A Model for Evaluating Popularity and Semantic Information Variations in Radio Listening Sessions

Lorenzo Porcaro
Music Technology Group, Universitat Pompeu Fabra
Barcelona, Spain
lorenzo.porcaro@upf.edu

Emilia Gómez
Music Technology Group, Universitat Pompeu Fabra
Barcelona, Spain
Joint Research Centre, European Commission
Seville, Spain
emilia.gomez@upf.edu

ABSTRACT

Listening to music radios is an activity that since the 20th century is part of the cultural habits for people all over the world. While in the case of analog radios DJs are in charge of selecting the music to be broadcasted, nowadays recommender systems analyzing users' behaviours can automatically generate radios tailored to users' musical taste. Nonetheless, in both cases listening sessions do not depend on the listener choices, but on a set of external recommendations received. In this preliminary study, we propose a model for estimating features' variation during listening sessions, comparing different scenarios, namely analog radios, personalized and not-personalized streaming radios. In particular, we focus on the analysis of track popularity and semantic information, features well-established in the Music Information Retrieval literature. The presented model aims to quantify the possible impacts of the sessions' variation on the user listening experience.

CCS CONCEPTS

• **General and reference** → **Evaluation**; • **Information systems** → **Recommender systems**.

KEYWORDS

recommender systems, evaluation, popularity, semantic information, variations

ACM Reference Format:
Lorenzo Porcaro and Emilia Gómez. 2019. A Model for Evaluating Popularity and Semantic Information Variations in Radio Listening Sessions. In *Proceedings of 1st Workshop on the Impact of Recommender Systems at the 13th ACM Conference on Recommender Systems (RecSys 2019) (ImpactRS 2019).* ACM, New York, NY, USA, 6 pages.

1 INTRODUCTION

Since the first part of the 20th century, the radio has been one of the main media thanks to which listeners can enjoy music. Its influence on music consumption has been enormous, due to its capacity to determine content diffusion and diversity [11]. Nowadays, streaming services have become the digital platforms where most music enthusiasts listen to music [23]. Within these services, recommender systems play a key role in helping users to explore and exploit the large music catalogue available, accordingly to their interests. For that goal, the inclusion of personalization techniques is considered fundamental for improving the quality of the recommendation provided [21]. Among the products available on streaming services, personalized radio stations can create a coherent sequence of songs to be played, starting from a specific track, artist or music genre.

However, the use of personalized services in online spaces is at the center of the scientific and public debate, due to its proven tendency in fostering polarization in society [5, 7] and in creating the so-called "echo chambers", virtual spaces where users are intensively exposed to content tailored to their profiles [10]. In particular, negative impacts of the use of recommender systems have already been proven, such as of homogenization of users' behaviour, i.e. similar users interacting with the same set of recommendations [8], or the objectification of personal tastes [14]. Recent discussions within the Music Information Retrieval (MIR) community are concerned about the impact of the developed technologies [2].

In this work, we aim at characterizing the listening session variations. We focus on two different aspects: popularity, feature historically related to the "long tail" problem [4], and semantic information, largely investigated in the MIR literature [15]. It is important to notice that we are not interested in *understanding how the system provides recommendations*, or in *comparing the performance of different systems.* By defining two indicators of variations of song sequences, our goal is to understand what can be the impact of listening to music radios on users' experiences in the long-term. Furthermore, we are interested in comparing personalized and non-personalized systems, analyzing the differences between these scenarios.

The paper is structured as follows. Section 2 provides an overview of previous work related to the research lines of this study. We then propose a methodology in Section 3, which includes a description of the model. Section 4 describes the case study examined, together with the results obtained. Finally, conclusions and future works are discussed in Section 5.

2 RELATED WORK

The limits of considering accuracy metrics as the yardstick for evaluating recommender systems have been discussed in the last two decades, since the problem of formulating user-centric measures has been approached and new metrics *beyond-accuracy* have started to be proposed [17]. Through the years, the strategy of "diversification" in recommender systems gained in visibility [24], especially for facing the problem of recommending only items highly similar to each other, known as "portfolio effect" [6]. Consequently, new methods began to be proposed, where next to accuracy metrics,

ImpactRS 2019, September 2019, Copenhagen, Denmark
© 2019

diversity, serendipity, novelty and coverage metrics started to be considered [13].

