Creating an Environment for Bisociative Music Discovery and Recommendation

Sebastian Stober
Data & Knowledge Engineering Group
Faculty of Computer Science
Otto-von-Guericke-University Magdeburg
D-39106 Magdeburg, Germany
sebastian.stober@ovgu.de

Stefan Haun
Data & Knowledge Engineering Group
Faculty of Computer Science
Otto-von-Guericke-University Magdeburg
D-39106 Magdeburg, Germany
stefan.haun@ovgu.de

Andreas Nürnberger
Data & Knowledge Engineering Group
Faculty of Computer Science
Otto-von-Guericke-University Magdeburg
D-39106 Magdeburg, Germany
andreas.nuernberger@ovgu.de

Audio Mostly 2011, September 7 – 9, Coimbra, Portugal.
ACM XXX-X-XXXX-XXXX-X/XX/XX

Abstract

Surprising a user with unexpected and fortunate recommendations is a key challenge for recommender systems. Motivated by the concept of bisociations, we propose ways to create an environment where such serendipitous recommendations become more likely. As application domain we focus on music recommendation using MusicGalaxy, an adaptive user-interface for exploring music collections. It leverages a non-linear multi-focus distortion technique that adaptively highlights related music tracks in a projection-based collection visualization depending on the current region of interest. While originally developed to alleviate the impact of inevitable projection errors, it can also adapt according to user-preferences. We discuss how using this technique beyond its original purpose can create distortions of the visualization that facilitate bisociative music discovery.

Keywords

Bisociation, Recommendation, Exploratory Search, User-Interfaces, Adaptive Similarity, Music Retrieval

ACM Classification Keywords

H.5.2 User Interfaces, H.3.3 Information Search and Retrieval, H.5.4 Hypertext/Hypermedia

Introduction

One of the big challenges of computer science in the 21st century is the digital media explosion. Online music stores already contain several millions of music tracks and steadily growing hard-drives are filled with personal music collections of which a large portion is simply collecting dust. Music recommender systems aim to help us cope with this amount of data and find new interesting music or rediscover once loved pieces we have forgotten about – a task also called "recomindation" [7]. One common problem that many recommender systems face is that their recommendations are often too obvious and thus not particularly useful when it comes to discovering new music. Especially, collaborative filtering approaches are prone to a strong popularity bias [1]. In fact, McNee et al. argue that there is too much focus on improving the accuracy of recommender systems. They identify several important aspects of human-recommender interaction of which serendipity is specifically related to the above phenomenon [6]. A serendipitous recommendation is unexpected and fortunate – something that is particularly hard to grasp and evaluate.

The MusicGalaxy User-Interface

MusicGalaxy is an interface for exploring music collections shown in Figure 1 that we have developed in previous work [11].[1] An overview of the entire collection is given by displaying few spatially well distributed music tracks as thumbnails (album covers) for orientation. The remaining tracks are displayed as points. The initial distribution of objects on the display is computed by multi-dimensional scaling (MDS). Users can enlarge interesting regions with a fish-eye lens called "primary focus" that allows more thumbnails to be displayed at bigger size. The surrounding space is compacted but not hidden to preserve overview.

facet	projection	distortion
dynamics	0.0	1.0
rhythm	1.0	0.0
timbre	0.0	1.0

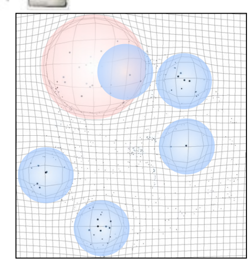

Figure 1: Top: MusicGalaxy (inverted color scheme for print). Bottom right: corresponding lens distortion resulting from (user-controlled) primary focus (red) & (adaptive) secondary lenses (blue). Bottom left: facet weights for the projection and distortion distance measures (values adaptable).

