화를 통해 동일 의미의 태그를 하나로 합쳐 중복성을 줄인다. 협업 태깅 시스템에서는 때때로 여러 태그가 동일한 의미를 갖는 경우가 있다. 이 문제를 태그의 중복성이라 하는데, 한 가지 의미에 대해 각각의 사용자가 생각하는 적절한 단어가 다를 수 있기 때문에 발생한다. 예를 들어, 'Christmas', 'Xmas'라는 태그는 모두 크리스마스 시즌에 듣기 좋은 음악을 묘사하기 위한 태그이다. 이처럼 실질적으로 동일한 의미를 갖는 태그를 서로 다른 태그로 취급할 경우 데이터의 희박성이 증가하여

시각화 과정에서 소수의 예외적인 태그 사용에 민감하게 영향 받을 수 있다. 또한 시각화했을 때 [그림 3]과 같이 사실상 동일한 분위기의 음악이 여러 군데 나타나 사용자의 선택을 어렵게 한다.

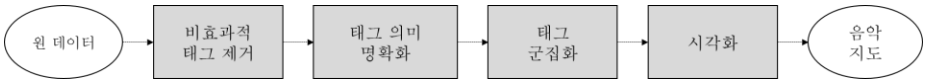


[그림 3] 'Christmas' 분위기의 음악을 듣기 위해 세 군데 영역을 선택해야 하는 경우

위의 세 가지 문제를 해결하면 이를 시각화 알고리즘의 입력 변수로 활용할 수 있다. 프레임워크의 마지막 단계인 시각화에서는 Self-Organizing Map을 사용한다. 또한 알고리즘의 결과물인 가중치 벡터를 통해 지도의 영역을 구분하고 재생 목록을 정의한다. 이는 비슷한 상황에 어울리는 음악을 가까이에 위치시켜 사람들이 재생 목록을 쉽게 구성하도록 하려는 본 연구의 목적에 부합하는 결과를 얻어내기 위함이다. 시각화 결과를 통해 재생 목록을 쉽게 탐색하기 위해서는 각 상황에 들을 법한 음악이 위치한 영역이 어디인지 명확히 나타나야 한다. 또한 하나의 음악이 여러 개의 재생 목록에 속할 수 있다는 점을 반영해야 한다. 예를 들어, 만약 'Lynden David Hall'의 'All you need is love'라는 음악이

‘Christmas’에 어울리는 음악을 모아놓은 재생 목록에는 존재하는데, ‘Lovely’한 음악을 모아놓은 재생 목록에는 속할 수 없다면 사용자들은 결과를 납득하기 어려울 것이다. 따라서 Self-Organizing Map에서 각 노드를 입력 변수 값으로 요약하는 가중치 벡터를 사용하여 위의 문제를 해결한다.

위의 네 단계를 포함한 음악 시각화 프레임워크는 [그림 4]와 같다. 먼저 태그의 비효과성을 해결하기 위해 비효과적 태그 제거를 수행한다. 다음으로 태그의 모호성을 해결하기 위한 태그 의미 명확화를 실시한다. 이 과정에서 모호한 태그를 가진 음악들을 사용된 맥락에 맞게 여러 군집으로 분리하고, 각각의 군집에 대해 원래의 태그 대신 새로운 태그를 부여한다. 다음으로 태그 군집화 과정을 통하여 중복성 문제를 해결한다. 동일한 의미로 사용된 태그들을 통합하여 데이터의 차원을 줄이는 것이다. 세 단계를 거친 후 모든 태그는 효과적이고 명확하며 분별력 있는 의미를 갖는다. 따라서 변형된 데이터를 가지고 시각화를 수행했을 때 보다 정확하고 해석이 용이한 음악 지도를 기대할 수 있다. 시각화 과정에서 Self-Organizing Map을 이용하는데, 가중치 벡터로 지도의 영역을 구분하고 재생 목록을 정의함으로써 음악 지도의 사용성을 높인다.



[그림 4] 음악 시각화 프레임워크

제 1 절 비효과적 태그 제거

분석의 대상이 되는 데이터는 의 형태로 구성되어 있다. 분석의 첫 단계인 비효과적 태그 제거에서는, 원 데이터로부터 비효과적인 태그를 제거하여 음악 지도의 각 영역이 실질적인 의미를 갖도록 한다. 또한 데이터 셋을 분석에 적합한 형태로 변환하는 과정을 거친다. 변환된 데이터는 아래의 [그림 5]와 같이 음악-태그 행렬의 형태로 나타난다.

	태그 1	태그 2	…	…	태그 n
음악 1	1	0	…	…	1
음악 2	0	3	…	…	0
…	…	…	…	…	…
…	…	…	…	…	…
음악 m	2	1	…	…	0

[그림 5] 음악-태그 행렬

제 2 절 태그 의미 명확화

비효과적인 태그를 제거한 후에는 남아있는 모든 태그가 음악 시각화에 있어 효과적이다. 하지만 이 중에는 모호한 태그들이 존재하는데, 이러한 태그들은 대중의 의견을 담고 있기 때문에 함부로 제거할 수는 없지만, 여러 맥락에서 사용되어 일관성이 떨어진다. 따라서 이 단계에서는 모호한 태그들을 맥락에 따라 분리하는 과정을 수행한다.

이 단계에서는 군집화에 기반한 태그 의미 명확화를 반복적으로 적용하여 데이터 셋에 존재하는 모든 모호한 태그를 분리한다. Au Yeung at al.(2009)는 군집화에 기반한 태그 의미 명확화가 원리가 단순하면서도 안정된 성능을 보이는 것을 확인하였다. 본 논문에서는 음악이 태그로 묘사되고 데이터의 희박성이 높다는 점에서 텍스트 데이터와 공통점을 갖는다는 것에 주목하여 토픽 모델링을 적용한 음악 네트워크를 구성하고 이를 군집화하였다.

태그 t 에 대한 태그 의미 명확화는 다음과 같은 절차로 진행된다.

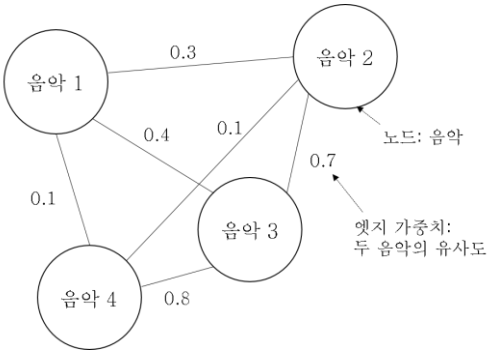
1) 주어진 태그 t 에 의해 묘사된 음악, 그리고 그 음악들을 묘사하는 태그들로 데이터 셋을 구성한다. 이 때, 데이터 셋에 포함되는 음악의 집합 M_t, 태그의 집합 T_t 는 각각 식 (1)과 식 (2)로 정의된다.

$$M_t = \{m | D_{m,t} > 0, m \in M\} \quad (1)$$

$$T_t = \{t' | D_{m,t'} > 0, m \in M_t, t' \in T\} \quad (2)$$

위 식에서 M은 전체 데이터 셋에서 음악의 집합을, T는 전체 데이터 셋에서 태그의 집합을 의미하며, D는 M을 행의 객체, T를 열의 객체로 갖는 2차원 행렬이다. D에서 각 원소 $D_{m,t}$는 음악 m에서의 태그 t 출현 빈도를 의미한다.

2) 정제된 데이터 셋에서 음악이 노드가 되고 각 노드 쌍은 음악 유사도를 엣지의 가중치로 갖는 음악 네트워크를 구성한다([그림 6]).



[그림 6] 음악 네트워크

음악 네트워크를 구성하기 위해서는 음악 간의 유사도를 계산해야 한다. 토픽 모델링을 적용한 유사도 계산을 위해 먼저 각 음악에 대하여 태그 빈도 벡터를 식 (3)과 같이 정의한다.

$$w_m = (D_{m,t_1}, D_{m,t_2}, \dots, D_{m,t_{|T_t|}}) \qquad (3)$$

그런데 태그 빈도 벡터는 태그 데이터의 특성을 제대로 반영하지 못한다. 태그 중에는 매우 포괄적인 의미를 가져서 출현 빈도는 크지만, 변별력이 떨어지는 태그들이 있기 때문이다. 태그 빈도 벡터에서는 이러한 태그들이 지배적으로 작용하기 때문에

태그의 다양성을 제대로 나타내지 못한다. 따라서 텍스트 마이닝에서 사용되는 TF-IDF Scheme(Salton and Buckley, 1988)을 적용하여 이를 보정한다. 식 (4)는 TF-IDF Scheme을 반영한 보정 태그 빈도 벡터이다.

$$w'_m = (S_{m,t_1}, S_{m,t_2}, \dots, S_{m,t_{|T_t|}}) \tag{4}$$

위 식에서 S의 원소 $S_{m,t_i} = D_{m,t_i} \times \log \frac{|M_t|}{|\{m|D_{m,t_i} > 0, m \in M_t\}|}$이다.

