POTENTIAL RELATIONSHIP DISCOVERY IN TAG-AWARE MUSIC STYLE CLUSTERING AND ARTIST SOCIAL NETWORKS

Dingding Wang
Center for Computational Science
University of Miami
Coral Gables, FL USA
d.wang1@miami.edu

Mitsunori Ogihara
Department of Computer Science
University of Miami
Coral Gables, FL USA
ogihara@cs.miami.edu

ABSTRACT

With the rapid growth of music information and data in today's ever changing world, exploring and analyzing music style has become more and more difficult. Traditional content-based methods for music style analysis and newly emerged tag-based methods usually assume music items are independent of each other. However, in real world applications, do there exist some relationships among them. In this paper, we construct the social relation graph among different music artists by extracting the friendship information from social media such as Twitter, and incorporate the generated social networking graph into tag-based music style clustering. Experiments on real data show the effectiveness of this novel integration of different information sources.

1. INTRODUCTION

As the rapid growth of music items on the Internet, music style analysis such as music classification and clustering has become increasingly prevalent in music information retrieval research. Traditional methods usually focus on audio feature extraction and acoustic content analysis. For example, Pampalk et al. [19] integrate different similarity sources based on fluctuation patterns and use a nearest neighbor classifier to categorize music items. Chen and Chen [3] apply both long-term and short-term features and uses support vector machines to classify music genres.

More recently, methods utilizing music social tags have emerged and have been receiving more and more attention. Social tags are free-text descriptions added by users to express their personal views and interests in music items such as songs, artists, albums, and playlists. The tags provide direct insights into user behavior and opinions and the retrieval methods using tags have been shown to be more effective than the traditional methods solely based on music content analysis [14,24,27]. For example, Bischoff et al. [2] demonstrate that different types of social tags can improve music search. Symeonidis et al. [23] propose a music recommendation algorithm using a user-tag-item tensor. Wang et al. [26] show the effectiveness of tag features by way of joint analysis of tags and contents.

Although the content-based and tag-aware methods are successful in many music information retrieval applications, they make a somewhat curious assumption that music items are independent of each other, which is not always true. In this paper, we assume that music items are related to each other and try to establish relations among them by discovering relationships among artists. To do this we look for the "following" information on Twitter and construct a linked graph to represent the artist social network. We then propose a novel tag-aware music style clustering system utilizing this network by way of matrix factorization. By assuming that the "follower" relationship as represented in the social network thus build is transitive, we can capture indirect relationships among the artists, which are usually ignored in the existing music style clustering methods.

The rest of this paper is organized as follows. Section 2 discusses the related work. Section 3 introduces our proposed approaches for constructing the artist social graph, generating artist relation matrix, and clustering using relation matrix based factorization. We conduct experiments on a real world data set and Section 4 presents the experimental results. Section 5 gives an conclusion and discusses the future work.

2. RELATED WORK

Automatic music analysis such as music item clustering, classification, and similarity search has been playing a central role in music information retrieval. Traditional automatic music analysis methods usually focus only on audio content analysis via audio feature extraction. Timbral texture features [25] are the most widely used features, which usually consist of Short-Term Fourier Transform (STFT) and Mel-Frequency Cepstral Coefficients (MFCC) [21]. Various data mining and statistical methods have been applied to such features for classification and clustering of music items, such as artists, songs, and albums [3, 4, 6, 12, 20, 25].

Analysis of music social tags is a subject quickly gaining popularity in music information retrieval research. Music social tags are free-text descriptions of any length (though in practice there sometimes is a limit in terms of number of characters) with no restriction on the words to be used. Because they are free texts, they are thought of as representing feelings of listeners on the music items (artists, songs, etc.) for which they leave tags. Also, because they are free texts, they range from a single character (e.g., "!") to a full sentence (e.g., "I love you baby, can I have some more?"). However, in many cases, they are one or two words, such as "Sad", "Happy", "Black Metal", "Loved it", and "Indie Pop". As can be easily seen social tags include words that do not necessarily appear as labels experts such as musicologists provide. Their amateurism notwithstanding, by collecting a large number of tags for one single piece of music item, an understanding can be obtained on how the general listeners appreciate the item. With that idea, work has been done to show the promise of using tags for music data analysis. For example, Lamere and Pampalk [15] use tags to enhance simple search, similarity analysis, and clustering of music items. Lehwark et al. [17] generate visual clustering of tagged music data. Karydis et al. [13] propose a tensor-based algorithm to cluster music items using 3-way relational data involving song, users, and tags. The effectiveness of tags may come from the fact that the distance between the original data source and the tag in terms of informativeness appears to be much smaller. There also exist a few efforts in combining content-based and tag-based analysis. For example, F. Wang et al. [27] attempts to integrate audio contents and tags for multi-label classification of music styles. D. Wang et al. [26] explores the integration of music content and tags in the problem of artist style clustering.