The MDS – as any other dimensionality reduction technique – introduces "projection errors" in the sense that objects that

[1] Demo videos are available at http://www.dke-research.de/aucoma

are very close in high-dimensional feature-space might be projected at large distances from each other and objects that are very dissimilar are placed next to each other respectively.[2] This effect is alleviated by automatically adapting a secondary focus consisting of additional fish-eye lenses in regions containing objects similar to those in primary focus. The resulting distortion brings separated nearest neighbors back together.

Additionally, it is possible to adapt the underlying similarity space. To this end, music similarity is represented as a distance measure that is a weighted linear combination of facet distances. Each facet covers a specific aspect of music similarity such as melody, harmony, rhythm, dynamics or lyrics and is defined by one ore more representative features and an appropriate distance measure. The importance of the individual facets can be adapted by changing their weights for the linear aggregation. As we have shown in recent experiments with simulated user-interaction [12], it is also possible to adapt the weights automatically based on (relative) preferences derived from user actions such as judging two objects to be more similar w.r.t. a third one. This way, a personalized music similarity measure can be learned.

Bisociations

A study was conducted to assess the usability and usefulness of the visualization technique for the exploration of large multimedia collections [10]. 30 participants had to select photos to represent several topics such as "birds" or "plants". As anticipated, some participants used the primary lens to skim through the photo collection in a rather continuous fashion. But surprisingly, there was also a group that browsed the collection mostly by moving (in a single click) the primary focus to some (previously) secondary focus region step-by-step – much like navigating an invisible neighborhood graph. Further, as it happened, one of the participants encountered a funny incident: While looking for photographs showing a lizard, he had an image of a *monitor lizard* in primary focus. To his surprise, the system retrieved an image showing the rock painting of a lizard. Interestingly, rock paintings were actually another topic to find photos for and the relevant photos were a lot harder to make out in the collection than the lizards. Bearing in mind that according to Isaac Asimov "the most exciting phrase to hear in science, the one that heralds new discoveries, is not 'Eureka!' (I found it!) but 'That's funny ...''', we decided to further investigate this phenomenon. What the participant encountered is called a *bisociation* – a bridging element between the two distinct domains: animals and rock paintings. While most associations are found between concepts of one domain, there are certain paths which either bridge two different domains or connect concepts by incorporating another domain. In his book *The Act of Creation*, Arthur Köstler, an Austrian publisher, coined the term *bisociation* for these types of associations and as it turns out, many scientific discoveries are in some way bisociations [3].

Admittedly, no one expects scientific discoveries from a music recommender application. However, the question persists whether we can leverage the effect of bisociations and create an environment where serendipitous recommendations become more likely. After all, the concept of bisociation is much easier to grasp than serendipity and can even be formalized by means of graph theory [4].

Bisociative Lens Distortions

How can MusicGalaxy be turned into an environment that supports bisociative music discovery? The general idea is to combine two distinct domain views into one visualization by using the secondary focus to highlight connections to nearest

[2] Note that it is impossible to fix these problems without causing damage elsewhere as the projection is in general already optimal w.r.t. the projection technique.

neighbors *in a different domain* than the one used for projection: The "primary domain" is directly visualized by the projection and contains the displayed tracks connected by neighborhood relations that are implicitly induced between each track and its neighbors in the projection.[3] On the contrary, the "secondary domain" which is used to identify nearest neighbors for the secondary focus distortion is not directly visible to the user. A bisociation occurs in this setting, if two tracks are not neighbors in the projection domain (i.e., close to each other in the display) but are connected in the secondary domain. In this case, the secondary focus will highlight this connection by focusing on the bisociated track.

