



Title	協調フィルタリングに基づく推薦システムにおける格付け寄与度の提案 (<特集>ソフトウェアエージェントとその応用論文)
Author(s)	山下, 晃弘; 川村, 秀憲; 鈴木, 恵二; 大内, 東
Citation	電子情報通信学会論文誌. D, 情報・システム, 92(11), 1902-1910
Issue Date	2009-11-01
Doc URL	<a href="http://hdl.handle.net/2115/52022">http://hdl.handle.net/2115/52022</a>
Rights	Copyright © 2009 社団法人 電子情報通信学会(IEICE).
Type	article
File Information	110007467231.pdf



[Instructions for use](#)

# 協調フィルタリングに基づく推薦システムにおける格付け寄与度の提案

山下 晃弘<sup>†,††</sup> 川村 秀憲<sup>†</sup> 鈴木 恵二<sup>†</sup> 大内 東<sup>†</sup>

## Proposal of Rating Contribution in Recommender Systems Based on Collaborative Filtering

Akihiro YAMASHITA<sup>†,††</sup>, Hidenori KAWAMURA<sup>†</sup>, Keiji SUZUKI<sup>†</sup>,  
and Azuma OHUCHI<sup>†</sup>

あらまし ショッピングサイトなどでは、個人の興味や嗜好に適應して積極的にアイテムを提示する推薦システムが実用化されている。協調フィルタリングは、近年最もよく用いられている推薦手法の一つで、アイテムに対する格付けをユーザ全体から収集し、それをもとに個人の興味や嗜好に適合したアイテムを推定して推薦を行う。一般的に、多くの格付けを収集すれば適合アイテムの推定精度が向上し、システム全体の有効性が向上する。そこで実運用においては、メールなどで格付けを促すほか、特典やポイントなどのコストをかけて格付け行為にインセンティブをもたせる試みがなされている。しかし、収集したすべての格付けが均一に推定に寄与するとは限らず、また同じ格付けを収集する場合を考えても、推定精度の向上に必要な格付けをより早い段階で重点的に収集した方が効率的である。本研究では、各々の格付けが推定に寄与する度合（格付け寄与度）を定義し、それを利用した効率的な推定精度の向上方法についてシミュレーションや実データに基づいて議論した。その結果、格付け寄与度が高い格付けを重点的に収集することで、より効率的に推定精度が向上することを示した。

キーワード 推薦システム, 協調フィルタリング, 格付け寄与度, エージェントベースシミュレーション

### 1. ま え が き

推薦システムは、商品や映画などの膨大なアイテムの中から、ユーザの興味や嗜好に適合したアイテムを推定し、積極的に提示するシステムである [1]。効果的な推薦により、ユーザは目的のアイテムに素早くたどり着け、また新商品を発見できるなど、利便性向上が期待できる。一方、運営側は宣伝効果による収益拡大が見込めるため、双方にメリットがある。既に Amazon.com<sup>(注1)</sup>, MovieFinder.com<sup>(注2)</sup>, Ebay<sup>(注3)</sup> など、実際に多くのサイトでその有効性が示されている [2], [3]。

協調フィルタリング (Collaborative Filtering, 以下 CF) は、推薦手法の一種であり、ユーザのアイテムに対する格付けをもとに推薦アイテムを決定する。

つまり、ユーザ-アイテムで表現される「格付け行列」を構築し、既に得た格付けから未知の格付けを推定し、推定値が高いアイテムを推薦する。一般的に、多くの格付けを得るほど推定精度は向上するため、格付けの収集は重要な課題である。

ユーザは、推薦結果に影響を受けながら次に購入するアイテムを決定し、格付けを行う。つまり推薦システムは、ユーザへの推薦と、その結果得られる格付けをフィードバックとしてその有効性を高めていく。したがって、ユーザと推薦システムは相互作用的であり、実運用で蓄積される格付けサンプルは、それ以前の推薦結果に影響を受けながら形成されていく。しかし従来研究では、格付けサンプルはそれが形成される過程とは独立と仮定し、推薦アルゴリズムの性能は、ある瞬間の推定精度に基づいて議論されていた。