In addition to social tags, much more social information has become available on the Internet. For instance, social networking sites, such as Facebook and MySpace, and a social medium Twitter can provide the friendship information among users by adding a friend on Facebook or following a tweet page on Twitter. Recent work by Anglade et al. [1] uses complex network theoretic analysis to group similar listeners. Jacobson et al. [11] and Fields et al. [8] study the influence of social networks for the music community detection and playlists generation. In this paper, we explore the effectiveness of the joint use of the analysis of the social networking graph and the tag-based music style clustering.

3. METHODOLOGY

3.1 Framework

Figure 3.1 shows the framework of our proposed music style clustering system that integrates tag and social graph analysis. Given a collection of representative music pieces from different artists, we first obtain the tags describing these music pieces to construct a music-tag matrix and generate the social networking graph among the artists, from which the artist relation matrix is created. We then perform matrix factorization on the music-tag matrix using artist relation matrix as the base. Upon the convergence of the factorization, we can obtain the music style indicator matrix and finally partition the music pieces into different style groups.

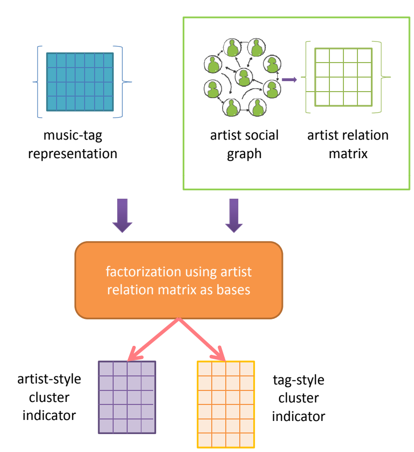

Figure 1. The framework of the proposed method.

3.2 Artist Social Graph Construction

In order to construct an artist social graph, we select 327 artists that are active users of Twitter. The genres covered these artists are Pop, Jazz, Rock, Hip Hop, and Country. Each node of the graph represents an artist. For these artists we extract the "following" information among these artists using the API provided by Twitter. If artist A_i is "following" the tweets of artist A_j, there will be a directed link from node A_i to node A_j. An example of the generated social graph is shown in Figure 2.

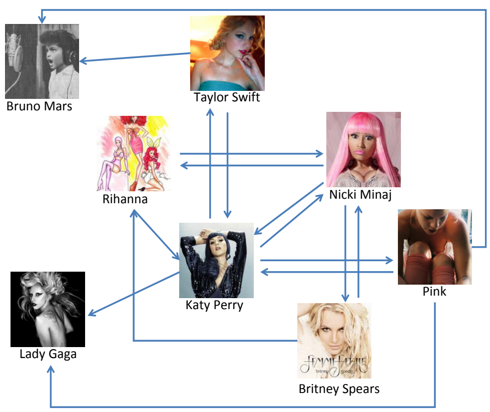

Figure 2. An Example Social Graph Generated from Twitter.

