

# ランキング情報を利用した楽曲推薦システムの構築

檜山 裕史<sup>†</sup> 手塚 太郎<sup>††</sup> 木村 文則<sup>†††</sup> 前田 亮<sup>††</sup>

<sup>†</sup> 立命館大学大学院 理工学研究科 〒525-8577 滋賀県草津市野路東 1-1-1

<sup>††,†††</sup> 立命館大学 情報理工学部 〒525-8577 滋賀県草津市野路東 1-1-1

E-mail: <sup>†,†††</sup> {cm003046, fkimura}@is.ritsumeai.ac.jp, <sup>††</sup> {tezuka, amaeda}@media.ritsumeai.ac.jp

**あらまし** 楽曲推薦システムにおいて多く利用されている手法として協調フィルタリングがある。膨大な情報の中から個人の嗜好に合った情報を得る手法として有効だが、嗜好に合わない楽曲を推薦してしまう場合が多々あり、推薦精度の向上が大きな課題となっている。そこで本稿では、協調フィルタリングに加えてランキング情報を利用する手法を提案する。ランキング情報としてCD売上ランキングと楽曲再生ランキングを利用することで、ユーザの好みに似た楽曲に加えて意外性のある楽曲も推薦できるようになると考えられる。楽曲を推薦するシステムを構築しユーザ実験を行い、従来の手法と比較した結果、推薦件数が増えるにつれて、ランキング情報利用による推薦精度の向上が確認された。

**キーワード** 楽曲推薦, 協調フィルタリング, ランキング情報

## Music Recommendation System Using Ranking Information

Yuji KASHIYAMA<sup>†</sup>, Taro TEZUKA<sup>††</sup>, Fuminori KIMURA<sup>†††</sup>, and Akira MAEDA<sup>††</sup>

<sup>†</sup> Graduate School of Science and Engineering, Ritsumeikan University

1-1-1 Noji-Higashi, Kusatsu, Shiga 525-8577, Japan

<sup>††,†††</sup> College of Information Science and Engineering, Ritsumeikan University

1-1-1 Noji-Higashi, Kusatsu, Shiga 525-8577, Japan

E-mail: <sup>†,†††</sup> {cm003046, fkimura}@is.ritsumeai.ac.jp, <sup>††</sup> {tezuka, amaeda}@media.ritsumeai.ac.jp

**Abstract** Collaborative filtering is a technique used in many music recommendation system. It is useful technique to get information that a user really want from huge amount of information. But this technique sometimes recommends music which is not suitable for user's taste, thus it is important to improve precision. In this paper, we present the method which use ranking information with collaborative filtering. We use CD selling ranking and music playing ranking as ranking information. Using these ranking, we can recommend the music which matches the user's taste and unexpected it. We implemented the music recommendation system and conducted user experiments. Comparing with the existing technique, we could confirm improving precision using ranking information as the recommendation number increased.

**Keyword** Music recommendation, Collaborative filtering, Ranking information

### 1. はじめに

近年、iPodなどの携帯音楽プレーヤーやiTunesミュージックストアなどの音楽配信サービスの普及により、音楽を楽曲単位で保存・再生することが一般的に行われている。また、新しい楽曲も次々に登場し楽曲数は日々増加している。最近では楽曲の歌詞を検索するサイトなど、音楽を対象とするサービスも増えてきており、楽曲を対象とする様々な技術が注目されている。

現在ユーザが楽曲を探す手段として、Web上でGoogle等の検索エンジンを用いて楽曲やアーティスト

を検索する方法がある。また、テレビや雑誌等から情報を得る方法や、実際にCDショップへ行き楽曲を探すといった方法もある。しかし、自分の好みに合う楽曲に辿りつくまで時間がかかる場合や、好みに合わない楽曲に辿りついてしまうことが多々ある。素早く簡単に自分の好みの楽曲を取得できることが重要であり、楽曲を探す手段としては理想である。

好みの楽曲を得るための方法として、個人の嗜好を考慮した楽曲推薦システムがある。これまで、楽曲推薦についての研究は多くされており、様々な手法が提

案されている。どの研究においてもシステムとしての有効性は確認されているが、十分に満足いく推薦精度とはいえないものが多く、推薦精度の向上が大きな課題となっている。

そこで本研究では、情報推薦においてよく利用されている協調フィルタリングの手法に加えて、ランキング情報を利用することで推薦精度を向上させる研究を行った。協調フィルタリングによって類似度の高い楽曲リストから嗜好に合う楽曲を決定し、さらにその中でランキングを重視することで、より嗜好に合う楽曲を推薦できるのではないかと考えた。本研究の手法の有効性を確認するために、協調フィルタリング、ランキング情報利用、双方を統合した3つのシステムの推薦精度の比較を行い検証した。

