Using Psychological Principles of Memory Storage and Preference to Improve Music Recommendation Systems

ANTHONY CHMIEL AND **EMERY SCHUBERT**

ABSTRACT

This paper proposes a novel approach to automated music recommendation systems. Current systems use a number of methods, although these are generally based on similarity of content, contextual information or user ratings. These approaches therefore do not take into account relevant, well-established models from the field of music psychology. Given recent evidence of this field's excellent capacity to predict music preference, we propose a function based on both the Ebbinghaus *forgetting curve* of memory retention and Berlyne's *inverted-U* model to inform recommendation systems through "collative variables" such as exposure/familiarity. According to the model, an intermediate level of these variables should generate relatively high preference and therefore presents significant untapped data for music recommendation systems.

Automated music recommendation systems have become increasingly sophisticated in recent years. These systems employ a number of methods for selecting dynamic, personalized music playlists that generally rely on principles of similarity [1,2]. If an individual selects a song, a recommendation can interrogate a number of features of that song—such as similarity of content, metadata, choice by other listeners and demographic—to form a list of recommendations. However, these systems do not take into account relevant advances made in the field of music psychology. By employing a theory of memory retention and Berlyne's inverted-U model of preference [3,4], the present work contains a broad overarching thesis on recommendation that can be applied to a variety of music stimuli. First, we outline the inverted-U model and review its efficacy and follow this with a summary of the state of the art in music recommendation systems. Next, we discuss how recommendation systems might benefit from the implementation of this psychological knowledge and propose a novel function for recommendation. Such an implementation is not intended as a replacement for current approaches but rather as a supplementary method to aid in the accurate contextual recommendation of songs the listener will enjoy.

EXPLANATION OF THE INVERTED-U MODEL

First published in 1960, Berlyne's inverted-U model is one of the most tested hypotheses of music preference [5]. The model proposes that preference is primarily related to *collative variables*—variables that can be mentally categorized and compared with one another. The most frequently tested collative variables are *complexity* and *familiarity/exposure* [6,7]. The model proposes that preference is related to collative variables in an inverted-U fashion, as outlined in Fig. 1. An intermediate level of a collative variable produces the highest level of preference and, as this collative level is increased or decreased, preference decreases from the optimal point. In cases where the range covered for a particular collative variable is narrow, the observed relationship may only produce a segment of the entire inverted-U curve. Therefore, in addition to a complete inverted-U pattern, the model can be expressed as segments of the curve [8,9], depicted in Fig. 1 with arrows. First, a positively correlated relationship representing the rising slope of the inverted-U curve can be expected in specific cases, such as when an unfamiliar stimulus is exposed only a handful of times. In this example, we might expect the optimal level of the collative variable not to be surpassed, whereas additional exposures could allow the entire inverted-U curve to appear. In contrast, decreasing preference could be expected for repeated exposures to a stimulus that is already well known.

Berlyne's model has received considerable support in the literature. Recently, we analyzed the results of 57 studies spanning 115 years investigating the relationship between music preference and one or more collative variables [10]. The reported trajectories of preference of 88% of the investigated studies were compatible with the segmented inverted-U model. In addition to this, 56% of the studies exclusively use

Anthony Chmiel (researcher), Empirical Musicology Lab, University of New South Wales, Sydney, Australia, 2052. Email: <anthony.chmiel@unsw.edu.au>. ORCID: 0000-0003-3294-0534.

Emery Schubert (researcher), Empirical Musicology Lab, University of New South Wales, Sydney, Australia, 2052. Email: <e.schubert@unsw.edu.au>. ORCID: 0000-0002-6541-8083.