In the field of MIR, the concept of diversity has been often associated with aspects of musical tastes and listening habits [12]. Starting from the users' listening histories, attempts have been made to understand how different degrees of diversity in musical tastes can affect recommendation models, but also the reverse process, hence how different levels of diversification in recommender systems can embrace different populations of users [16]. The relevance of musical taste as a proxy for determining other social factors has been also investigated, where measurements of diversity have been used to study different facets related to personal interests [19].

Alongside, from different points of view, temporal dimensions of the listening experience have been object of study. For instance, a dynamic model of the listening experience has been proven to be effective for measuring the evolution of musical taste over time [18]. Furthermore, it has been shown how including tag information with temporal dynamics of user interaction is useful for generating personalized music recommendations [25]. However, it is also important to note that, from a user perspective, the temporal dimension may not always be perceived as relevant, as observed in [22].

In conclusion, the study of the relationships between sequentially ordered objects is part of the *sequence-aware* recommender system framework [20], which is gaining attention thanks to recent advancements in deep learning and reinforcement learning techniques. The applications of this framework are particularly relevant in the music field, considering the sequential nature of music consumption (e.g. playlist) [3].

3 METHODOLOGY

The main goal of our model is to characterize indicators of variations for a recommendation list. Specifically, we focus on music recommendations, creating two metrics, one for the track popularity, and another one for the semantic content represented by a set of tags describing the track music genre.

3.1 Model

We define a recommendation list R as an ordered set of n tracks $R = \{r_1, r_2, ..., r_n\}$, where each track r_i has two different attributes: 1) a set of semantic tags s_i, where $s_i = \{tag_1, tag_2, ..., tag_n\}$; 2) a popularity index p_i, computed as the sum of the track and artist popularity values $p_i = pop_{r_i} + pop_{artist(r_i)}$. To compare lists generated in different contexts, we study how track attributes vary. In detail, we adopt two temporal strategies for a sequential analysis of the tracks: we first consider single tracks and we then form groups according to their temporal location.

3.1.1 Single Track Level. A music recommendation list can be considered as a set of songs potentially listened in sequence. Our first goal is to model the deviation of each track with respect to previous ones in terms of semantic information and popularity. Therefore, we define two recursive functions, which at each step return an indicator i containing information about these variations.

Let's consider S_n as the union of all the tags observed until the track r_n and i_n as the number of tags of the n-th track which have been not yet observed in previous tracks. Analytically, this process can be described with the following formula:

$$\begin{cases} S_0 = s_0 \\ i_0^{(s)} = 0 \end{cases} \begin{cases} S_n = \bigcup_{i=1}^{n} s_i \\ i_n^{(s)} = \frac{|s_n \setminus S_{n-1}|}{|s_n|} \end{cases} \quad (1)$$

where $|S|$ represents the cardinality of a set S, and the $\setminus$ operation between sets represents the relative complement i.e. objects that belong to the first set and not to the second. Considering that multiple tags can be associated with the same track, the relative complement is divided by the cardinality of s_n, giving the same weight to every tag. As the space of tags is finite, when $n \to \infty \Rightarrow i_n^{(s)} \to 0$. This relationship can be imagined as a hypothetical session, where a user listens to a radio without ever stopping, and where eventually the variations in terms of semantic tags will become null at some point.

In the case of popularity, at each track is associated a value p such that $p \in [0, 1]$. At step n, we estimate the variation as a weighted sum of powers of two with the previous track popularities:

$$\begin{cases} P_0 = p_0 \\ i_0^{(p)} = 0 \end{cases} \begin{cases} P_n = \frac{p_n + P_{n-1}}{2} \\ i_n^{(p)} = \frac{|p_n - P_n|}{2} \end{cases} \quad (2)$$

We observe that by the nature of this iteration, the n-th variation will be influenced by all the previous tracks popularity in the list, but giving more relevance to the closest tracks. In this case, the convergence to zero is not guaranteed when n tends to infinite. The only case when variations decrease until reaching zero is when the popularity attribute tends to a constant value, hence if $p_n \to c$ for $n \to \infty \Rightarrow i_n^{(p)} \to 0$.