이와 더불어 음악 간의 잠재적 관계도 고려한다. w'_m를 통해 유사도를 구할 경우 두 음악이 공통으로 갖는 태그가 없으면 두 음악의 유사도는 0으로 취급된다. 하지만 'Christmas'와 'Xmas'를 각각 태그로 갖는 두 음악이 실질적으로 유사함은 명백하다. 따라서 이를 보정하기 위해 Latent Semantic Indexing(LSI)(Deerwester et al., 1990)를 적용한다. Singular Vector Decomposition(SVD)를 통해 수행되는 LSI는 희박성이 큰 단어-문서 행렬 혹은 문서-단어 행렬을 적은 차원의 행렬로 변환하는 기법으로, 텍스트 마이닝과 협업 태깅 시스템 분석에서 잠재적 관계를 반영하기 위한 수단으로 사용된다(Lamere, 2008; Zelikovitz and Hirsh, 2001). 이 단계에서 w'_m을 행으로 하는 음악-태그 행렬을 구성하고, 기존의 LSI 프레임워크에 따라 차원 축소를 실시한다(식 (5), 식 (6)).

$$\begin{bmatrix} w'_{m_1} \\ w'_{m_2} \\ \vdots \\ w'_{m_{|M_t|}} \end{bmatrix} = X_{|M_t| \times |T_t|} \cong U_{|M_t| \times k} \Sigma_{k \times k} V^T_{k \times |T_t|} \tag{5}$$

$$X'_{|M_t|\times k} = U_{|M_t|\times k}\Sigma_{k\times k} \tag{6}$$

축소된 행렬 $X'_{|M_t|\times k}$ 에서 각각의 행 x'_{m_i} 은 w'_{m_i} 의 적절한 근사가 된다.

차원 축소 후 얻은 보정 태그 빈도 벡터 x'_{m_i}, x'_{m_j} 를 통해 음악 m_i , m_j 의 유사도를 식 (7)과 같이 정의한다. 유사도 척도로는 코사인 유사도를 사용한다.

$$a_{m_i m_j} = sim\left(x'_{m_i}, x'_{m_j}\right) = \frac{x'_{m_i} \cdot x'_{m_j}}{\left\|x'_{m_i}\right\| \times \left\|x'_{m_j}\right\|} \tag{7}$$

3) 구성된 네트워크 상에서 군집화를 수행하고, 각각의 군집을 해당 태그의 맥락으로 취급한다. 본 논문에서는 군집화 수단으로 네트워크 분석에서 빠른 속도와 안정된 성능을 보이는 것으로 알려진 Louvain Method(Blondel et al., 2008)를 사용한다. 유통망(Raeder and Chawla, 2011)과 통신망(Greene et al., 2010) 등의 분석에 활용된 바 있는 이 알고리즘은 네트워크 군집화에서 통용되는 성능 지표인 Modularity(식 (8))(Newman, 2004)를 최대화하는 방식으로 구동된다. 가중 네트워크에서의 Modularity는 군집 내부의 엣지 밀집도가 군집 사이의 엣지 밀집도가 클수록 커지며 0에서 1 사이의 값을 갖는다.

$$\mathrm{Q} = \frac{1}{2m}\sum_{i,j}\left[A_{ij} - \frac{k_i k_j}{2m}\right]\delta(c_i, c_j) \tag{8}$$

위 식에서 A_{ij}는 노드 i,j를 연결하는 엣지의 가중치를, k_i는 노드 i에 연결된 모든 엣지의 가중치 합, c_i는 노드 i가 속한 군집을 의미하고, $\delta(u,v)$는 $u=v$이면 1, $u \neq v$이면 0으로 정의되는 함수이다. 또한 $\mathrm{m} = \frac{1}{2}\sum_{ij} A_{ij}$이다.

Louvain Method는 두 단계로 이루어지는데, 이들이 반복적으로 수행되면서 Modularity를 최대화한다(Blondel et al., 2008). 첫 번째 단계에서는 네트워크를 구성하는 각 노드에 각각 다른 군집 레이블을 할당하고 Modularity가 국지 최대에 이를 때까지 노드에 할당된 레이블을 변경한다. 이 때, 각 노드에 대하여 존재하는 모든 레이블을 할당해보고 Modularity Gain이 양수이면서 최대가 되도록 하는 레이블을 선택한다. 노드 i를 군집 C에 할당했을 때의 Modularity Gain은 아래의 식 (9)와 같이 계산된다.

$$\Delta \mathrm{Q} = \left[\frac{\sum_{in} + 2k_{i,in}}{2m} - \left(\frac{\sum_{tot} + k_i}{2m}\right)^2\right] - \left[\frac{\sum_{in}}{2m} - \left(\frac{\sum_{tot}}{2m}\right)^2 - \left(\frac{k_i}{2m}\right)^2\right] \quad (9)$$

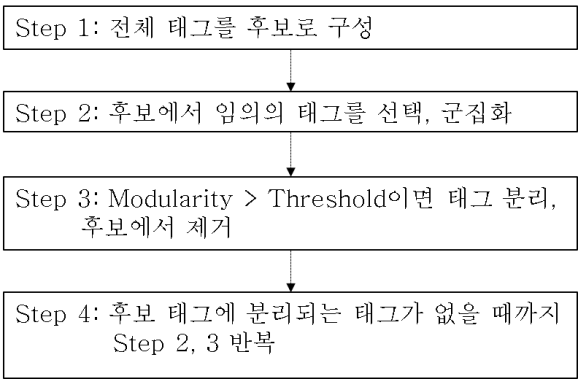
위 식에서 $\sum_{in}$은 군집 C 내부에 존재하는 엣지 가중치의 합을, $\sum_{tot}$는 군집 C 내부의 노드와 외부 노드를 연결하는 엣지 가중치 합을 의미한다. 또한 k_i는 노드 i와 연결된 모든 엣지의 가중치 합을, $k_{i,in}$은 노드 i와 군집 C 내부 노드를 연결하는 엣지 가중치 합을, m은 네트워크 내의 모든 엣지 가중치의 합을 의미한다.

두 번째 단계에서는 첫 번째 단계의 결과물로부터 각 군집이 노드 역할을 하는 새로운 네트워크를 구성한다. 군집에 속한 모든 노드를 하나로 통합하고, 군집간 엣지 가중치는 각 군집에 속한 노드들끼리의 엣지 가중치 합으로 한다. 같은 원리로 각 군집은

군집 내부의 엣지 가중치 합을 Self-loop의 엣지 가중치로 갖게 된다. 두 번째 단계를 수행하고 나면 여기에 다시 첫 번째 단계를 수행하게 되며 Modularity가 더 이상 증가하지 않을 때까지 이 과정을 반복한다. 알고리즘 수행이 완료되면 목적함수인 Modularity를 기준으로 최적의 군집들을 얻게 된다.

군집화 단계에서 음악 노드로 구성된 네트워크를 좀 더 명확히 하기 위해 Pruning을 적용한다. 음악 사이의 유사도를 엣지의 가중치로 하는 네트워크를 구성하면 그 결과물은 완전 연결 네트워크이다. 이 때, 유사도가 매우 낮은 값이어도 두 노드가 연결되기 때문에, 군집화 결과가 매우 보수적으로 나타나고 군집화에 소요되는 시간이 길어진다. 따라서 본 실험에서는 구성된 네트워크에서 가중치가 Threshold보다 작은 엣지를 제거한다. 이와 관련하여 Bernhard(2010)의 논문에서는 Threshold에 따른 군집화 결과의 Purity를 측정하였다. 실험 결과 Threshold가 증가할 때 Purity가 증가하며, 특히 Threshold가 0~0.3 사이에서 변화할 때 다른 구간에서보다 Purity가 빠르게 증가하여 적은 양의 정보 손실로도 충분한 개선 효과를 얻을 수 있음을 보였다.

위에서 언급한 바와 같이 본 연구에서는 하나의 태그가 아닌 데이터 셋의 모든 모호한 태그를 분리해야 하기 때문에, 위의 과정을 반복적으로 수행한다. 전체 태그에 대한 태그 의미 명확화 절차는 [그림 7]과 같다.



[그림 7] 태그 의미 명확화 프로세스

이 과정에서 태그 의미 명확화의 대상을 하나가 아닌 여러 태그로 확대함에 따라 문제가 발생하는데, 바로 어떤 태그를 모호하다고 취급하여 분리할 것인지에 대한 것이다. 여기에서는 기본적으로 모든 태그를 태그 의미 명확화의 대상으로 하되, 군집화 결과 각 군집이 명확하게 구분되는 경우만을 모호하다고 판단하여 분리한다. 만약 각 군집이 명확히 구분되지 않는다면, 이는 태그가 사용되는 맥락이 명확하게 구분되지 않는데 억지로 분리하는 것을 의미한다. 이러한 태그를 분리하면 중복성이 증가하여, 그 결과를 사용하는 것이 전체 프레임워크의 성능을 저하한다. 이러한 이유로 [그림 7]의 Step 3에서 Modularity가 Threshold 이상인 경우에만 태그를 분리한다.

데이터 셋에서의 태그 분리 과정은 [그림 8]에서 예를 들어 설명하였다. 태그 분리 과정에서 해당 태그에 대한 열은 군집의 개수로 분리되는데, 각 열은 대응되는 군집에 속한 음악에 대해서만 출현 빈도 값을 원소로 갖게 된다. [그림 8]의 예는 음악이 두 개의 군집으로 구분됨에 따라 태그 역시 두 개로 분리되는 경우이다. 군집화는 연속적으로 모든 태그에 대해 수행하며, 데이터 셋이

갱신됨에 따라 이전 단계에서 Threshold보다 낮은 Modularity를 가졌던 태그가 이후에 Threshold를 넘는 Modularity를 가질 수 있으므로 모든 태그가 Threshold 미만의 Modularity를 가질 때까지 동일한 과정을 반복한다.

