RECSYS CHALLENGE 2018: AUTOMATIC PLAYLIST CONTINUATION

Markus Schedl[1] **Hamed Zamani**[2] **Ching-Wei Chen**[3]
Yashar Deldjoo[4] **Mehdi Elahi**[5]
[1] Department of Computational Perception, Johannes Kepler University Linz, Austria
[2] Center for Intelligent Information Retrieval, University of Massachusetts Amherst, USA
[3] Spotify USA Inc., New York, USA
[4] Department of Computer Science, Politecnico di Milano, Italy
[5] Department of Computer Science, Free University of Bozen-Bolzano, Italy
markus.schedl@jku.at, zamani@cs.umass.edu, cw@spotify.com

ABSTRACT

In recent years, considerable attention has been given to studies on the role of playlists in music consumption. A study carried out in 2016, by the Music Business Association [6], showed that playlists accounted for 31% of music listening time among listeners in the USA. Another study, conducted by MIDiA [1], revealed that as many as 55% of streaming music service subscribers create playlists. Accordingly, music streaming services such as Spotify currently host over 2 billion playlists [9]. This evidence may indicate the growing importance of playlists as a mode for music consumption, and indeed the crucial necessity of developing algorithms for *automatic playlist continuation*, which is the focus of the ACM Recommender Systems Challenge 2018 [7]. In this paper, we — the organization team of this challenge — briefly discuss the particular task we defined for the participating teams. We also provide some information on the overall Challenge process.

ACM Recommender Systems Challenge

The Recommender Systems Challenge is a yearly competition focusing on creating the best-performing recommendation approach for a specific task and a specific scenario. From 2010 to 2017, the competition has drawn diverse participants from academia and industry [10, 11, 15]. Today, the Recommender Systems Challenge has become a key part of the ACM Conference on Recommender Systems series, the leading conference in recommender systems research. The Recommender Systems Challenge has followed a similar structure since its inception: (1) a real-world problem is presented with a corresponding dataset, (2) researchers and developers form teams and sign up for participation, (3) participating teams submit their solutions prior to a deadline, (4) top participating teams submit papers that outline their approaches, (5) during a workshop at the ACM RecSys conference, accepted papers are presented and the winning teams are announced.

In MIREX, a few attempts to launch music recommendation [3] or personalized radio stationing [4] tasks were performed in the past, but they were never realized to the best of the authors' knowledge. In comparison, by organizing the task in 2018 at the ACM Recommender Systems conference, we will reach a broader audience and with the full support of a major music streaming company, Spotify, we are sure to attract a remarkable number of participants. The MIREX task probably closest to ours is the Audio Music Similarity and Retrieval task [5], which has been run in 2016 for the last time. However, the organizers explicitly state that the task is "not playlist generation or music recommendation", rather pure similarity aspects. Another related challenge was the Million Song Dataset Challenge [2], which featured a traditional recommendation task: based on a part of the user's listening history, another, withheld part needed to be predicted. In contrast, our task explicitly considers hand-curated playlists, not listening histories.

Automatic Playlist Continuation

The task for the Recommender Systems Challenge in 2018 will be automatic playlist continuation (APC). This task consists of adding one or more tracks to a music playlist (of arbitrary length) in a way that fits the target characteristics of the original playlist [12, 16]. APC is a useful feature for music streaming services not only because it can extend listening session length, but also because it can increase engagement of users on their platform by making it easier for users to create playlists that they can enjoy and share.

As part of this challenge, Spotify will be releasing a public dataset of roughly 1 million user-created playlists. The dataset will include the title of each playlist, as well as the list of tracks, and some associated metadata, for each playlist. A separate evaluation set will consist of a set of playlists from which a number of tracks have been withheld. The task will then be to predict the missing tracks in those playlists, and participating teams will be required to submit their predictions for those missing tracks. An appropriate accuracy metric will be defined, which will then be used to evaluate the performance of each of the submissions.

The dataset and associated evaluation metrics are scheduled to be released by the end of 2017. Up-to-date information can be found on the Challenge website. [1]

[1] http://2018.recsyschallenge.com

Current Challenges

Limitations of current research in APC, which will be important to address by participants of the challenge, are manifold and briefly summarized in the following. While some work on APC highlights the special characteristics of playlists, i.e., their *sequential order*, it is not well understood to which extent and in which cases taking into account the order of tracks in playlists helps create better models for recommendation. For instance, in [18] Vall et al. recently demonstrated on two datasets of hand-curated playlists that the song order seems to be negligible for accurate playlist continuation when a lot of popular songs are present. On the other hand, the authors argue that order does matter when creating playlists with tracks from the long tail. Another study by McFee and Lanckriet [14] also suggests that transition effects play an important role in modeling playlist continuity. In another recent user study [17] conducted by Tintarev et al., the authors find that many listeners do not care about the order of tracks in recommended playlists, sometimes they do not even notice that there is a particular order. However, this study was restricted to 20 participants who used the Discover Weekly service of Spotify [8].