Finally, for deriving a central tendency of the variations for each list R we compute the median of the obtained values:

$$I_R^{(s)} = median(i_{j=1:n}^{(s)}), \quad I_R^{(p)} = median(i_{j=1:n}^{(p)}), \quad (3)$$

However, the proposed modelization represents a situation which is far from being real. Indeed, it is hard to imagine a person who turns on the radio and listens to the same station without stopping or changing a song. If we consider temporal relationships between tracks played in very distant time spans of a listening session, we may include information which is not relevant. In order to avoid that, we present in the next section an approach to estimate how a recommendation list can embed different degrees of variations, considering group of tracks instead of single ones.

3.1.2 Group Track Level. Starting from a recommendation list R, we divide it into uniform groups of M tracks. As a first step, we have to define how the track attributes, semantic tags and popularity, are aggregated to create a group-level representation.

With respect to semantic information, we combine all the tags related to the tracks in the group. If $g_i = \{r_{i*M}, ..., r_{(i+1)*M-1}\}$, we define s_{g_i} for the group as:

$$s_{g_i} = \bigcup_{j=i*M}^{(i+1)*M-1} s_j \quad , \forall i = 0, 1, ... \quad (4)$$

Table 1: Seed tracks and related features.

track ID	Title	Artist	Year	Popularity	Tags
A	Strange Way	Firefall	1979	0.42	album rock, classic rock, country rock
B	My Sharona	The Knack	1979	0.68	album rock, power pop
C	Sugar Walls	Sheena Easton	1985	0.45	mellow gold, minneapolis
D	Like A Virgin	Madonna	1985	0.76	dance pop
E	We Got A Love Thang	CeCe Peniston	1992	0.41	diva house, hip house, vocal house
F	Under The Bridge	Red Hot Chili Peppers	1992	0.83	alternative rock, funk metal, funk rock
G	Sick of Being Lonely	Field Mob	2003	0.44	atl hip hop, dirty south rap, gangster rap
H	In Da Club	50 Cent	2003	0.77	east coast hip hop, gangster rap, hip hop
I	Stay The Night	Zedd feat. Hayley Williams	2014	0.73	complextro, dance pop, edm
J	Happy	Pharrell Williams	2014	0.80	pop, pop rap

In terms of popularity, we define the group popularity as the mean track popularity:

$$p_{g_i} = \frac{1}{M} \sum_{j=i*M}^{(i+1)*M-1} p_j \quad , \forall i = 0, 1, \ldots \tag{5}$$

After computing the attributes at a group level, we define a recursive process for estimating the variation between groups:

$$\begin{cases} S_{g_0} = s_{g_0} \\ i_{g_0}^{(s)} = 0 \end{cases} \begin{cases} S_{g_n} = s_{g_n} \\ i_{g_n}^{(s)} = \frac{|S_{g_n} \setminus S_{g_{n-1}}|}{|S_{g_n} \cup S_{g_{n-1}}|} \end{cases} \tag{6}$$

$$\begin{cases} P_{g_0} = p_{g_0} \\ i_{g_0}^{(p)} = 0 \end{cases} \begin{cases} P_{g_n} = p_{g_n} \\ i_{g_n}^{(p)} = \frac{|P_{g_n} - P_{g_{n-1}}|}{2} \end{cases} \tag{7}$$

Using this formulation, we remove long-term dependencies as we only evaluate the relationship between consecutive groups. The convergence of the variation to zero is only guaranteed if the groups' attributes tend to similar values, which happens when the average track popularity remains constant between different groups, or the group tags do not vary between a group and the following one. After obtaining the variations between consecutive groups, we extract an overall attribute indicator by computing a median value:

$$I_R^{(s)} = median(i_g^{(s)}), \quad I_R^{(p)} = median(i_g^{(p)}), \quad \forall g \in R \tag{8}$$

The proposed group model represents a scenario nearer to the scheduling philosophies behind the music broadcasted in analog radios. Even if there are no indications on optimal values of the group size, we empirically choose to separate recommended lists in group of 5 consecutive tracks, to assure a temporal proximity within the group. As observed in [1], the optimal choice of the programming can be influenced by the targeted audience (*mass* or *niche*), and the music radio style, factors not considered in this work.

4 CASE STUDY

We compute the indicators defined in Section 3.1 on several recommendation lists, comparing how starting from the same tracks, personalized and non-personalized systems create listening sessions with different variation characteristics.

In our study, we select as seed tracks, for each of the five last decades, one of the top songs and one of the last songs from the Billboard Top 100 charts[1]. Table 1 shows the selected tracks. Popularity and tag features have been computed using the Spotify API[2].