そこで我々は、ユーザと推薦システムの相互作用を

<sup>†</sup> 北海道大学大学院情報科学研究科, 札幌市  
Graduate School of Information Science and Technology,  
Hokkaido University, Sapporo-shi, 060-0814 Japan

<sup>††</sup> 日本学術振興会, 東京都  
Japan Society for the Promotion of Science, Tokyo, 102-8472  
Japan

(注1): <http://www.amazon.com/>

(注2): <http://www.moviefinder.com/>

(注3): <http://www.half.ebay.com/>

考慮するため、エージェントモデルに基づく新たな評価手法を提案し、評価実験を行った[4]。その評価実験では、相互作用が推定精度に及ぼす影響の検証が目的であったため、推薦システムは格付けの推定値が最も高いアイテムを推薦し、エージェントは必ずそのアイテムを格付けするという単純なモデルを用いた。

しかし実際は、アイテム購入時に必ず格付けを行うとは限らず、格付けの収集に工夫が必要である。実際のサイトでは、メールなどで格付けを促すほか、ポイントや特典で格付け行為にインセンティブをもたせるなどの試みもあるが、実施にはコストが必要である。

もし、まだ得ていない格付けが、推定精度の向上にどの程度寄与するのかを予測できれば、大きく寄与する特定の格付けを重点的に収集でき、コストを抑え効率的にシステム全体の有効性を高められる。

そこで本研究では、ある格付けが推定精度に寄与する割合を、推定アルゴリズムの観点から検討し、「格付け寄与度」として定義、定量化を行った。また、格付け寄与度に基づいた収集の有効性について、エージェントシミュレーション及び実データによる分析を行った。

本論文では、まず協調フィルタリングのアルゴリズムに基づき、格付け寄与度を定義、定量化する。その後、シミュレーションと実データの双方による評価実験とその結果について述べ、格付け寄与度の利用法と有効性について考察する。

## 2. 関連研究

推薦アルゴリズムは、「内容に基づくフィルタリング」とCFに大別され、更にCFは、格付け行列全体を用いる「メモリベース方式」と、格付け行列をもとに構築したモデルを用いる「モデルベース方式」に分類できる[5]。GroupLensプロジェクトのユーザ間CF[6]や、Amazon.comなどで用いられるアイテム間CF[3]は、メモリベース方式の代表例である。一方、クラスタリングを用いた手法[7]などはモデルベース手法の例であり、両手法の融合も提案されている[8]。

モデルベース手法は、計算時間の面で比較的優れているが、新たに得た格付けをモデルに反映するためには更新が必要であり、また、各々の格付けが推定精度にどの程度寄与するかはモデルに依存する。

本研究では、逐次的に格付けが与えられ、それに応じて推薦結果が変化し、その推薦が次に格付けされるアイテムの選択に影響するという、ユーザと推薦システム間の相互作用に着目する。その上で、ある格付け

が全体の推定精度向上に寄与する割合を格付け寄与度として定義し、有効性を検証することを目的とする。したがって、本研究では格付けが直ちに推薦結果に反映されるメモリベースのCFについて議論する。

## 3. 協調フィルタリングと格付け寄与度

### 3.1 ユーザ間協調フィルタリング

CFは、1992年にその基本概念が初めて示され[9]、その後、k-nearest neighbor法を初めて導入したGroupLensプロジェクト[6]など、多くの改良、分析、評価がなされてきた[10]~[13]。中でもユーザ間CFは、最も基本的なアルゴリズムの一つであり、現在提案されている多くの推薦アルゴリズムの基礎になっている。そこで、本論文ではユーザ間CFを対象とした格付け寄与度について議論する。

