Journal of Convergence for Information Technology
Vol. 11. No. 9, pp. 214-220, 2021
e-ISSN 2586-4440
DOI : https://doi.org/10.22156/CS4SMB.2021.11.09.214

Design and implementation of a music recommendation model through social media analytics

Kyoung-Rock Chung[1], Koo-Rack Park[2*], Sang-Hyock Park[3]
[1]Student, Department of Computer Engineering, Kongju National University
[2]Professor, Department of Computer Science & Engineering, Kongju National University
[3]Student, Department of Computer Engineering, Kongju National University

소셜 미디어 분석을 통한 음악 추천 모델의 설계 및 구현

정경록[1], 박구락[2*], 박상혁[3]
[1]공주대학교 컴퓨터공학과 박사과정, [2]공주대학교 컴퓨터공학과 교수, [3]공주대학교 컴퓨터공학과 박사과정

Abstract With the rapid spread of smartphones, it has become common to listen to music everywhere, just like background music in life, so it is necessary to create a music database that can make recommendations according to individual circumstances and conditions. This paper proposes a music recommendation model through social media. Since emotions, situations, time of day, weather, etc. are included in hashtags, it is possible to build a social media-based database that reflects the opinions of various people with collective intelligence. We use web crawling to collect and categorize different hashtags from posts with music title hashtags to use real listeners' opinions about music in a database. Data from social media is used to create a music database, and music is classified in a different way from collaborative filtering, which is mainly used by existing music platforms.

Key Words : Music Database, Web Crawling, Recommendation System, Collective intelligence, Hash Tag

요 약 스마트폰이 빠르게 보급되면서 음악을 생활 속의 배경음악처럼 항상 모든 곳에서 듣는 것이 일반화되어 개인의 상황과 조건에 맞는 추천을 할 수 있는 음악 데이터베이스를 필요하다. 본 논문에서는 소셜 미디어를 통한 음악추천 모델을 제안한다. 소셜 미디어의 데이터를 사용하여 음악 데이터베이스를 작성하고 기존의 음원 제공 플랫폼이 주로 사용하는 협업필터링과는 다른 방식으로 음악을 분류한다. 웹크롤링으로 음악 제목이 해시 태그로 달린 게시글을 찾아 해당 글에 함께 달린 다른 해시 태그들을 수집하고 분류하여 실제 청취자의 음악에 관한 의견을 데이터베이스에 사용한다. 소셜 미디어를 작성할 때의 감정, 상황, 시간대, 날씨 등 많은 조건이 해시 태그에는 포함되어 있으므로 다양한 사람의 의견이 집단지성으로 반영된 소셜 미디어 기반 데이터베이스를 구축할 수 있다.

주제어 : 음악 데이터베이스, 웹 크롤링, 추천 시스템, 집단 지성, 해시 태그.

1. Introduction

Before digital sound sources were invented, acquiring a device to listen to music, collecting records, and maintaining both required considerable cost and effort. Listening to music as a hobby had to have adequate space and time. However, with the advancement of science and technology, most of the music has been changed from analog to digital, and the price of devices

*Corresponding Author : Koo Rack Park(ecgrpark@kongju.ac.kr)
Received July 31, 2021
Revised September 2, 2021
Accepted September 20, 2021
Published September 28, 2021

that play digital music has become cheaper. As smart phones became widely distributed and applied technologies diversified, they even played a role as a personal audio device. With the expansion of the Internet network, data communication and Wi-Fi can be used freely, and music can be listened to on a smartphone at any time. If you use Internet streaming on your smartphone, you can listen to songs from old to new songs without purchasing a record as before [1]. Therefore, the use of digital sound sources continues to increase[2]. Music listening on a smartphone can be used in two forms: on-demand and streaming by using the Internet [3]. People who didn't listen to music as a hobby but started to listen to music newly with the spread of smart phones came to find and listen to songs that suit each individual situation rather than listening to a favorite song of a certain singer. Due to such a trend, the music service market environment is rapidly changing[4]. Previously, listening to music, which was a noble hobby, became a part of life, and the lifestyle of listening to music as a background in everyday life has become common. The proportion of using music as an auxiliary role that uses music as a background in everyday life has increased than the role whose purpose is to listen to music itself. When (Time), where (Place), and what (Object) in life, music suitable for listening is becoming the standard for music use. Platform sites that provide sound sources use different algorithms to recommend and provide music that consumers want. Collaborative filtering algorithms are widely used as a way to recommend music. It is a method of recommending music to other people based on the past data of each platform site user[5]. However, people who listen to music these days want the music to be recommended according to their situation, time, and purpose. The music recommendation method based on each user's past data is different from the taste of music these days, so it is difficult to recommend the latest music with little information on the sound source. This paper proposes a music recommendation model that utilizes social media hashtags and uses the latest popular music data to meet the needs of music users. If the model proposed in this paper is used, it is expected that it will provide users with music that is recently popular, and that various people have listened to for their purpose.

