

発見性を考慮した協調フィルタリングアルゴリズムに関する基礎検討

清水 拓也[†] 土方 嘉徳[†] 西田 正吾[†]

現在、商用 Web サイトで利用されている推薦システムの多くは、協調フィルタリングを用いている。従来の協調フィルタリングによる推薦方法では、推薦の精度を向上することに重点を置いていたが、すでに知っているアイテムが多く推薦されるという問題があった。この推薦は、精度を考えたときにはよい推薦と言えらるが、ユーザの満足度を考えたときには発見性が欠如していることから、必ずしも良い推薦とは言えない。本研究では、ユーザがどのようなアイテムを知っているかという観点で、ユーザ間、アイテム間の類似度を計算し、ユーザが知らないであろうアイテムを推測する。この方法と従来の嗜好に基づく協調フィルタリングを組み合わせることで、ユーザの知らない好みのアイテムを推薦しユーザの満足度を向上することを目指す。

A Basic Study on Discovery-oriented Algorithm for Collaborative Filtering

TAKUYA SHIMIZU,[†] YOSHINORI HIJIKATA[†] and SHOGO NISHIDA[†]

A number of recommender systems employed in commercial websites use collaborative filtering. The main goal of traditional techniques of collaborative filtering is improving the accuracy of the recommendation. Though, they have a problem that they include many items the user has already known. When we consider only the accuracy, these recommendations appear good. On the other hand, when we consider users' satisfactions, they are not necessarily good because of the lack of discovery. In our work, we infer items which a user does not know by computing the similarity of users or items based on what items the user has already known. We try to recommend items which the user likes and does not know combining this method and the most popular method of collaborative filtering. We hope that users' satisfactions will improve.

1. はじめに

インターネット上には大量の情報が存在している。しかし、人々は自分の興味のあるコンテンツや商品(以降、これらをアイテムと呼ぶ)を探すことが困難になる情報洪水の問題に直面している。近年、この情報洪水に対処する有効な手法の一つとして推薦システムが目ざされている。推薦システムとは、ユーザの過去の嗜好、購入履歴、デモグラフィック情報(年齢や性別、地域などの人口統計的な情報)に基づいてアイテムを推薦するシステムである¹⁾。推薦システムを実現する手法には、コンテンツに基づくフィルタリングと協調フィルタリングがある^{2),3)}。このうち協調フィルタリングは、コンテンツを解析しなくてもすむ処理の手軽さから、多くの方式が開発され⁴⁾⁵⁾⁶⁾⁷⁾、また Amazon.com などの多くの商用サイトで用いられている⁸⁾。

伝統的に、協調フィルタリングの研究は、精度、再現率や平均絶対誤差(MAE)を用いて、推薦の正確さを向上することに重点を置いてきた。しかし、協調フィルタリングには、その精度の高さが逆にわざわざして、対象のユーザが既知っているアイテムやどれも似たようなアイテムばかりが推薦されるという問題がある。そのため、ユーザはすぐに推薦に飽きてしまい、その推薦を受けなくなってしまう可能性がある。すなわち、推薦の精度が高いとしても、発見性や内容の多様性が欠落していることから、ユーザが推薦に満足しない問題が起きている⁹⁾。我々が本稿で用いる推薦の発見性とは、ユーザの知らないアイテムがどれだけ推薦されるかを示し、推薦の多様性とは、アイテムに元々つけられているジャンルなどの分類の相違性が推薦リスト内においてどれだけ高いかを示す。例えば、あるユーザが主にモーニング娘のアルバム CD を買っていたとする。すると推薦にはモーニング娘のシングル CD や DVD、ミニモニやぶっちモニ(それぞれモーニング娘のサブグループ)の CD がリストアップされる。このとき、純粋に精度だけを考えると推薦

[†] 大阪大学大学院基礎工学研究科
Graduate School of Engineering Science, Osaka
University

は良いものにみえるが、新しく得る情報の少なさを考えると、ユーザは推薦に満足しないと思われる。これは特にユーザが他のアーティストや他のジャンルにも興味があった時に、より顕著となる。推薦の発見性や多様性の欠落の問題を解決することで、ユーザがより満足する推薦を行うことができると考えられる。

最近に行われた研究には、推薦に対するユーザの満足度を向上することを目的としているものがある。Zieglerら⁹⁾は、元々定義されているアイテムの分類を用い、推薦リスト内のトピックを多様化することでユーザの満足度を向上させている。