ユーザ間CFは、まず、被推薦者 $a \in U$ の類似ユーザ(neighbors)を求める。本論文では、被推薦者 $a$ のneighborsを $S(a) = \{s_{a,1}, \dots, s_{a,N_a}\}$ と記す。その後、neighborsの格付けから、まだ被推薦者が格付けしていないアイテム $i \in I$ に対する格付けを推定し、その推定値が最も高いアイテムを推薦する。

$S(a)$ は、被推薦者 $a$ と、その他の各ユーザ $u \in U$ 間の類似度 $sim(a, u)$ から求める。類似度として、ピアソン相関係数[6]や、コサイン距離[5]などが用いられるが、一般的にはピアソン相関係数の精度が良いため[5]、本論文でもピアソン相関係数を用いる(式(1))。

$$sim(a, u) = \frac{\sum_{i \in I_a \cap I_u} (r_{a,i} - \bar{r}_a)(r_{u,i} - \bar{r}_u)}{\sqrt{\sum_{i \in I_a \cap I_u} (r_{a,i} - \bar{r}_a)^2} \sqrt{\sum_{i \in I_a \cap I_u} (r_{u,i} - \bar{r}_u)^2}} \quad (1)$$

ただし、 $r_{u,i}$ はユーザ $u$ のアイテム $i$ に対する格付け、 $I_u$ は、ユーザ $u$ が格付けしたアイテム集合、 $\bar{r}_u$ は、ユーザ $u$ の格付けの平均値である。

neighborsの選択は、類似度の高い $k$ 人のユーザを選択するk-Nearest neighbor法や、あらかじめ設定したしきい値以上の類似度をもつユーザを選択する方法が提案されている[14],[15]。本論文では、 $sim(a, u) \geq 0.5$ であるすべてのユーザをneighborsとした。

格付けの推定値 $\hat{r}_{a,i}$ は、式(2)で計算する。ただし、 $U_i$ はコンテンツ $i$ を格付けしたユーザ集合であり、 $U_i \cap S(a) = \phi$ の場合は推定できない。

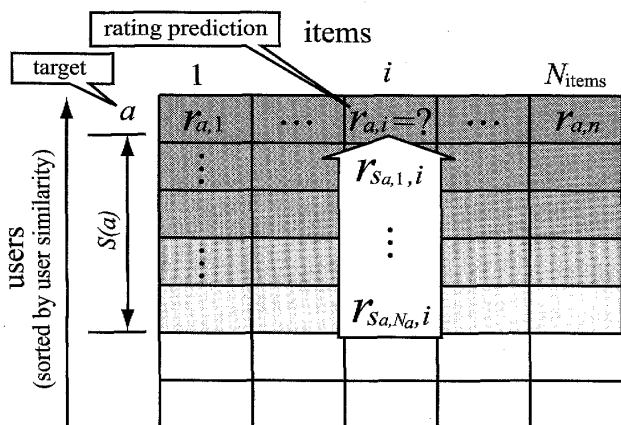


図1 ユーザ間CFに基づく格付けの推定

Fig. 1 Rating prediction based on user-based CF.

最後に推定値  $\hat{r}_{a,i}$  が最大となるアイテムを推薦する.

$$\hat{r}_{a,i} = \bar{r}_a + \frac{\sum_{u \in U_i \cap S(a)} sim(a,u)(r_{u,i} - \bar{r}_u)}{\sum_{u \in U_i \cap S(a)} |sim(a,u)|} \quad (2)$$

### 3.2 ユーザ間CFにおける格付け寄与度

本節では、格付け寄与度をユーザ間CFに基づいて定義する。式(2)によると、格付けの推定値  $\hat{r}_{a,i}$  は、ユーザ  $u \in U_i \cap S(a)$  の格付けから計算される(図1)。