See <mitpressjournals.org/toc/lmj/-/28> for supplemental files associated with this issue.

 doi:10.1162/LMJ_a_01045

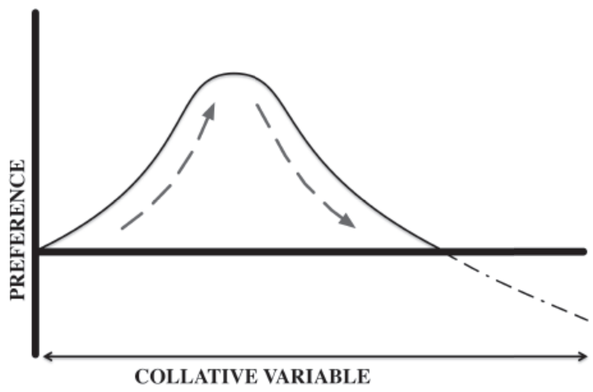

Fig. 1. Appropriation of the inverted-U relationship described by Berlyne [34]. The two segments of the curve are depicted with arrows. (© Anthony Chmiel)

linear analysis methods, whereas only 25% of the studies incorporate nonlinear analysis methods; therefore, the studies applying linear-only analysis may have hidden inverted-U results. With this in mind, when other noncollative variables (such as those associated with meaningful and personal associations, so-called ecological variables [11]) are held constant, the inverted-U model can be used to predict a considerable amount of variance in preference.

DESCRIPTION OF EXISTING MUSIC RECOMMENDATION SYSTEMS

There are five common approaches used by automated music recommendation systems: (1) *collaborative filtering*, in which recommendations are based on previous user-item relationships such as user ratings; this is the most common approach used by music recommendation systems [12]; (2) *content-based filtering,* in which the system examines descriptive information on the music, such as metadata, and compares this information with the user's listed preferences to create a recommendation; other systems also analyze the audio signal, comparing the timbral, temporal and tonal content with the signal content of other songs; (3) *demographic filtering*, in which the system creates demographic stereotypes for specific songs, artists or styles (based on variables such as age, gender and cultural or social traits) and recommends music based on the preferences of other users of a similar demographic; (4) *context-based filtering*, in which information is gathered to characterize the listening situation, which may entail gathering information about user habits and preferences, such as by examining sequential patterns of usage in a playlist [13], or about the mood, occasion, social setting or specific task at hand [14]; (5) *hybrid approaches*, in which two or more approaches are combined in an effort to increase the scope and accuracy of predictions. A hybrid approach may, for example, combine audio analysis with content-based or contextual approaches [15–17] or integrate user-rating data with contextual information [18,19]. As noted in the literature [20–22], hybrid approaches are effective at bypassing the various limitations to which singular approaches are prone. As one of many examples of this, collaborative filtering is regarded as being prone to popularity bias, or the "rich get richer" principle, in which the more an item is rated and purchased, the more it is recommended, and therefore additionally likely to continue being rated and purchased [23]. Collaborative filtering is conversely unable to recommend items that have no user ratings.

Most approaches to music recommendation include *similarity* as a central component from which their recommendations are based [24]. Similarity-based systems are often used for recommendations of books, television shows and films, in which systems are generally only required to suggest a single item at a time due to their longer consumption time and as they are not typically repeatedly consumed. This approach is not ideal for music recommendation systems, in which songs can be repeatedly consumed—sometimes in multiple successive listenings. We therefore suggest that music recommendation systems may be overlooking an important variable. We specifically propose the consideration of collative variables, such as familiarity, by tracking the previous number of exposures and the length of time between exposures of the same song, artist or genre.

APPLICATION OF THE INVERTED-U MODEL

While some research [25] has investigated the inverted-U model in terms of music sales, none has been cited that explicitly uses collative variables as a predictor of preference in a recommendation system. However, one existing music recommendation system, by Hu and Ogihara [26], has inspired our own attempt to address this gap. Their system uses five "perspectives" (genre, recording year, time pattern, favor and freshness) upon which recommendations are based. *Genre* and *recording year* were determined by time series analysis based on the latest 16 songs that were played for at least half of their duration, whereas *time pattern* is informed by the user's typical listening habits for that time of the day. A *favor* rating is also produced, with the intent of giving recommendation priority to tracks that have previously been played in full. The final perspective used by Hu and Ogihara is *freshness*, defined as the "strength of strangeness or the amount of experience forgotten" for a piece [27]. This variable is therefore related to memory as well as to familiarity. Their system applies the Ebbinghaus *forgetting curve* to calculate freshness, as depicted below [28]:

$$R = e^{-\frac{t}{S}} \qquad \text{Equation 1}$$

in which R is memory retention, S is the relative strength of memory, and t is time elapsed. S can also be thought of as an exposure event to a song in an individual's personal, mental library; we therefore express it as a positive integer. We propose that freshness or favor can be interpreted as the "boost" the R retention curve receives when a new exposure occurs (which we notate as exposure number S, which is incremented by one from the previous exposure, S-1, for $S \geq 2$) for the piece in question. For a recommendation system, this would be the modeled boost if S were incremented at time t. That is, it would be $R(t,S) - R(t,S\text{-}1)$. This can be represented in a simplified form as:

$$F(t,S) = e^{-\frac{t}{S}} - e^{-\frac{t}{S-1}} \qquad \text{Equation 2}$$

Here, F is the "freshness" or "favor" that would be gained if the song k in the library were to be played at time t. An array of F_k is generated at any point in time for all k songs in the individual's library, with the items of the array scoring the highest F values being the ones that are more likely to be recommended, along with consideration of other approaches as noted above.

Note that S is an integer ≥ 2 because at least one exposure is required against which to compare the next proposed exposure, and t is the time elapsed since the last exposure to piece k. The calculation of a value for F at a particular point in time is demonstrated in Fig. 2, and the consequent function that emerges in terms of t and S is shown in Fig. 3. If a vertical plane parallel to the F-S plane is used to form two-dimensional slices of the function (of Fig. 3), with $S > 2$, an inverted-U curve emerges, with longer time delays between exposures (slicing plane placed further along the time-axis) producing a "gentle," more slowly falling F value over exposure number than at shorter time delays (slicing, vertical plane placed closer to the time-axis origin). The slice at short time delays is consistent with preference behavior for pieces undergoing massed exposure that reach a peak F quickly, but then soon go out of favor, as is the case with high-rotation music playlists [29]. Inverted-U results also occur when the slicing plane is tilted, emanating from the time-exposure origin (still parallel to the F axis), where an increase in t and S still traces out an inverted-U shape in two dimensions. For simplicity, we have not included any constants in the proposed equations (e.g. α and β in

$$\alpha\left(e^{-\frac{t}{S}}\right) - \beta\left(e^{-\frac{t}{S-1}}\right)),$$

and it is assumed that each piece already exists in the individual's mental music library. In his review of the literature, Finnäs surmises that six to nine exposures may be sufficient to move the preference curve to its optimal point [30]. However, these matters need to be dealt with using empirical data, should the system be implemented, and we therefore recommend further study in this regard.

Additionally, future approaches may serve to extend our use of exposure to other collative variables. A brief discussion of this in terms of two variables follows (although this is a nonexhaustive list in terms of potential applications of collative variables): First, future systems may recommend music based upon a listener's perceived level of complexity, either through similarity recommendations based upon the variable or with S in the above Equation 2 being replaced by a "complexity rating." Second, a system could use familiarity for *styles* of music (stylistic familiarity) by storing a library of styles (e.g. rock, classical, jazz, etc.) and using the F value for those (call it array F_j) to provide an additional level of sophistication for inverted-U recommendation, building on systems such as that reported by Cai et al. [31], in which contextual information such as tagged metadata is used to predict a sequence of songs to be recommended next.

In short, the proposed system can continue to create recommendations according to the dynamic consideration of

