International Journal of Scientific Research in Computer Science, Engineering and Information Technology
 | Volume 3 | Issue 3 | ISSN : 2456-3307

A Survey of Music Recommendation System

Puja Deshmukh*, Dr. Geetanjali Kale

Department of Computer Engineering, Pune University, Pune, Maharashtra, India

ABSTRACT

A recommendation system is a system that tries to predict the rating or preference that a user would give to any item. Recommender systems reduces human efforts by recommending items and hence have become increasingly popular in recent years and are utilized in a variety of areas including movies, music, books, research articles, social tags etc. Various techniques have been proposed for performing recommendation, including content based, collaborative, knowledge based and other techniques. To improve the performance, these methods sometimes have been combined to form hybrid recommenders. This paper surveys a general framework on the music recommendation system and various recommendation model. The main motivation behind music recommendation systems is the rapid expansion of digital music formats, managing and searching for songs relevant to the user's requirements from a huge collection of music available since decades.

Keywords : Recommendation System, Feature Extraction, Collaborative Filtering, Content Based Filtering.

I. INTRODUCTION

Music is universal and subjective. It not only can convey emotions but also can modulate listener's mood. Music is easy to listen, but hard to find. What makes it more difficult is that the tastes in music varies from person to person. Also, there is a huge collection of various genres of music composed by many singers. With so many options it is often difficult to know what song to hear next. In the last few decades the access to multimedia content was limited because of less availability of resources, but now due to the huge advancement in the network internet has become the major source of retrieving multimedia information as music, video, books etc.

Music is a very important aspect of human life. Research has also shown that people prefer to listen to music more often than any other activity [1] (i.e. watching television, reading books and watching movies). With the internet technology, a huge amount of music content has become available to millions of users around the world. In recent years, personal music consumption behaviour has also changed dramatically. Due to vast availability of music, user's personal music collection contains plenty of songs. Efficient searching and labelling techniques are required to manage this huge collection [2].

There is a huge collection of various genres of music since decades. With millions of artists and songs in the market it is becoming increasingly difficult for people to search for music. The problem with such huge amount of music is to organize and manage millions of music titles produced by the society. To overcome these problems music recommender systems are used. Music recommender system helps users to filter and select songs as per their tastes. A good recommendation system should be automatically able to detect user preference and generate playlist accordingly.

Music recommendation systems are decision support tools that reduces the information overload by retrieving only those items that are estimated as relevant to users based on the user profile that is generated by the recommendation system [3]. This paper surveys a general framework of the algorithms that have been used by the music recommendation system.

Popular algorithms: Content based model(CBM), collaborative filtering(CF), Context based model, Emotion based model, Hybrid model and Meta data based model.

Based on user's listening behaviour and historical ratings, collaborative filtering algorithm has been found to perform well [4]. Combined with the use of content-based model, the user can get a list of similar songs by low-level acoustic features such as rhythm, pitch or high-level features like genre, instrument etc. [5]. An emotion based model and a context based model have also been proposed [6][7]. Emotion based model recommends music based on the mood of the user [8]. Whereas the context based model collects other contextual information such as comments, music review, or social tags to generate the playlist.
Figure 1 represents a general music recommendation system. The initial phase of any music recommendation system is to collect the information about the music. The second phase consists of collecting information about the user to better understand the user preference and provide a personalized recommendation experience. Feedback can be provided by the users to rate the level of satisfaction from the proposed recommendation system.

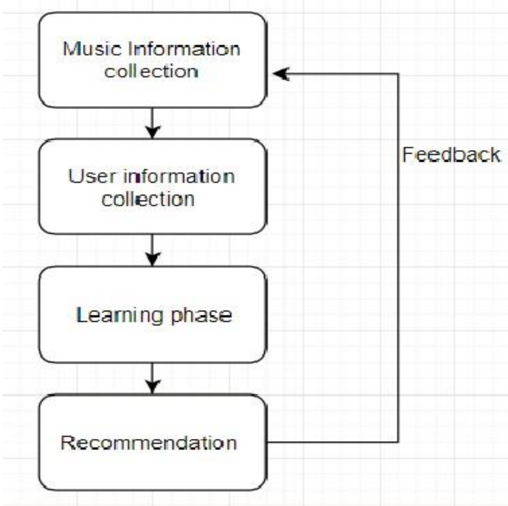

Figure 1. Architecture of music recommendation system

II. RETRIEVING MUSICAL DATA

The following section describes the three main methods of retrieving musical data.