ここで、アイテム  $i$  を格付けしていないユーザ  $u' \notin U_i \cap S(a)$  を考える。あるとき、ユーザ  $u'$  がアイテム  $i$  の格付け  $r_{u',i}$  を新たに入力し、 $u' \in U_i \cap S(a)$  になったと仮定すると、その格付け  $r_{u',i}$  によって、格付けの推定値  $\hat{r}_{a,i}$  が変化する。これは、格付け  $r_{u',i}$  が  $\hat{r}_{a,i}$  の推定に「寄与した」ととらえることができる。

このとき、もとの推定値  $\hat{r}_{a,i}$  と、格付け  $r_{u',i}$  が入力された後の推定値  $\hat{r}'_{a,i}$  の差分が推定値の変化量であり、推定に寄与した度合と考えられる。計算すると、

$$\hat{r}_{a,i} - \hat{r}'_{a,i} = \begin{cases} c_{u',i}(a,i) \{ \hat{r}_{a,i} - \bar{r}_a \} - (r_{u',i} - \bar{r}_{u'}) & (sim(a,u') \geq 0) \\ c_{u',i}(a,i) \{ -(\hat{r}_{a,i} - \bar{r}_a) - (r_{u',i} - \bar{r}_{u'}) \} & (sim(a,u') < 0) \end{cases}$$

where

$$c_{u',i}(a,i) = \frac{sim(a,u')}{\sum_{u \in U_i \cap S(a)} (|sim(a,u)| + |sim(a,u')|)} \quad (3)$$

となる(計算過程は付録に示す)。

ここで、 $r_{u',i}$  は実際にはまだ得られておらず、 $(\hat{r}_{a,i} - \bar{r}_a) - (r_{u',i} - \bar{r}_{u'})$  は計算できないが、 $c_{u',i}(a,i)$

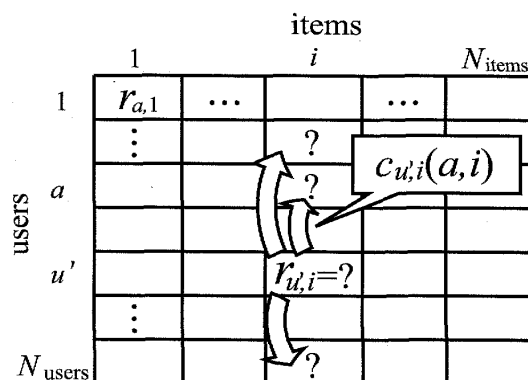


図2 格付け寄与度の概念図

Fig. 2 Concept of rating contribution.

は推定値の変化量全体の係数であり、本論文では  $c_{u',i}(a,i)$  を、 $\hat{r}'_{a,i}$  に対する  $r_{u',i}$  の格付け寄与度と定義する。

一般的に、 $r_{u',i}$  は、ユーザ  $u'$  を neighbors に含み、まだアイテム  $i$  を格付けしていないすべてのユーザ  $\{a | a \notin U_i, u' \in S(a)\}$  の格付け推定に寄与する(図2)。したがって、格付け  $r_{u',i}$  が寄与する全ユーザに対して式(3)を計算し、その合計値を  $r_{u',i}$  の格付け寄与度  $c_{u',i}$  と定義する(式(4))。

$$c_{u',i} = \sum_{a \in \bar{U}_i \cap S(a)} c_{u',i}(a,i) \quad (4)$$

### 3.3 推定精度

協調フィルタリングの有効性を表す指標として、一般的に正解率(accuracy)、精度(precision)、再現率(recall)、平均絶対値誤差(MAE)などの推定精度が用いられる。本論文では、多数の先行研究で用いられているMAEを推定精度とした。すべての格付けサンプルを訓練用とテスト用に分け、訓練用の格付けを用いて協調フィルタリングにより格付けの推定値を計算した場合、そのMAEは、テスト用の格付けとその推定値の誤差の平均値である。ただし、推定値  $\hat{r}_{a,i}$  は実数値であるが、MAEを計算する際にはあらかじめ四捨五入により離散化した値を用いた。