2. Related Works

2.1. Web crawling

Web crawlers, also known as bots or spiders, are computer programs that search the web in an automated way, and most web search engines use crawlers[6-8]. The user provides a URL or set of URLs to the scheduler and requests the downloader to download the page with the desired URL. After downloading the page, the downloader transmits the page content to the HTML parser, filters the content, and delivers the output to the scheduler. The scheduler stores metadata in the database. The database stores a list of URLs for specific pages in the queue. When a user requests a search, a keyword is provided, delivered to the search agent, and the final result is output using the stored information, the web crawling system collects web pages by performing tasks such as DNS caching, protocol, and encoding[9,10]. Most crawling is based on the link structure[11]. The following Fig. 1 shows structure of a web crawler.

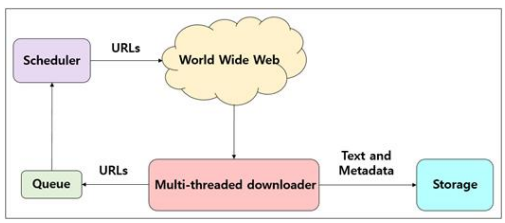

Fig. 1. The structure of a web crawler

2.2. Recommendation System

The recommendation system is a system that provides information desired by users using filtering, and can be divided into content-based filtering and collaborative filtering. Content-based filtering measures the similarity between item content and user information for recommendation, and ranks the results and displays them. Since it is recommended depending on the user profile, only items similar to the items already evaluated by the user are excessively recommended[12,13].

Collaborative filtering collects users' item preferences and recommends items using information of other users who have similar preferences to a specific user. Even if the user has not searched for, the favorite items of other users with similar tastes to the user can be recommended. There is a disadvantage that it is difficult to predict with a collaborative filtering technique because the items already used by some of the users become the basis for recommendation, and new items cannot be included in the recommendation items[14].

2.3. Hash tag

Hash tags include an informative hash tag and a meme hash tag. Informational hash tags are used for the purpose of re-searching posted articles later, and are mainly composed of one word that can represent the posted article. It usually contains useful information, so it is widely shared and used consistently with many users. Meme hash tags encourage users to write articles that include hashtags. It is used in a way to write a text that matches the tag posted by another person, and it is mainly composed of attached sentences that describe the situation. It suddenly becomes popular among users, and then quickly disappears. Particularly, meme hash tags increase in usage when exposed to real-time trends[15].

3. Proposed Model

3.1. Hash tag-based database music recommendation model

Web crawlers, also known as bots or spiders, are computer programs that search the web in an automated way, and most web search engines use crawlers[6].

The following Fig. 2.proposed model configuration. Users can use the Hashtag-based database music service by connecting via mobile. The system server handles crawling, database sorting, and configuration. Hash tags and music-related data are stored in a database.

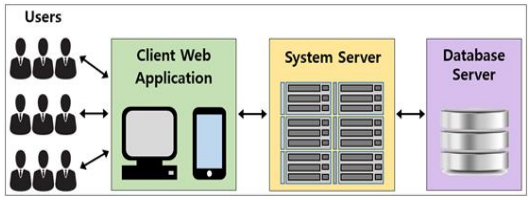

Fig. 2. Proposed model configuration

3.2. System Process

The following Fig. 3. is the structure of the Hash tag-based database music recommendation model system, which is largely divided into three parts.

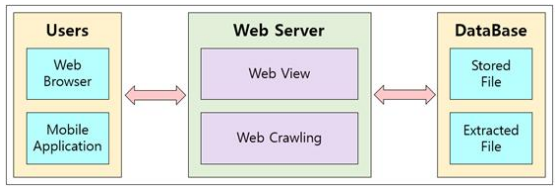

Fig. 3. Proposed system architecture

3.3. Construction of Database

The model proposed in this paper utilizes hashtag data and music ranking chart data. Hash tag data utilizes hash tags on social media. As

data for recommending music, social media hash tags were applied to exclude subjective judgments of specific individuals. As a method of using collective intelligence, the music was classified by collecting hash tags along with the articles related to music posted on social media. The music ranking chart uses the Billboard chart's single ranking, Hot 100, as recommended music data. The Billboard chart, which releases a new ranking every week based on a survey by Nielsen, an information research agency in the United States, is suitable as the data of the latest popular songs objectively verified. As the process of the Hashtag-based database music recommendation model, the process of constructing a music source database is shown in Fig. 4.

Step 1 is Construct a song title database by extracting song titles by web crawling from the Hot 100 on Billboard charts to get the latest trending music data. Step 2 is Collect post data by searching social media for song titles in the song title database by web crawling. Step 3 is Extract all hash tags of the collected post data.

