A Study on Combinative Value Creation in Songs Selection

Hiroko Shoji[1], Jun Okawa[2], Ken Kaji[2], and Ogino Akihiro[3]

[1] Chuo University, Japan
[2] Graduate School, Chuo University, Japan
[3] Kyoto Sangyo University, Japan
hiroko@indsys.chuo-u.ac.jp,
kxexn69@gmail.com,
ogino@cc.kyoto-su.ac.jp

Abstract. Recently, advances in information and communications technology have allowed us to easily download our favorite songs from the Internet. A song in general is more often played in sequence with other various ones than listened separately. The evolution of devices, however, has caused an increased number of portable songs and thus frequent difficulties in nicely combining multiple songs from a flood of songs to make a satisfactory playlist. There are many existing research works on songs search and retrieval, such as a songs each system using affective words and a songs recommendation system in consideration for the user's preference. These existing researches, however, are intended for "selecting a single song suited to the user's image", and never takes into consideration a combination of multiple songs. Therefore, it is difficult that existing systems automatically generate a desired playlist.

Keywords: combination value, playlist, recommendation, onomatopoeia.

1 Introduction

1.1 Combination Value

In the modern information society, we can access information with little to no difficulty. What is difficult, however, is creating beneficial combinations from this vast quantity of available information. Constantly deliberating about the most advantageous combinations would involve strenuous and time-consuming effort. To take a simple example of this kind of combination, every day we combine, out of the clothes we own, a set of clothing to wear. Clothes are not typically worn independently; they are combined with other clothes or accessories such as shoes or bags. The overall value of a set of clothes depends on how well they are coordinated, in addition to the individual value of each item of clothing. Likewise, the overall value of information changes depending on how the individual pieces of information are combined. Combination value is defined as the difference (should there be one) between the sum of

M. Kurosu (Ed.): Human-Computer Interaction, Part V, HCII 2013, LNCS 8008, pp. 372–380, 2013.

the individual values of all pieces of information and the overall value of the information once combined. Combination value may be expressed, using "individual value" and "overall value", as follows:

$$\text{Overall value} = \sum \text{individual value} \pm \text{combination value} \quad (1)$$

In our study, we focused on combination value, as expressed by Equation 1, and narrowed our scope to the subject of music. People do not typically listen to a single piece of music, but rather to a combination of musical pieces, demonstrating their interest in the combination value of the respective pieces. With respect to music, we intend to enhance overall value by increasing the combination value, and propose, in this paper, a method for doing so.

1.2 Background and Purpose of the Study

A playlist created by a user of music technology tends to reflect that user's mood, as it commonly takes them a significant amount of time to create a combination of music to their liking. In other words, such a playlist typically has a high combination value. However, given that advances in technology have increased the number of songs available to users, carefully selecting playlists that reflect user mood from such a large selection of music may involve significantly more effort on the user's behalf. It might be wiser to use the random play function in order to save time. However, the random play function simply plays music from a designated selection of songs, and does not consider the positional relationship of these songs, or the music's overall flow or concept. One of the tools available on iTunes for creating playlists is an automatic playlist generating function (Smart Playlist). However, this function recommends musical pieces from the user's play record, typically representing a large selection of songs, and can produce a playlist that is not in accord with the user's mood. Studies have already been conducted on music recommendation approaches that consider each piece of music individually: there are music retrieval systems that use sensitive indicators, and music recommendation systems that consider user preferences. However, these existing systems do not consider song combinations, making it difficult for them to automatically generate playlists that meet user expectations. Therefore, our study analyzes how to combine songs based on user mood, and proposes a method for generating playlists that eventually become more similar to those manually produced by the user.

2 How Users Combine Songs

2.1 Combination Analysis Experiment

Questionnaire on Music Impressions. As a preliminary step in the experiment, a questionnaire survey on music impressions was conducted. A set of questions was given to 14 subjects, asking them to answer, using a checklist, how they felt about each piece of music. In our study, onomatopoeias were used as "impression words": words that convey the impression a subject felt or received when they listened to a particular piece of music. Music is made of sound, and thus we believe that onomatopoeias were the most appropriate way to express how a person felt about the music. A total of 52 onomatopoeias were used, including both the onomatopoeias used by Ogino et al. in their recommendation system[1], and onomatopoeias selected from a Japanese onomatopoeia dictionary[2]. The music selected for our study consisted of 135 musical pieces from various genres. One or more impression words were assigned to each of the 135 pieces of music, based on the results of the questionnaire survey.

