Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 91 (2016) 959 – 965
Procedia Computer Science

Information Technology and Quantitative Management (ITQM 2016)

Genre-based Link prediction in Bipartite Graph for Music Recommendation

Daozhen ZHAO[a,b] , Lingling ZHANG[a,b],Weiqi ZHAO[c,*]

[a] *School of Economics and Management, University of Chinese Academy of Sciences, Beijing, 100190, China*
[b] *Key Laboratory of Big Data Mining and Knowledge Management，Chinese Academy of Sciences，Beijing, 100190, China*
[c] *National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, BeiJing, 100190, China*

Abstract

The bipartite graph method of link prediction can apply in many fields of recommendations, with the nodes (users and items) and links (interactions between users and items). However, that links cannot represent the users' dual preferences (like and dislike). Some researchers improved that limits by complex number representations, but still not consider the influence of users' similarity recommendation performance. Here, we proposed an improved method to cope with this deficiency, build the relational dualities by complex number representations and computing the users' similarity by genres weight relations. In experiments with the Xiami.com music dataset, the proposed music genre weight-based music recommendation model (MGW) performances better than the CORLP method.

Peer-review under responsibility of the Organizing Committee of ITQM 2016

Keywords: Music recommendation; Link prediction; Music genres

1. Introduction

The study of music recommendation has been become the focus of academic circles and the business community. Music recommendation save a lot of time for user to find their own preferences from the massive songs. At the same time, the music recommend subscription service brought huge profits to merchants [1]. In China, with the development of internet and intelligent mobile phone industry, reported music application has been become the longest time used application for users in all entertainment application. At the same time, more and more domestic music web site (such as douban.fm, music.163.com) recommend music to users through a variety of recommended technology like private station and recommendation playlist. User can hear the song they will like but unable to find from the vast amounts of songs on the internet. Among the numerous

* Corresponding author. Tel.: +8618500240788.
E-mail address: zdz_86@126.com.

Peer-review under responsibility of the Organizing Committee of ITQM 2016
doi:10.1016/j.procs.2016.07.121

recommendation algorithms, link prediction research has raised wide attention in recent years and involved in many fields such as computer science, biology and physic, et al [2]. The most prominent contribution of link prediction is to find potential interested products for users in electronic commerce [3]. Also in the field of social networks, link prediction can help user to find new friends they may know or want to know [4]. Link prediction takes the number of common neighborhoods into account, which influences on user preferences in different path of lengths. Zhihao Wu [5] has improved the limits of collaborative filtering algorithms which only compared two users for each other. Bipartite graph as a kind of representation in the link prediction, can simplify the complex network as binary link relations like user-item link relations. However, the traditional bipartite graph can only represent the unitary of the links between the nodes but cannot represent the dual links [6]. For example, in traditional bipartite graph algorithm, we only know the links between the user and songs, but don not the links is affirmative or negative. Feng Xie [7] modeling the relational dualities using complex number which fixed the above problem. They represent complex number $i, -i$ as affirmative and negative preference of user, and from their experiment, the complex number representation link prediction method (CORLP) achieved the best performance in the three power of the adjacency matrix. The CORLP consider the dual user's preference to improve the traditional bipartite graph method, but still not consider the similarity of the users in the real music community.

In this paper, we proposed music genre weight-based music recommendation model (MGW) based on CORLP method. We extract the genre preference weight features of users to improve the similarity of users combined with the music genre classification from the music domain knowledge. In the experiment part, we evaluate the validity and efficiency of the MGW model at last.

2. Model

To similar with the most offline music recommendations, our model generate a recommendation list to users that according to extract music features [8]. Figure 1 shows the process of music genre weight-based music recommendation model (MGW) model. The core idea of this process is the combination of the music genre preference weight feature of users and CORLP method, improve the similarity of users. The CORLP method can describe both positive and negative music preference of the user, improved the traditional link prediction which only can represent the unitary music preference of the user. For example, user not only have their favorite music list but also have blacklist, all of these factors should be considered in the real music community.

The music genre preference weight feature regard as a user attribute can describe the degree of music preference from different users to the same music genre, which can improve the affection of music recommendation. For example, two users may both like the rock genre, but the similarity of two users will different if their preference degree of the rock genre is not the same.

To construct the model, we proposed the process including the extraction of music genre preference feature from user, discovery the association between music genre preference feature and CORLP method. Given the music preference and music genre preference weight from user, the model will generate a music recommendation list.

