Modeling the Music or the User? A review of Music Recommender Systems

Nishant Mohan
19300436

Trinity College Dublin, The University of Dublin
mohanni@tcd.ie

Abstract. With expanding reach of internet, popularity of entertainment in audio-visual format has risen tremendously. There seems to be a renewed interest in online music streaming in the past decade rather than using DVDs or radio for music. Applications such as Spotify, Apple Music and Amazon Music have seen tremendous rise in their userbase. Popular video streaming service YouTube launching its music application is a further proof of the same. A successful application is one that can tap the nerve of the user by providing best user experience. In order to be successful an application should be usable as well as useful. While usability can be improved with superior and smooth application design, one of the most important factors in improving usefulness, in case of music applications, is whether the users can easily find the music of their choice. There has been much research in the field of Recommender Systems focusing on music. This paper presents an overview of two major algorithms in Music Recommender Systems- based on collaborative filtering and based on content. We explain the difference between these and review some state-of-the-art recommendation algorithms based on these methods, along with other hybrid methods.

Keywords: Music Recommender, Information Retrieval, Collaborative Filtering, Content-based Filtering

1 Introduction

In the last few decades, internet has become our primary source of entertainment. Be it watching movies, television series or listening to music, we are consuming more and more multimedia content. Though this is true for the entertainment industry, music industry in particular has seen an interesting pattern. The revenue of global recorded music industry noticed a gradual decline beginning this millennium till 2014, after which it slightly increased for the next 5 years. Despite irregular trend in overall revenue, the share generated by streaming music has been continuously increasing from 2005. While the statistics show little share of streamed music prior to 2005, in 2018 it was 46.9% of the total revenue. In 2018 alone, the growth in paid music streaming revenue was 32.9%, global revenue growth being 9.7%. This translates to an incredible 34% growth in overall streaming revenues [1]. Therefore, streaming seems to be the next big thing in music industry. While Spotify takes credit for providing superior user experience, YouTube Music which was launched in May 2018, has made heads turn too. Among all the glittery numbers, however, the reports also point out that the users feel that all the streaming services are too similar [2].

With Spotify, Apple Music and Amazon Music leading the brigade, users played 91 billion songs in 2018 [2]. Owing to the evident demand, there has been considerable interest in the field of Music Recommender Systems (MRSs) recently [3]. With such huge volume of music available, organizing, managing and recommending titles to a user is a daunting task. From users' perspective, music needs to be a seamless experience in the sense that the integration of music in their lives should be effortless. Users need finer and even more personalized experience. Recommender systems perform precisely this task.

In this paper, we survey some state-of-the-art approaches in music recommendation. Further paper is structured as follows. Section 2 explains in detail about recommendation systems and music recommenders. In section 3, we discuss some popular state-of-the-art approaches towards recommending music to the user. This article may therefore serve as an entry point for young researchers in the field of Music Recommendation Systems.

2 Music Recommender Systems

The aim of a general recommender system is to provide users with personalized recommendations of an online product or service. This is to handle the increasing online information overload problem while improving customer relationship management at the same time [4]. Burke et. al. [10] provide a summary of recommendation techniques as provided in Table 1.

Table 1. Recommendation Techniques. Assume that I is the set of items over which recommendations might be made, U is the set of users whose preferences are known, u is the user for whom recommendations need to be generated, and i is some item for which we would like to predict u's preference. [10]

Technique	Background	Input	Process
Collaborative	Ratings from U of items in I.	Ratings from u of items in I.	Identify users in U similar to u, and extrapolate from their ratings of i.
Content-based	Features of items in I	u's ratings of items in I	Generate a classifier that fits u's rating behavior and use it on i.
Demographic	Demographic information about U and their ratings of items in I.	Demographic information about u.	Identify users that are demographically similar to u, and extrapolate from their ratings of i.
Utility-based	Features of items in I.	A utility function over items in I that	Apply the function to the items and determine i's rank.

		describes u's preferences.	
Knowledge based	Features of items in I. Knowledge of how these items meet a user's needs.	A description of u's needs or interests.	Infer a match between I and u's need.

In MRSs, two methods are broadly used, among others, to identify user taste and suggest recommendations. These are collaborative filtering and content-based music information retrieval [3]. All other methods either use a hybrid algorithm or use some addition on top of these algorithms.

Collaborative filtering (CF) refers to recommendation techniques wherein a recommendation is based on shared interests [4]. Item-based CF approach refers to recommendation based on preferences as seen in the past, while user-based CF refers to the recommendations based on likings of other users [4,5,6]. Though domain-agnostic, CF have been successfully applied to music recommenders [7,8, 11,12,13].