## 4. 検証実験のためのエージェントモデル

本論文では、エージェントモデルと実データの双方に基づいて格付け寄与度の有効性を検証する。エージェントモデルとしては、既に我々が提案した「推薦と格付けの相互作用に基づく評価モデル」[4]を用いた。本章ではモデルの詳細について説明する。

#### 4.1 構成要素

検証実験に用いたモデルは、ユーザをモデル化したエージェント、アイテム、推薦システムで構成される。推薦システムは、エージェント集合  $U = \{u | u = 1, 2, \dots, N_{user}\}$  の各エージェントに対して、アイテム集合  $I = \{i | i = 1, 2, \dots, N_{item}\}$  の中から適合アイテムを推定し、推薦する。各エージェントは、嗜好をもち、嗜好ベクトル  $\mathbf{p}_u$  として表現する。同様に、各アイテムは特徴をもち、特徴ベクトル  $\mathbf{v}_i$  として表現する。ただし、 $\mathbf{p}_u$  と  $\mathbf{v}_i$  はともに  $d$  次元のベクトルであり、本論文では  $d = 5$ 、及び各要素は  $[-1, 1]$  の範囲の実数とした。

エージェント  $u$  は、アイテム  $i$  を推薦されたとき、 $\mathbf{p}_u$  と  $\mathbf{v}_i$  間の距離をもとに格付け  $r_{u,i}$  を決定し、推薦システムに入力する。推薦システムは、格付け  $r_{u,i}$  を、格付け行列に保持する。

#### 4.2 嗜好ベクトル $\mathbf{p}_u$ の分布

$\mathbf{p}_u$  を様々な分布に従って生成することで、全体の嗜好のばらつき具合（トレンド）が変化する。

実世界は、様々な嗜好をもつユーザが偏りなく存在する状況や、似た嗜好をもつ多数のユーザと、ごく少数のニッチからなるロングテール構造など様々な状況が想定され、一概に決めることはできない。本論文では、格付け寄与度の有効性検証を目的とし、次の三つの分布をモデルとして用いてその違いを比較した。

一様分布 (Uniform distribution) :  $\mathbf{p}_u$  の各要素独立に  $[-1, 1]$  の一様乱数で与える。この分布は、各分布の中で最もトレンドが弱い分布となる。

多変量正規分布 (Multivariate normal distribution) : 多変量正規分布に基づくベクトル乱数  $\mathbf{x}$  は、次元数  $d$  のとき、各要素間で独立な  $d$  変量正規乱数  $\mathbf{z} = (z_1, \dots, z_d)$  と、分散共分散行列  $\Sigma$  をコレスキー分解  $\Sigma = LL^T$  した行列  $L$  の積で求まる。つまり、平均ベクトルを  $\mu$  とすれば、 $\mathbf{x} = L\mathbf{z} + \mu \sim N(\mu, \Sigma)$ 、で計算できる。本論文では、 $\mu = \mathbf{0}$  とし、 $\Sigma$  は、対角要素がすべて  $v$  で、それ以外すべて  $0$  という単純な行列  $\Sigma_v$  を用いた。したがって、 $v$  が小さい程偏りが強くなる。

二つの頂点をもつ分布 (Two-peak distribution) 嗜好の偏りは、一つに集中するとは限らず、グループの形成も想定される。そこで、二つの頂点をもつ分布を生成し検証を行った。二つの頂点をもつ分布は、分散共分散行列  $\Sigma$  が等しく平均ベクトル  $\mu$  が異なる二つの多変量正規分布を作成し、等確率でどちらかの分布に従

うベクトル乱数により発生させた。本論文では、二つの平均ベクトルとして  $\mu_1 = (0.5, 0.5, 0.5, 0.5, 0.5)^T$ ,  $\mu_2 = (-0.5, -0.5, -0.5, -0.5, -0.5)^T$ 、及び分散共分散行列として  $\Sigma_{0.05}$  を使用した。