本研究では、ある程度の推薦精度を維持し、なおかつ推薦の発見性を向上させることによりユーザの満足度を向上させることを目指す。具体的には、嗜好と発見性を考慮した新たな協調フィルタリングアルゴリズムを提案する。我々の手法と従来の協調フィルタリングの間で最も異なる点は、従来の協調フィルタリングではユーザの嗜好のプロファイルのみを利用していたが、我々の手法ではそれに加えて、ユーザがどのようなアイテムを知っているかについてのプロファイル(以降、これを既知・不既知のプロファイルと呼ぶ)を利用することである。このプロファイルは、ユーザに事前にアイテムに既知・不既知の評価(既知なら1, 不既知なら0)をつけてもらうことにより獲得する。ユーザに嗜好のみでなく既知・不既知の評価をつけてもらうためユーザの労力は大きくなるが、満足度の高い推薦が受けられるのであればある程度受け入れられると期待している。

既知・不既知のプロファイルを用いることによる本研究の貢献は以下の通りである。

- **知らないアイテムの予測** 我々はユーザが知らないアイテムを予測する方法を提案する。ユーザがアイテムにつけた既知・不既知の評価から、ユーザ間、アイテム間の類似度を計算し、その類似度を用いて未評価のアイテムの既知・不既知の予測値を計算する。この予測値を基に順位付けされた知らないアイテムのリストを作成し、その予測精度を調べる。
- **嗜好と発見性を考慮した推薦** 嗜好の予測値で順位付けされたリストと、既知・不既知の予測値で順位付けされたリストを結合することを考える。この考えに基づいた、ユーザの知らない好みのアイテムを多く推薦するアルゴリズムを提案する。ユーザベース⁵⁾¹⁰⁾とアイテムベース¹¹⁾¹²⁾それぞれの協調フィルタリングにおいて、このアルゴリズムの効果を検証する。

本稿の構成は以下の通りである。2章では、協調フィルタリングの2つのアルゴリズムについて説明する。3章では、過去に研究されてきた一般的な評価指標とユーザの満足度に関連する評価指標について説明する。4章では、知らないアイテムの予測方法と、嗜好と発見性を考慮した協調フィルタリングアルゴリズムを示す。5章では、本アルゴリズムの嗜好と発見性の組み合わせに対する有効性を検証する。6章では、本研究の関連研究を示す。7章では、今後の課題を示す。

2. 協調フィルタリング

協調フィルタリングは推薦システムを作る際に最もよく使われている方法である。その基本的な考えは、ユーザと好みの似たユーザグループが好きなアイテムをそのユーザに推薦するというものである。つまり、ユーザがアイテムに対して付けた評価値の統計量に基づいて、そのユーザと近い好みを持つユーザグループを特定し、それらのユーザが高く評価しているアイテムを推薦するというものである。用いられる評価値には明示的なものと暗黙的なものがある。明示的なものはユーザにアイテムについての評価を明示的に示してもらうことにより得られるものである。暗黙的なものは、購入という行動やアイテムについて言及した回数などを評価の表現として使うものである。暗黙的な評価値は一般的に収集が容易であるが、収集された情報はノイズを多く含む¹³⁾。

2.1 ユーザベースの協調フィルタリング

ユーザベースの協調フィルタリングは過去10年間で非常に多くの研究が行われてきた。推薦の質が優れていることから、最もよく用いられる推薦アルゴリズムになっている。

ユーザ集合を $A = \{a_1, a_2, \dots, a_n\}$ 、アイテム集合を $B = \{b_1, b_2, \dots, b_m\}$ とし、ユーザ a_i がアイテム b_k につけた評価値を $r_i(b_k)$ とする。

ユーザベースの協調フィルタリングの処理手順は次の2段階に分けることができる。

- **近傍形成** a_i を注目しているユーザ (active user) としたとき、全ての $a_o \in A \setminus \{a_i\}$ に対する類似度 $s(a_i, a_o)$ が、 r_i と r_o の類似度に基づいて計算される。一般的には、 $s(a_i, a_o)$ の計算にはピアソン相関かコサイン距離が用いられる。最も似ているユーザ上位 M 人が a_i の近傍メンバーになり、その集合を $neighbor(a_i) \subseteq A$ と表す。
- **評価値予測** a_i の近傍メンバー $a_o \in neighbor(a_i)$ が評価をつけており、かつ a_i が未評価であるアイテム b_k 全てに対して、嗜好の予測値 $p_i(b_k)$ が

計算される。予測値 $p_i(b_k)$ の値は以下のように計算される。

$$p_i(b_k) = \bar{r}_i + \frac{\sum_{a_o \in A'_i} s(a_i, a_o) * (r_o(b_k) - \bar{r}_o)}{\sum_{a_o \in A'_i} |s(a_i, a_o)|} \quad (1)$$

$$A'_i := \{a_o | a_o \in neighbor(a_i)\}$$

$$\bar{r}_i = \sum_{i=1}^n r_i(b_k) / n$$

最終的に、予測評価値 p_i に基づいて上位 N 個の推薦リスト $L_{p_i} : \{1, 2, \dots, N\} \rightarrow B$ が計算される。関数 L_{p_i} は最も高い予測値をもつアイテムを1位とした降順の推薦ランキングを示す。