3.3 Artist Relation Matrix Generation

Based on the artist social graph, we can generate the artist relation matrix which considers both the direct and indirect relationships using the method proposed in [9]. Suppose that artist A_i is followed by a set of artists F_i, a matrix S to represent the direct relationships among the artists can be computed in this way:

$$S_{ji} = \begin{cases} 1/|F_i| & if A_j \in F_i \\ 0 & otherwise \end{cases},$$

where $|F_i|$ is the size of set F_i. To capture the indirect relationships, we perform a random walk on the directed graph denoted by S. An artist can be identified as a related one if the random walk stops at the node representing him/her. A parameter α is used to specify the probability that the random walk stops at the current node which is set to 0.99 in the experiments, and based on the properties of random walk, the relation matrix can be computed as

$$B = (1-\alpha)(1-\alpha S)^{-1}.$$

3.4 Factorization with Artist Relation Base Matrix

3.4.1 the Model

In order to obtain the music style clusters, we perform matrix tri-factorization [7] using the artist relation matrix as the base matrix. The problem can be treated as an optimization problem with the following objective:

$$\min_{B\geq 0, U\geq 0, V\geq 0} ||X - BUV^T||, s.t. U^TU = I, V^TV = I,$$

where X denotes the artist-tag matrix, and B is the generated artist relation matrix as described in Section 3.3. From U, we can obtain the artist-style clusters, and from V we can get the tag-style clusters. To solve this optimization problem, we use an algorithm similar to the tri-factorization [7] and nonnegative matrix factorization (NMF) [16] to iteratively update U and V as follows:

$$U_{as} \leftarrow U_{as}[CB^TV]_{as}$$

$$V_{ts} \leftarrow V_{ts}[BD^TU]_{ts},$$

where $C_{ij} = X_{ij}/[UV^TB]_{ij}$, and $D_{ij} = X_{ij}/[BUV^T]_{ij}$. Different with the traditional tri-factorization approach, here we use the social relation matrix as the base matrix to incorporate social networking information among the artists, and the base matrix is fixed during the updates of the other two matrices. The benefit of using the base matrix is that the artist relations obtained from the social media can be naturally incorporated to guide the factorization procedures.

3.4.2 Computational Algorithm

In the algorithm derivation, we follow the Expectation-Maximization (EM) procedure to maximize the marginalized likelihood of observations by iteratively updating the artist-style and tag-style matrices until convergence. The computational algorithm is described in Algorithm 1.

3.4.3 Algorithm Correctness

Now we prove the loss $\ell(U, V)$ is nonincreasing under the update rules.

Proof Let $\alpha_{iklj} = B_{ik}\widetilde{U}_{kl}\widetilde{V}_{jl}/[B\widetilde{U}\widetilde{V}^T]_{ij}$. Applying Jensen's inequality, we obtain

$$\begin{aligned}\ell(U,V) &= \sum_{ij}(\sum_{kl} B_{ik}U_{kl}V_{jl} - X_{ij}\ln(\sum_{kl} B_{ik}U_{kl}V_{jl})) \\ &\leq \sum_{ij}\sum_{kl}(B_{ik}U_{kl}V_{jl} - X_{ij}\ln\frac{B_{ik}U_{kl}V_{jl}}{\alpha_{iklj}}) \\ &= -\sum_{ijkl} C_{ij}B_{ik}\widetilde{U}_{kl}\widetilde{V}_{jl}\ln(U_{kl}V_{jl}) \\ &\stackrel{\text{def}}{=} \mathcal{Q}(\mathcal{U},\mathcal{V};\widetilde{\mathcal{U}},\widetilde{\mathcal{V}}).\end{aligned} \quad (1)$$

The equality holds when $U = \widetilde{U}$ and $V = \widetilde{V}$. Instead of minimizing ℓ, we minimize $\mathcal{Q}$ without the non-negative

Algorithm 1 Factorization given an artist relation base.

Input: X: artist-tag matrix.
B: artist-artist matrix;
Output: U: artist-style matrix;
V: tag-style matrix.
begin
1. **Initialization:**
Randomly initialize U and V.
2. **Iteration:**
repeat
2.1 Compute $C_{ij} = X_{ij}/[UV^TB]_{ij}$;
2.2 Assign $U_{as} \leftarrow U_{as}\left[B^TCV\right]_{st}$,
and normalize each column to 1;
2.3 Compute $D_{ij} = X_{ij}/[BUV^T]_{ij}$;
2.4 Assign $V_{ts} \leftarrow V_{ts}\left[D^TBU\right]_{dt}$,
and normalize each row to 1;
until convergence
3. **Return** U**,** V
end

constraints. Later on, we find that the update rules satisfy the non-negative constraints. The Lagrangian of $\mathcal{Q}$ is

$$\mathcal{L}(\mathcal{U},\mathcal{V};\boldsymbol{\xi}) = \mathcal{Q}(\mathcal{U},\mathcal{V};\widetilde{\mathcal{U}},\widetilde{\mathcal{V}}) + \boldsymbol{\xi}^T(\mathcal{U}^T\mathbf{1}-\mathbf{1}). + \boldsymbol{\zeta}^T(\mathcal{V}\mathbf{1}-\mathbf{1}). \quad (2)$$