4.1 Dataset

We consider for our experiment different kinds of recommendation lists: personalized and non-personalized lists from streaming radio, and analog radio lists. For each list, we consider the first 25 recommended tracks.

4.1.1 Streaming Radio Personalized Lists. 10 users of Spotify and Youtube Music generated their personalized lists. Of these, 6 participants are male, and 4 female, with mean age of 29 years (standard deviation of 5 years). Each participant was asked to search for each seed track and automatically create radio lists using the functionalities of each platform. In total, we collected lists from 10 users for 10 tracks, for 2 platforms, hence 200 personalized lists.

4.1.2 Streaming Radio Non-Personalized Lists. Using the same set of seed tracks, we created non-personalized lists with two methods. First, we considered radio lists in YouTube Music without logging in, using a page in incognito mode. Second, we used Spotalike[3], a service for generating playlists starting from a track.

4.1.3 Analog Radio Lists. We randomly selected 10 lists from a dataset of radio playlists from hundreds of radio stations in the United States[4], presented in [9]. Due to the different seed tracks of the analog radio list considered, it was not possible to compare those to the personalized and non-personalized correspondent lists. Consequently, we use them to create a baseline for comparison of streaming radio lists. For doing so, for each track attribute and level of analysis separately, we averaged the indicators obtained for every list, achieving a single representative value for this type of lists.

[1] http://billboardtop100of.com
[2] https://developer.spotify.com/
[3] https://spotalike.com
[4] https://www.cs.cornell.edu/~shuochen/lme/data_page.html

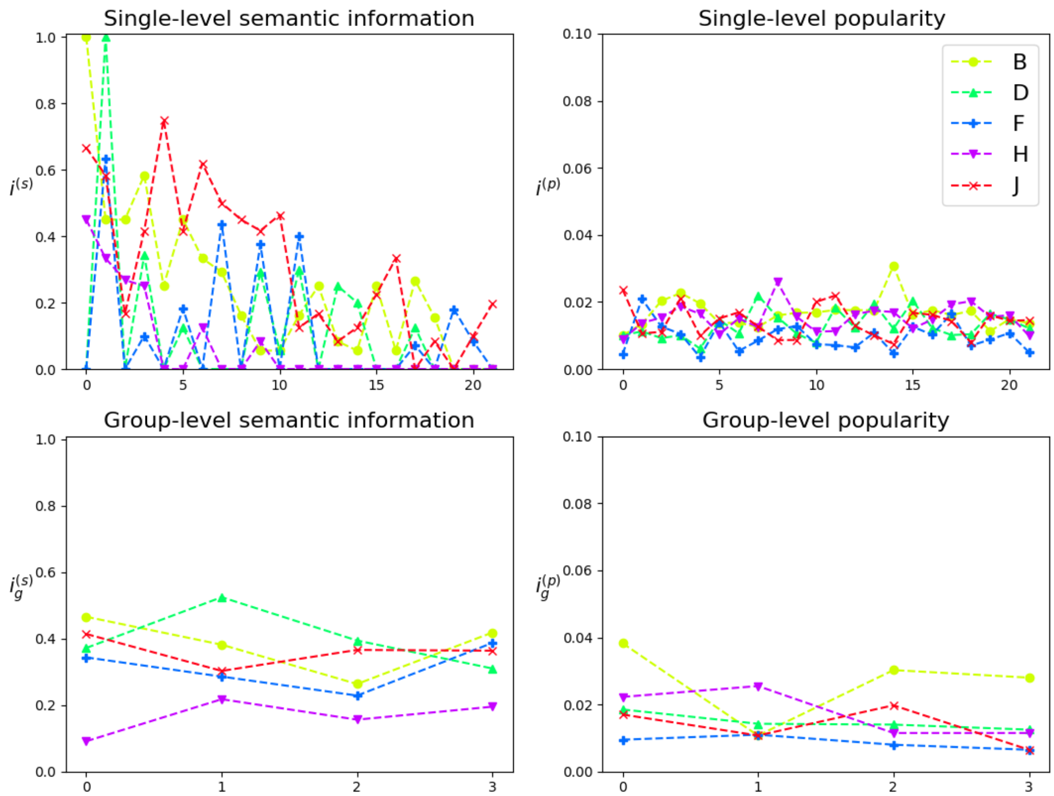

Figure 1: Variations computed using the personalized recommendation lists created starting from the seed tracks B, D, F, H, J. Results shown are obtained with the single level model described in Section 3.1.1 (semantic information (*top left*), popularity (*top right*)) and group level model described in Section 3.1.2 (semantic information (*bottom left*), popularity (*bottom right*))

4.2 Results

According to the model described in Section 3.1, we compute indicators of variations using as input the recommendation lists presented in Section 4.1, and the results are presented in Figure 1 and 2. For both features, values are normalized between 0 and 1. Given the small sample size and the nature of the experiment, these results can only be considered as descriptive.