2.2 Music Combination Experiment

An experiment was conducted to examine what combinations users create. This experiment involved 10 subjects who had each listened to the 135 pieces used in the questionnaire, and was conducted according to the following procedure:

1. Each subject selected 10 or more songs from those evaluated with the word "shimi-jimi (feelingly)" to describe the listeners' impression of the music.
2. Each subject prepared a playlist of 10 songs of their own choice from the songs selected in Step 1.
3. Each subject repeated Steps 1 and 2 for songs evaluated with "wai-wai (excitedly or uproariously)"

Subjects were allowed to listen to the selected songs (hooks only) as many times as they liked during the course of the experiment.

2.3 Experimental Results

Our results indicate that music combinations can largely be classified into three patterns: conceptual, dispersed and intermediate. Each pattern is explained below:

Conceptual Pattern. Some of the songs that subjects included in combinations classified as having a conceptual pattern are shown in Table 1.

This table shows that with respect to songs selected from the group of songs originally evaluated with the word "shimi-jimi", subjects tended to combine songs described by the three onomatopoeias shown in the table. Likewise, a similar conceptual pattern was shown with respect to combinations created from songs that were originally evaluated with the word "wai-wai". Therefore, we conclude that songs combined by users emphasizing conceptual music combinations tend to be uniform in terms of the impressions generated by the respective songs, with user mood typically reflected throughout the entire combination.

Dispersed Pattern. Dispersed pattern combinations are composed of songs that tend to generate a variety of often contradictory impressions, in contrast to the more uniform impressions generated by songs in the conceptual pattern combinations. Table 2 shows examples of songs included in combinations classified as having a dispersed pattern, and that were selected from the group of songs originally evaluated with the word "shimi-jimi".

This table shows that songs frequently evaluated as "wai-wai (excitingly)" and "yuttari (leisurely)" tended to be included in combinations classified as having a dispersed pattern. We infer that a combination of diverse musical pieces allows the user to remain interested, and combinations following the dispersed pattern typically have a relatively balanced mixture of songs associated with a variety of impressions.

Table 1. Conceptual pattern combinations

Song	Shimi-jimi (feelingly)	Jin-wari (gradually)	Shin-miri (quietly)
Itsuka Dokokade (Sometime Somewhere)	4	4	5
Mirai Yosozu II (Future Forecast Diagram II) – 2007 Version	3	3	3
I for You	2	3	3
Shiawasena Ketsumatsu (Happy Ending)	3	3	3
Ienaiyo (Can't Say)	5	6	3
Sobaniiruyo (I'll Be There for You)	3	2	5
Single Bed	4	2	4
Mou Koinante Shinai (Won't Fall in Love Again)	5	4	7
Invitation	1	1	1
Heya to Y shatsu to watashi (Room, Y-shirt and Me)	1	3	2

Table 2. Dispersed pattern combinations

Song	Shimi-jimi (feelingly)	Yuttari (leisurely)	Wai-wai (excitingly)
Love Somebody	1	1	3
KNOCKIN' ON YOUR DOOR	1	0	2
Booty Music	1	2	4
Bye Bye (So So Def Remix)	1	3	0
Lay Up	2	3	0
One Last Try	2	1	1
You Gotta Be	2	2	1
If I Ain't Got You	4	4	0
Take Me Back	2	2	0
Take A Bow	5	3	0