Content-based (CB) Music Information Retrieval (MIR) refer to the strategies that enable access to music collections as per the search of the user [9, 14, 15]. These strategies may provide information about the song at different levels through metadata (artist, album, etc.) or signal analysis.

A combination of CF, CB and other methods is generally used to suggest a track to the user. Popular music streaming services such as Spotify, Apple Music, SoundCloud, Shazam use sophisticated version of these algorithms in one form or the other.

3 State-of-the-art approaches in Music Recommender Systems

Apart from CF and CB, Music Recommender Systems can use several methods such as Metadata information retrieval, Emotion-based models and Context-based models to suggest users with a song [19]. Metadata information retrieval-based models simply use available editorial information provided such as artist, genre and lyrics to target a song. Emotion-based models use the energy, rhythm, spectral and temporal features in songs to identify the emotion contained in the song. Context-based models utilize the social network data through text-mining to obtain public opinions of different songs [19].

In this paper we limit ourselves to the most popular methods CF and CB, being applied in music recommendation in recent years.

3.1 Collaborative filtering

Collaborative Filtering (CF) methods provide recommendations based on the feedback given to tracks from other users. There are two approaches within CF- memory-based and model-based.

Memory-based CF methods work by identifying nearest neighbors using preferred items. This method performs good when recent changes in preferences are to be updated

[11]. Model-based CF methods improve scalability which is a shortcoming of memory-based CF. Using item rating provided by users, they find neighboring items rather than neighboring users [11]. However, both methods do not perform well on Gray-sheep users. Gray-sheep are those users whose preference in music is eccentric. Therefore, not many neighbors can be found for these users.

Sánchez-Moreno et. al. [11] proposed use of artist and user playing coefficients to determine the degree to which a user is a gray-sheep. They derived the user ratings for CF algorithms from counts of times the gray-sheep user plays a track. Their algorithm significantly outperforms other CF methods.

Neural networks and Deep learning have also been applied in the field of MRS to learn interaction function in data. He et. al. [12] presented a novel framework using Neural networks for collaborative filtering (NCF) to model user-item interactions. Their work focuses on implicit feedback received from user such as clickstream data and purchasing products, rather than explicit feedback such as ratings and reviews.

Dias et. al. [13] studied temporal patterns in a session by characterizing it explicitly as well as implicitly. Explicitly, they modeled temporal data based on properties such as time of the day, weekday, or month. On the other hand, implicit characteristics are modeled based on the session diversity measured by the ratio of distinct songs and total songs played within one session. Their results show up to 200% better accuracy as compared to standard session-based CF algorithms.

Though used extensively, cold-start is a major drawback of CF. Cold-start problem occurs when there is no previous usage data. Therefore, CF fails to recommend new and unpopular songs. CB methods provide a solution to this problem, as discussed in the next section.

3.2 Content-based filtering

CB filtering helps identify and extract features such as genre and mood from the audio signals. Content-based music description can help users identify what kind of music they want. In fact, these techniques work even in situations when user does not know what to search specifically [9]. For instance, with CB methods, users can identify a track if they do not remember the lyrics but only the melody. Services such as shazam work using these algorithms.

Oord et. al. [14] assert that recent advances in deep learning work very well for MRS. Their deep convolutional neural network successfully establishes that latent features can very well be predicted from music audio. In fact, their model performed better than the more conventional bag-of-words model, which is frequently used in MIR.

Soleymani et. al. [15] contend that using genre for music recommendation could be ambiguous as one song could fit into tens of genres. They proposed to use a five-factor model named as MUSIC by psychologists after Mellow, Unpretentious, Sophisticated, Intense and Contemporary music preference factors. The authors used the five- factor model to detect CB attributes using auditory modulation features. Their research shows that such a model works well in cold start problems. Also, this CB method performs better than different methods focusing on genre or artists.

3.3 Hybrid Models

Considering limitations in individual usage of both CF and CB, researchers have often used a combination of one or more methods [16,17,18]. Apart from CB and CF, these may also use context-based filters, emotion-based filtering and others. Different methods may be combined based on weighted, feature combination, cascading, feature augmentation or meta-level methods.

Vall et. al. [16] introduced a hybrid playlist continuation model which analyzed if a song fits to given playlist, by treating a playlist-song pair exclusively as a feature vector. This is aimed to prevent leaning towards a particular playlist and song. Therefore, the model helps in extending a playlist by picking songs from a pool, which was not seen at the time of training.

Chiliguano et. al. [17] proposed a deep-learning model that used user information and high-level information from audio data. Their convolutional neural network method performs similar to the existing genre classification MIR methods.

Schedl et. al. [18] proposed three computational features, namely diversity, mainstreaminess and novelty in the user's music taste. They employed several variations of standalone and hybrid recommenders using these three as well as other features based on popularity and location. Authors concluded that hybrid methods generally outperform standalone recommenders across user groups.