Figure 1 shows the variations i computed using the formulas (1), (2), (6), and (7). We only display the variations of personalized recommended lists generated starting from 5 seed tracks (B, D, F, H, J). In the case of semantic information, we notice how single level variations describe a decreasing trajectory, converging to zero, in line with the hypothesis discussed in Section 3.1.1. For the popularity feature, the variations tend to remain constant. In addition, we do not observe any difference between lists created from different seed tracks. Analyzing the variations at a group level, we see how removing the long-term dependencies the variations behave independently, not following a specific tendency but depending more consistently from the conformation of the groups. In this case, we observe similar behaviours for both features.

Figure 2 shows the obtained indicators described in formulas (3) and (8). In terms of semantic information, we observe that the indicators produced by the single level model describe sessions more homogeneous than in the case of the group level. Indeed, the values of the indicators are lower for individual tracks than the correspondent values obtained at a group level, showing how the long-term dependencies impact the variations, leading to a median value equal to zero in several cases. In particular, for the non-personalized lists, the median of the variations is zero in half of the cases. It can be considered as a consequence of extensive use of semantic relationships while building the recommendations, being not tailored to specific users. In terms of popularity, we observe that, when comparing tracks from the same decade (A-B, C-D, E-F, etc.), the lists created starting from the tracks with lower popularity (A, C, E, etc.) present more variations than the ones with greater popularity, both in personalized and non-personalized settings. These results partially reflect the impacts of the well-known popularity bias in recommender systems, describing a situation where starting from a popular track, other popular tracks tend to be recommended, and variations are limited.

From the group level perspective, computing the Pearson correlation coefficient r we notice that there is a higher correlation between non-personalized and personalized lists than in the case of the single level model. Indeed, for single level indicators we have r=0.15 for the semantic information feature and r=0.44 for the popularity, while for the group level we have r=0.55 for semantic