A. String based method

Monophonic music (music with single vocal melody) can be represented by a string of characters. These strings can be used to represent interval sequences, gross contour, sequence of pitches. The occurrence of one string in another, longest common subsequence or the editing distance can be calculated by various string matching algorithms to determine the matching melodies relevant to the user requirement. The music retrieval is done by checking exact match of query string in the database entries. "Musipedia" is a music retrieval system that makes use of this technique. Indexing to database entries is done by standard indexing methods as B-trees, inverted files etc. Sometimes there is lack of words in the music composition, in such cases the melodies are cut into n-grams and then indexed [9].

B. Set based method

These methods are mainly used for polyphonic music (more than one parts, means simultaneous notes). In these methods music is viewed as a set of events with properties like note onset time, pitch and duration.

The onset times are quantized and the music is segmented into measures, Clausen et al. (2000) [10] uses inverted files to perform indexing. Distances to a fixed set of objects are pre-calculated for each database entry. Queries then only need to be compared with similar distance to objects.

C. Probabilistic model

This method aims at determining the properties of the candidate sets in the database and comparing these properties with the corresponding properties of the queries. GUIDO [11] is the system that generates Markov model for music retrieval. Indexing in probabilistic models can be done using trees.

III. MUSIC FEATURES

A feature is a characteristic that helps to distinguish one thing from another (or one group of things from another group of things). There are various properties that can be selected as features in recommendation system. In the last, few decades' internet tags that were assigned by the users to the music were used as a feature to extract and recommend music. Although this method has a limitation, that the new songs cannot be recommended until they have been annotated. Hence features such as loudness, pitch, rhythm, etc. are considered for music recommendation. The music is segmented into short frames to extract features and to measure distance between the candidate pieces in the database [12].

Music features such as melodies, rhythms and chords from the music data are used to generate indices that will help in faster retrieval of relevant music quickly [13]. Music features can be divided into three parts as frequency features, audio perceptual feature and statistical characteristics of beat. Based on frequency domain and time domain signal processing, wavelet analysis and singular value decomposition [14] a method for music feature extraction has been proposed. This method shows a high precision of 95.33%.

Other techniques, such as Fourier transform, Short Time Fourier Transform (STFT) , Wavelet transform , constant-Q transform (CQT) , Gammatone filter bank (GTFB) and multiscale spectro-temporal modulations (MSTM) [15,16,17,18,19] are used for music feature extraction.

IV. ELEMENTS OF MUSIC RECOMMENDATION SYSTEM

A successful music recommendation system must satisfy all the user requirements. The performance of the music recommendation system depends on the performance of its components. There are 3 key components in the system which includes: User profile, Item profile and Query type.

A. User profile

User profile is created by collecting basic information about the user as age, gender, geographical region etc. Music is recommended to the user based on their profile. However, obtaining user information is a very expensive task and requires lots of human efforts [20]. Music preference also depends on the user's personality and intelligence. Rentfrow and S Gosling [21] have proposed the relationship between the user's music preference and BFI (openness, conscientiousness, extraversion, aggregable and neuroticism). Clema [22] proposed that the user profile can be categorized as demographic (age, marital status etc.), geographic (location, city etc.) and psychographic (mood, attitude etc). Shao [23] stated that there are diverse types of listeners, categorized as savants, enthusiasts, casuals and indifferent. Listening patters and access behaviours are also important parameters to be considered during user profiling.

B. Item profile

Item profile gives information about the music. Music can be divided into three categorized as: Editorial metadata, cultural metadata and acoustic metadata. Editorial metadata includes the data that has been provided by the music editor, which includes composer of music, title, genre etc. Cultural metadata includes the data obtained after the analysis of textual information which is obtained from the public sources. As an example, cultural metadata might include similarity between music items. Acoustic metadata is obtained after the analysis of audio signals. Acoustic metadata includes beat, pitch, temp etc.