		태그 1	…	태그 i	…	태그 n
군집 1	음악 1	1	…	2	…	0
	음악 2	2	…	3	…	1
	음악 3	1	…	1	…	0
군집 2	음악 4	0	…	2	…	1
	음악 5	0	…	1	…	2
	음악 6	1	…	4	…	0

	태그 1	…	태그 i-1	태그 i-2	…	태그 n
음악 1	1	…	2	0	…	0
음악 2	2	…	3	0	…	1
음악 3	1	…	1	0	…	0
음악 4	0	…	0	2	…	1
음악 5	0	…	0	1	…	2
음악 6	1	…	0	4	…	0

[그림 8] 태그 분리 예시

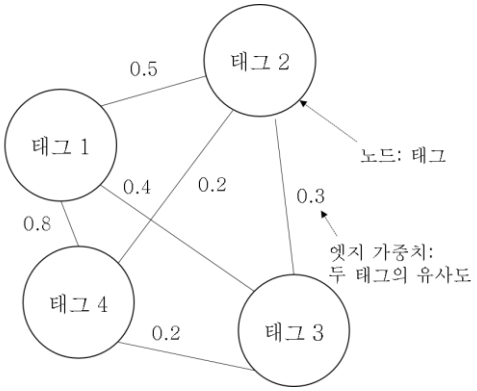
제 3 절 태그 군집화

태그 의미 명확화까지 마친 데이터 셋에서 모든 태그는 효과적이고 명확한 의미를 갖는다. 하지만 태그들 중에는 중복 의미를 갖는 태그들이 존재한다. 이러한 태그들은 데이터의 차원을 증가시키고 희박성을 높이기 때문에 이 단계에서 태그 군집화을 통해 중복되는 태그들을 하나로 통합해주는 과정을 수행한다.

태그 군집화와 태그 의미 명확화는 유사한 방식으로 진행된다. 태그 군집화에 있어서 Bagelman et al.(2006)의 방법이 유용한 것으로 알려져 있는데 이 방법은 태그 의미 명확화에서와 그 원리가 같다. 다만 태그 군집화에서는 군집화의 대상이 태그이기 때문에 태그가 노드인 네트워크를 구성한다는 점에서 차이가 있다. 또한 태그 의미 명확화에서는 각각의 태그에 대해서 데이터 셋을 정의하고 음악 군집화를 수행하였는데, 태그 군집화에서는 전체 데이터 셋에 대하여 태그를 군집화한다는 차이점이 있다. 본 논문에서는 태그가 여러 음악에 등장한다는 점과 데이터의 희박성이 높다는 특성을 고려하여 토픽 모델링을 적용한 태그 네트워크를 구성하고 이를 군집화하였다.

전체 태그 집합 T에 대한 태그 군집화는 다음과 같은 절차로 진행된다.

1) 태그를 노드로 하고, 각 노드 쌍이 태그 유사도를 엣지의 가중치로 갖는 태그 네트워크를 구성한다([그림 9]).



[그림 9] 태그 네트워크

태그 네트워크를 구성하기 위해서는 태그 간의 유사도를 계산해야 한다. 토픽 모델링을 적용한 유사도 계산을 위해 먼저 각 태그에 대하여 음악 빈도 벡터를 식 (10)과 같이 정의한다.

$$w_t = (D_{m_1,t}, D_{m_2,t}, \dots, D_{m_{|M|},t}) \qquad (10)$$

위 식에서 D는 음악의 집합 M을 행의 객체, 태그의 집합 T를 열의 객체로 갖는 2차원 행렬이다. 행렬 D에서 각 원소 $D_{m,t}$는 음악 m에서의 태그 t 출현 빈도를 의미한다.

태그 의미 명확화에서와 같이 출현 빈도는 크지만 변별력이 작은 음악의 영향력을 줄이기 위하여 식 (11)과 같이 보정 음악 빈도 벡터를 정의한다.

$$w'_t = (S'_{m_1,t}, S'_{m_2,t}, \dots, S'_{m_{|M|},t}) \qquad (11)$$

위 식에서 S'의 원소 $S'_{m_i,t} = D_{m_i,t} \times \log \frac{|T|}{|\{t | D_{m_i,t} > 0, t \in T\}|}$이다.

태그 군집화에서도 태그 간의 잠재적 관계를 고려해야 타당한

결과를 얻을 수 있으므로 LSI(Deerwester, 1990)를 적용한다. 이 단계에서는 w'_t를 행으로 하는 태그-음악 행렬을 구성, 기존의 LSI 프레임워크에 따라 차원 축소를 실시한다(식 (12), 식 (13)).

$$\begin{bmatrix} w'_{t_1} \\ w'_{t_2} \\ \vdots \\ w'_{t_{|T|}} \end{bmatrix} = X_{|T|\times|M|} \cong U_{T\times k}\Sigma_{k\times k}V^T_{k\times|M|} \tag{12}$$

$$X'_{|T|\times k} = U_{|T|\times k}\Sigma_{k\times k} \tag{13}$$

축소된 행렬 $X'_{|T|\times k}$에서 각각의 행 x'_{t_i}은 w'_{t_i}의 근사 벡터이다.

차원 축소 후 얻은 보정 음악 빈도 벡터 x'_{t_i}, x'_{t_j}를 통해 태그 t_i, t_j의 유사도를 식 (14)로 정의한다. 유사도 척도로는 코사인 유사도를 사용한다.

$$a_{t_i t_j} = sim\left(x'_{t_i}, x'_{t_j}\right) = \frac{x'_{t_i} \cdot x'_{t_j}}{\left\|x'_{t_i}\right\| \times \left\|x'_{t_j}\right\|} \tag{14}$$

2) 구성된 태그 네트워크에서 군집화를 수행하고 같은 군집에 속한 태그들을 하나로 통합한다. 군집화 수단으로 태그 의미 명확화에서와 같이 Louvain Method(Blondel et al., 2008)를 사용한다.

이 때 태그 네트워크가 완전 연결 네트워크에 가깝기 때문에 군집화 결과가 보수적으로 나타나고 소요 시간이 길어진다. 따라서 Threshold보다 작은 값을 가중치로 갖는 엣지를 제거하여 네트워크 구조를 명확히 한 후 군집화를 수행한다.

제 4 절 시각화

앞의 과정을 통해 효과적이고 명확하며 분별력 있는 의미를 가진 태그로 구성된 데이터 셋은 시각화 모델의 입력 변수로 적합하다. 이 단계에서는 비교사 학습 기법인 Self-Organizing Map을 통해 음악을 2차원 평면에 시각화한다.

Self-Organizing Map은 Input Vector로부터 각 노드를 대표하는 가중치 벡터를 만들어낸다. 이를 위해 매 Iteration s마다 각각의 행 벡터 $x(s)$에 대해 가장 가까운 가중치 벡터(식 (15))를 찾고, Competitive Learning(식 (16))을 통해 모든 가중치 벡터를 보정해나간다. 이 과정에서 Neighborhood Function(식 (17))을 사용하여 원 데이터의 위상이 2차원 평면에서도 보존되게 한다. 학습이 끝나면 각각의 행 벡터는 가장 가까운 가중치 벡터를 갖는 노드(Best Matching Unit)에 할당된다.

$$\mathrm{c} = \arg min_k \|x(s) - w_k(s)\| \tag{15}$$

$$w_k(s+1) = w_k(s) + \alpha(\mathrm{s})\lambda_{ck}(\mathrm{s})\big(\mathrm{x}(\mathrm{s}) - w_k(\mathrm{s})\big) \tag{16}$$

$$\lambda_{ck}(s) = exp\left(-\frac{\|r_c - r_k\|^2}{2\sigma(s)^2}\right) \tag{17}$$

위 식에서 $w_k(\mathrm{s})$ 는 Iteration s 에서 k 번째 가중치 벡터를 의미하며, $\alpha(\mathrm{s})$ 는 Learning Rate를, $\lambda_{ck}(\mathrm{s})$ 는 Neighborhood Function을 의미한다. 또한 r_k는 2차원 좌표 상에서 w_k의 위치를, $\sigma(s)$는 Neighborhood Function의 Width Parameter를 나타낸다.

Self-Organizing Map의 결과로 각 음악은 2차원 평면의 노드에 할당되고 각 노드는 원 데이터의 근사 벡터인 가중치

벡터를 할당 받는다. 이를 이용하여 각각의 태그에 대한 재생 목록을 정의하고, 시각화를 통해 지도에서 재생 목록의 영역을 확인할 수 있도록 한다. 노드에 할당된 가중치 벡터가 노드에 속한 음악들의 특성을 요약하는 정보이기 때문이다.