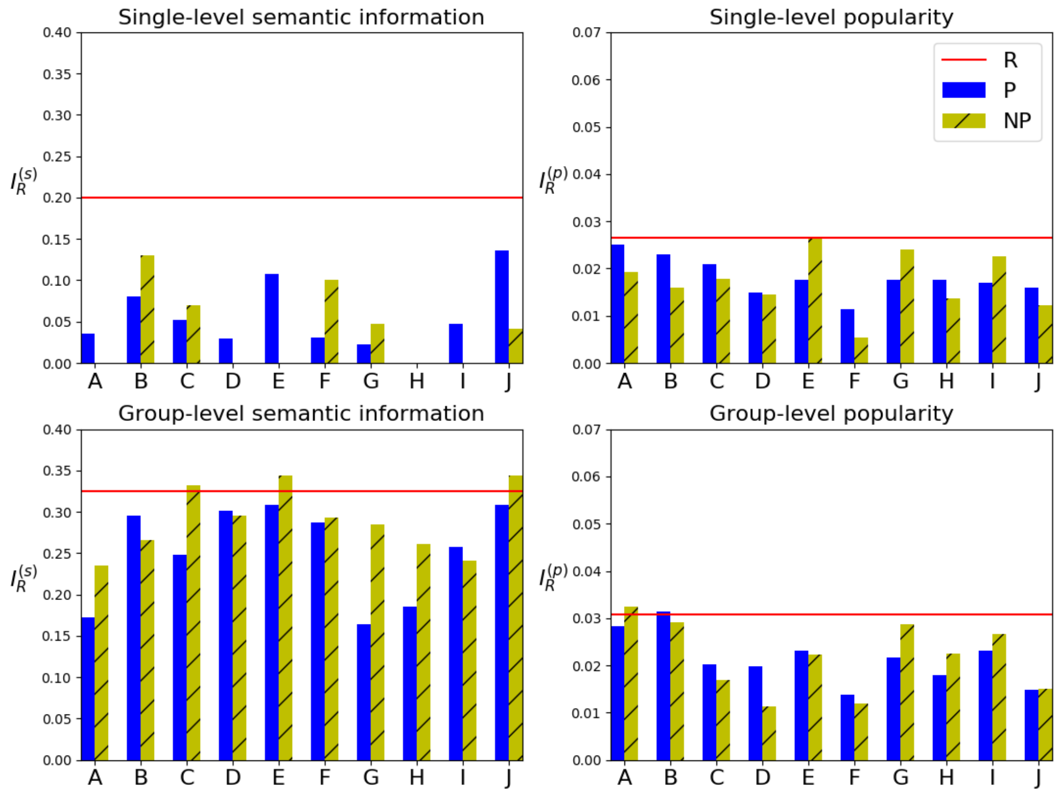

Figure 2: Semantic tag indicator of variations for the single track level (*top left*, Pearson correlation coefficient *r*=0.15) and group track level (*bottom left*, *r*=0.55), and popularity indicator of variations for the single track level (*top right*, *r*=0.44) and group track level (*bottom right*, *r*=0.79). P=Personalized lists; NP=Non-personalized lists; R=Analog radios baseline.

information while r=0.79 for popularity. Even if these results are sensitive to the group size, we imagine that joining different tracks' features in a single group attribute, as defined in Section 3.1.2, can lead to a mitigation of variations, hence to obtain results more correlated between personalized and non-personalized systems. Furthermore, the removal of long-term dependencies in the group level model impacts the degree of variations observable, leading to higher median values for both features. In this case, no particular effects have been noticed according to the seed tracks considered.

In conclusion, apart from the case of group level semantic information, we observe that the overall indicators of analog radios result to have higher values on average than the personalized and non-personalized recommendation lists. However, the baseline build is weak, given the few numbers of analog radio considered, hence to make a proper comparison it is needed to increase the amount of data analyzed, as discussed in the following section.

5 CONCLUSIONS AND FUTURE WORK

We have presented a model for analysing the variations of two items' features in a recommendation list: popularity and semantic information. In our experiment, we have compared several music recommendation lists generated in different contexts: personalized and non-personalized streaming radios, and analog radios.

Several limitations of the obtained results are related to the scarcity of data. Indeed, the lists taken into account cannot be considered as representative, and a larger amount of data is needed for validating empirically our model. In this preliminary stage, we do not have evidence about the advantages or disadvantages of using personalized or non-personalized systems for generating diverse listening sessions, but comparing against radio playlists generated by humans, we believe that differences with algorithmic-driven listening experiences can emerge.

Moreover, being the detail of the recommender systems' architecture used in this study unknown, it is difficult to make comparisons between personalized and non-personalized lists, and also to link observed effects to a specific cause. For instance, the popularity bias may potentially be reinforced more in non-personalized popularity-based recommender systems than in personalized settings.

Finally, for properly evaluating the impact of personalization in music recommender systems, we need to include in the model additional information about the users' tastes and listening behaviours. In the presented work, personalization has been analyzed as a *de facto* phenomena, but further work needs to be done for evaluating its relationship to the recommendation provided. As an alternative line of research, users' perceptions of listening sessions variations can be considered, comparing them with the indicators designed in this work.

6 ACKNOWLEDGMENTS

This work is partially supported by the European Commission under the TROMPA project (H2020 770376). The authors would like to thank Jacopo Temperini, Elena Congeduti, and the anonymous reviewers for their comments and feedback.

REFERENCES

[1] Jarl A. Ahlkvist. 2001. Programming philosophies and the rationalization of music radio. *Media, Culture and Society* 23, 3 (2001), 339–358.

[2] Mark Coeckelbergh Andre Holzapfel, Bob L. Sturm. 2018. Ethical Dimensions of Music Information Retrieval Technology. *Transactions of the International Society for Music Information Retrieval* 1, 1 (2018), 44–55.

[3] Christine Bauer. 2019. The potential of the confluence of theoretical and algorithmic modeling in music recommendation. In *Proceedings of the ACM CHI 2019 Workshop on Computational Modeling in Human-Computer Interaction.*

[4] Christine Bauer and Markus Schedl. 2019. Global and country-specific mainstreaminess measures: Definitions, analysis, and usage for improving personalized music recommendation systems. *PLOS ONE* 14, 6 (2019), 1–36.