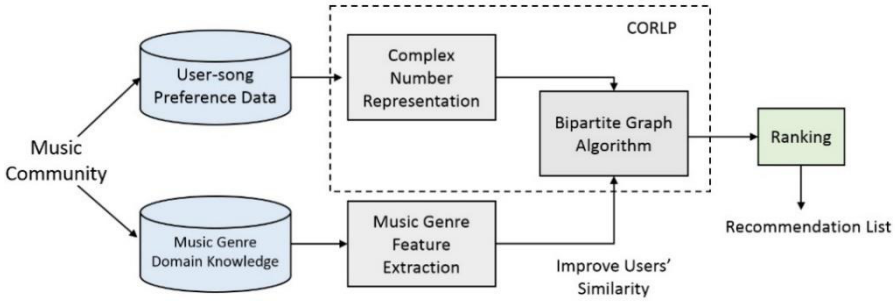

Fig 1. The framework of the MGW model

2.1. CORLP method

The CORLP (Complex Representation-based Link Prediction) method models relational dualities using complex number instead of unitary in the past between users and items. This method achieved significant performance compared with state-of-the-art methods.

Input data are modelled as a directed graph $G = (V, E, w)$ where the set of nodes V consists of all users U and items I, E is set of links that represent various relations, and w includes all of the links' weight. Then, set $w_{like} = i$, $w_{dislike} = -i$, $w_{similar} = 1$ and $w_{dissimilar} = -1$, where i is imaginary number whose square is negative one. Hence, we can convert directed graph $G = (V, E, w)$ to $G' = (V', E')$ which is an unweighted and undirected network. The adjacency matrix A of the user-item graph G' has the following property:

$$A(x,y) = \begin{cases} 1, & if\ \ x\ similar\ y \\ -1, & if\ \ x\ dissmilar\ y \\ i, & if\ \ x\ likes\ y\ or\ y\ dislikes\ x \\ -i, & if\ \ x\ dislikes\ y\ or\ y\ likes\ x \\ 0, & if\ (x,y) \notin E \end{cases} \tag{1}$$

Where $A(x, y)$ is the value in row x and column y of matrix A. Matrix A can be denoted as $\begin{bmatrix} A_{UU} & A_{UI} \\ A_{IU} & A_{II} \end{bmatrix}$, where A_{UU}, A_{II} are the similarity matrix of users and items, A_{UI}, A_{IU} are the preference matrices of users. Ignoring the initial relations between users and items, the G' is a bipartite graph and the adjacency matrix A can be simplified to $\begin{bmatrix} 0 & A_{UI} \\ -A_{UI}^T & 0 \end{bmatrix}$, and we can further convert A to $\begin{bmatrix} 0 & iB \\ -iB^T & 0 \end{bmatrix}$,where B is a real matrix.Based on path counting in the unweighted and undirected networks, the path counting for paths of length k can be derived similarly using A^k as following equation 2:

$$A^k = \begin{cases} \begin{bmatrix} (BB^T)^n & 0 \\ 0 & (B^TB)^n \end{bmatrix} & ,k = 2n \\ i \cdot \begin{bmatrix} 0 & (BB^T)^n B \\ -(B^TB)^n B^T & 0 \end{bmatrix}, & k = 2n+1 \end{cases} \tag{2}$$

Hence, any sum of the powers of the adjacency matrix A can be split into even and odd part. Furthermore,

we can using the sum of odd parts for recommendation combined with the matrix exponential.

2.2. Feature Extraction

Music has a lot feature, which include melody, rhythm, tempo, mode, key, harmony, dynamics and tone-color. [9][10]. Based on the users' preferences, music genre can distinguish the users' favour of music. The music genre preference weight of users can reflect the similarity of users, so we regard the music preference weight as the user features to measure the differences between the users. Then we improve the accuracy of music recommendation to users with modified similarity of the users.

In the real music community, each user has his favourite list of music artist, and every artist has at least one or more music genres such as Pop or Rock genre. Table.1 Shows twenty-three music genre from the real music community (www.xiami.com) which put all the music into twenty-three categories.