태그 t 에 대하여, 재생 목록 노드 집합 P_t 는 식 (18)로 정의된다. 전체 노드 집합에서, 태그 t 에 대한 가중치가 다른 노드들에 비해 높게 나타난 노드들에 속한 음악을 재생 목록으로 구성하는 것이다.

$$P_t = \{n | w_{n,t} > m_t + 2\sigma_t\} \tag{18}$$

위 식에서 $w_{n,t}$ 는 노드 n 에서 태그 t 의 가중치를, m_t 와 σ_t 는 각각 전체 노드에서 태그 t 에 대한 가중치의 평균과 표준편차를 의미한다.

이 단계에서 정제된 음악-태그 행렬의 행 벡터가 Self-Organizing Map의 반복적인 학습을 통해 군집화된다. 시각화 결과에서 각각의 음악은 하나의 점으로 표현되며, 유사한 음악은 가까이에, 상이한 음악은 멀리에 위치한다. 또한 지도에서 각 영역에 속한 음악들은 하나의 재생 목록으로 구성되며 각기 다른 사용자 인식을 의미한다.

제 4 장 실험 결과

제 1 절 데이터 수집

본 논문에서 분석한 데이터는 대표적인 음악 협업 태깅 시스템인 Last.fm [①] 에서 수집하였다. 이 서비스는 사용자들이 자신이 즐겨 듣는 음악을 태그로 묘사하는 방식으로 이루어진다. 본 연구의 목적은 사람들이 재생 목록을 쉽게 구성할 수 있도록 하는 것이므로, 분석에 사용하는 데이터는 사람들이 재생 목록을 구성하는 데에 주로 고려하는 상황과 분위기에 연관된 음악으로 한정하였다.

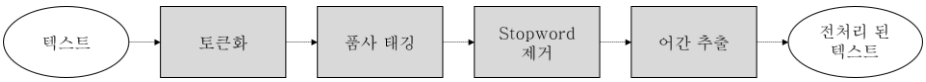
본격적인 분석에 앞서 재생 목록에 대한 데이터로부터 주요 상황과 분위기를 파악하는 과정이 필요하다. 음악 협업 태깅 시스템에서 가장 자주 등장하는 태그들을 살펴보는 방법도 있지만 음악 협업 태깅 시스템에서의 태그는 대부분 '재생 목록을 구성하기 위한 태그'가 아닌 '해당 음악으로부터 가장 쉽게 떠올릴 수 있는 태그'라는 문제를 갖는다. 이는 사용자들이 재생 목록이 아닌 각각의 음악에 태그를 달기 때문이며, 등장 횟수 기준 상위 태그의 대부분이 장르나 포괄적인 단어로 이루어진 것에 대한 이유이기도 하다.

이 점을 고려하여 본 실험에서는 대표적인 재생 목록 공유 사이트인 Playlists.net [②] 에 나타난 재생 목록의 주제 동향을

---

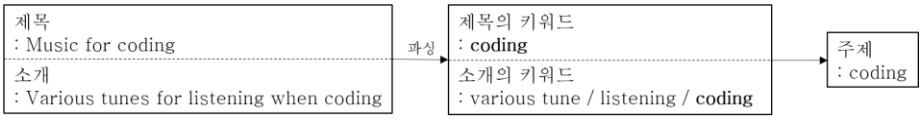
① http://last.fm
② http://playlists.net

탐색하였다. 재생 목록은 제목과 소개를 포함하고 있는데, 여기에 간단한 텍스트 마이닝 기법을 적용함으로써 재생 목록 구성에 사용되는 주요 상황을 효율적으로 탐색할 수 있다. 이 단계에서는 전통적인 텍스트 마이닝 전처리 과정([그림 10])을 통해 재생 목록의 제목과 소개에서 주요 단어를 추출해낸 후 제목과 소개에 공통으로 등장하는 어휘를 재생 목록의 주제로 추정하는 방식을 사용하였다.



[그림 10] 텍스트 마이닝 전처리 과정

전처리의 첫 단계에서는 텍스트를 토큰화하여 각각의 단어로 쪼갠 후, 품사 태깅을 통해 각각의 단어가 문법적으로 어떤 역할을 하는지 파악하였다. 이후 'is', 'an'과 같이 문법적인 기능을 담당하는 단어를 포함하여 실질적으로 의미를 갖지 않는 단어를 Stopword로 지정 및 제거하고, 어간을 추출하여 다양한 형태로 등장한 단어를 하나의 어형으로 통합해주는 작업을 수행하였다. [그림 11]의 예에서는 전처리의 결과로 'coding'이라는 단어가 재생 목록의 제목과 소개에 공통으로 등장하였고, 재생 목록을 잘 요약하는 주제어로 취급된 것이다.



[그림 11] 재생 목록 탐색 예시

전처리를 위해 Probabilistic Context Free Grammar(PCFG)에 기반한 Stanford NLP Parser(Klein, 2003)를 사용하였는데, 이 Parser는 이전 연구에서 트위터 분석(Agarwal et al., 2011), 뉴스 분석(Hagenau et al., 2012) 등의 용도로 사용된 바 있다. 주제어를 추출할 때는 Unigram에 대해서 명사, 동사, 형용사, 부사만을, Bigram에 대해서 영어의 일반적인 Collocation 형태인 '부사 + 형용사', '형용사 + 명사', '명사 + 명사', '명사 + 동사', '동사 + 명사', '동사 + Particle(동사구를 이루는 부사나 전치사)'만을 수집하여 의미 있는 어휘만 걸러내었다. 이와 더불어 재생 목록이라는 도메인의 특성에 맞게 'Playlist', 'Music', 'Song' 등을 Stopword로 지정하여, 주제어의 성격을 갖는 어휘만 수집되도록 하였다. 2014년 1월 23일 기준, Playlists.net에 공유된 재생 목록 중에 제목과 소개를 모두 갖고 재생 횟수 기준 상위 78,206개의 재생 목록을 탐색 대상으로 하였으며, 결과로 재생 목록에 가장 많이 등장한 상황과 분위기 50개를 추출하였다([표 2]).

[표 2] 재생 목록에 나타난 주요 상황 및 분위기

party, dance, summer, love, christmas, chill, festival, workout, club, chill out, night, happy, beat, running, work, driving, relax, sad, halloween, weekend, christian, study, rainy day, soft, morning, sleep, winter, wedding, calm, reading, ambient, friday, spring, sunday, late night, dream, beach, gym, road trip, slow, break up, holiday, coding, fun, mellow, autumn, car, dark, easy listening, sweet

이제 [표 2]에서 제시한 주요 상황과 분위기에 연관된 음악은 음악 시각화 대상으로 적합하다. 따라서 음악의 태그 정보가 있는 Last.fm으로부터 각각의 상황과 분위기에 대한 데이터를 추출하였다. Last.fm에는 [표 2]와 동일한 형태의 태그가 존재하며, 분석을 위해 50개 각각의 태그에 대해 태그 횟수 기준 상위 50개의 음악과 그에 달린 태그 그리고 태깅 횟수를 수집하였다. 몇몇 태그의 경우 50개 미만의 음악에 태깅된 경우가 있어 존재하는 만큼만 데이터를 수집하였다. 그 결과 2014년 1월 23일 기준, 27,424개의 태그와 2,429개의 음악으로 구성된 데이터 셋을 확보하였다.

제 2 절 제안 기법 적용 단계

제 1 항 비효과적 태그 제거

비효과적 태그 제거의 첫 단계로 전체 데이터 셋의 1%인 25개 미만의 음악에만 등장한 태그를 비효과적인 것으로 간주하고 제거하였다. 다음 단계로 음악의 특징을 나타내지 않는 태그를 제거하였다. 제거된 태그에는 'Albums I own', 'Favorite' 등이 있다. 그 결과 234개의 태그와 2,429개의 음악으로 구성된 음악-태그 행렬을 얻었다.

제 2 항 태그 의미 명확화

효과적인 태그 중에서 모호한 태그를 맥락에 따라 분리하기 위해 태그 의미 명확화를 수행하였다. 앞에서 언급한 바와 같이 태그 의미 명확화 모든 모호한 태그에 대해 반복적으로 이루어지며, 군집화 결과 Modularity가 일정 기준 이상인 태그에 대해서만 태그를 분리하게 된다.

태그 의미 명확화의 첫 단계로 각 태그와 연관된 음악들끼리의 토픽 유사도를 계산하고 네트워크를 구성한다. 본 실험에서는 LSI를 적용하여 원래의 25% 수준으로 차원을 축소하였다. 이어서 네트워크 구조의 명확화를 위해 0.3 이하를 가중치로 갖는 엣지를 제거하는 Pruning을 수행하였다.

태그 분리의 기준이 되는 Modularity Threshold는 0.5로 설정하였다. 그런데 실험 과정에서 출현 빈도가 매우 높은 태그들은 다른 태그와의 관계가 매우 복잡하여 비교적 낮은 Modularity를

나타내었다. 이러한 태그들이 분리되지 않으면 출현 빈도가 높지 않은 태그들도 몇몇 모호한 주요 태그와 관계를 맺게 되어 제대로 분리되지 않는 현상이 발생한다. 따라서 전체 데이터 셋의 10%에 해당하는 243곡 이상에 등장한 태그에 대해서 Modularity Threshold를 0.2로 하향 조정하고, Threshold 조정으로 과도하게 분리된 태그는 다음 단계인 태그 군집화에서 다시 통합하는 방법을 사용하였다. 지나치게 세분화 되는 것을 방지하기 위해 한 번 분리된 태그는 분리 대상에서 제외하였다.