以下の本稿の構成は、第2章では楽曲推薦システムの関連研究について述べる。第3章ではシステムの概要とその詳細について説明する。第4章ではシステムの評価実験、考察を行い、第5章でまとめを述べる。

## 2. 関連研究

これまでに、多くの楽曲推薦システムの研究が行われている。武内らの研究[1]では、個人の音楽嗜好データを OWL(Web Ontology Language)で表現し、OWL 構造を利用した類似度計算を行い、推薦楽曲を決定している。また、土方らの研究[2]では、テンポや音長などの音響的特徴と、楽曲に対する評価値から決定木を構築し、決定木と推薦楽曲の特徴量を比較し推薦するかどうかを決定している。この研究の問題点は、有効な音響的特徴を抽出するのが困難であり、必要な特徴量をうまく抽出できない場合がある。どちらの手法もシステムの有効性は確認されているが、十分に満足いく推薦精度とは言えず、推薦精度の向上が課題となっている。

2つの異なる推薦手法を連動する方法として黒瀬らの研究[3]がある。この研究では、協調フィルタリングと内容ベースフィルタリングの2つの手法を連動させることで双方の問題点を補い、推薦精度の向上を狙っている。連動させることによる有効性は確認されたが、双方の推薦手法が共に有効であるときに初めて有効に働くため、それぞれの手法の推薦精度が十分でないと有効に機能しない。

## 3. 提案手法

提案手法による楽曲推薦システムの概要図を図1に示す。ユーザが自分の好みの楽曲リスト（以下、利用者楽曲リストと表記）をシステムに入力すると、システムは楽曲リストを受け取り、ランキング情報利用手法・協調フィルタリング利用手法による推薦楽曲を決

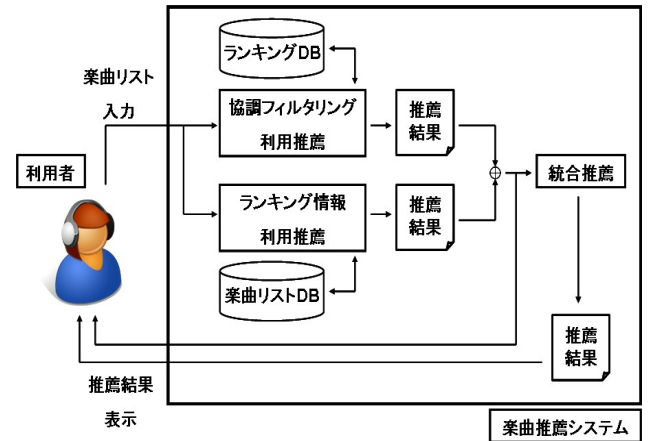


図1 本研究における楽曲推薦システムの概要

定し、推薦結果をユーザに返す。以下に、それぞれの手法の詳細について述べる。

### 3.1. ランキング情報利用推薦

ランキング情報とは、対象となるものをあるテーマにおいて順位付けをしたものであり、本研究では、楽曲を対象とする順位付けされた情報をランキング情報と定義する。ORICON STYLE[4]で公開されているCD売上ランキングと、mixi ミュージック[5]からアーティスト別の楽曲再生ランキングを取得し、推薦に使用する。

ランキング情報利用推薦では、利用者楽曲リストから特徴抽出を行い、抽出した特徴とランキング情報を利用して推薦する楽曲を決定する。

最初に、受け取った利用者楽曲リストに対して特徴抽出を行う。ここでは、アーティストの特徴・年代的特徴の2つの特徴抽出を行う。特徴抽出には、以下の式(1)、(2)を用いる。

$$\frac{F_{artist}(A_i)}{N} \geq \frac{1}{m} \quad (1)$$

$$\frac{F_{year}(Y_j)}{N} \geq \frac{1}{n} \quad (2)$$

利用者楽曲リスト中の楽曲数を  $N$ 、アーティストを  $A_i(i=1,2,\dots,m)$ 、発売年を  $Y_j(j=1,2,\dots,n)$  とし、 $F_{artist}(A_i)$ 、 $F_{year}(Y_j)$  はそれぞれ、Artist Frequency、Year Frequency で、利用者楽曲リスト中のアーティストが  $A_i$ 、発売年が  $Y_j$  である楽曲数を表す。(1)、(2)式が成り立つ  $A_i$ 、 $Y_j$  を特徴として抽出する。