[5] Engin Bozdag. 2013. Bias in algorithmic filtering and personalization. *Ethics and Information Technology* 15, 3 (2013), 209–227.

[6] Robin Burke. 2002. Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction* 12, 4 (2002), 331–370.

[7] L. Elisa Celis, Sayash Kapoor, Farnood Salehi, and Nisheeth Vishnoi. 2019. Controlling Polarization in Personalization: An Algorithmic Framework. In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19).* ACM, 160–169.

[8] Allison J. B. Chaney, Brandon M. Stewart, and Barbara E. Engelhardt. 2018. How Algorithmic Confounding in Recommendation Systems Increases Homogeneity and Decreases Utility. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys '18).* ACM, 224–232.

[9] Shuo Chen, Josh L. Moore, Douglas Turnbull, and Thorsten Joachims. 2012. Playlist Prediction via Metric Embedding. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '12).* ACM, 714–722.

[10] Pranav Dandekar, Ashish Goel, and David T Lee. 2013. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences* 110, 15 (2013), 5791–5796.

[11] John Dimmick and Daniel G McDonald. 2001. Network Radio Oligopoly, 1926-1956: Rivalrous Imitation and Program Diversity. *Journal of Media Economics* 14, 4 (2001), 197–212.

[12] Katayoun Farrahi, Markus Schedl, Andreu Vall, David Hauger, and Marko Tkalčič. 2014. Impact of Listening Behaviour on Music Recommendation. In *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR 2014).* 1–6.

[13] Marius Kaminskas and Derek Bridge. 2016. Diversity, Serendipity, Novelty, and Coverage: A Survey and Empirical Analysis of Beyond-Accuracy Objectives in Recommender Systems. *ACM Transactions on Interactive Intelligent Systems* 7, 1 (2016), 1–42. https://doi.org/10.1145/2926720

[14] Nedim Karakayali, Burc Kostem, and Idil Galip. 2018. Recommendation Systems as Technologies of the Self: Algorithmic Control and the Formation of Music Taste. *Theory, Culture & Society* 35, 2 (2018), 3–24.

[15] Paul Lamere. 2008. Social tagging and music information retrieval. *Journal of New Music Research* 37, 2 (2008), 101–114.

[16] Feng Lu and Nava Tintarev. 2018. A diversity adjusting strategy with personality for music recommendation. In *Proceedings of the 5th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems co-located with ACM Conference on Recommender Systems (RecSys 2018).* 7–14.

[17] Sean M. McNee, John Riedl, and Joseph A. Konstan. 2006. Being Accurate is Not Enough: How Accuracy Metrics Have Hurt Recommender Systems. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems (CHI EA '06).* ACM, 1097–1101.

[18] Joshua F Moore, Shuo Chen, Joachims Thorsten, and Douglas Turnbull. 2013. Taste Over Time: The Temporal Dynamics of User Preferences. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2013).* 401–406.

[19] Minsu Park, Ingmar Weber, Mor Naaman, and Sarah Vieweg. 2016. Understanding Musical Diversity via Online Social Media. In *Proceedings of the 10th International AAAI Conference on Web and Social Media (ICWSM'16).*

[20] Massimo Quadrana, Paolo Cremonesi, and Dietmar Jannach. 2018. Sequence-Aware Recommender Systems. *ACM Comput. Surv.* 51, 4, Article 66 (July 2018), 36 pages.

[21] Markus Schedl, Hamed Zamani, Ching-Wei Chen, Yashar Deldjoo, and Mehdi Elahi. 2018. Current Challenges and Visions in Music Recommender Systems Research. *International Journal of Multimedia Information Retrieval* 7, 2 (2018), 95–116.

[22] Nava Tintarev, Christoph Lofi, and Cynthia C.S. Liem. 2017. 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.* 391–392.

[23] Patrick Vonderau. 2019. The Spotify Effect: Digital Distribution and Financial Growth. *Television and New Media* 20, 1 (2019), 3–19.

[24] Cong Yu, Laks Lakshmanan, and Sihem Amer-Yahia. 2009. It Takes Variety to Make a World: Diversification in Recommender Systems. In *Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology (EDBT '09).* ACM, 368–378.

[25] Ervine Zheng, Gustavo Yukio Kondo, Stephen Zilora, and Qi Yu. 2018. Tag-aware dynamic music recommendation. *Expert Systems with Applications* 106 (2018), 244–251.