태그 의미 명확화 결과 234개의 태그가 465개의 태그로 분리되었다. 분리된 태그로는 'Indie', 'Melancholy', 'Relax', 'Cute', 'Summer' 등 58개가 있다.

제 3 항 태그 군집화

동일한 의미를 갖는 여러 태그들과 태그 의미 명확화 과정에서 과도하게 분리된 태그들을 하나의 태그로 통합하기 위해 태그 군집화를 수행하였다. 이를 위해 태그를 노드로 하고 태그간의 토픽 유사도를 엣지의 가중치로 하는 태그 네트워크를 구성하였다. 본 실험에서는 LSI를 통해 원래의 25% 수준으로 차원을 축소하였으며 가중치가 작은 엣지를 제거하는 Pruning을 통해 네트워크의 구조를 명확하게 하였다. 이 때 Threshold에 따라 변하는 태그 군집의 수를 고려하여 Threshold를 0.5로 설정하였다. Threshold가 너무 낮을 경우에 태그가 거대한 소수의 군집으로 통합되어 태그의 다양성이 크게 소실되고, 통합된 태그는 다시 모호성을 갖게 된다. 반면, Threshold가 너무 클 경우, 네트워크의 엣지가 대부분 소실되어 태그 군집화의 목적인 중복성 감소 효과를 얻을 수 없게

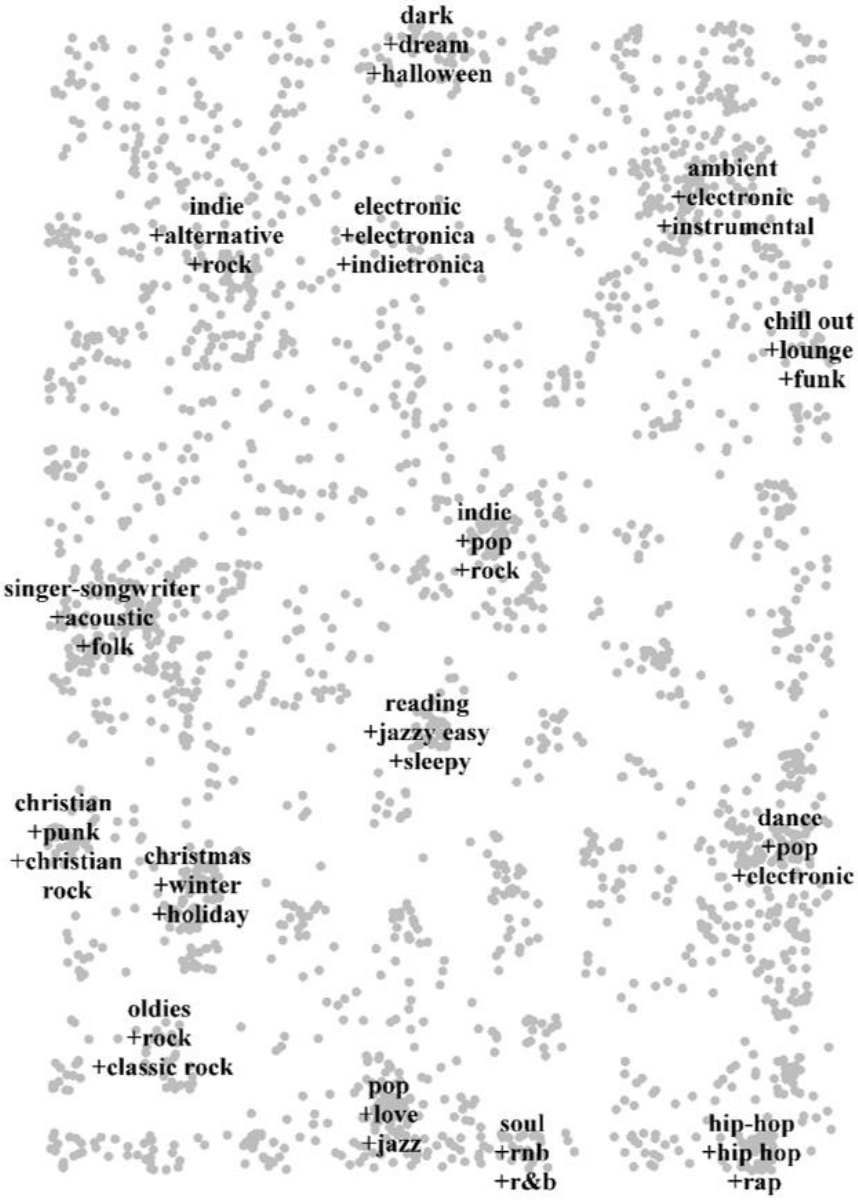
된다.

태그 군집화 결과 앞 단계에서 465개로 분리되었던 태그가 62개로 통합되었다. 해석의 용이성을 고려하여 각 군집에서 TF-IDF 점수 합이 가장 큰 3개의 태그를 추출하여 이들을 새롭게 통합된 군집의 레이블로 정하였다.

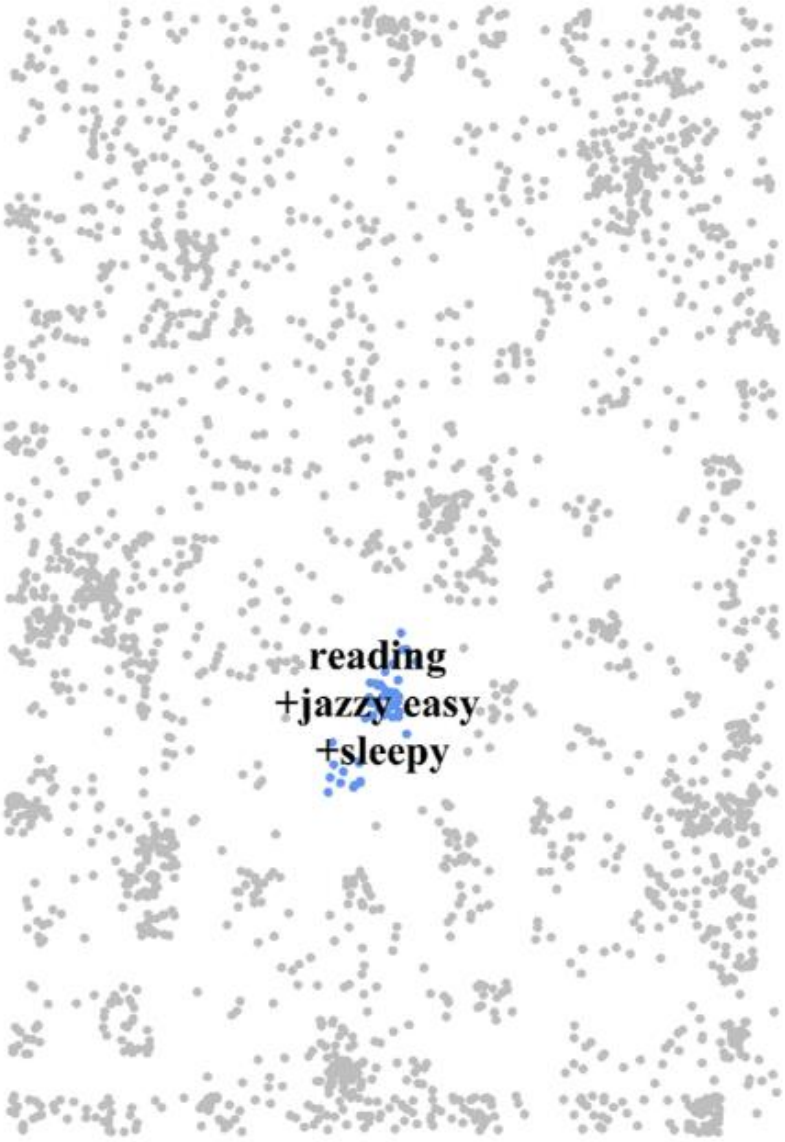
제 4 항 시각화

비효과성과 모호성 그리고 중복성 문제를 해결한 데이터 셋에서 이제 모든 태그는 효과적이고 명확하며, 분별력 있는 의미를 갖기 때문에 시각화 기법의 입력 변수로 적합하다. 따라서 이 데이터를 Self-Organizing Map을 사용하여 시각화하였다. MATLAB 환경에서 개발된 SOM Toolbox(Vesanto et al., 1999)를 사용하였고, 노드의 수는 시각화 대상인 음악의 수를 고려하여 30*15로 결정하였다.