The Karush-Kuhn-Tucker (KKT) conditions are

$$\partial\mathcal{L}U_{kl} = -\frac{1}{U_{kl}}\widetilde{U}_{kl}\left[B^TC\widetilde{V}\right]_{kl} + \xi_l = 0, \quad (3)$$

$$\partial\mathcal{L}V_{jl} = -\frac{1}{V_{jl}}\widetilde{V}_{jl}\left[D^TB\widetilde{U}\right]_{jl} + \zeta_j = 0, \quad (4)$$

$$\partial\mathcal{L}\boldsymbol{\xi}_l = \sum_k U_{kl} - 1 = 0, \quad (5)$$

$$\partial\mathcal{L}\boldsymbol{\zeta}_j = \sum_l V_{jl} - 1 = 0 \quad (6)$$

We derive the update rule from the KKT conditions. We can verify that the update rules keep U and V non-negative.

4. EXPERIMENTS

4.1 Data Collection

For experimental purpose, we select 327 most popular artists of the following 5 styles: Pop (91 artists), Rock (67 artists), Country (55 artists), Jazz (48 artists), and Hip Hop (66 artists). We use the API provided by Twitter to check if there is a "following" relationship among these artists. The style information and tags of the artists are collected from Last.fm (http://www.last.fm).

4.2 Implemented Baselines

We implement the following baselines to compare them with our proposed method which integrating the social tags and the social networking graph.

- K-means - performs standard K-means clustering on the artist-tag matrix.
- Normalized Cuts (Ncut) [28] - conducts graph-based spectral clustering using normalized cuts.
- Nonnegative Matrix Factorization (NMF) [16] - performs nonnegative matrix factorization on the artist-tag matrix to obtain the artist-style matrix from which the artist cluster assignments can be obtained.
- Tri-factorization (Tri-fac) [7] - performs tri-factorization on the artist-tag matrix.
- Probabilistic Latent Semantic Indexing (PLSI) [10] - performs PLSI on the artist-tag matrix.
- PLSI+PHITS [5] - combines the tag-based analysis with social graph using PLSI plus Probabilistic Hyperlink-Induced Topic Search (PHITS).

These baseline methods that we use in the experiments are most widely used clustering algorithms and some new emerged methods combing content and link analysis in data mining, information retrieval, and social network analysis areas. We aim to compare our proposed models with the state-of-the-art methods for artist clustering.

4.3 Evaluation Methods

To measure the artist style clustering performance, we use accuracy and normalized mutual information (NMI) as performance measures.

- Accuracy measures the relationship between each cluster and the ground truth class. It sums up the total matching degree between all pairs of clusters and classes. Accuracy can be represented as:

  $$Accuracy = Max\,(\sum_{C_k,L_m} T(C_k,L_m))/N,$$

  where C_k denotes the k-th cluster, and L_m is the m-th class. $T(C_k,L_m)$ is the number of entities which belong to class m and are assigned to cluster k. Accuracy computes the maximum sum of

$T(C_k, L_m)$ for all pairs of clusters and classes, and there is no overlap among these pairs. It is obvious that the greater accuracy, the better clustering performance.

- NMI [22] measures the amount of statistical information shared by two random variables representing cluster assignment and underlying class label. Suppose entry n_{ij} denotes the amount of data items belonging to cluster i and class j. NMI is then computed as:

$$NMI = \frac{\sum_{i=1}^{c}\sum_{j=1}^{k}\frac{n_{ij}}{n}log\frac{n_{ij}n}{n_{i.}n_{.j}}}{\sqrt{(\sum_{i=1}^{c}\frac{-n_{i.}}{n}log\frac{n_{i.}}{n})(\sum_{j=1}^{k}\frac{-n_{.j}}{n}log\frac{n_{.j}}{n})}},$$

where $n_{i.} = \sum_{j=1}^{k} n_{ij}$, $n_{.j} = \sum_{i=1}^{c} n_{ij}$, n, c, k denote the total number of data objects, the number of clusters, and the number of classes, respectively. Based on our prior knowledge of the number of classes, we set the number of clusters equal to the true number of classes, i.e., $c = k$.