Table 1. Music genres from xiami.com

No.	Genres	No.	Genres
1	Pop	13	New Age
2	Rock	14	Stage & Screen & Entertainment
3	Folk	15	Country
4	Electronic	16	Reggae
5	R&B	17	Classical
6	Jazz	18	Singer-Songwriter
7	Easy Listening	19	Latin
8	Hip Hop	20	Chinese Characteristic
9	ACG	21	Experimental
10	Blues	22	Children
11	Metal	23	Audio Book
12	World Music		

2.3. Recommendation

The improved CORLP method is consider the influence of music preference weight of users to the similarity of users in a real situation. We modify the second power of the adjacency matrix A combined with music genre domain knowledge, and recommend song to users that have not heard before.

According to the section 3.2, the second power of the adjacency matrix A can be represented as following equation 6 based on CORLP method:

$$A^2 = \begin{bmatrix} BB^T & 0 \\ 0 & B^TB \end{bmatrix} \tag{3}$$

Where B is real matrix. After we calculated the similarity of users again combined with music genre, the top-left matrix BB can be converted to matrix C. we can convert A^2 to $(A^2)'$, as following as:

$$(A^2)' = \begin{bmatrix} C & 0 \\ 0 & B^TB \end{bmatrix} \tag{4}$$

Where C is a modified matrix which consider the music genre preference weight of the users. We can infer the modified $(A^3)'$ as:

$$(A^3)' = \begin{bmatrix} C & 0 \\ 0 & B^T B \end{bmatrix} \cdot A = \begin{bmatrix} 0 & -iC \cdot B \\ -iB^T B B^T & 0 \end{bmatrix} \tag{5}$$

Where the $-iC \cdot B$ part of the $(A^3)'$ matrix is music preference matrix of users. That part considers the relations between users using A^3 which represented as path counting for paths of length 3. Therefore, we can compare $(A^3)'$ with$A$, analysing and resulting the music recommendation list to the users.

3. Experiment

To evaluate the effectiveness of our proposed music recommendation approach, we use the dataset contains 4,879 ratings including like or dislike by 20 users on 4,363 songs on real-world datasets.

Users' music preferences and music genres were collected during the four-month period from November 15st, 2015 through February 31st. The 20 Users were randomly chosen from an artist circle of fans, which have more than 2,000 users. The 23 music genres feature were collected through the existing classification of music genres from the professional music recommendation web site (xiami.com). The ratings data was dual music preferences which the users like the songs or dislike it.

In our data set, ratings divided into two subsets, training and test. The training subset was randomly chosen 60% of the data set, and the test subset was randomly chosen 40% of the data set.

The recommendation performance is measured by Top-N method, which are defined as:

$$Precision = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |T(u)|} \tag{6}$$

$$Recall = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |R(u)|} \tag{7}$$

$$F = \frac{Precision}{Recall} \tag{8}$$

Where $R(u)$ is the number of the recommended songs, $T(u)$ is number of the songs that user liked. Where F is a synthetically index to evaluate the Top-N recommended method.