시각화 결과는 [그림 12]와 같다. [그림 12]에서는 음악의 분포를 알아볼 수 있도록, 사용자에 의해 가장 많이 태깅된 15개의 재생 목록을 지도에 위치시켰다. [그림 13]에서는 가중치 벡터를 통해 정의한 재생 목록의 영역을 나타내었다. 두 예시에서 각 태그에 대한 재생 목록이 지도의 한 영역에 위치함을 확인할 수 있다.


dark
+dream
+halloween
ambient
+electronic
+instrumental
indie
+alternative
+rock
electronic
+electronica
+indietronica
chill out
+lounge
+funk
indie
+pop
+rock
singer-songwriter
+acoustic
+folk
reading
+jazzy easy
+sleepy
christian
+punk
+christian
rock
christmas
+winter
+holiday
dance
+pop
+electronic
oldies
+rock
+classic rock
pop
+love
+jazz
soul
+rnb
+r&b
hip-hop
+hip hop
+rap

[그림 12] 음악 지도



[그림 13] 재생 목록 시각화 결과

제 3 절 성능 평가

본 논문에서 제안한 프레임워크가 유용한지를 확인하기 위해서는 이를 검증하는 과정이 필요하다. 따라서 검증을 위해 태그의 모호성과 중복성을 해결하기 전, 후의 데이터 셋을 각각 시각화하여 둘을 비교한다. 여기에서 프레임워크의 주요 이슈 중 하나인 태그의 비효과성에 대해서는 다루지 않는다. 비효과적인 태그는 시각화 성능과는 별개로 해당 태그가 실질적인 의미를 갖지 않는 것이기 때문이다. 즉 효과적인 의미를 갖지 않는 태그의 경우 이미 확립된 기준에 의해 제거된 것이고, 따라서 아래에서는 모호성과 중복성에 대해서만 확인한다. 이와 더불어 시각화 결과가 음악 지도로서의 가치를 갖는지도 평가한다. 평가는 정성적 평가와 정량적 평가 두 가지 관점에서 이루어진다.

제 1 항 정성적 평가

프레임워크의 효과를 정성적으로 확인하기 위해 다음의 세 가지 질문을 구성하고, 각각에 대하여 답한다.

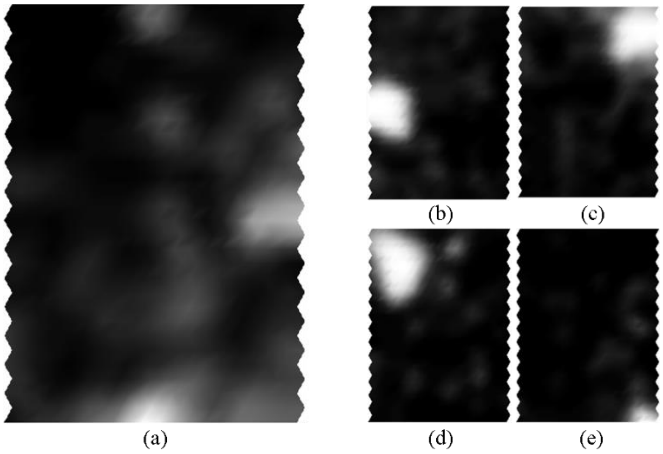
[표 3] 정성적 평가를 위한 질문 구성

태그 의미 명확화를 통해 분리된 태그는 과연 ‘명확한’ 맥락에서 사용되었는가?
태그 군집화를 통해 통합된 태그는 과연 ‘유사한’ 맥락의 태그인가?
시각화 결과에서 유사한 상황의 음악들이 가까이에 위치하였는가?

첫 번째, 두 번째 질문에서 태그 의미 명확화와 태그 군집화의 효과를 확인하고, 개선된 음악 시각화 결과가 이전보다 명확한지를 마지막 질문에서 확인한다.

먼저, 태그 의미 명확화를 통해 분리된 태그가 명확한 맥락에서 사용되었는지 확인하기 위해 Self-Organizing Map의 분석에서 자주 활용되는 기법인 Component Plane을 사용한다. Vesanto(1999)에 따르면, Component Plane은 원 데이터의 근사 벡터인 가중치 벡터를 이용하여 특정 변수의 영향력 분포를 2차원 평면에 시각화한 것인데, 변수 간의 관계를 파악하고자 할 때 유용한 시각화 기법이다. Alhoniemi et al.(1999)와 Xiao et al.(2003)를 비롯한 몇몇 연구에서 변수의 영향력을 확인하기 위해 Component Plane을 사용한 바 있다. Component Plane은 가중치 벡터에서 각각의 원소가 노드에 대한 변수의 영향력을 의미한다는 점을 이용한 것이다. 가중치 벡터에서 변수에 대응되는 값이 크면 해당 노드에 대한 변수의 영향력이 큰 것이고, 반대의 경우 영향력이 작은 것인데, 이를 색깔로 코딩하면 Self-Organizing Map 결과에서 변수의 영향력을 쉽게 이해할 수 있다. 본 논문에서는 특정 영역에서의 영향력이 클수록 명암의 구별이 확실히 나타나게끔 하였다.

개선 효과를 확인하기 위해 비효과적 태그 제거만 마친 데이터 셋과 태그 의미 명확화 및 태그 군집화까지 마친 데이터 셋에 대하여 Self-Organizing Map을 수행하였다. 편의를 위하여 각각을 Map_{base}, $Map_{proposed}$라 하자.



[그림 14] 태그 'Chill'의 Component Plane
(a): $\boldsymbol{Map_{base}}$, (b), (c), (d), (e): $\boldsymbol{Map_{proposed}}$

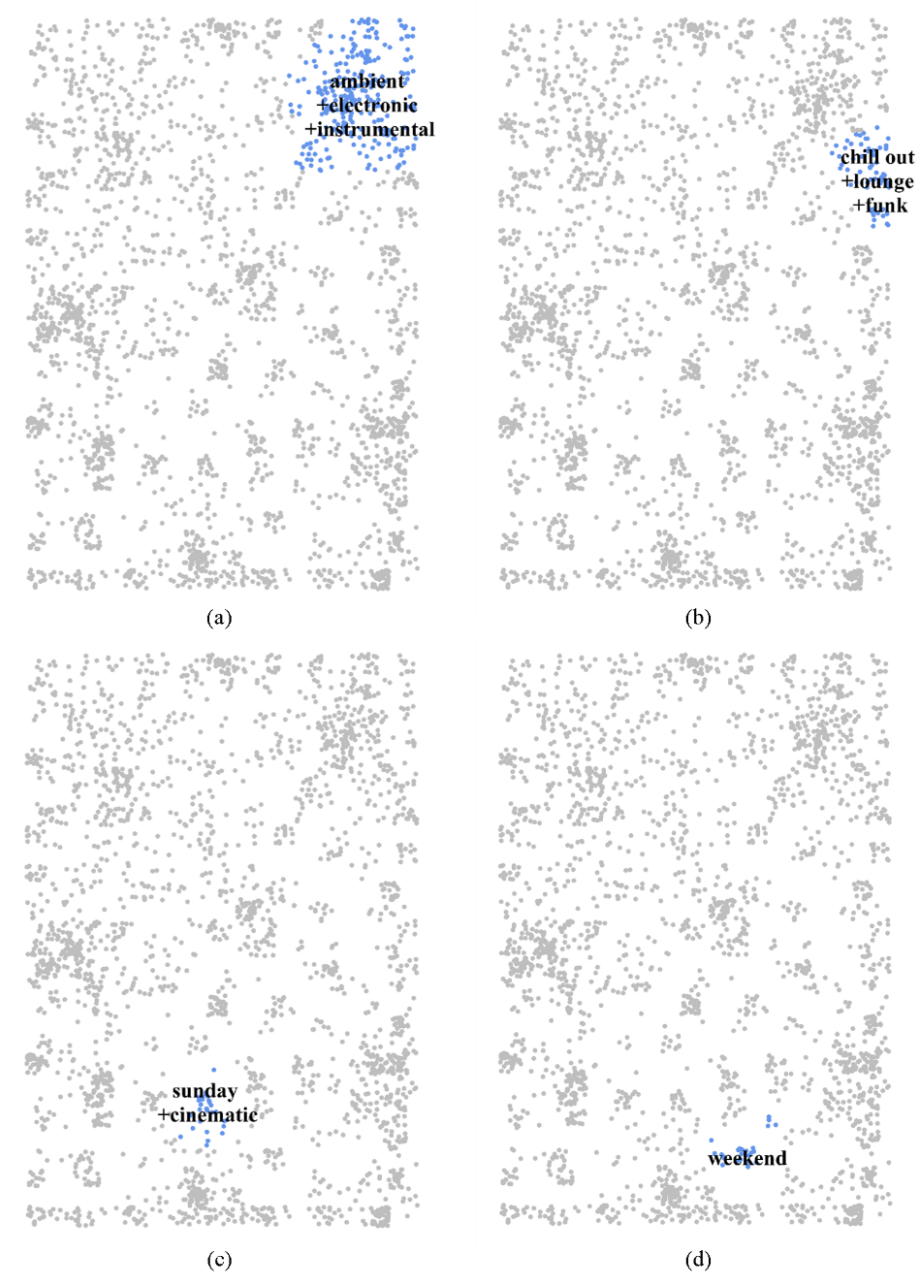
[그림 14]에서 (a)는 Map_{base}에서 태그 'Chill'의 Component Plane이다. 그림에서 이 태그가 강한 영향력을 발휘한 영역이 여러 군데인 것을 통해 태그가 여러 맥락에서 사용되었음을 쉽게 파악할 수 있다. 태그 'Chill'은 태그 의미 명확화 과정에서 4개의 태그로 분리되었고, 태그 군집화에서 각각 다른 태그들과 통합되었다. 통합된 태그는 이제 하나의 맥락을 의미하는데, 이에 대한 Component Plane을 통해 개선 효과를 확인한다. [그림 14]의 (b), (c), (d), (e)는 각각 $Map_{proposed}$에서 태그 'Chill 1', 'Chill 2', 'Chill 3', 'Chill 4'가 통합된 태그의 Component Plane이다. Map_{base}에서 모호한 태그 'Chill'이 지도의 여러 영역에서 강한 영향력을 발휘한 것에 반해, $Map_{proposed}$에서는 분리된 태그 각각이 명확한 의미를 가지고 지도의 한 영역에서 강한 영향력을 발휘하는 것을 확인할 수 있다. 이를 통해 태그가 맥락에 따라 잘 분리되었음을 직관적으로 확인할 수 있다. 또한 태그의 분리가 변수의 특성을 명확히 하여 시각화 모델의 성능을 향상 시킬 것을 예상할 수 있다.