4.4 Experimental Results

Figure 3 and Figure 4 show the accuracy and NMI results of different clustering methods respectively.

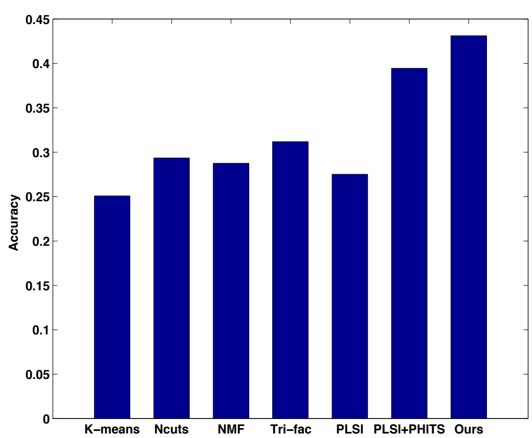

Figure 3. The accuracy results of different clustering methods.

The clustering results of our proposed method outperforms the state-of-the-art methods significantly. From the results, we have the following observations.

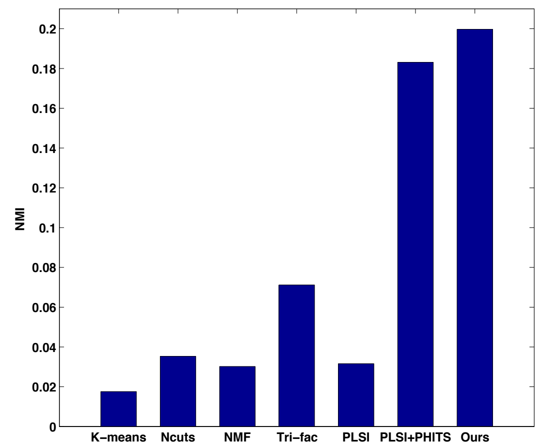

Figure 4. The NMI results of different clustering methods.

- (1) Graph-based and factorization based tag-aware clustering methods outperform traditional clustering methods such as K-means.
- (2) Methods incorporating social networking graph analysis (such as PLSI+PHITS and Ours) demonstrate more promising performance than the methods using only social tag information, which shows the effectiveness of the integration of the different information sources.
- (3) Our factorization with given artist relation bases outperforms PLSI+PHITS which is one of the most widely used combination methods because our method takes the indirect relationships into consideration and naturally incorporates it into the algorithm.

5. CONCLUSION AND FUTURE WORK

In this paper, we explore the potential benefits of integrating tags and social networking graphs in music style clustering. Given a collection of artists and their representative music pieces, social tags of free languages are extracted to describe the music pieces. The direct and indirect relationships among the artists are also discovered from the artist social networking graph, which is generated from popular social media sites, such as Twitter. Then a factorization based algorithm is derived to make use of both the two types of information. Experimental results on real world data demonstrate the effectiveness of the proposed method.

This is a pilot study of incorporating social networking analysis into music style clustering, and the initial results show the promising future of research in this direction. In the future work, large-scale data sets will be collected and further experiments will be performed on them. We will also discover other meaningful and useful types of information and examine if they can facilitate the task of music style analysis.

6. ACKNOWLEDGMENTS

This work is in part supported by NSF Grant CCF-0948590.

7. REFERENCES

[1] A. Anglade, M. Tiemann, and F. Vignoli: "Complex-network theoretic clustering for identifying groups of similar listeners in p2p systems," *Proceedings of RecSys*, 2007.

[2] K. Bischoff, C. Firan, W. Nejdl, and R. Paiu: "Can all tags be used for search?," *Proceedings of CIKM*, 2008.

[3] S. Chen and S. Chen: "Content-based music genre classification Using timbral feature vectors and support vector machine," *Proceedings of ICIS*, 2009.