Orthogonal Similarity Measures

A simple way to create such a setting is to use orthogonal similarity measures (i.e., defined on non-overlapping facet sets) for the two domains by choosing the facet weights accordingly. For instance, in Figure 1 the tracks in secondary focus are very different in rhythm (large distance in projection) but very similar in dynamics and timbre w.r.t. the track in primary focus. This approach could also be used in different applications. To illustrate the possibilities, imagine a user wants to explore a collection of world-music as, e.g., addressed by mHashup [5]. In such applications, a straightforward way for the arrangement of the tracks would be according to their geographical origin, i.e., mapping the tracks on a common world map. Using this primary domain instantly gives the user an overview of the geographic distribution of the tracks in the collection. With the primary fish-eye lens, the user could magnify a region he is interested in. This would allow to display the local distribution of tracks in more detail and differentiate smaller (sub)regions. Note that in this special case, the arrangement of the tracks is perfect in the sense that all distances can be displayed distortion-free (except for the neglectible mapping of the earth's surface to a plane) because there is no dimensionality reduction involved. The secondary focus in its original setting would be unnecessary here anyway and it could therefore be freely used to highlight regions with nearest neighbors w.r.t. other aspects addressed by the secondary domain – e.g., acoustic similarity as a combination of several respective facets. Further, analyzing the interaction with the user, the system can over time learn which (acoustic) facets (of the secondary domain) are particularly important for the user and personalize the similarity measure for nearest neighbor retrieval accordingly. This has already been described and evaluated in [12].

Generalization to Domain Graphs

The above example uses an orthogonal similarity measure for the secondary domain. This is, however, only a very special case. Generally, the secondary domain might be any graph that contains at least the tracks as concepts (nodes) and allows to find neighboring tracks by some way of traversing relations between the concepts. An orthogonal similarity measure as described above induces such a graph: In this case, the graph contains only the tracks as concepts plus relations between tracks that are nearest neighbors and finding nearest neighbors for a track means simply returning all directly related tracks. An alternative way to construct such a sparse neighborhood graph for the secondary domain is to use any (black-box) system that recommends similar tracks for a seed track or even a combination of several such systems. However, the graph does not need to be confined to tracks. In fact, it may be arbitrarily complex – e.g., contain also artists, albums plus respective relations and possibly allowing multiple paths between tracks. For instance,

[3] This is rather an artificial mental model a user perceives as no connections are visualized explicitly. Due to possible distortions introduced by dimensionality reduction, it only approximates the one derived from the actual distances in the original space.

[4] http://musicbrainz.org/

from the freely available data from MusicBrainz[4], a user maintained community music metadatabase, a large graph can be constructed containing more than 10M tracks, 740K albums, 600K artists and 48K labels.[5] Between these entities, common relationships exist that, e.g., link tracks to artists and albums as well as albums to artists and labels. Apart from this, a large variety of advanced relationships links (ARL) exists. They are particularly interesting as they go beyond trivial information, such as links from tracks and albums to mastering and recording engineers, producers and studios (in total more than 281K artist-album and 786K artist-recording ARLs), how artists are related with each other (more than 135K ARLs), or which tracks contain samples of others (more than 44K recording-recording ARLs).[6]

Nearest neighbors for a track in primary focus can be found by traversing the MusicBrainz graph in breadth-first order collecting paths to other tracks. Graph traversal stops when either the traversal depth or the number of reached track nodes exceeds a predefined threshold. As only the most relevant tracks can be highlighted by the secondary focus, some relevance measure is required to rank the retrieved tracks. Because increasing serendipity is the main objective, the relevance measure should capture how likely a track will be a lucky surprise for the user. This is however all but trivial. Possible simple heuristics are:

- Prefer tracks that are projected far away from the primary focus (and thus most likely sound very different).
- Prefer tracks that the user has not listened to a lot or for a long time (and probably is no longer aware of).
- Prefer tracks of different artists and/or albums.