C. Query type

Traditional methods for querying music is based on textual queries or queries regarding the metadata of music. However, in the past decade a technique called as "Query by humming/singing" system (QBHS) was introduced [24]. This system allowed the user to find the songs by humming or singing. Although this method requires lots of human efforts.

V. MUSIC RECOMMENDATION TECHNIQUES

An efficient music recommendation system should automatically recommend personalized music to the listener [25]. What makes music recommendation difficult is that unlike movies or books, the length of music is shorter and people tend to listen to their favourite songs more than once. So far, many music discovery websites such as Last.fm, AllMusic, Pandora, Audiobaba6, Mog7, Musicovery 8, Spotify9, Apple "Genius" have aggregated millions of users, and the development is explosive [26].

A. Metadata based information retrieval

The most common and traditionally used method for music retrieval is based on the metadata of the music. This method uses editorial information such as title of the song, artist name, lyrics, movie name etc. It is the simplest way to search for music. [27]. Even so this method has certain limitations that the user must know the editorial information about their favourite songs, the recommendation is also poor as none of the user's information is considered. Metadata based information retrieval is also time consuming.

B. Collaborative Filtering (CF)

This approach is based on the concept of recommend music to the user based on the choice of other similar users. Collaborative filtering assumes that if the users' A and B rate n items similarly then they will also rate other items similarly. Collaborative filtering is further divided into three categories: memory based, model based and hybrid collaborative filtering [28].

- Memory based CF: Memory based collaborative method is based on the observation that people tend to trust the recommendations from their similar minded friends. Memory based CF models apply nearest neighbour like algorithms to determine the similarity between the users. They work by grouping the users according to similar interests such that whenever new item arrives, nearest neighbour algorithm can be used to determine massive number of explicit user votes [29].
- Model based CF: Model based CF first generates a descriptive model of user preferences and then uses this model for predicting ratings [30]. Many of these models are based on neural network classifier, Bayesian network, item based CF, association rule mining linear classifiers.
- Hybrid model: Hybrid collaborative filtering model works by combining all the other CF approaches. Hybrid CF model gives better results than any individual CF model [31].

Although CF works well it faces some key problems such as cold start, popularity bias and human efforts [32]. Due to popularity bias, the song that is popular gets higher ratings however the song that is not so popular but good might not be considered for recommendation. Sometimes at an early stage very few ratings are provided to the songs, this results in poor recommendation results. This problem is known as the cold start/data sparsity problem.

C. Content Based music retrieval

In content based approach the music recommendation is done based on the analysis of the sound track. In this approach, songs are recommended to the users based on their listening history. Lots of research has been done to determine the acoustic features of the song. Mostly used features are timbre and rhythm. Based on these features the similarity between the songs is determined. Most common methods to compute similarity are: K-means clustering [33] , Expectation-Maximization with Monte Carlo sampling , and Average Feature Vectors with Euclidean Distance.

Songs can be expressed by humming and singing. In 1990's using the content based retrieval model a query by humming system(QBSH)was proposed [34]. A better performance using content based model can be achieved by embedding lyrics and enhancing the main voice along with the melody.

Since content-based model largely depends on acoustic features, the number of selected features needs to be further considered, this is main limitation of content based music retrieval. Moreover, other user information and non-acoustic information should be included for future modification and augmentation.

D. Emotion based model

Music is believed to trigger human emotions. Music emotion has appealed huge research and it has become the main trend for music recommendation. "Musicovery", a web service uses a fundamental emotion model (2D Valence-Arousal). It allows the users to locate their emotions in a 2D space, Valence (positive or negative) and Arousal (exciting and calming). Human emotions are also dependent on different acoustic cues. Different music features such as energy, rhythm, harmony have been used to determine emotions.

Research has also shown that user's mood also plays a key role in selecting the songs. With the advancement in technology, people use social networking sites [35] to express themselves through posts having positive, negative and neutral emotions. These sentences can be used to determine the mood of the user and then recommend him songs accordingly. Major advantage of emotion based model is that user emotions are considered while recommending songs to them.