[4] R. Cilibrasi, P. Vitányi, and R. Wolf: "Algorithmic clustering of music Based on string compression,"*Computer Music Journal* 28:4, 2004.

[5] D. Cohn and T. Hofmann: "The missing link - a probabilistic model of document content and hypertext connectivity," *NIPS*, 2000.

[6] H. Deshpande, R. Singh, and U. Nam: "Classification of music signals in the visual domain," *Proceedings of the the COST-G6 Conference on Digital Audio Effects*, 2001.

[7] C. Ding, T. Li, W. Peng, and H. Park. "Orthogonal nonnegative matrix tri-factorizations for clustering," *SIGKDD*, 2006.

[8] B. Fields, K. Jacobson, C. Rhodes, and M. Casey: "Social Playlists and Bottleneck Measurements: Exploiting Musician Social Graphs Using Content-Based Dissimilarity and Pairwise Maximum Flow Values", *Proceedings of ISMIR*, 2008.

[9] Z. Guo, S. Zhu, Y. Chi, Z. Zhang, and Y. Gong. "A latent topic model for linked documents," *Proceedings of SIGIR*, 2009.

[10] T. Hofmann: "Probabilistic latent semantic indexing," *SIGIR*, 1999.

[11] K. Jacobson, B. Fields, and M. Sandler: "Using Audio Analysis and Network Structure to Identify Communities of On-Line Social Networks of Artists," *Proceedings of ISMIR*, 2008.

[12] Y. Liu, Y. Wang, A. Shenoy, W. Tsai, and L. Cai: "Clustering music recordings by their keys," *ISMIR*, 2009.

[13] I. Karydis, A. Nanopoulos, H. Gabriel, and M. Spiliopoulou: "Tag-aware spectral clustering of music items," *ISMIR*, pp. 159–164, 2009.

[14] P. Knees, T. Pohle, M. Schedl, D. Schnitzer, K. Seyerlehner, and G. Widmer: "Augmenting text-based music retrieval with audio similarity," *ISMIR*, 2009. (Tutorial)

[15] P. Lamere and E. Pampalk: "Social tags and music information Retrieval," *ISMIR*, 2008.

[16] D. Lee and H. Seung: "Algorithms for non-negative matrix factorization," *NIPS*, 2001.

[17] P. Lehwark, S. Risi, and A. Ultsch: "Visualization and Clustering of Tagged Music Data," in *GFKL*, 2007.

[18] M. Levy and M. Sandler: "Learning latent semantic models for music from social tags" *Journal of New Music Research*, 37:137–150, 2008.

[19] E. Pampalk, A. Flexer, and G. Widmer: "Improvements of audio-based music similarity and genre classification," *ISMIR*, 2005.

[20] D. Pye: "Content-based methods for managing electronic music," *ISCASSP*, 2000.

[21] L. Rabiner and B. Juang: *Fundamentals of Speech Recognition*, Prentice-Hall, NJ, 1993.

[22] A. Strehl and J. Ghosh: "Clustering ensembles - a knowledge reuse framework for combining multiple partitions," *Journal of Machine Learning Research*, 3:583-617, 2003.

[23] P. Symeonidis, M. Ruxanda, A. Nanopoulos, and Y. Manolopoulos: "Ternary semantic analysis of social tags for personalized music Recommendation," *ISMIR*, 2008.

[24] D. Turnbull, L. Barrington, M. Yazdani, and G. Lanckriet: "Combining audio content and social context for semantic music discovery," *SIGIR*, 2009.

[25] G. Tzanetakis and P. Cook: "Musical Genre Classification of Audio Signals," *IEEE Transactions on Speech and Audio Processing*, 10:5, 2002.

[26] D. Wang, T. Li, and M. Ogihara: "Are tags better than audio features? The effect of Joint use of tags and audio content features for artistic style clustering," *ISMIR*, 2010.

[27] F. Wang, X. Wang, B. Shao, T. Li, and M. Ogihara: "Tag integrated multi-label music style classification with hypergraph," in *ISMIR*, pp. 363–368, 2008.

[28] J. Shi and J. Malik: "Normalized Cuts and Image Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(8):888–905, 2002.