Intermediate Pattern. Combinations classified as having this pattern principally consist of pieces that conform to the combination’s core concept impression; however, a smaller number of pieces associated with a wider variety of impressions are also included. Of the song combinations generated during this experiment, the largest number fell into the intermediate pattern category. Table 3 shows examples of songs included in intermediate pattern combinations that were selected from the group of songs originally evaluated with the word "shimi-jimi". The table shows that songs associated with onomatopoeias similar to "shimi-jimi (feelingly)" and "shin-miri (quietly)" were included in the combinations, and that songs associated with quite different onomatopoeias like "nori-nori (willingly)" were selected as well. The choice of songs associated with impressions dissimilar to "shimi-jimi" varied depending on the subject. It was observed that subjects frequently included in their combinations songs described with "wai-wai" or "nori-nori" (which represent impressions nearly opposite to that represented by "shimi-jimi"); however, the balance between uniformity and diversity varied from person to person. Therefore, any method for reproducing intermediate pattern combinations will need to be able to appropriately balance these two factors before it can be included in our music combination system.

Table 3. Intermediate pattern combinations

Song	Shimi-jimi (feelingly)	Shin-miri (quietly)	Nori-nori (willingly)
1_3 No Junjo Na Kanjo (Pure Sentiment of 1_3)	1	1	3
Fragile	4	2	1
M	5	5	0
Mirai Yosozu II (Future Forecast Diagram II) - Version 2007	3	3	0
Zoo	3	4	0
Mou Koinante Shinai (Won't Fall in Love Again)	5	7	0
Watarase Bashi (Watarase Bridge)	4	3	0
Single Bed	4	4	0
Ienaiyo (Can't Say)	5	3	0
Aitai (Wanna See You)	5	5	0

3 Implementation of the Proposed Methods

3.1 Definition of Distance

In order to analyze the combinations created by the subjects, and simulate our proposed methods, each musical piece is assumed to be a point set of the 52nd dimension. The distance between songs is calculated to determine the degree of similarity between them. Euclidean distance is used for distance in this calculation. The distance between a piece i and another piece j, L_{ij} (i = 1,2,3…135, j = 1,2,3…135), is calculated as follows (using the evaluation x_{ik} of the onomatopoeia k of the song i, and the evaluation x_{jk} of the onomatopoeia k of the song j):

$$L_{ij} = \sqrt{\sum_{k=1}^{52} (x_{ik} - x_{jk})^2} \tag{2}$$

3.2 In Our Study, It Is Understood That Songs Closer to Each Other Have Greater Similarity. Proposal of a Playlist Generation Method

Proposed Method. In Section 2 we reported the details of an experiment on how users combine music. The results revealed that there were users who created combinations with a strong, conceptual pattern, and other users who created combinations of slight diversity with weak conceptual patterns. Section 3 proposes three methods (using the concept of distance defined in 3.1) that allow us to obtain results similar to the combinations created by the subjects. Each method aims at generating combinations that reproduce one of the three patterns identified by the experiment.

- Minimum distance method: This method selects songs closest in distance to those initially selected by users. It is expected to produce combinations similar to the conceptual variety.
- Maximum distance method: This method selects songs farthest in distance from those initially selected by users, and is conceptually opposite to the minimum distance method. It is expected to produce combinations similar to the dispersed variety.
- Hub selection method + minimum distance method (Hub selection method): In this combination of methods, two or more hubs are selected, and then a few songs are selected around the hubs, based on the minimum distance method. In our study, a song of higher centrality is called a "hub". Matrix A (obtained by deleting all songs, from the 135, which have a mutual distance of 8.5 or less) is used to calculate information centrality (C_{inf}). Songs of greater centrality have an increased probability of selection. Information centrality is a centrality indicator that considers the shortest route between vertices contained in the network or the length of each route.

From matrix A = a_{ij}, we then determine matrix B = b_{ij} as diagonal component b_{ii}, which is equal to the sum of the values of sides connected to vertex i plus one:

$$b_{ij} = \begin{cases} 1 & \text{when vertex i and vertex j do not lie side by side} \\ 1 - a_{ij} & \text{when vertex i and vertex j lie side by side} \end{cases}$$

Then, using Matrix B, Cinf is calculated as follows:

$$C_{inf} = {}^{n}/_{(nT + S_T - 2R)} \tag{3}$$

where, the inverse of matrix B having been determined, the vector of the diagonal component of the obtained inverse matrix is T, the trace of the inverse matrix is S_T, and the row sum of any given inverse matrix is *R[3]*. In this study, three hubs (representing three songs) are chosen, and then two more pieces close to the hubs are selected. The hub selection method is expected to produce combinations similar to the intermediate variety.