Although emotion based model provides greater user satisfaction as compared to any other model it still has certain limitations. First limitation includes data collection, emotion based model requires huge number of datasets that requires huge human efforts [36]. Second limitation includes ambiguity and granularity. The same feeling experienced by different people might result into different emotions. Hence it is not easy to define emotions.

E. Context based model

Context based model considers the public opinions to recommend music. Public opinions can be obtained through social networking sites such as Facebook, Twitter, YouTube etc. in the form of comments, reviews, posts etc. Context based model uses web document mining techniques to obtain important information. Research has suggested that this model performs well due to collection of social information.[37] proposes an efficient approach to use the music contents and context together. Music can also be efficiently recommended considering the social influence, which gives information regarding the similar interests of the users. [38] suggests an efficient approach to develop a model that considers the social impact while recommending music. [39] shows an approach in which the users previous listening behaviour can be used to recommend them music in real time. The results show that this method behaves very efficiently on the sparse data. [40] shows the effective of considering the context of music during recommendation, how considering previous listening records improves the efficiency of recommending the next song. [41] proposes a multidimensional context model for music recommendation. This model tries to eliminate the issue of how the user preference varies in a mobile environment. [42] proposes a music recommendation system named Just-for-Me which considers user location related context and then provides recommendation accordingly.

F. Hybrid model

Hybrid model combines two or more of the above stated models to improve the performance of the music recommendation system. Broke [43] proposed that various methods such as weighted, switching, mixed, feature combination and cascade can be used to develop a hybrid recommendation model. [44] proposes a hybrid model that combines content and collaborative approach. [45] shows a deep learning approach in which it combines the process of extracting music content and recommendation using deep network and probabilistic model.

This approach not only outperforms the traditional hybrid recommendation model but also improves the performance of CF recommendation. [46] shows an approach to combine the tags that are generated from the context of music and how it improves the recommendation of music for queries that are well tagged as well as sparsely tagged. [47] proposes a hybrid approach of combining the audio features extracted from the songs and user personalities using the support vector machines. Another important context for music recommendation is the user's mood, the model proposed in [48] determines the mood of the user from micro blogs and then combines this mood to recommend music to the users.

VI. CONCLUSION

In this paper, we have explained the components of a music recommendation system and various models that can be used for recommendation such as content based recommendation, collaborative recommendation, metadata based recommendation, emotion based recommendation etc. Collaborative recommendation model has achieved great success but it has drawbacks as popularity bias, human efforts etc. On the other hand, hybrid music recommendation model has less disadvantages and better performance but, its complexity has not yet been studied fully. Emotion based and Context based model highly improves the quality of recommendation by considering the social information. The research regarding these models is still in an early stage and increasing rapidly.

The recommenders from last few years have been improved to provide more personalized and subjective music recommendation. A great amount of work in recent years have been done in music perception, psychology, neuroscience which studies the relationship between music and the impact of human behaviour. Music is an important part of human life and now with the advancement in technology we have an easy access to it.

Designing a personalized music recommendation system is complicated as it is difficult to properly understand the user requirements and meet their needs. Therefore, future recommendation systems should allow the users to reasonably choose music. Along with music recommendation if the system generates an automatic playlist it will result in a greater user satisfaction. To the end, we hope this study will help to build a bridge between isolated researches in all the other disciplines.

VII. REFERENCES

[1]. Rentfrow, Peter J., and Samuel D. Gosling. "The do re mi's of everyday life: The structure and personality correlates of music preferences." Journal of personality and social psychology 84.6 (2003): 1236.

[2]. Song, Yading, Simon Dixon, and Marcus Pearce. "A survey of music recommendation systems and future perspectives." 9th International Symposium on Computer Music Modeling and Retrieval. Vol. 4. 2012.

[3]. Ricci, Francesco, Lior Rokach, and Bracha Shapira. "Introduction to recommender systems handbook." Recommender systems handbook. springer US, 2011. 1-35.