2.2 アイテムベースの協調フィルタリング

アイテムベースの協調フィルタリングは Sarwar ら¹¹⁾によって提案され、ここ5年ほどで活発に研究されるようになってきている。その理由には計算の複雑さの点で有利なことと、計算処理モデルを実際の予測をすることから切り離していることがあげられる。このアルゴリズムは Amazon.com が提供する推薦システムでも用いられている⁸⁾。

アイテムベースの協調フィルタリングは、定義と予測精度の点でユーザベースに似ているが、ユーザ間ではなくアイテム間の類似度 s が計算される。2つのアイテム b_k, b_e に対して各ユーザが近い評価値を付けているとき、これらのアイテムを似ているとし類似度 $s(b_k, b_e)$ は大きい値を示す。類似度の計算にはコサイン距離を用いることが多い。各 b_k に対して最も似ているアイテム上位 M 個が近傍 $neighbor(b_k) \subseteq B$ と定義される。予測値 $p_i(b_k)$ は以下のように計算される：

$$p_i(b_k) = \frac{\sum_{b_e \in B'_k} (s(b_k, b_e) \cdot r_i(b_e))}{\sum_{b_e \in B'_k} |s(b_k, b_e)|} \quad (2)$$

$$B'_k := \{b_e | b_e \in neighbor(b_k)\}$$

上位 N 個の推薦リスト L_{p_i} の最終的な計算は、ユーザベースの協調フィルタリングの手順に従う。

3. 評価指標

評価指標は推薦システムの質と性能を判断するために重要である。多くの評価指標は正確さの測定に関するものであり、他の要因、例えば推薦の新規性や掘り出し物を見つける性能、推薦リスト内のアイテムの多様性や発見性は考慮していない。以下の節で一般的によく使用される評価指標の概観を示す。なお、正確さの指標については¹⁴⁾を参考にしてている。

3.1 正確さの指標

正確さの指標は目的の違いから2つに分けられる。一つ目は、個々の予測の正確さを判断するために用いられ、アイテム b_k の予測値 $p_i(b_k)$ が、 a_i の実際の評価値 $r_i(b_k)$ とどれだけ違っているかを測定する。二つ目は、予測値を基にして作成した推薦リストを評価するためのものである。一般的に2値の嗜好を想定している。

3.1.1 予測の正確さの指標

予測の正確さの指標は、予測評価値が実際のユーザの評価値にどれだけ近いかを測定する。平均絶対誤差 (MAE) は、アイテム集合 B_i の予測 $p_i(b_k)$ の正確さを統計的に測定する有効な手段であり、最も広く使われている¹⁰⁾¹⁵⁾。MAE は以下の式で計算される。

$$|\bar{E}| = \frac{\sum_{b_k \in B_i} |r_i(b_k) - p_i(b_k)|}{|B_i|} \quad (3)$$

MAEに関連したものに平均2乗誤差 (MSE) があり、これは加算の前に誤差を2乗する。そのため、大きい誤差は小さい誤差より目立つようになる。

予測の正確さの指標は、システム全体における予測の性能を評価することには適しているが、上位 N 個の推薦リストの質を評価するには不適切である。なぜなら、ユーザは高いランクのアイテムの誤差については気にするが、低いランクのアイテムの誤差には興味がないということが分かっているからである。

3.1.2 決定支援指標

精度と再現率は情報検索システムの評価指標としてよく知られている。これらは予測と実際の評価値のずれを表したものではなく、むしろ推薦リストの集合が、注目しているユーザにとって必要かどうかを表したものである。

これらの指標を使う前に、ユーザ a_i によって評価付けされた全てのアイテムの集合を R_i とし、 R_i をなるべく等サイズの互いに素な K 個のスライスに分割する。この結果、ランダムに選ばれた $K-1$ 個のスライスが a_i の訓練事例 R_i^* を形成する。これらの評価値は、推薦を計算するために使われる a_i のプロフィールとなる。 a_i の残りのスライス ($R_i \setminus R_i^*$) は保持され、予測のためには使われない。このスライスはテストセットを構成する。つまり、推薦システムが予測しようとするアイテムとなる。テストセット中のアイテムの中で、ユーザが好きであるアイテムの集合を T_i^* と表す。評価値の範囲が1-5で与えられる場合は、評価値を2値に変換する必要がある。一般的には、4または5である全ての評価値を”好き”、1-3である全ての評価値を”嫌い”として2値化する。