3.3 Comparative Evaluation

This paragraph details a comparative evaluation of proposed methods, using the results of the combinations created by the subjects. The differences in average distance between the playlists created by the subjects and the corresponding playlists generated by the proposed methods are shown in Table 4. With respect to the minimum distance method, the minimum average distance was recorded for subject G. With the songs initially selected by subject G as respective "centers", the playlist of G, and the corresponding playlist produced by the minimum distance method, are presented graphically in Fig. 1.

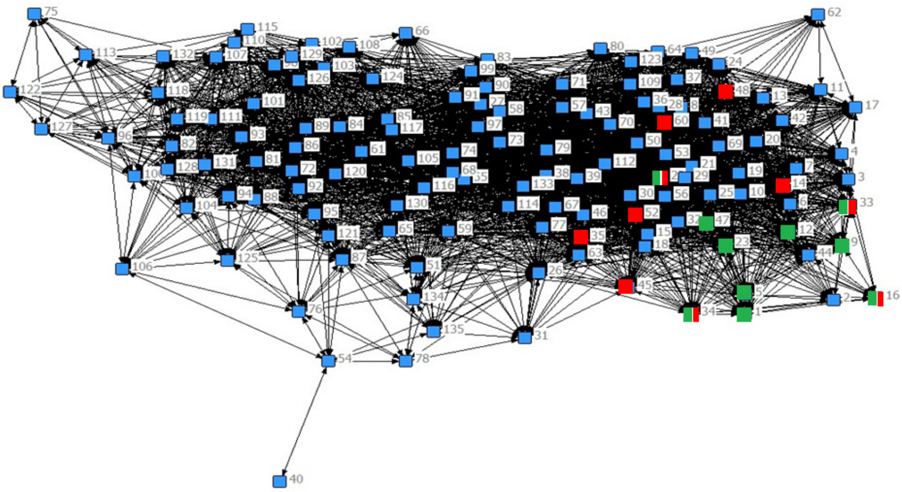

Fig. 1. Comparison between subject G's playlist and the corresponding playlist generated by the minimum distance method

In this figure, red nodes represent the combination created by G, while green nodes represent songs selected by the minimum distance method. It can be seen that the selected songs are typically close to those initially selected by G.

Table 4. Differences in average distance between the subjects' results and those of our proposed methods

Subject	**Minimum distance method**	**Maximum distance method**	**Hub selection method**
A	-2.322	3.992	-1.776
B	-1.97	2.325	0.239
C	-2.189	2.681	-0.702
D	-2.03	3.048	-1.67
E	-5.482	0.623	-5.102
F	-2.844	2.025	-1.069
G	-0.94	3.99	1.255
H	-1.401	1.757	-1.454
I	-1.794	3.266	-0.76
J	-2.203	2.875	-1.852

Table 4 shows that the maximum distance method produced results closest to those of subject E. In Fig. 2, songs included in subject E's playlist are represented in red, while those selected by the maximum distance method are represented in green. Fig. 2 also shows that the combination of songs selected by the subject is similar to the combination produced by the proposed maximum distance method. It is therefore reasonable to conclude that the proposed minimum and maximum distance methods are able to successfully reproduce the conceptual pattern and dispersed pattern combinations, respectively.

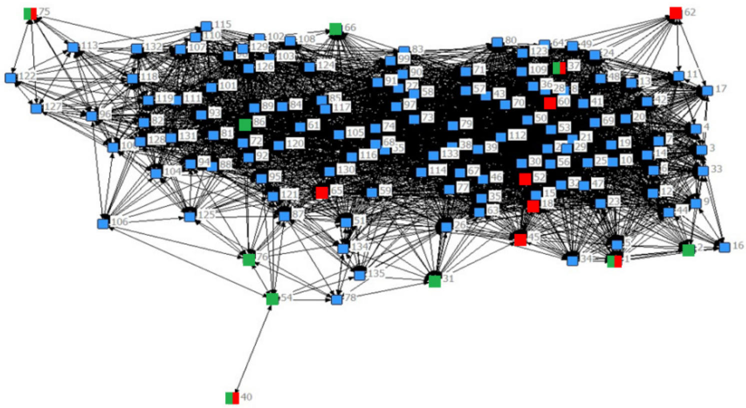

Fig. 2. Comparison between subject E's playlist and the corresponding playlist generated by the maximum distance method