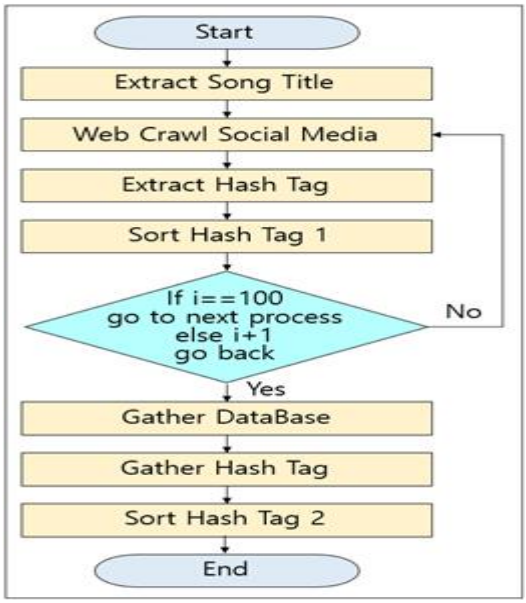

Fig. 4. Design of construction of database

Step 4 is Store the extracted hash tags that are scored in the order of relevance, sorted by ranking. Step 5 is Collect data for the rest of the songs in the song title database through the same process of step 2 through step 4. Step 6 is Build a database that ranks the stored databases with hash tags collected by song title. Step 7 is Build a database that integrates all hash tags Step 8 is Complete a database of songs for each hashtag. Using the completed database, users can view a list of songs by selecting a hashtag.

3.4. Web crawling

Web crawling was used to collect the latest sound source data and hashtags. Trendy music titles were collected by crawling Hot100 song titles on the Billboard charts. The following Fig. 5. is a part of the algorithm to get the title from the billboard chart.

```
def setUpClass(cls):
cls.chart = billboard.ChartData("hot-100")
cls.expectedTitle = "The Hot 100"
cls.expectedNumEntries = 100
.
.
.
>>> song = chart[0]  # Get no. 1 song on chart
>>> song.title
'All I Want For Christmas Is You'
```

Fig. 5. Crawling algorithm

Get and save all song titles on the Hot100 chart for that week. The following Table 1. is a part of the extracted title.

Table 1. Music title

No.	Song Title
1	All I Want For Christmas Is You
2	Rockin' Around The Christmas Tree
3	Jingle Bell Rock
4	A Holly Jolly Christmas
5	Circles
6	Roxanne
7	It's The Most Wonderful Time Of The Year
#	

When searching, the social media address changes according to the song title, so urllib.parse.quote is used so that the search term can be used in the URL. After finding out the address value of the post, put the list in which the address is stored in the for loop and run it, and import and save the ID and hash tag with Beautifulsoup. The following Table 2. is a part of the extracted hashtags.

Table 2. Hash tag

No.	Hash Tag
1	newyear
2	newyearsday
3	newyearseve
4	holiday
5	happynewyear
6	xmas
7	Christmas
#	.....

3.5. Main Screen

The mobile application screen consists of three buttons: chart view, hash tag search, and recommended hash tag. For example, if you select and press the recommended hash tag button, the screen moves to the next recommended hash tag screen, and if you select the desired hash tag from the recommended hash tag screen, you can see the song title associated with the hash tag. The following Fig. 6. is a step-by-step screen of a mobile application that selects a hashtag to view song titles.

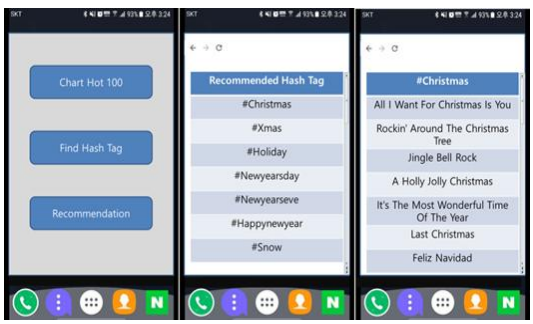

Fig. 6. Mobile application screen

The order of finding the music the user wants is as follows. First, on the first screen, press the recommendation button to select a recommended hash tag. Second, select the hash tag of interest (#Chistmas) of the user from the recommended hash tag. Finally, see a list of recommended music sources based on the hash tag (#Chistmas).

4. Conclusion

The model proposed in this paper was surveyed for a week and evaluated by surveying users' system satisfaction and efficiency. The following Table 3. shows general matters of the evaluation survey.

Table 3. Survey contents

Division	Contents
Survey Method	Hashtag-based database music recommendation model
Test Period	2020.0105~2020.01.11
Target Person	Mobile phone user, Music platform user

A questionnaire survey was conducted for 30 ordinary users who listen to music using the internet network using a mobile device or computer in their daily lives to answer the satisfaction and efficiency of the system in five steps. The following Table 4. is the result of the survey.