Sarwar¹⁶⁾ は、再現率をテストセット中の好きなアイテムの総数 $|T_i^p|$ に対する、推薦リスト L_i^p 中に含まれる好きなアイテム $b \in T_i^p$ の割合として定義している：

$$Recall = 100 \cdot \frac{|T_i^p \cap \mathcal{S}L_i^p|}{|T_i^p|} \quad (4)$$

記号 $\mathcal{S}L_i^p$ は画像 L_i^p の像であり、推薦リストの全アイテムを示す。

次に、精度は推薦リストの大きさに対する L_i^p 中に含まれる好きなアイテム $b \in T_i^p$ の割合として定義されている：

$$Precision = 100 \cdot \frac{|T_i^p \cap \mathcal{S}L_i^p|}{|\mathcal{S}L_i^p|} \quad (5)$$

Breese ら¹⁵⁾ は推薦リスト内での順位を考慮にいれた再現率を提案した。それは重み付け再現率または Breese スコアとして知られている。この指標は、推薦が間違っている場合には、その順位に応じたペナルティーを科すことで順位を考慮にいれている。このペナルティーは下位になるにつれ指数関数的に減衰していくので、推薦リストの下位の方で生じた間違った推薦にはあまり厳しいペナルティーを科さない。

3.2 正確さ以外の指標

推薦の正確さは実用性に大きな影響を与える。しかし、ユーザの満足度に影響を与える特徴には、正確さ以外のものも考えられる。正確さ以外の特徴に関する指標は、これまでの研究ではわずかしかが使われていない。

3.2.1 Coverage

正確でない評価指標の中で、Coverage が最も頻繁に使われている¹⁰⁾¹⁸⁾。Coverage は、システムがどれだけアイテムを予測可能であるかを測定する指標である。システムがより多くのアイテムの予測を行えることは、そのシステムがユーザの好みのアイテムをより多く見つける可能性があることを示している。Coverage は、全アイテム数に対する予測がなされるアイテムの個数の割合として測定される。

3.2.2 Novelty と Serendipity

Novelty と Serendipity は推薦が当たり前でないことを示す特徴である¹⁴⁾。推薦されたアイテムがユーザの知らない好みのものであるとき、この推薦は Novelty であるという。Novelty の精度と再現率をテストセット中の知らない好みのアイテムの集合 C_i^p を用いて式で表すと以下ようになる：

$$精度 (Novelty) = 100 \cdot \frac{|C_i^p \cap \mathcal{S}L_i^p|}{|\mathcal{S}L_i^p|} \quad (6)$$

$$再現率 (Novelty) = 100 \cdot \frac{|C_i^p \cap \mathcal{S}L_i^p|}{|C_i^p|} \quad (7)$$

推薦が Serendipity であるということは、ユーザ自身で探すことができなかつたであろう好みのアイテムが推薦されることを意味する。Herlocker らは Serendipity の検出のためには、推薦されたアイテムがユーザをどの程度引き付け、驚きを与えたかを測定すればよいと述べている。しかし、これらを測定することは非常に難しいため、測定するためのよい指標は考えられていない。

3.2.3 リスト内の類似度

リスト内の類似度は推薦リストの多様性を測定する目的で考えられた指標である⁹⁾。リスト内のアイテム間のトピックの類似度を計算し、それを合計したものをリスト内の類似度として考える。アイテム間のトピックの類似度は、ジャンルや作者、その他の特性に基づいて計算される。リスト内の類似度が高いことは、多様性が低いことを示す。

3.3 発見性

我々は推薦リスト内に知らないアイテムがどれだけあるかを測定するために、発見性という評価指標を提案する。発見性をリスト内のアイテム数に対するリスト内の知らないアイテム数の割合として定義する。テストセット中の知らないアイテムの集合 D_i^p を用いて発見性を式で表すと以下ようになる：

$$発見性 = 100 \cdot \frac{|D_i^p \cap \mathcal{S}L_i^p|}{|\mathcal{S}L_i^p|} \quad (8)$$

発見性が高いことはリスト内に知らないアイテムが多いことを示す。

4. 嗜好と発見性を考慮した推薦

4.1 知らないアイテムの予測

ユーザが知らないであろうアイテムを予測するために、1章で述べた既知・不既知のプロファイルを用いる。提案する方法は、知っているアイテムが似ている他のユーザが知っているアイテムは、注目しているユーザも知っているだろうという考えに基づく。この考えは協調フィルタリングと同様の考えであるので、既知・不既知のプロファイルに協調フィルタリングアルゴリズムを適用することにより、ユーザの知っているアイテムを予測できると考える。協調フィルタリングの処理手順に従い、注目するユーザと似た既知・不既知のプロファイルを持つユーザ集合を特定し、そのユーザ集合が知らないアイテムを予測するという方法を提案する。