[4]. Bogdanov, Dmitry, and Perfecto Herrera. "How Much Metadata Do We Need in Music Recommendation? A Subjective Evaluation Using Preference Sets." ISMIR. 2011.
[5]. Burke, Robin. "Hybrid recommender systems: Survey and experiments." User modeling and user-adapted interaction 12.4 (2002): 331-370.
[6]. Song, Yading, Simon Dixon, and Marcus Pearce. "A survey of music recommendation systems and future perspectives." 9th International Symposium on Computer Music Modeling and Retrieval. Vol. 4. 2012.
[7]. Kim, Youngmoo E., et al. "Music emotion recognition: A state of the art review." Proc. ISMIR. 2010.
[8]. Eerola, Tuomas, and Jonna K. Vuoskoski. "A comparison of the discrete and dimensional models of emotion in music." Psychology of Music 39.1 (2011): 18-49.
[9]. Content Based Music Retrieval - Music Formats, Retrieval tasks, searching symbolic music, Searching musical audio, Feature extraction, Audio Fingerprinting, Concluding Remarks. Information on http://encyclopedia.jrank.org/articles/pages/6698/Content-Based-Music-Retrieval.html.
[10]. Clausen, Michael, et al. "PROMS: A Web-based Tool for Searching in Polyphonic Music." ISMIR. 2000.
[11]. Kornstadt. Themefinder: A web-based melodic search tool. In W. Hewlett and E. Selfridge-Field, editors, Melodic Similarity: Concepts, Procedures, and Applications, Computing in Musicology, volume 11. MIT Press, Cambridge, 1998.
[12]. S. Downie. Evaluating a simple approach to music information retrieval: Conceiving melodic n-grams as text. PhD thesis, University of Western Ontario, London, Ontario, Canada, 1999.
[13]. Michael TANG, YIP Chi Lap and Ben KAO, "Selection of Melody Lines from Music Databases," IEEE Multimedia, Page 243-248,2000.
[14]. Gao, Zhanchun, Yuting Liu, and Yanjun Jiang. "An effective method on content based music feature extraction." Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2015 IEEE. IEEE, 2015.
[15]. Cataltepe, Zehra, and Berna Altinel. "Hybrid music recommendation based on different dimensions of audio content and an entropy measure." Signal Processing Conference, 2007 15th European. IEEE, 2007.
[16]. Wang, Fei, et al. "Tag Integrated Multi-Label Music Style Classification with Hypergraph." ISMIR. 2009.
[17]. Panagakis, Yannis, and Constantine Kotropoulos. "Music classification by low-rank semantic mappings." EURASIP Journal on Audio, Speech, and Music Processing 2013.1 (2013): 13.
[18]. Qi, Jun, et al. "Auditory features based on gammatone filters for robust speech recognition." Circuits and Systems (ISCAS), 2013 IEEE International Symposium on. IEEE, 2013.
[19]. Mesgarani, Nima, Malcolm Slaney, and Shihab A. Shamma. "Discrimination of speech from nonspeech based on multiscale spectro-temporal modulations." IEEE Transactions on Audio, Speech, and Language Processing 14.3 (2006): 920-930.
[20]. Turnbull, Douglas, Luke Barrington, and Gert RG Lanckriet. "Five Approaches to Collecting Tags for Music." ISMIR. Vol. 8. 2008.
[21]. Rentfrow, Peter J., and Samuel D. Gosling. "Message in a ballad: The role of music preferences in interpersonal perception." Psychological science 17.3 (2006): 236-242.
[22]. O. Celma Herrada. Music Recommendation and Discovery in the Long Tail. PhD Thesis, 2009.

[23]. Shao, Bo, et al. "Music recommendation based on acoustic features and user access patterns." IEEE Transactions on Audio, Speech, and Language Processing 17.8 (2009): 1602-1611.

[24]. Ghias, Asif, et al. "Query by humming: musical information retrieval in an audio database." Proceedings of the third ACM international conference on Multimedia. ACM, 1995.

[25]. Lamere, Paul. "Social tagging and music information retrieval." Journal of new music research 37.2 (2008): 101-114.

[26]. Cano, Pedro, Markus Koppenberger, and Nicolas Wack. "Content-based music audio recommendation." Proceedings of the 13th annual ACM international conference on Multimedia. ACM, 2005.