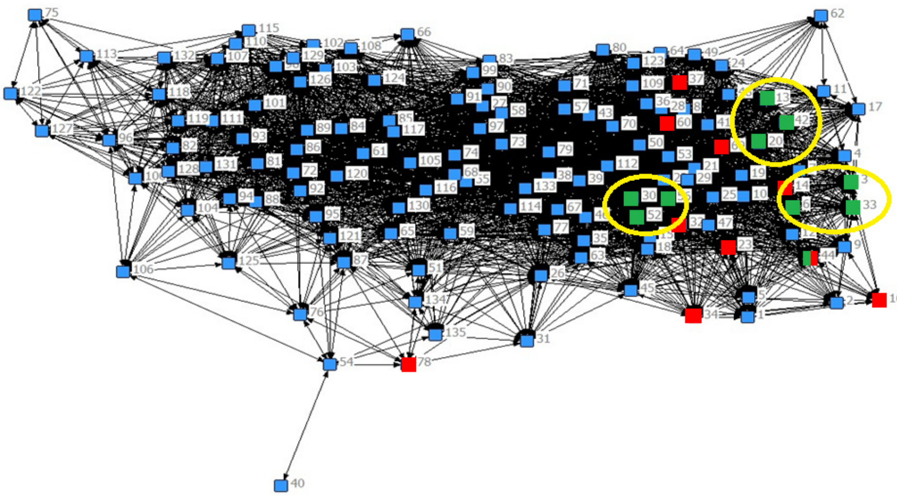

Fig. 3. Comparison between subject B's playlist and the corresponding playlist generated by the hub selection method

The hub selection method produced results similar to those of subjects B, C and I. Fig. 3 shows the songs selected by the hub selection method and those selected by subject B, in different colors. The circled areas show the hubs and the songs selected based on the hubs. It remains unclear which hubs will be selected by the hub selection method, and it can be reasonably concluded that differences in hub selection will lead to varying results.

4 Summary and Future Tasks

This study calculated distances between songs based on an evaluation method using impression words, and proposed a playlist generation method based on song distances. The experimental results could demonstrate that music combinations created by users could be categorized as conceptual pattern combinations (combinations of songs that generate similar impressions) and dispersed pattern combinations (combinations of songs that generate different impressions). Our study attempted to develop a means of creating music combinations with similar patterns using the concept of distance between musical pieces. It is concluded that the proposed minimum distance and maximum distance methods are capable of reproducing the conceptual pattern and diverse pattern combinations. On the other hand, although the hub selection method produced results similar to those of some subjects, this method is designed to stochastically select songs with high centrality, and there is a lack of clarity with respect to the number of hubs, and/or songs close to the hubs, the method selects. In addition, it is necessary to conduct simulations when songs farther from the hubs are selected. Therefore, a future task is to improve the hub selection method, such that it can flexibly reproduce combinations with patterns similar to those generated by users.

References

1. Yoshida, N., Ogino, M.: Proposal of a Music Selection Support System to Conform to User Image. Collection of Papers Prepared for the Japan Society of Kansei Engineering for 13th Convention vol. 13 (2011)
2. Ono, M.: Onomatopoeias and Mimetic Words 4500 Japanese Onomatopoeia Dictionary. Shogakukan Publishing (2007)
3. Suzuki, T.: Network Analysis. Data Science Learned with R, vol. 8. Kyoritsu Shuppan Co., Ltd (2008)
4. Okawa, H., Kaji, M., Nishimura, T., Shoji, Y., Ogino, A.: Experiment and Analysis Concerning the Creation of Music Combination Value. Collection of Papers Prepared for the 14th Convention of the Japan Society of Kansei Engineering (2012)