次に、抽出した特徴を基に、楽曲のスコアを計算する。スコアの計算には以下の式(3)、(4)、(5)を用いる。

$$S_A(x) = \frac{P_{play}(x)F_{artist}(A_x)}{N} \quad (3)$$

$$S_Y(x) = \frac{P_{rank}(x)F_{year}(Y_x)}{N} \times \alpha_{imp} \quad (4)$$

$$S_R(x) = S_A(x) + S_Y(x) \quad (5)$$

ある楽曲  $x$  に対し,  $P_{play}(x)$ ,  $P_{rank}(x)$ はランキング情報から付与した楽曲自身もつポイントである. CD 売上情報からのポイント  $P_{rank}(x)$ は, ランキングの 1~30 位の楽曲に 30~1 ポイントを付与し楽曲ごとに加算したものである. また, 楽曲再生ランキングからのポイント  $P_{play}(x)$ の計算には式(6)を用いる.

$$P_{play}(x) = \frac{count(x)}{avg_{count}(A_x)} \times \alpha_{avg} \quad (6)$$

$count(x)$ は楽曲  $x$  の再生数,  $avg_{count}(A_x)$ は楽曲  $x$  のアーティスト  $A_x$ の 1 楽曲あたりの平均再生数,  $\alpha_{avg}$ はアーティスト  $A_x$ における  $P_{rank}$ の 1 楽曲あたりの平均値である. また,  $F_{artist}(A_x)$ ,  $F_{year}(Y_x)$ を係数として掛けることで, ランキングにおけるアーティスト・年代の重要度をスコアに反映している.  $\alpha_{imp}$ は利用者楽曲リストにあるアーティストと同じ楽曲のスコアを他より高くする補正係数で, 本研究では通常は  $\alpha_{imp}=1$  とし, 利用者楽曲リストにあるアーティストと同じ楽曲の場合には  $\alpha_{imp}=1.2$  とした.

抽出した特徴とランキング情報から楽曲のスコアを計算し, スコアの上位楽曲を推薦結果として出力する.

### 3.2. 協調フィルタリング利用推薦

協調フィルタリングとは, 嗜好が類似する他ユーザを発見し, そのユーザが好む情報を推薦してもらう手法である. 本研究ではこの考え方を応用し, 楽曲リスト間の類似度を算出し, 類似度の高い楽曲リストを用いて推薦する楽曲を決定する. 本研究では, 他ユーザの楽曲リストとして, mixi ミュージック [5]の機能の一つである「お気に入りリスト」から入手し利用した.

楽曲リスト間の類似度を求める前に, タイトル, アーティストからの類似ポイントを計算する. 類似ポイントの計算には式(7), (8)を用いる.

$$sim(t) = F_{title}(t) \log \left( \frac{N_{list}}{F_{list}(t)} \right) \quad (7)$$

$$sim(a) = F_{artist}(a) \log \left( \frac{N_{list}}{F_{list}(a)} \right) \quad (8)$$

利用者楽曲リスト中の楽曲  $x$  のタイトルを  $t$ , アーティストを  $a$  とするとき,  $t$  を含む他の楽曲リストの類似ポイントを  $sim(t)$ ,  $a$  を含む他の楽曲リストの類似ポイントを  $sim(a)$ とする.  $F_{title}(t)$ は Title Frequency で, 利用者楽曲リスト中の楽曲名が  $t$  である楽曲数を表す.

$F_{artist}(a)$ は 3.1 節のものと同じである.  $N_{list}$  は対象とする楽曲リストの総数,  $F_{list}(t)$ ,  $F_{list}(a)$ は List Frequency で  $t, a$  を含む楽曲リストの数を表す.

楽曲リスト  $L$  の楽曲集合を  $X$  としたとき, 利用者楽曲リストと  $L$  との間の類似度  $A(L)$ を以下の式(9)で求める.

$$A(L) = \sum_{t \in X} sim(t) + \sum_{a \in X} sim(a) \quad (9)$$

類似度  $A(L)$ の高い楽曲リスト  $L$  をスコア計算に用いる. 本研究では, 上位 50 件の楽曲リストを利用する. 楽曲  $x$  のスコア  $S_C(x)$ の計算には以下の式(10)を用いる.

$$S_C(x) = \sum_{\{L|x \in L\}} A(L) \quad (10)$$

$L$  は使用する上位 50 件の楽曲リストである. 式(10)によりスコアを計算し, スコアの上位楽曲を推薦結果として出力する.