ユーザ a_i がアイテム b_k につけた既知・不既知の評

```

procedure discovery ( $L_{p_i}, \alpha$ ) {
   $B_i \leftarrow \mathcal{G}L_{p_i}$ ;
   $b \in B_i$  : compute  $p_i^{know}(b)$ ;
  compute  $L_{p_i^{know}} : \{1, 2, \dots, |B_i|\} \rightarrow B_i$  using  $p_i^{know}$ ;
  for all  $b \in B_i$  do
     $L_{p_i^{unknow}}^{-1}(b) \leftarrow |B_i| - L_{p_i^{know}}^{-1}(b)$ ;
     $p_i^*(b) \leftarrow L_{p_i}^{-1}(b) \cdot (1 - \alpha) + L_{p_i^{unknow}}^{-1}(b) \cdot \alpha$ ;
  end do
  compute  $L_{p_i^*} : \{1, 2, \dots, |B_i|\} \rightarrow B_i$  using  $p_i^*$ ;
  return  $L_{p_i^*}$ ;
}

```

図 1 嗜好と発見性を考慮した協調フィルタリングアルゴリズム

価値を $h_i(b_k)$ とする。この評価値に対して 2 章で述べたユーザベース、またはアイテムベースの協調フィルタリングアルゴリズムを適用すると、ユーザの既知・不既知の評価の予測値 $p_i^{know}(b_k)$ を計算することができる。アイテムを予測値 $p_i^{know}(b_k)$ により順位付けしたリスト $L_{p_i^{know}}$ を作成する。 $L_{p_i^{know}}$ はユーザが知っているとして予測された順でアイテムを順位付けしたリストであるので、これを逆順にソートすることで、知らないとして予測された順でアイテムを順位付けしたリスト $L_{p_i^{unknow}}$ を得ることができる。

4.2 嗜好と発見性を考慮した推薦アルゴリズム

本節では、嗜好と発見性を考慮した協調フィルタリングアルゴリズムについて、まず基本的な考え方を説明し、その後でアルゴリズムの詳細を説明する。嗜好のプロファイルに協調フィルタリングアルゴリズムを適用して作成されるアイテムリスト L_{p_i} と、既知・不既知のプロファイルに協調フィルタリングアルゴリズムを適用して作成されるアイテムリスト $L_{p_i^{unknow}}$ を結合することを考える。

まず、それぞれのリストが、結果として生じるリストにどれだけ影響を及ぼすかを表す重みを定める。 $L_{p_i^{unknow}}$ にかける重みを $\alpha \in [0, 1]$ とすると、 L_{p_i} にかける重みは $(1 - \alpha)$ とすることができる。次に、リスト内の各アイテムについて、それぞれのリストでの順位に重みをかけたものの和を計算する。この値が小さい順にアイテムをソートすることで、リスト $L_{p_i^*}$ を作成することができる。 α が大きい場合は知らないであろうアイテムを多く推薦するリストを作成し、 α が小さい場合は元の嗜好に基づく推薦リストに近い推薦リストを作成する。

上記の考え方に基づき考案した推薦アルゴリズムを

図 1 に示す。嗜好の評価値に協調フィルタリングアルゴリズムを適用することで得られる上位 N 個のリスト L_{p_i} と、先に定義した重み α を入力とし、新しい推薦リスト $L_{p_i^*}$ を出力とする。まず、リスト L_{p_i} 内の全てのアイテム $b \in B_i$ について、既知・不既知の評価の予測値 $p_i^{know}(b)$ を計算する。この予測値 $p_i^{know}(b)$ の順にアイテムを順位付けしたリスト $L_{p_i^{know}}$ を作成する。 $L_{p_i^{know}}$ を逆順にソートすることで、知らないとして予測される順にアイテムを順位付けしたリスト $L_{p_i^{unknow}}$ を得る。ここまでの手順は 4.1 節で提案した手法である。続いて、各アイテム b に対して $L_{p_i^{unknow}}$ での順位に重み α をかけたものと L_{p_i} での順位に $(1 - \alpha)$ をかけたものの和 $p_i^*(b)$ を計算する。 $p_i^*(b)$ の小さい順にソートすることで、最終的なリスト $L_{p_i^*}$ を作成する。アルゴリズムでの推薦対象となるアイテムを、嗜好の予測値に基づく上位 N 個のリスト L_{p_i} 内に含まれるアイテムに限定しているのは、明らかに嫌いとして予測されるアイテムが推薦されることを防ぐためである。実際にユーザに提示する推薦リストは、上位 N 個のリスト $L_{p_i^*}$ の上位 N' 個である。ただし、入力である上位 N 個のリスト L_{w_i} は、ユーザに提示する上位 N' 個のリストよりかなり大きくなければいけないことに注意しなければならない。後の実験では、ユーザに提示する上位 10 個の推薦のために上位 50 個の入力リストを用いた。