[27]. Downie, J. Stephen. "Music information retrieval." Annual review of information science and technology 37.1 (2003): 295-340.

[28]. Su, Xiaoyuan, and Taghi M. Khoshgoftaar. "A survey of collaborative filtering techniques." Advances in artificial intelligence 2009 (2009): 4

[29]. Nilashi, Mehrbakhsh, Othman Ibrahim, and Karamollah Bagherifard. "A recommender system based on collaborative filtering using ontology and dimensionality reduction techniques." Expert Systems with Applications 92 (2018): 507-520.

[30]. Jiang, Zhengshen, et al. "Recommendation in Heterogeneous Information Networks Based on Generalized Random Walk Model and Bayesian Personalized Ranking." Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. ACM, 2018.

[31]. Appalla, Padmaja, et al. "Hybrid Fuzzy Recommendation System for Enhanced E-learning." Advances in Systems, Control and Automation. Springer, Singapore, 2018. 21-32.

[32]. Herlocker, Jonathan L., et al. "Evaluating collaborative filtering recommender systems." ACM Transactions on Information Systems (TOIS) 22.1 (2004): 5-53.

[33]. Ferretti, Stefano. "Clustering of Musical Pieces through Complex Networks: an Assessment over Guitar Solos." IEEE MultiMedia (2018).

[34]. Wang, Chung-Che, Jyh-Shing Roger Jang, and Wennen Wang. "An Improved Query by Singing/Humming System Using Melody and Lyrics Information." ISMIR. 2010.

[35]. Chen, Jinpeng, Pinguang Ying, and Ming Zou. "Improving music recommendation by incorporating social influence." Multimedia Tools and Applications (2018): 1-21.

[36]. Janto Skowronek and M McKinney. Ground Truth for Automatic Music Mood Classification. In Proc. ISMIR, pages 4–5, 2006.

[37]. Katarya, Rahul, and Om Prakash Verma. "Efficient music recommender system using context graph and particle swarm." Multimedia Tools and Applications 77.2 (2018): 2673-2687.

[38]. Chen, Jinpeng, Pinguang Ying, and Ming Zou. "Improving music recommendation by incorporating social influence." Multimedia Tools and Applications (2018): 1-21.

[39]. Wang, Dongjing, Shuiguang Deng, and Guandong Xu. "Sequence-based context-aware music recommendation." Information Retrieval Journal (2017): 1-23.

[40]. Vall, Andreu, et al. "The Importance of Song Context in Music Playlists." Proceedings of the Poster Track of the 11th ACM Conference on Recommender Systems (Rec-Sys), RecSys. Vol. 17. 2017.

[41]. Sassi, Imen Ben, Sadok Ben Yahia, and Sehl Mellouli. "User-based context modeling for music recommender systems." International Symposium on Methodologies for Intelligent Systems. Springer, Cham, 2017.

[42]. Schedl, Markus, and Dominik Schnitzer. "Location-aware music artist recommendation." International Conference on Multimedia Modeling. Springer, Cham, 2014.

[43]. Burke, Robin. "Hybrid recommender systems: Survey and experiments." User modeling and user-adapted interaction 12.4 (2002): 331-370.

[44]. Katarya, Rahul, and Om Prakash Verma. "Efficient music recommender system using context graph and particle swarm." Multimedia Tools and Applications 77.2 (2018): 2673-2687.

[45]. Wang, Xinxi, and Ye Wang. "Improving content-based and hybrid music recommendation using deep learning." Proceedings of the 22nd ACM international conference on Multimedia. ACM, 2014.

[46]. Horsburgh, Ben, Susan Craw, and Stewart Massie. "Learning pseudo-tags to augment sparse tagging in hybrid music recommender systems." Artificial Intelligence 219 (2015): 25-39.

[47]. Cheng, Rui, and Boyang Tang. "A music recommendation system based on acoustic features and user personalities." Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, Cham, 2016.

[48]. Deng, Shuiguang, et al. "Exploring user emotion in microblogs for music recommendation." Expert Systems with Applications 42.23 (2015): 9284-9293.