### 3.3. 統合推薦

協調フィルタリングは自分の現在の好みに似た曲が多く推薦されるが, 意外性のある楽曲が得られないことが多い. ランキング情報は, 大量のユーザの好み が反映された情報源であり, 自分の現在の好みとは別 のであり, 意外性のある楽曲が得られる. 協調フィルタリングとランキング情報を統合することで, ユーザの好みに似た楽曲に加えて意外性のある楽曲も推薦できる ようになると考えられる.

2 つの手法による推薦結果を統合し, 最終的な推薦結果として出力する. それぞれの手法で異なるスコア付けが行われているので, そのままでは統合することは難しい. そこで, 同じスケールで比較するためにスコアの正規化を行う. 正規化を行うことで, 同じスケールでの比較ができるようになるため, スコアの統合が可能となる.

正規化後のスコア  $S_{Rf}(p)$ ,  $S_{Cf}(q)$ を, 以下の式(11), (12)で求める.

$$S_{Rf}(p) = S_R(p) \times \frac{n \times \alpha_{rel}}{\sum_{k=1}^{n \times \alpha_{rel}} S_R(k)} \quad (11)$$

$$S_{Cf}(q) = S_C(q) \times \frac{m}{\sum_{k=1}^m S_C(k)} \quad (12)$$

$S_R(p)$ ,  $S_C(q)$ は, それぞれの手法による推薦楽曲  $p, q$  のスコア,  $n, m$  は推薦楽曲数である.  $\alpha_{rel}$  はユーザの嗜好とランキングとの関連度を表す係数で, 式(13)のように定義する.

$$\alpha_{rel} = \frac{(\text{ランキングに登場したことがある楽曲数})}{(\text{楽曲リストの楽曲数}) + (\text{協調フィルタリングによる推薦楽曲数})} \quad (13)$$

楽曲リストの楽曲と協調フィルタリングによる推薦楽曲のうち、何曲がランキングに登場したのかを割合で表す。これによってランキング利用による結果をどの程度反映するかを決定する。この係数を用いることで、ランキングに相関がある利用者にはランキング利用による結果が重視され、そうでない利用者には協調フィルタリングによる結果が重視されるようになる。これによって、利用者の好みの傾向をスコアに反映させることができると考えられる。

すべての楽曲に対し正規化を行った後、式(14)を用いてスコアの統合を行い、楽曲  $x$  のスコアを決定する。

$$S_{final}(x) = S_{Rf}(x) + S_{Cf}(x) \quad (14)$$

統合後のスコア  $S_{final}(x)$  の上位楽曲を統合推薦の推薦結果として出力する。

#### 4. 評価実験

第3章で説明した推薦手法の評価実験を行う。実験を行うにあたり、74019件の楽曲リスト、1990年から2007年までのCD売上ランキング、2828アーティストの楽曲再生ランキングを用意し、実験に利用する。

##### 4.1. 実験手順

被験者14人に楽曲リストを作成してもらい、作成した利用者楽曲リストをシステムに入力し、得られた推薦結果に対して「1.嫌い-5.好き」の5段階で主観評価をしてもらう。この評価結果を用いて推薦精度を調べる。評価値が4以上の楽曲を好きな楽曲であると判断し、得られた推薦結果のうち、何曲が好きな楽曲であったかを推薦精度として算出する。

##### 4.2. 実験結果と考察

それぞれの手法によって得られた推薦結果のうち、スコアの高い上位  $n$  件 ( $Prec@n$ ) での推薦精度を算出し、各手法との比較を行う。表1, 2, 3にそれぞれの手法によって得られた推薦結果をまとめたものを示す。

2つの手法を単独で実行した場合、ランキング利用よりも協調フィルタリングのほうが良い結果であることがわかる。協調フィルタリングは、利用する情報が少ない段階では類似する情報をうまく発見できず、有効な推薦ができない。しかし今回の実験ではあらかじめ大量の楽曲リストを用意したことで、類似する楽曲リストを求める際により嗜好の近いものを取得できたため、協調フィルタリングが有効に機能したためと考えられる。また、いずれの結果においても約半数以上が好きな楽曲であるという評価が得られたので、システムとしてある程度有効に機能していると考えられる。