태그 군집화의 효과를 확인하기 위해서는 각 군집에 속한 태그들이 의미적으로 유사한지를 확인한다. 이를 위해 [표 4]에서 결과를 제시하였다. 군집 2를 예로 들면, 군집에 속한 'Christmas', 'Winter', 'Holiday', 'Blues', 'Xmas'는 모두 크리스마스 분위기를 연상시키는 단어이다. 또 다른 예로 군집 5의 태그인 'Workout', 'Powerful'을 보면, 사람들은 운동할 때에 주로 빠르고 힘찬 노래를 듣는다는 점에서 설득력을 갖는다. 이외에도 여러 군집이 음악적 맥락이 유사한 태그를 포함하고 있어 태그 군집화의 결과가 의미적으로 합당하다고 할 수 있다. 따라서 태그 군집화가 중복성을 감소시켜 시각화 모델의 성능을 향상시킴을 예상할 수 있다.

[표 4] 태그 군집화 결과

군집	태그	레이블
군집 1	Ambient, Electronic, Instrumental, Chillout, Downtempo etc.	Ambient + Electronic + Instrumental
군집 2	Christmas, Winter, Holiday, Blues, Xmas etc.	Christmas + Winter + Holiday
군집 3	Christian, Punk, Christian rock, Worship, Alternative etc.	Christian + Punk + Christian rock
군집 4	Dark, Dream, Halloween, New Wave, Ethereal etc.	Dark + Dream + Halloween
군집 5	Reading, Jazzy easy, Sleepy	Reading + Jazzy easy + Sleepy
군집 6	Workout, Powerful	Workout + Powerful

시각화의 결과를 평가하기 위해 의미적으로 유사한 태그에 대하여 재생 목록을 [그림 15]에 시각화하였다. (a)와 (b)에서 잔잔한 음악과 사람을 진정시켜주는 음악들이 지도의 우측 상단에 위치하고 (c)와 (d)에서는 주말에 어울리는 음악들이 가까이에 위치한다. 이는 음악 지도에서 유사한 상황과 분위기에 어울리는 음악들이 가까이에 위치함을 보여주는 결과이다. 따라서 제안한 프레임워크를 통해 음악을 시각화했을 때, 결과물이 음악 지도로서의 역할을 수행한다고 할 수 있다.



[그림 15] 의미적으로 유사한 재생 목록의 인접성 확인

제 2 항 정량적 평가

본 연구에서 사용한 시각화 기법인 Self-Organizing Map의 두드러지는 장점 중 하나는 시각화의 성능을 정량적으로 확인할 수 있다는 점이다. Self-Organizing Map의 성능을 평가하는 지표로는 Quantization Error(QE)와 Topographic Error(TE)가 주로 사용된다(Pölzlbauer, 2004).

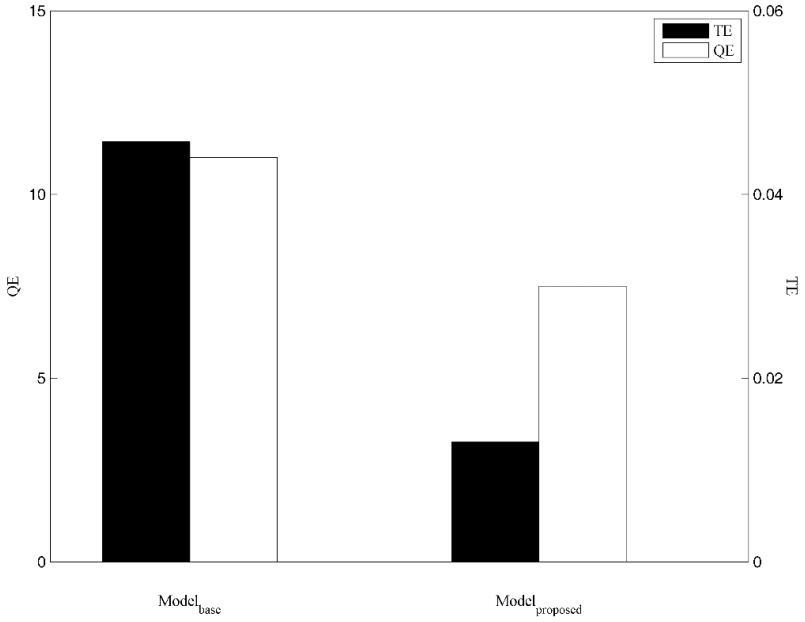
QE는 각 행 벡터로부터 가장 가까운 가중치 벡터까지 거리의 평균으로 정의된다. QE는 데이터를 설명하기에 노드의 수가 충분한지를 나타내는 척도로 해석할 수 있다. 일반적으로 노드의 수가 부족하면 적은 수의 노드로 데이터 전체를 설명해야 하기 때문에 QE가 커지는 것으로 알려져 있다. TE는 Best Matching Unit(BMU)를 이용하는데, 전체 데이터에 대하여 1st BMU와 2nd BMU가 이웃하지 않은 데이터의 비율로 정의된다. TE는 원 데이터의 위상이 2차원 평면에서 제대로 보존되었는지를 나타내는 척도이다. 일반적으로 QE와 TE가 작을수록 시각화가 잘 되었다고 할 수 있다. 그러나 Self-Organizing Map에서 Vector Quantization과 Projection이 서로 상충 효과를 갖는다는 특성 때문에, 일반적으로 QE를 감소시키는 방향으로 데이터를 처리하면 위상 왜곡이 일어나 TE가 증가하는 것으로 알려져 있다(Pölzlbauer, 2004).

정성적 평가에서와 마찬가지로 제안한 프레임워크의 효과를 확인하기 위해 비효과적 태그 제거만을 마친 데이터 셋과 태그 의미 명확화 및 태그 군집화까지 마친 데이터 셋에 대해 Self-Organizing Map을 수행하였다. 편의를 위해 각각을 Map_{base} ,

$Map_{proposed}$라 하자. [표 5]와 [그림 16]에서 결과를 제시하였는데, 실험 결과 QE와 TE 모두에서 오류가 감소하였다. 특히 QE의 경우 원래의 29% 수준으로 감소해 그 효과가 매우 크다. QE와 TE 사이에 상충 효과가 존재함에도 불구하고 최종 단계에서는 두 값이 모두 감소하였기 때문에 태그 의미 명확화와 태그 군집화를 통해 음악 시각화 결과가 보다 정확해졌다고 할 수 있다.

[표 5] QE, TE 결과 표

Map	QE	TE
Map_{base}	11.431	0.044
$Map_{proposed}$	3.260	0.030



[그림 16] QE, TE 결과 그래프

제 5 장 결 론

본 논문에서는 음악을 협업 태깅 시스템에 나타난 사용자 인식에 기반하여 음악을 시각화하고 재생 목록을 구성하는 프레임워크를 제안하였다. 제안하는 프레임워크는 크게 네 단계로 구성된다: 비효과적 태그 제거, 태그 의미 명확화, 태그 군집화, 시각화. 프레임워크의 첫 세 단계를 거치면서 각각 태그의 비효과성과 모호성, 그리고 중복성을 해결하고, 마지막 단계에서는 개선된 데이터 셋을 이용해 유사한 음악이 가까이에 위치하도록 시각화하여 사용자가 상황과 분위기에 맞게 음악을 고를 수 있도록 하였다.

본 연구는 협업 태깅 시스템에서 태그가 갖는 문제를 체계적으로 해결하였다는 점에서 의의를 가지며, 실제 데이터를 이용한 실험을 통해 그 과정을 구체화하였다. 또한 프레임워크의 효과를 정성적 및 정량적인 방법으로 평가하였다. 이러한 시도는 음악에 대한 접근성이 커짐에 따라 발생하는 재생 목록 구성에 대한 수요를 충족시키는 하나의 대안이 될 것이다.