5. Novelty に関する実験

提案するアルゴリズムが Novelty を向上させるのに有効か否かを検証するために評価実験を行った。我々は次の 3 つの点に注目した。

- 知らないアイテムの予測は可能であるか？
既知・不既知のプロファイルを用いることで、知らないアイテムをどの程度予測できるかを調べる。知らないアイテムの予測精度は、提案するアルゴリズムに大きな影響を与える要因である。
- 嗜好による予測と既知・不既知による予測の間の相関はどうであるか？
提案するアルゴリズムは、嗜好の評価値の予測により作成されるリストと既知・不既知の評価値の予測により作成されるリストの間に、それほど高い相関がないことを前提としている。好きなアイテムは必ず知っているもので、嫌いなアイテムは必ず知らないものであれば、提案するアルゴリズムにより Novelty が向上することは期待できない。そのため、提案するアルゴリズムの評価をする前に、2 つのリストの相関を調べた。

● Novelty は向上するか？

提案するアルゴリズムは、従来の協調フィルタリングアルゴリズムよりも Novelty が向上するかどうかを調べる。さらに、提案するアルゴリズムにおける、嗜好と発見性のどちらを重視するかを表すパラメータである α を変化させて、Novelty と、嗜好の精度・再現率、発見製との関係を調べる。これにより最適な α がどの程度の値なのかを明らかにする。

5.1 データセット

本実験のために、20000 の評価データを収集した。具体的には、合計 1000 曲からなる音楽データベースに対して、100 人のユーザに評価付けをもらった。本節では、音楽データベースの構築方法、評価データの収集方法について説明する。

筆者らは、協調フィルタリングアルゴリズムの評価を目的として、音楽データベースを構築した。本データベースには、1000 曲の音楽データを収録した。音楽データベースには、各楽曲に対する曲名、アーティスト名、リリース年、音楽ジャンル、試聴先の URL の情報を入力した。実験での音楽の使用許諾は JAS-RAC¹⁹⁾ から受け、実験での試聴データの使用許諾は、大手の商用音楽サイトから受けている。1000 曲のジャンル構成は、市場の流通量の大きさを考慮した結果、J-Pop : 700 曲、演歌 : 75 曲、アニメ : 75 曲、洋楽 : 150 曲とした。また J-Pop や洋楽に関しては、1960 年代から現在に至るまでの幅広い楽曲を選択した。

評価データの収集には、ユーザ 100 人 (10 代:13 人, 20 代:51 人, 30 代:5 人, 40 代:18 人, 50 代以上:13 人) に協力してもらった。それぞれのユーザから、嗜好のプロファイルと既知・不既知のプロファイルを収集した。各ユーザごとに、音楽データベースから 200 曲をランダムに選択し、それらの曲に対して嗜好の 5 段階評価 (1~5) と既知・不既知の評価 (1 or 0) をつけてもらった。知らない曲に対しては試聴してもらった後に嗜好の評価をつけてもらった。一人のユーザに付き、試聴の時間を含めて、これら評価付けにかかった時間は、平均で約 3 時間である。

各ユーザが評価付けを行った 200 曲を、150 曲の訓練事例と 50 曲のテストセットに分けることで実験を行った。

5.2 知らないアイテムの予測

4.1 節で説明した手法により、知らないであろうアイテム上位 10 個のリストを作成し、その発見性を調べた。表 1 に結果を示す。この手法では、ユーザベースの協調フィルタリングを用いた場合とアイテムベ

表 1 知らないアイテムの予測

	ユーザベース	アイテムベース
精度	85.7	98.8

表 2 嗜好と既知・不既知の相関

	ユーザベース	アイテムベース
相関	0.573	0.243

スの協調フィルタリングを用いた場合との 2 通りを試した。この結果から、既知・不既知のプロファイルを、ユーザベースあるいはアイテムベースの協調フィルタリングに適用すれば、知らないアイテムを非常に高い精度で予測できることが分かる。

5.3 嗜好と既知・不既知の相関

嗜好の評価値の予測により作成されるリストと既知・不既知の評価値の予測により作成されるリストの間の相関を調べた。表 2 は、ユーザベースとアイテムベースの協調フィルタリングそれぞれにおける、2 つのリストの相関係数である。ユーザベースの協調フィルタリングでは少しの相関が観測されたが、既知・不既知の予測からの効果は得られると思われる。

5.4 嗜好の精度、発見性、Novelty

提案するアルゴリズムにおいて $\alpha = 0$ としたとき、生成される推薦リストは嗜好のみを考慮しているため従来の協調フィルタリングによる推薦とみなすことができる。提案手法と従来手法の比較のために、 $\alpha = 0$ での上位 10 個の推薦リストと $\alpha \in [0.1, 0.2, \dots, 0.9, 1.0]$ での上位 10 個の推薦リストの Novelty の精度・再現率を比較した。また、上記のように α を変化させることで最適な α の値を調べた。さらに、嗜好の精度・再現率、発見性と Novelty の関係について調べた。

ユーザベースの協調フィルタリングに対して α を変化させたときに作成される上位 10 個の推薦リストの嗜好の精度、発見性、Novelty の精度の変化を図 2 に、嗜好の再現率、Novelty の再現率の変化を図 3 に示す。同様にアイテムベースの協調フィルタリングについては図 4、図 5 に示す。