表1 推薦結果 (Prec@3)

推薦手法	評価値	曲数	割合	評価平均値
協調フィルタリング 利用推薦	1~2	1	4.2%	4.29
	3	4	16.6%	
	4~5	19	79.2%	
ランキング情報 利用推薦	1~2	7	19.5%	3.69
	3	8	22.2%	
	4~5	21	58.3%	
統合推薦	1~2	3	7.1%	4.10
	3	11	26.2%	
	4~5	28	66.7%	

表2 推薦結果 (Prec@5)

推薦手法	評価値	曲数	割合	評価平均値
協調フィルタリング 利用推薦	1~2	2	5.0%	4.05
	3	11	27.5%	
	4~5	27	67.5%	
ランキング情報 利用推薦	1~2	10	16.6%	3.68
	3	16	26.7%	
	4~5	34	56.7%	
統合推薦	1~2	10	14.3%	3.81
	3	19	27.1%	
	4~5	41	58.6%	

表3 推薦結果 (Prec@10)

推薦手法	評価値	曲数	割合	評価平均値
協調フィルタリング 利用推薦	1~2	14	17.5%	3.63
	3	23	28.7%	
	4~5	43	53.8%	
ランキング情報 利用推薦	1~2	29	24.2%	3.39
	3	33	27.5%	
	4~5	58	48.3%	
統合推薦	1~2	22	15.7%	3.70
	3	39	27.9%	
	4~5	79	56.4%	

次に、統合前と統合後の推薦精度の比較をする。協調フィルタリングと統合後を上位3件、5件で比較すると、協調フィルタリングのほうが統合後よりも推薦精度が高い結果となり、統合することによって逆に推薦精度が下がっている。この理由として、協調フィルタリングが有効に機能する状況だったため、統合することでランキング利用の精度が影響して推薦精度が下がったものと考えられる。しかし、協調フィルタリングは推薦楽曲が多くなるにつれて推薦精度の下落が大きいのにに対し、統合後では同じように推薦精度は下落するが、下落幅が小さく上位10件での推薦精度は協調フィルタリングを上回る結果が得られた。推薦件数が増えるにつれて好みに合わない楽曲が増え、全体とし

て推薦精度は減少していくが，そういった状況になるとランキング利用の結果が有効に反映されるようになり，推薦精度の下落を抑えることにつながったと考えられる．

## 5. まとめ

本研究では，個人の嗜好に合った楽曲を推薦するシステムに注目し，推薦精度の向上を目標に，協調フィルタリングによる推薦手法とランキング情報を利用する推薦手法を統合する新たな手法を提案した．システムの有効性を検証するために，被験者による評価実験を行い，推薦精度の比較を行った．その結果，ある程度のシステムの有効性が確認できた．しかし，今回得られた結果は十分に満足いく結果とは言いがたく，更なる推薦精度の向上が必要である．

今後の課題としては，売上ランキング・楽曲再生ランキング以外のランキング情報の利用が考えられる．Web 上には様々なランキングが存在し，その種類も多様である．ユーザによって使用するランキングを切り替えることで，より多くのユーザの嗜好に対応できるものと考えられる．また，現在のシステムは同じ楽曲リストからは同じ結果が出力されるといった，静的なシステムであるので，以前に評価した結果をフィードバックし推薦に利用するなど，動的なシステムにすることも今後検討していく．そして，推薦結果をどういった形で利用者に提示すべきか，といったインタフェース部分についても改良を検討していく．

## 文 献

- [1] 武内裕一，青野雅樹，“OWL を用いた音楽嗜好データの表現と音楽情報推薦への応用”，情報処理学会研究報告，DBS-143，pp.277-281，2005
- [2] 土方嘉徳，岩濱数宏，西田正吾，“決定木を用いた音楽情報フィルタリングシステムとその有効性の検証”，情報処理学会研究報告，DBS-133，pp.17-24，2004
- [3] 黒瀬崇弘，梶川嘉延，野村康雄，“視聴履歴を用いた楽曲推薦システム”，電子情報通信学会技術研究報告，HIP2003-84，pp.41-44，2003
- [4] “ランキング - ORICON STYLE”，ORICON STYLE，<http://www.oricon.co.jp/rank/>
- [5] “[mixi] ミュージック”，mixi( ミク シ イ )，<http://music.mixi.jp/>