#### 4.3 特徴ベクトル $\mathbf{v}_i$ の分布

本論文では、好まれるアイテムと、そうでないアイテムの偏りを把握しやすくするため、特徴ベクトル  $\mathbf{v}_i$  の分布は一様分布とし、嗜好ベクトル  $\mathbf{p}_u$  の分布のみをパラメータとした。

#### 4.4 格付け

エージェント  $u$  は、嗜好ベクトル  $\mathbf{p}_u$  に近い特徴ベクトル  $\mathbf{v}_i$  をもつアイテム  $i$  を好みとし、高い格付けを与える。本論文では5段階による格付けを用いた。

実世界では、高い格付けに偏ったユーザが存在するなど、格付けの基準はユーザに依存する。しかし、本論文では単純化のため、嗜好ベクトルと特徴ベクトルの距離  $|\mathbf{p}_u - \mathbf{v}_i|$  を五つに分割するしきい値を設け、すべてのエージェントがそのしきい値に基づいて格付けを行うものとした。各しきい値の具体的な値は、 $\mathbf{p}_u$  及び  $\mathbf{v}_i$  を一様分布で与えた場合に、5段階の格付けが同じ頻度で行われる値を予備実験により求め、その値をすべての実験における固定値として設定した。したがって、 $\mathbf{p}_u$  の分布が変われば5段階の各格付けの出現頻度は変化する。

### 5. 検証実験

#### 5.1 実験の流れ

実験は、以下の処理1から処理5を、1シミュレーションステップとして繰り返す。

処理1: 未入力の全格付けの格付け寄与度を計算する。

処理2: 計算した格付け寄与度に基づき次に格付けすべきエージェント  $u$  と対象のアイテム  $i$  を決定する。

処理3: エージェント  $u$  にアイテム  $i$  の格付け  $r_{u,i}$  を入力してもらう。

処理4: 未入力の全格付けに対する推定値  $\hat{r}_{a,i}$  をユーザ間 CF に基づき計算する。格付けが少なく推定値が計算できない場合は、平均値の3を推定値とする。

処理5: 推定精度 (MAE) を計算し、処理1に戻る。ただし、処理2では、格付け寄与度に基づくボルツマン選択を用いて決定した。つまり、式(5)に示す確率  $x_{u,i}$  に基づくルーレット選択である。

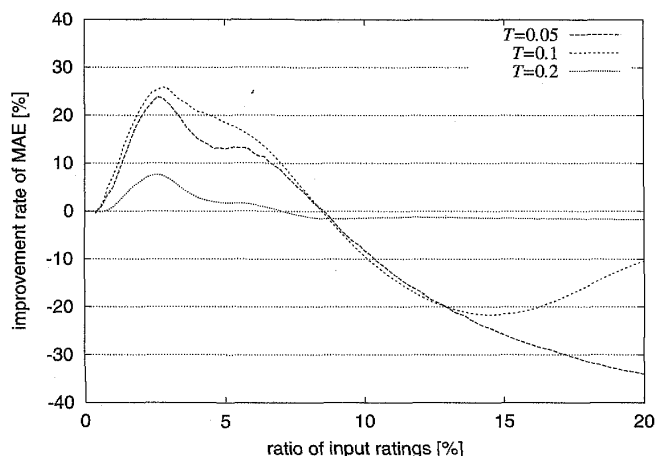


図3 エージェントモデルを用い、 $T = \{0.05, 0.1, 0.2\}$ ,  $\mathbf{p}_u$ を一様分布で与えたときのランダム選択に対するMAEの改善率

Fig.3 Improvement rate of MAE over random selection by using agent model in setting  $T = \{0.05, 0.1, 0.2\}$  and  $\mathbf{p}_u$  is generated based on uniform distribution.

$$x_{u,i} = e^{\frac{c'_{u,i}}{T}} \quad \text{where} \quad c'_{u,i} = \frac{c_{u,i}}{\max_{u',i'} c_{u',i'}} \quad (5)$$