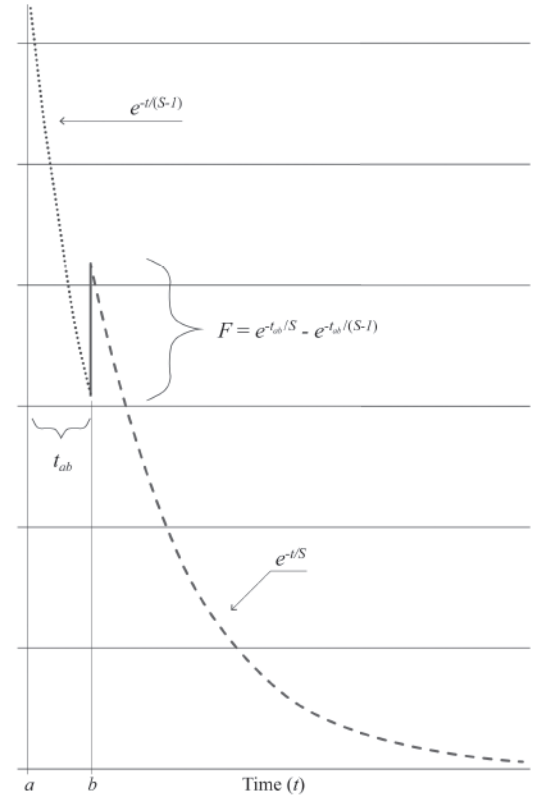

Fig. 2. Visualization of $F(t,S)$ calculation given a stimulus exposure number S occurring t_{ab} seconds after an earlier exposure to the same stimulus. Note that time a to b ($t = t_{ab}$) is the time elapsed from exposure number S-1 to the next exposure (S). F is the "favor/freshness" output used for recommendation analysis. The greater the value of F, the greater the recommendation value. The two curves apply the Ebbinghaus *forgetting curve* $R(t,S)$. Equations and plots are shown in simplified form. (© Anthony Chmiel)

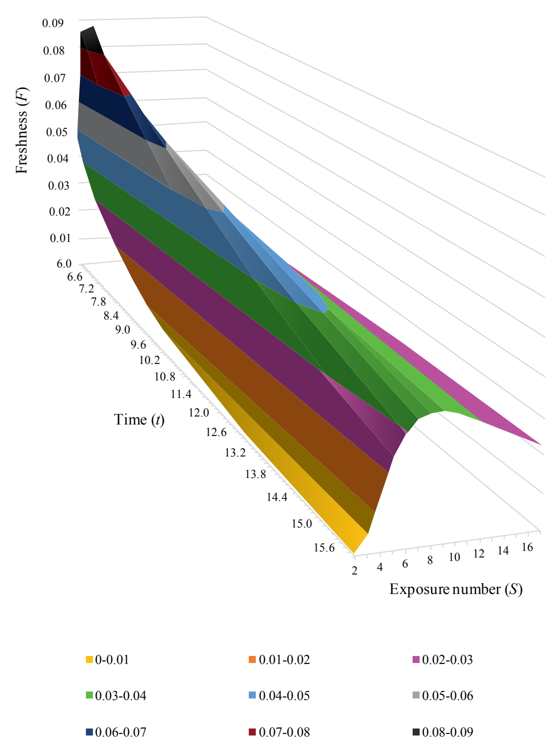

Fig. 3. $F(t, S)$. Scale values on the z-axis are arbitrary, based on the function. (© Anthony Chmiel)

collative variables such as familiarity/exposure in an effort to predict the setting of those variables that maximize enjoyment according to the psychological principles we have employed. Thus, we could expect music recommendations generated by a system based on collative information to be quite different from the recommendations made by current systems.

CONCLUSION

This paper proposed that automated music recommendation systems can be improved by incorporating well-grounded psychological principles of preference for optimal levels of collative properties of music (such as complexity and familiarity) and thus applying the inverted-U model, which describes the trajectory of preference for these variables. We implemented the inverted-U model through another well-established psychological principle from mental memory storage literature. In contrast to our approach, a large number of conventional systems repeatedly recommend similar songs and styles without explicitly taking additional variables, such as oversaturation (too many exposures over too small a time period), into account. We propose that systems informed by collative variables and the inverted-U model would therefore produce significantly different, and arguably more contextual, recommendations than existing systems. The advantages afforded by the use of these traditional, well-studied variables of music psychology cannot be overstated. We have already found elements of such an approach in existing systems [32,33], yet not in a manner that is theoretically informed by the inverted-U model. The inverted-U shape is more conceptual than mathematical. Our study may be the first psychologically plausible mathematical representation of the inverted-U. Of note is our proposal of collative variables as a new set of parameters for recommendation to coexist alongside existing parameters, such as similarity of genre and recording year. In the present work, we aimed to develop this new direction, from which engineers could implement the inverted-U model into future systems.

References

1 O. Celma, *Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space* (Berlin: Springer-Verlag, 2010).

2 M. Schedl et al., "Music Recommender Systems," in F. Ricci et al., eds., *Recommender Systems Handbook* (New York: Springer, 2015) pp. 453–492.