Another challenge for the APC task is the high subjectivity of what makes a great playlist [13], including factors such as intent of the creator or listener. Important criteria when creating or judging a playlist include track similarity/coherence, variety/diversity, but also the user's personal preferences and familiarity with the tracks, as well as the intention of the playlist creator. Unfortunately, current automatic approaches to APC are agnostic of the underlying psychological and sociological factors that influence the decision of which songs users choose to include in a playlist. Since knowing about such factors is vital to understand the intent of the playlist creator, we believe that algorithmic methods for APC need to learn such aspects from manually created playlists and integrate respective intent models. However, we are aware that in today's era where billions of playlists are shared by users of online streaming services [9], a large-scale analysis of psychological and sociological background factors is impossible. Nevertheless, in the absence of explicit information about user intent, a possible starting point to create intent models might be the metadata associated with user-generated playlists, such as the title and the list of tracks.

Conclusion

This paper briefly investigated the growing importance, as well as remaining challenges, of the automatic music playlist continuation task. The authors, who are the organizers of the Recommender Systems Challenge 2018, decided to focus on this particular, real-world problem of the music industry and they are currently preparing the details of the actual task.

1. REFERENCES

[1] Announcing MIDIAS State of the Streaming Nation 2 Report. https://www.midiaresearch.com/blog/announcing-midias-state-of-the-streaming-nation-2-report. Accessed: 2017-10-16.

[2] Million Song Dataset Challenge. https://www.kaggle.com/c/msdchallenge#description. Accessed: 2017-10-20.

[3] MIREX 2009: Music Recommendation. https://www.music-ir.org/mirex/wiki/2009:Music_Recommendation. Accessed: 2017-10-20.

[4] MIREX 2009: Music Recommendation Personalized Radio. https://www.music-ir.org/mirex/wiki/2009:Music_Recommendation_Personalized_Radio. Accessed: 2017-10-20.

[5] MIREX 2016: Audio Music Similarity and Retrieval. https://www.music-ir.org/mirex/wiki/2016:Audio_Music_Similarity_and_Retrieval. Accessed: 2017-10-20.

[6] Music Biz Consumer Insights. https://musicbiz.org/resources/tools/music-biz-consumer-insights/consumer-insights-portal. Accessed: 2017-10-17.

[7] RecSys Challenge 2018, the ACM Conference Series on Recommender Systems. https://recsys.acm.org/recsys18/challenge. Accessed: 2017-10-17.

[8] Spotify homepage. https://www.spotify.com/discoverweekly. Accessed: 2017-10-17.

[9] Spotify Press homepage. https://press.spotify.com/us/about. Accessed: 2017-10-17.

[10] Fabian Abel, Yashar Deldjoo, Mehdi Elahi, and Daniel Kohlsdorf. Recsys challenge 2017: Offline and online evaluation. In *Proceedings of the Eleventh ACM Conference on Recommender Systems*, RecSys '17, pages 372–373. ACM, 2017.

[11] David Ben-Shimon, Alexander Tsikinovsky, Michael Friedmann, Bracha Shapira, Lior Rokach, and Johannes Hoerle. Recsys challenge 2015 and the yoochoose dataset. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 357–358. ACM, 2015.

[12] Geoffray Bonnin and Dietmar Jannach. Automated generation of music playlists: Survey and experiments. *ACM Computing Surveys (CSUR)*, 47(2):26, 2015.

[13] Jin Ha Lee. How Similar is Too Similar? Exploring Users' Perceptions of Similarity in Playlist Evaluation. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, Miami, FL, USA, October 2011.

[14] Brian McFee and Gert Lanckriet. Hypergraph Models of Playlist Dialects. In *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR)*, Porto, Portugal, 2012.

[15] Alan Said. A Short History of the RecSys Challenge. *AI Magazine*, 37(4), 2016.

[16] Markus Schedl, Hamed Zamani, Ching-Wei Chen, Yashar Deldjoo, and Mehdi Elahi. Current Challenges and Visions in Music Recommender Systems Research. volume abs/1710.03208, 2017.

[17] Nava Tintarev, Christoph Lofi, and Cynthia C.S. Liem. Sequences of diverse song recommendations: An exploratory study in a commercial system. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, UMAP '17, pages 391–392, New York, NY, USA, 2017. ACM.

[18] Andreu Vall, Massimo Quadrana, Markus Schedl, Gerhard Widmer, and Paolo Cremonesi. The Importance of Song Context in Music Playlists. In *Proceedings of the Poster Track of the 11th ACM Conference on Recommender Systems (RecSys)*, RecSys '17, Como, Italy, 2017.