본 연구의 결과물이 복잡하고 변화가 빠른 현실 세계에서 제대로 작동하기 위해서는 추후 연구에서 개선해야 할 사항이 두 가지 있다. 첫 번째 개선 사항은 데이터 전처리 과정에서 태그의 어형을 통합하는 자동화 과정을 도입하는 것이다. 본 연구에서는 어형 통합 과정을 따로 거치지 않고, 태그 군집화 과정에서 비슷한 태그들이 통합되도록 하였는데, 데이터의 희박성이 높은 경우 의미가 비슷한 태그라도 연관고리가 전혀 없으면 통합되지 않을 수

있다. 따라서 이를 데이터 전처리 과정에서 보완하는 것이 필요하다. 이에 대한 해결책으로 Edit Distance(Levenshtein, 1966)나 WordNet(Miller, 1995)에 기반한 방법을 고려할 수 있다. Edit Distance는 텍스트 마이닝에서 어형 통합의 방법으로 사용되는데, 태그 또한 단어의 형태라는 점에서 적용 가능하다. 다만 대부분의 단어가 매우 짧기 때문에 어형 통합의 기준을 구성하는 데에 있어 보수적인 접근이 이루어져야 할 것이다. WordNet은 영어 단어에 대한 온톨로지인데, 어형 통합과 동의어 검출 등의 용도에서 주로 사용된다. 위의 방법을 통해 어형 통합을 수행할 경우, 실질적으로 같은 태그가 다르게 취급되어 생기는 노이즈를 제거할 수 있다.

두 번째 개선 사항은 태그 의미 명확화 과정에서 계산량을 줄이는 문제이다. 본 논문에서는 임의로 분리 대상이 되는 태그를 선택하고 군집화를 수행하였다. 하지만 이 방식의 경우 분리 기준을 충족시키는 태그가 발견될 때까지 같은 과정을 반복해야 한다는 단점이 있다. 만약 분리될 가능성이 높은 태그를 사전에 찾아낼 수 있다면 효율을 높일 수 있을 것이다. 이를 위해 텍스트 마이닝에서 변수 선택을 위해 사용되는 필터 접근법을 차용할 수 있다.

이 외에도 프레임워크 내의 모델에 사용되는 파라미터를 조정하는 문제 등을 고려할 수 있다. 위의 문제를 해결할 경우 본 논문에서 다룬 데이터보다 더 큰 규모의 데이터나 개인이 소장한 음악 데이터 각각에 대하여 자동화된 프레임워크를 적용할 수 있을 것이다.

참고 문헌

Agarwal, A., Xie, B., Vovsha, I., Rambow, O., Passonneau, R., 2011. Sentiment analysis of twitter data. In: Proceedings of the ACL 2011 Workshop on Languages in Social Media. pp. 30-38.

Alhoniemi, E., Hollmén, J., Simula, O., Vesanto, J., 1999. Process monitoring and modeling using the self-organizing map. Integrated Computer-Aided Engineering 6 (1), 3-14.

Au Yeung, C., Gibbins, N., Shadbolt, N., 2009. Contextualising tags in collaborative tagging systems. In: Proceedings of the 20th ACM Conference on Hypertext and Hypermedia. pp. 251-260.

Begelman, G., Keller, P., Smadia, F., 2006. Automated tag clustering: Improving search and exploration in the tag space. In: Proceedings of the Collaborative Web Tagging Workshop at WWW2006. pp. 15-33.

Bernhard, D., 2010. Query expansion based on pseudo relevance feedback from definition clusters. In: Proceedings of the 23rd International Conference on Computational Linguistics. pp. 54-62.

Blondel, V., Guillaume, J., Lambiotte, R., Lefebvre, E., 2008. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 2008 (10), 10008.

Dattolo, A., Eynard, D., Mazzola, L., 2011. An integrated approach to discover tag semantics. In: Proceedings of the 2011 ACM Symposium on Applied Computing. pp. 814–820.

Deerwester, S., Dumais, S., Furnas, G., Landauer, T., Harshman, R., 1990. Indexing by latent semantic analysis. Journal of the American Society for Information Science 41 (6), 391–407.

Gemmell, J., Ramezani, M., Schimoler, T., Christiansen, L., Mobasher, B., 2009. The impact of ambiguity and redundancy on tag recommendation in folksonomies. In: Proceedings of the Third ACM Conference on Recommender Systems. pp. 45–52.

Greene, d., Doyle, D., Cunningham, P., 2010. Tracking the evolution of communities in dynamic social networks. In: Proceedings of the 2010 International Conference on Advances in Social Networks Analysis and Mining. pp. 176–183.

Gupta, M., Li, R., Yin, Z., Han, J., 2010. Survey on social tagging techniques. SIGKDD Explorations Newsletter 12 (1), 58–72.

Hagenau, M., Liebmann, M., Hedwig, M., Neumann, D., 2012. Automated news reading: Stock price prediction based on financial news using context-specific features. In: Proceedings of the 2012 45th Hawaii International Conference on System Sciences. pp. 1040-1049.

Kiang, M., Hu, M., Fisher, D., 2006. An extended self-organizing map network for market segmentation—a telecommunication example. Decision Support Systems 42 (1), 36-47.

Kiviluoto, K., 1998. Predicting bankruptcies with the self-organizing map. Neurocomputing 21 (1), 191-201.

Klein, D., Manning, C., 2003. Accurate unlexicalized parsing. In: Proceedings of the 41st Annual Meeting on Association for Computational Linguistics - Volume 1. pp. 423-430.

Kohonen, T., 1990. The self-organizing map. Proceedings of the IEEE 78 (9), 1464-1480.

Lagus, K., Kaski, S., Kohonen, T., 2004. Mining massive document collections by the WEBSOM method. Information Sciences 163 (1), 135-156.

Lamere, P., 2008. Social tagging and music information retrieval.

Journal of New Music Research 37 (2), 101-114.

Laurier, C., Sordo, M., Serrà, J., Herrera, P., 2009. Music mood representations from social tags. In: Proceedings of the 10th International Society for Music Information Retrieval Conference. pp. 381-386.

Lee, S., Kim, B., Huh, M., Park, J., Kang, S., Cho, S., Lee, D., Lee, D., 2014. Knowledge discovery in inspection reports of marine structures. Expert Systems with Applications 41 (4), 1153-1167.

Lehwark, P., Risi, S., Ultsch, A., 2008. Visualization and clustering of tagged music data. Data Analysis, Machine Learning and Applications. pp. 673-680.

Levenshtein, V., 1966. Binary codes capable of correcting deletions, insertions and reversals. Soviet Physics Doklady 10, 707.

Miller, G., 1995. WordNet: A lexical database for English. Communications of the ACM 38 (11), 39-41.

Mörchen, F., Ultsch, A., Nöcker, M., Stamm, C., 2006. Visual mining in music collections. From Data and Information Analysis

to Knowledge Engineering. pp. 724-731.

Newman, M., 2004. Analysis of weighted networks. Physical Review E 70 (5), 056131.

Pölzlbauer, G., 2004. Survey and comparison of quality measures for self-organizing maps. In: Proceedings of the Fifth Workshop on Data Analysis. pp. 67-82.

Raeder, T., Chawla, N., 2011. Market basket analysis with networks. Social network analysis and mining 1 (2), 97-113.

Rauber, A., Pampalk, E., Merkl, D., 2003. The SOM-enhanced JukeBox: Organization and visualization of music collections based on perceptual models. Journal of New Music Research 32 (2), 193-210.

Salton, G., Buckley, C., 1988. Term-weighting approaches in automatic text retrieval. Information Processing & Management 24 (5), 513-523.

Vembu, S., Baumann, S., 2005. A self-organizing map based knowledge discovery for music recommendation systems. Computer Music Modeling and Retrieval. pp. 119-129.

Vesanto, J., 1999. SOM-based data visualization methods. Intelligent Data Analysis 3 (2), 111-126.

Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J., 1999. Self-organizing map in Matlab: the SOM toolbox. In: Proceedings of the Matlab DSP Conference. pp. 35-40.

Weinberger, K., Slaney, M., Van Zwol, R., 2008. Resolving tag ambiguity. In: Proceedings of the 16th ACM International Conference on Multimedia. pp. 111-120.

Xiao, L., Wang, K., Teng, Y., Zhang, J., 2003. Component plane presentation integrated self-organizing map for microarray data analysis. FEBS Letters 538 (1), 117-124.

Zelikovitz, S., Hirsh, H., 2001. Using LSI for text classification in the presence of background text. In: Proceedings of the Tenth International Conference on Information and Knowledge Management. pp. 113-118.

Abstract

User Perception Based Music Map: Resolving Issues in Collaborative Tagging System

Songwon Han
Department of Industrial Engineering
The Graduate School
Seoul National University

People listen to music in various contexts. In each contexts, different kinds of music are preferred by people, which led to the emergence of playlist. However, with the dramatic increase of the number of available music, people have trouble in generating playlist from available music.

This research proposes the music visualization framework that represents music as user perception based music map by utilizing collaborative tagging system. Tag, in collaborative tagging system, is another representation of user perception. However, lack of coherence of usage pattern is the biggest barrier to utilizing in data mining analysis. Proposed framework resolves ineffectiveness, ambiguity and redundancy of tag based on categorization of tag characteristic and clustering algorithm.

In the next step, the framework represents music on the two-dimensional map with Self-Organizing Map. This research also shows that the framework is effective for music visualization with quantitative and qualitative evaluation. Experiment on real data is conducted for evaluation.

With the results of this research, people can choose music which fit specific context easily by selecting specific region on the music map. This research will be an alternative to 'paradox of choice' which means that too many choices could decrease people's satisfaction.

Keywords : Music Visualization, Ineffective Tag Filtering, Tag Sense Disambiguation, Tag Clustering, Self-Organizing Map
Student Number : 2013-21089