3 D. Berlyne, *Conflict, Arousal and Curiosity* (New York: McGraw Hill, 1960).

4 D. Berlyne, *Aesthetics and Psychobiology* (New York: Appleton-Century-Crofts, 1971).

5 D. Hargreaves and A. North, "Experimental Aesthetics and Liking for Music," in P.N. Juslin and J. Sloboda, eds., *Handbook of Music and Emotion: Theory, Research, Applications* (New York: Oxford Univ. Press, 2010) pp. 515–546.

6 For an extended list of additional collative variables, see Berlyne [3].

7 See also Berlyne [4].

8 R. Heyduk, "Rated Preference for Musical Compositions as It Relates to Complexity and Exposure Frequency," *Perception and Psychophysics* **17**, No. 1, 84–91 (1975).

9 E. Walker, "Psychological Complexity and Preference: A Hedgehog Theory of Behavior," in D.E. Berlyne and K.B. Madsen, eds., *Pleasure, Reward, Preference* (New York: Academic Press, 1973) pp. 65–97.

10 A. Chmiel and E. Schubert, "Back to the Inverted-U for Music Preference: A Review of the Literature," *Psychology of Music* **45**, No. 6, 886–909 (2017).

11 Berlyne [4] p. 69.

12 Schedl et al. [2] p. 454.

13 Celma [1].

14 R. Cai et al., "Scalable Music Recommendation by Search," *Proceedings of the 13th ACM International Conference on Multimedia* (Singapore, 2007).

15 J. Bu et al., "Music Recommendation by Unified Hypergraph: Combining Social Media Information and Music Content," *Proceedings of the 18th ACM International Conference on Multimedia* (Firenze, 2010).

16 F. Vignoli and S. Pauws, "A Music Retrieval System Based on User Driven Similarity and Its Evaluation," *Proceedings of the 6th International Conference on Music Information Retrieval* (London, 2005).

17 B. Shao et al., "Music Recommendation Based on Acoustic Features and User Access Patterns," *IEEE Transactions on Audio, Speech, and Language Processing* **17**, No. 8, 1602–1611 (2009).

18 K. Yoshii and M. Goto, "Continuous pLSI and Smoothing Techniques for Hybrid Music Recommendation," *Proceedings of the 10th International Society for Music Information Retrieval Conference* (Kobe, 2009).

19 K. Yoshii et al., "Hybrid Collaborative and Content-Based Music Recommendation Using Probabilistic Model with Latent User Preferences," *Proceedings of the 7th International Conference on Music Information Retrieval* (Victoria, 2006).

20 Celma [1].

21 O. Celma and P. Cano, "From Hits to Niches? Or How Popular Artists Can Bias Music Recommendation and Discovery," in *2nd Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition* (Las Vegas, 2008).

22 Schedl et al. [2].

23 Celma [1].

24 Schedl et al. [2].

25 See, for example, A. North et al., "Energy, Typicality, and Music Sales: A Computerized Analysis of 143,353 Pieces," *Empirical Studies of the Arts* **35**, No. 2, 214–229 (2017).

26 Y. Hu and M. Ogihara, "Nextone Player: A Music Recommendation System Based on User Behavior," *Proceedings of the 12th International Society for Music Information Retrieval Conference* (Miami, 2011).

27 Hu and Ogihara [26] p. 105.

28 See also H. Ebbinghaus, *Memory: A Contribution to Experimental Psychology* (New York: Teachers College, 1913).

29 See, for example, C. Martindale, "The Pleasures of Thought: A The-

ory of Cognitive Hedonics," *Journal of Mind and Behavior* 5 (1984) pp. 49–80.

30 L. Finnäs, "How Can Musical Preferences Be Modified? A Research Review," *Bulletin of the Council for Research in Music Education* 102 (1989) pp. 47–54.

31 Cai et al. [14].

32 Hu and Ogihara [26].

33 North et al. [25].

34 Berlyne [4] p. 89.

Manuscript received 30 November 2017.

ANTHONY CHMIEL *is a PhD candidate at the University of New South Wales in Sydney, Australia, as well as an active musician. His primary research focus is on empirical investigations into predictors of music preference.*

EMERY SCHUBERT *is a professor of music at the University of New South Wales and head of the UNSW Empirical Musicology Laboratory.*