● 本手法と従来手法の比較

α がどの値であっても提案手法における Novelty の精度・再現率は、従来手法 ($\alpha = 0$) よりも高い値を示していることが分かる。ユーザベースにおいて Novelty が最も向上しているのは、 $\alpha = 0.6$ のときであり、従来手法と比較して精度は約 0.15、再現率は約 0.25 高い値を示している。アイテムベースにおいては、 $\alpha = 0.5$ のときに、精度は約 0.16、再現率は約 0.31 高い値を示している。

● Novelty と嗜好の精度・発見性の関係

α が増加するに従い嗜好の精度は減衰し、発見性

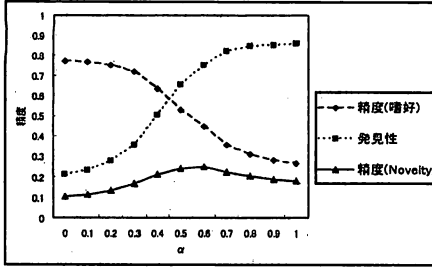


図 2 ユーザベース協調フィルタリングの精度の推移

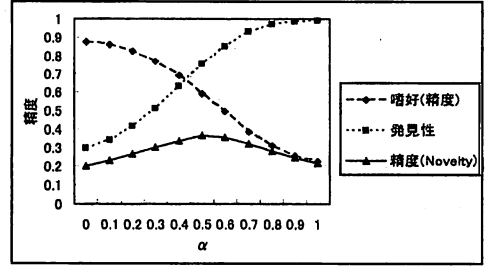


図 4 アイテムベース協調フィルタリングの精度の推移

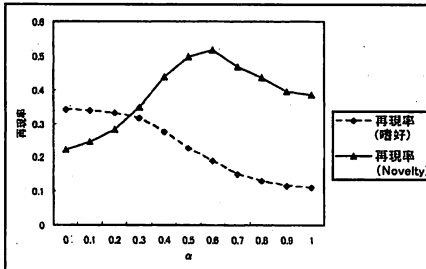


図 3 ユーザベース協調フィルタリングの再現率の推移

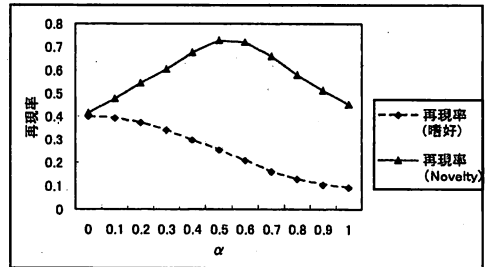


図 5 アイテムベース協調フィルタリングの再現率の推移

は上昇している。Novelty は嗜好の精度と発見性のバランスがとれているときに高い値を示す。これは、嗜好の評価値予測と既知・不既知の評価値予測をバランスよく組み合わせることが Novelty の向上につながることを意味する。

6. 関連研究

本研究では、推薦の正確さの他に発見性を考慮している。推薦の正確さと他の評価指標との両方を考慮した研究はいくつかある。Ziegler ら⁹⁾は、協調フィルタリングによる推薦の多様性の欠如の問題に取り組んでいる。推薦リストを多様化する方法は、従来の協調フィルタリングにより作成されるリストと、リスト内のトピックの類似度により順位付けされたリストを組み合わせるといったものである。トピックの類似度は、作者やジャンルによる分類表を用いて計算される。

また加藤らは、潜在的に興味を持っているアイテムが推薦されにくいという、従来の推薦システムの問題点を解決するために、推薦の正確性と意外性のバランスを考慮した推薦方法を提案している²⁰⁾。この方法では、まず商品の特徴を利用してアイテム特徴ベクトルとユーザー特徴ベクトルをカテゴリに分類している。そして、アイテム特徴ベクトルとユーザー特徴ベクトルの比較処理において、同じカテゴリ内に属するベクトル

を比較する回数と、異なるカテゴリに属するベクトルを比較する回数を確率的に変化させることにより、正確性と意外性の調和のとれた推薦を実現している。

また本研究では、発見性を考慮するために既知・不既知の評価値を利用している。他の評価指標を考慮するため、嗜好の評価値以外の評価値を利用した協調フィルタリングの研究に Massa ら²¹⁾の方式がある。これは、アイテムに対する嗜好の評価だけでなく、ユーザが他のユーザに対してつけた信頼度を利用している。ユーザは他のユーザがアイテムに付けている評価などを見て、そのユーザの信頼度を評価することができ、これにより作成される信頼のネットワークから他のユーザの信頼度を予測することができる。この信頼度が高いユーザを近傍ユーザとして選択し、そのユーザとの類似度からアイテムの予測評価値を計算することで推薦を行う Trust-Aware 推薦システムを提案している。この方法は従来の協調フィルタリングと比較したとき、coverage が大きく、予測誤差が小さくなっている。