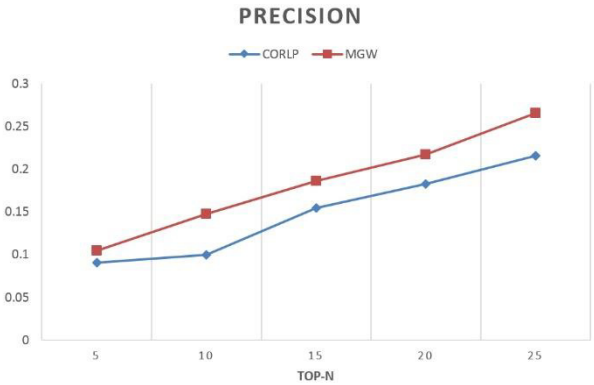

Fig 2. The performance of the accuracy

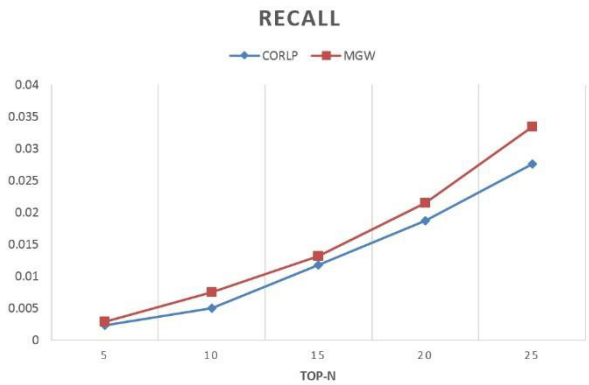

Fig 3. The performance of the recall

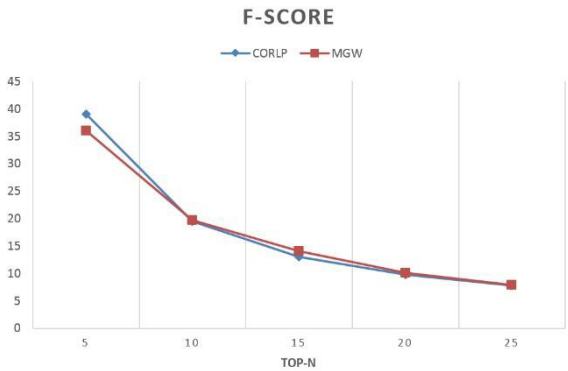

Fig 4. The performance of the F-score

The comparison between our proposed MGW algorithm and MAG algorithm are shown in Fig.2, 3 and 4. Fig.2 and fig.3 show that the experimental results of *accuracy* and *recall*, as the number of items recommended for users ranges from 5 to 25 for the xiami.com dataset. The two algorithms share the property that the *accuracy* and *recall* increase as the growth of the number of recommended items, and the MGW method achieved a higher *accuracy* and *recall* than the CORLP method. The F-score is one of the aggregative indicator to verify the validity of the Top-N method, in fig4, MGW performances better than the CORLP method only except for the Top-5.

4. Conclusion

In this paper, we presented a MGW model to recommend music based on the link prediction. The core of our proposed approach is to construct the recommendation model based on the CORLP method, and combined with the similarity of users by extracting the genre weight feature of users' preference. In the experiment part, 20 user is too small to disturb the stability of the recommendation results. In the future, we plan to collect more preferences data of the users to overcome the sparsely problems of data.

Acknowledgments

We gratefully thank the anonymous reviewers for their constructive feedback. We appreciated Professor Lingling Zhang in this paper. This work was partially supported by National Natural Science Foundation of China (Grant No.71471169、71071151), and the Major Research plan of the National Natural Science Foundation of China (Grant No. 91546201). This work was also supported by the Open Project of Key Laboratory of Big Data Mining and Knowledge Management, Chinese Academy of Sciences.

References

[1] Wang, X., Rosenblum, D., & Wang, Y. Context-aware mobile music recommendation for daily activities. *In Proceedings of ACM international conference on multimedia*, 2012; 20: 99-108.

[2] Z. Huang, X. Li, and H. Chen, "Link prediction approach to collaborative filtering," *in Proceedings of the ACM/IEEE-CS joint conference on Digital libraries*, 2005; 5: 141-142.

[3] T. Zhou, J. Ren, M. Medo, and Y.-C. Zhang, "Bipartite network projection and personal recommendation," *Physical Review E*, 2007;**76**:46-115.

[4] L. M. Aiello, A. Barrat, R. Schifanella, C. Cattuto, B. Markines, and F. Menczer, "*Friendship prediction and homophily in social media*", ACM Transactions on the Web (TWEB), 2012;**6**:9.

[5] Zhihao Wu, Youfang Lin, Jing Wang, Steve Gregory. Link prediction with node clustering coefficient. *Physica A: Statistical Mechanics and its Applications*, 2016; **452**:1-8.

[6] Dias, R., & Fonseca, M. J. Improving music recommendation in session-based collaborative filtering by using temporal context. *In Proceedings of international conference on tools with artificial intelligence*, 2013; 32: 783-788.

[7] Feng Xie, Zhen Chen and Jiaxing Shang. A link prediction approach for item recommendation with complex number. *Knowledge-Based Systems*, 2015; **81**:148-158.

[8] P. Lops, M. de Gemmis, G. Semeraro, *Recommender Systems Handbook*, 3rd ed. New York: Springer Verlag, 2011.

[9] D.N. Sotiropoulos, A.S. Lampropoulos, G.A. Tsihrintzis, Musiper: a system for modeling music similarity perception based on objective feature subset selection, *User Modeling and User-Adapted Interaction*, 2008; **4**: 315-348.

[10] Man-Kwan Shan, Fang-Fei Kuo, Meng-Fen Chiang, Suh-Yin Lee, Emotion-based music recommendation by affinity discovery from film music.*Expert Systems with Applications*, 2009; **36**: 7666-7674.