Table 4. Results

Division	Satisfaction	Efficiency
Excellent	11	8
Good	9	14
Normal	7	3
Bad	3	5
Very Bad	0	0

As a result of the questionnaire survey, 90% showed above average satisfaction, and 84% showed above average efficiency. By applying the model proposed in this paper, it is possible to create data that allows users to see music in

fields of interest by collecting and providing songs that are related to hash tags. In addition, by using charts updated every week and articles on social media, new sound source data is always created, and recommendations that meet the conditions desired by users with songs that are popular in real time can be provided. The latest trending music list and hash tag-based data reflecting the opinions of social media users can be used to help users select the music they want. However, it is not possible to use information such as opinions and comments about music on SNS, targeting only single word hash tags.

REFERENCES

[1] J. R. Park. (2017). An Analysis of Music Listening Behavior of Korean through Media in the Smartphone Era. *Journal of Music and Theory, 29*, 142-168.

[2] G..Yu. (2014). Advancement Research of Digital Music Contents Service based on Smart Media. *Journal of Korea Design Knowledge, 29,* 259-268.

[3] J. Ahn & S. W. Lee. (2016). The substitutability between streaming radio service and previous digital music service. *Information Society & Media, 17(1),* 31-56.

[4] A. N. Hagen. (2015). The playlist experience: Personal playlists in music streaming services. *Popular Music and Society, 38(5),* 625-645. DOI : 10.1080/03007766.2015.1021174.

[5] S. J. Lee. (2019). *A study of content-based feature exploration using lyrics and music signals for music recommendation.* Master's thesis Seoul National University, Seoul.

[6] A. Heydon & M. Najork. (1999). Mercator: A scalable, extensible web crawler. *World Wide Web, 2(4),* 219-229. DOI : 10.1023/A:1019213109274

[7] C. Castillo. (2004). *Effective Web Crawling*, (Online). http://chato.cl/research/crawling_thesis

[8] A. Kausar, V. S. Dharka & S. K. Singh. (2013). Web Crawler: A Review, *International Journal of Computer Applications, 63(2),* 31-36.

[9] V. Shkapenyuk & T. Suel. (2002). Design and implementation of a high performance distributed web crawler. *In Proceedings of the 18th International Conference on Data Engineering* (ICDE), 357-368.

[10] S. Chakrabarti, B. Dom, S. R. Kumar, P. Raghavan, S. Rajagopalan & A. Tomkins. (1999), Mining the web's link structure, *IEEE Computer, 32,(8),* 60-67.

[11] M. Jamali, H. Sayyadi, B. B. Hariri & H. Abolhassani. (2006). A method for focused crawling using combination of link structure and content similarity. *In Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence,* 753-756.

[12] M. Balabanović & Y. Shoham. (1997). Fab: content-based, collaborative recommendation. *Communications of the ACM, 40(3),* 66-72. DOI : 10.1145/245108.245124.

[13] F. E. Walter, S. Battiston & F. Schweitzer. (2008). A model of a trust-based recommendation system on a social network, Autonomous *Agents and Multi-Agent Systems, 16,* 57-74.

[14] R. Burke. (2002). Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction 12(4),* 331-370. DOI : 10.1023/A:1021240730564.

[15] M. N. Jeon, S. H. Jun, S. M. Rho & E. J. Hwang. (2013). Music recommendation scheme based on the analysis of Twitter hash tags and tweets. *Korea information science society,* 1257-1259

정 경 록(Kyoung-Rock Chung) [정회원]

· 1995년 2월 : 단국대학교 경영학과 (경영학 학사)
· 2019년 2월 : 공주대학교 컴퓨터공학과(공학석사)
· 2019년 2월 ~ 현재 : 공주대학교 컴퓨터공학과 박사과정
· 관심분야 : 웹크롤링, 데이터베이스
· E-Mail : Rokk1969@kongju.ac.kr

박 구 락(Koo-Rack Park) [정회원]

· 1986년 2월 : 중앙대학교 전기공학과(공학사)
· 1988년 2월 : 숭실대학교 전자계산학과(공학석사)
· 2000년 2월 : 경기대학교 전자계산학과(이학박사)
· 1991년 ~ 현재 : 공주대학교 컴퓨터공학부 교수
· 관심분야 : IT 컨버전스, 정보통신, 머신러닝, 전자상거래
· E-Mail : ecgrpark@kongju.ac.kr

박 상 혁(Sang-Hyock Park) [정회원]

· 2009년 2월 : 성결대학교 컴퓨터공학과(공학사)
· 2019년 2월 : 공주대학교 컴퓨터공학과(공학석사)
· 2019년 3월 ~ 현재 : 공주대학교 컴퓨터공학과 박사과정
· 관심분야 : AI, 블록체인
· E-Mail : jobis114@naver.com