7. むすび

本研究では、嗜好と発見性を考慮した協調フィルタリングアルゴリズムを提案した。そして、提案アルゴリズムが推薦の Novelty を向上させるのに有効であ

るかの評価実験を行った。実験の結果、提案アルゴリズムにより Novelty の向上を確認することができた。

今後は、実際にユーザに推薦を行い、本手法による推薦に対する満足度を調査するための実験を行う。そして、推薦の嗜好、発見性、及び Novelty の精度について、各指標と推薦に対するユーザの満足度との相関を調査する。また、従来の手法と本手法で上記の関連性を比較する。

参 考 文 献

- 1) Resnick, P. and Varian, H.: Recommender systems, *Communications of the ACM*, Vol.40, No.3, pp.56-58 (1997).
- 2) N. Ramakrishnan: PIPE: Web Personalization by Partial Evaluation, *IEEE Internet Computing*, Vol.4, No.6, pp.21-31 (2000).
- 3) D. Riecken.: Personalized Views of Personalization, *Comm. of the ACM*, Vol.43, No.8, pp.26-158 (2000).
- 4) Goldberg, D., Oki, B.M. and Terry, D.: Using Collaborative Filtering to Weave an information Tapestry, *Communications of the ACM*, Vol.35, No.12, pp.61-70 (1992).
- 5) Resnick, P., Iacovou, N., Suchak, M., Bergstorm, P. and Riedl, J.: GroupLens: an open architecture for collaborative filtering of netnews, in *Proceedings of the ACM 1994 Conference on Computer Supported Cooperative Work*, ACM, pp.175-186 (1994).
- 6) Shardanand, U. and Maes, P.: Social Information Filtering: Algorithms for Automating "Word of Mouth" , *Proc. of CHI'95*, pp. 210-217 (1995).
- 7) Hill, W., Stead, L. Rosenstein, M. and Furnas, G.: Recommending and Evaluating Choices in a Virtual Community of Use, *Proc. of CHI'95*, pp. 194-201 (1995).
- 8) Linden, G., Smith, B., and York, J. : Amazon.com recommendations: Item-to-item collaborative filtering, *IEEE Internet Computing*, Vol.4, No.1 (2003).
- 9) Ziegler, C., McNee, S.M., Konstan, J.A. and Lausen, G.: Improving Recommendation Lists Through Topic Diversification, *WWW2005*, pp.22-32, (2005).
- 10) Herlocker, J., Konstan, J., Borchers, A. and Riedl, J.: An algorithmic framework for performing collaborative filtering, in *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM Press, pp230-237 (1999).
- 11) Sarwar, B., Karypis, G., Konstan, J. and Riedl, J.: Item-based collaborative filtering recommendation algorithms, in *Proceedings of the Tenth International World Wide Web Conference* (2001).
- 12) Deshpande, M. and Karypis, G.: Item-based top-n recommendation algorithms, *ACM Transactions on Information Systems*, Vol.22, No.1, pp.143-177 (2004).
- 13) Nichols, D.: Implicit rating and filtering, in *Proceedings of the Fifth DELOS Workshop on Filtering and Collaborative Filtering* (Budapest, Hungary, 1998), ERCIM, pp.31-36 (1998).
- 14) Herlocker, J., Konstan, J., Terveen, L., and Riedl, J.: Evaluating collaborative filtering recommender systems, *ACM Transactions on Information Systems*, Vol.22, No.1, pp.5-53 (2004).
- 15) Breese, J., Heckerman, D. and Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering, in *Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufmann, pp.43-52 (1998).
- 16) Sarwar, B., Karypis, G., Konstan, J., and Riedl, J.: Application of dimensionality reduction in recommender systems, in *ACM WebKDD Workshop*(2000)
- 17) Balabanovic, M. and Shoham, Y.: Fab - content-based, collaborative recommendation, *Communications of the ACM*, Vol.40, No.3, pp.66-72 (1997).
- 18) Middleton, S., Shadbolt, N. and De Roure, D.: Ontological user profiling in recommender systems, *ACM transactions on Information Systems* , Vol.22, No.1 pp.54-88 (2004).
- 19) 社団法人日本音楽著作権協会 JASRAC, <http://www.jasrac.or.jp/>.
- 20) 加藤由花, 川口賢二, 箱崎勝也: オンラインショッピングを対象とした正確性と意外性のバランスを考慮したリコメンダシステム, *情報処理学会論文誌*, Vol.46, No. SIG13(TOD 27), pp.53-64 (2005).
- 21) P. Massa and P. Avesani.: Trust-aware Collaborative Filtering for Recommender Systems, in *Proceedings of International Conference on Cooperative Information Systems* (2004).