The result of using either heuristic or a combination thereof will most likely surprise the user but at the same time the risk is high that the connection to the primary focus is too far fetched. Therefore, paths need to be judged according to their interestingness. Platt [8] defines discrete edge distances depending on the type of relationships for a similar graph created on a dataset from the All Music Guide [2]. Similar weightings can be applied here. Alternatively, weights could be assigned to common path patterns instead – possibly penalizing longer paths. For instance, some path patterns are straightforward such as *track-artist-track* (same artist) or *track-album-track* (same album) where the latter is more interesting in terms of serendipity because it could be a compilation that also contains tracks of other artists. Both weighting approaches require empirical tuning of the respective weights. Another option is to count the frequencies of occurring path patterns and boost infrequent and thus remarkable patterns which can be interpreted as analogy to the idf weights used in text retrieval. This favors patterns containing ARLs. If multiple paths between two tracks are found, their weights can be aggregated, e.g., using the maximum, minimum or average. More sophisticated methods are currently developed to facilitate bisociations on text collections [9] and could also be applied here to further increase the chances of bisociative recommendations from complex domain graphs. This is currently studied more thoroughly as the impact of the different heuristics and the values of their respective parameters are not yet fully clear.

Conclusions

This paper described an approach to increase the chance of serendipitous recommendations in an exploratory music retrieval scenario. Instead of addressing serendipity directly, we proposed to exploit the related concept of bisociations that can be formalized by means of graph theory. We demon-

[5] Figures as of January 2011 when the graph was created.

[6] Full list available at: http://wiki.musicbrainz.org/Category:Relationship_Type

strated how separating the underlying similarity measures for projection and distortion in the MusicGalaxy interface makes is possible to link two distinct domain views on a music collection – creating a setting that promotes bisociations where serendipitous recommendations become more likely. We hope that this paper can contribute to the ongoing discussion of improving the serendipity of recommendations and at the same time spreads the awareness of the bisociation concept.

Acknowledgments

This work was supported in part by the German National Merit Foundation, the German Research Foundation (DFG) project AUCOMA and the European Commission under FP7-ICT-2007-C FET-Open, contract no. BISON-211898.

References

[1] O. Celma and P. Cano. From hits to niches?: or how popular artists can bias music recommendation and discovery. In *Proc. of NETFLIX'08*, 2008.

[2] D. P. W. Ellis, B. Whitman, A. Berenzweig, and S. Lawrence. The quest for ground truth in musical artist similarity. In *Proc. of ISMIR'02*, 2002.

[3] A. Köstler. *The Act of Creation*. Macmillan, 1964.

[4] T. Kötter, K. Thiel, and M. Berthold. Domain bridging associations support creativity. In *Proc. of ICCC'10*, 2010.

[5] M. Magas, M. Casey, and C. Rhodes. mhashup: fast visual music discovery via locality sensitive hashing. In *SIGGRAPH '08 demos*, 2008.

[6] S. M. McNee, J. Riedl, and J. A. Konstan. Making recommendations better: an analytic model for human-recommender interaction. In *Proc. of CHI'06*, 2006.

[7] C. Plate, N. Basselin, A. Kröner, M. Schneider, S. Baldes, V. Dimitrova, and A. Jameson. Recomindation: New functions for augmented memories. In *Proc. of AH'06*, 2006.

[8] J. C. Platt. Fast embedding of sparse music similarity graphs. In *Advances in Neural Information Processing Systems (NIPS'03)*, volume 16, 2004.

[9] M. Segond and C. Borgelt:. Selecting the links in bisonets generated from document collections. *Advances in Intelligent Data Analysis IX*, 6065:196–207, 2010.

[10] S. Stober, C. Hentschel, and A. Nürnberger. Evaluation of adaptive springlens – a multi-focus interface for exploring multimedia collections. In *NordiCHI'10*, 2010.

[11] S. Stober and A. Nürnberger. MusicGalaxy – an adaptive user-interface for exploratory music retrieval. In *Proc. of SMC'10*, 2010.

[12] S. Stober and A. Nürnberger. Similarity adaptation in an exploratory retrieval scenario. In *Proc. of AMR'10*, 2010.