4 Conclusion

We presented a comprehensive review of some state-of-the-art methods in music recommenders. Specifically, we looked at some methods researchers have used in past three to four years in collaborative filtering, content-based music information retrieval as well as some hybrid approaches. It can be concluded that hybrid approaches perform better than any individual method as they help to incorporate features from across domains.

There is a growing interest towards using deep-learning convolutional neural networks to identify latent features within music. Also, researchers are enthusiastic about identifying new ways to model user's expectations when it comes to music recommendation. However, there is still a dearth of trustworthy unbiased data that can be used to validate many hypotheses. Since we are gradually moving towards a connected world with the Internet Of Things (IOT), it can be expected that in future we would have more and more relevant user data that can be used for improved MRS.

We hope that the present paper helps young researchers who are new to the field of music recommenders in getting started. Music recommendation is a field which involves human emotions and personalities, which are quite subjective and difficult to model. However, there is still scope for immense improvement and innovation in this field.

References

1. Global Music Report 2019, International Federation of the Phonographic Industry, https://www.ifpi.org/news/IFPI-GLOBAL-MUSIC-REPORT-2019, last accessed 15 Dec. 2019.
2. Savage, M: Music streaming market 'needs more choice', https://www.bbc.com/news/entertainment-arts-50472906, last accessed 15 Dec. 2019.
3. Schedl M., Knees P., McFee B., Bogdanov D., Kaminskas M.: Music recommender systems. In: Recommender systems handbook, chap. 13, 2nd edn. Springer, Berlin, pp 453–492 (2015).
4. Jie L., Dianshuang W., Mingsong M., Wei W., Guangquan Z.: Recommender system application developments: A survey. Decision Support Systems, ISSN: 0167-9236, Vol: 74, Page: 12-32 (2015).
5. Deshpande M., Karypis G.: Item-based top-N recommendation algorithms. ACM Transactions on Information Systems (TOIS) 22 143–177 (2004).
6. Sarwar B., Karypis G., Konstan J.,Riedl J.: Item-based collaborative filtering recommendation algorithms. In Proceedings of the 10th International Conference on World Wide Web, ACM, pp. 285–295 (2001).
7. Shardanand U., Maes P.: Social information filtering: algorithms for automating “word of mouth”. In Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 210–217 (1995).
8. Slaney M., White W.: Similarity Based on Rating Data. In Proceedings of the 8th International Conference on Music Information Retrieval. Vienna, Austria (2007).
9. Casey M. A., Veltkamp R., Goto M., Leman M., Rhodes C., Slaney M.: Content-Based Music Information Retrieval: Current Directions and Future Challenges. In Proceedings of the IEEE, vol. 96, no. 4, pp. 668-696 (2008).
10. Burke R.: Hybrid recommender systems: survey and experiments. User Modeling and User-Adapted Interaction 12 331–370 (2002).
11. Sánchez-Moreno, D., Gil González, A. B., Muñoz Vicente, M. D., López Batista, V. F., & Moreno García, M. N.: A collaborative filtering method for music recommendation using playing coefficients for artists and users. Expert Systems with Applications, 66, 234–244 (2016).
12. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T.S.: Neural Collaborative Filtering. Proceedings of the 26th International Conference on World Wide Web - WWW ’17 (2017).
13. Dias, R., & Fonseca, M. J.: Improving Music Recommendation in Session-Based Collaborative Filtering by Using Temporal Context. IEEE 25th International Conference on Tools with Artificial Intelligence (2013)
14. Oord, A., Dieleman, S. Schrauwen, B.: Deep content-based music recommendation. Advances in Neural Information Processing Systems. ISSN 1049-5258 (2013).
15. Soleymani, M., Aljanaki, A., Wiering, F., Veltkamp, R. C.: Content-based music recommendation using underlying music preference structure. IEEE International Conference on Multimedia and Expo (ICME) (2015).

16. Vall, A., Dorfer, M., Schedl, M., Widmer, G.: A hybrid approach to music playlist continuation based on playlist-song membership. Proceedings of the 33rd Annual ACM Symposium on Applied Computing (2018).
17. Chiliguano, P., Fazekas, G.: Hybrid music recommender using content-based and social information. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (2016).
18. Schedl, M., Hauger, D.: Tailoring Music Recommendations to Users by Considering Diversity, Mainstreaminess, and Novelty. Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR (2015).
19. Song Y, Dixon S., Pearce M.: A Survey of Music Recommendation Systems and Future Perspectives. In Proceedings of International Symposium on Computer Music Modeling and Retrieval (CMMR), UK (2012).