ここで、 $T$ は温度パラメータと呼ばれ、 $T = \infty$ では $c_{u,i}$ に依存しないランダム選択となり、 $T = 0$ では、常に最大の $c_{u,i}$ が選択されるGreedy選択となる。 $T$ を変化させることで、格付け寄与度が次の格付けの選択に影響する割合を制御できる。

本論文では、すべての格付け行為に偏りなくポイントなどのインセンティブをもたせて収集した場合は、ランダムな順序( $T = \infty$ )で格付けされるものと仮定し、それを基準として格付け寄与度に基づいた収集の有効性を検証した。具体的な温度パラメータとしては、 $T = \{0.05, 0.1, 0.2\}$ のそれぞれで実験を行った。

なお、本論文では、格付けサンプルとして正解が得られている全格付け数に対して、シミュレーションにおいて推薦システムに入力された格付けの割合を格付けの「収集率」と定義する。MAEは、正解が得られている格付けサンプルの中で、まだ推薦システムに入力されていない格付けに基づき計算される。

## 5.2 エージェントモデルに基づく実験結果

エージェント数1000、アイテム数1000とし、全100万個の格付けを格付けサンプルとして検証実験を行った。全く格付けが入力されていない状態から、20万個の格付け(収集率20%)が入力されるまで推薦と格付けを繰り返し、MAEの推移を検証した。

図3は、嗜好ベクトル $\mathbf{p}_u$ の分布を一様分布で与え、

$T = \{0.05, 0.1, 0.2\}$ とした場合の結果である。ただし、縦軸はランダム選択( $T = \infty$ )での $MAE_{T=\infty}$ に対するMAEの改善率 $(MAE_{T=\infty} - MAE)/MAE_{T=\infty}$ を示す。つまり、値が正であれば、格付け寄与度に基づく収集がランダムな収集よりも推定精度が良いことを示している。また、横軸は収集率である。

格付け寄与度 $c_{u,i}$ は各格付けに対して独立に計算される。しかし $c_{u,i}$ は、ユーザ $u$ をneighborsに含むすべてのユーザに対する格付け寄与度の合計値である。したがって、より多くのユーザのneighborsに含まれるユーザは、どのアイテムの格付け寄与度も比較的高い値になりやすい。つまり、比較的多くの格付けを入力したユーザは他のユーザのneighborsに含まれやすく、その影響で他の格付け寄与度も高くなる。したがって、格付け寄与度に基づいた選択では、 $T$ が小さいほどユーザごとの格付け回数に差が生じる。格付け回数の多いユーザは、他の多くのユーザの格付け推定に寄与するため、推薦システムにおける中心的なユーザといえる。

図3によると、 $T = 0.2$ では、改善率がほぼ0であり、推定精度はランダム推薦に近い。一方、 $T = \{0.05, 0.1\}$ とした場合、収集率が約8%以下では、格付けの改善率が正の値となり、無策に格付けを促してランダムに収集するより、寄与度に基づく収集が有効であることを示している。これは、比較的数量多く格付けを行う中心的なユーザの存在が、多くの格付け推定に寄与し、推定精度を向上させていると考えられる。

一方、収集率が10%以上では、改善率が負の値である。ランダムな選択は、収集率が一定以上であれば、各エージェントから比較的偏りなく格付けを収集でき、網羅的な推定が可能である。しかし、格付け寄与度に基づく収集では、ユーザごとの格付け回数の差が原因となり、収集率が高くても推定できない格付けが多く存在し、改善率の低下につながる。もし、多くの格付けの収集が見込める場合は、格付け寄与度に基づく収集においても、特定のユーザの格付けに偏ることなく、全体から格付けを収集すべきであると考えられる。

したがって、最終的に得られる格付け数が予測できれば、どの程度格付けの寄与度を考慮すべきか判断可能である。しかし、運用時において、格付けはユーザの負担であり、実際の格付けの数は格付け行列全体の1%以下ともいわれている[2],[16]。つまり、より少ない格付けから高い推定精度を実現することが求められる。

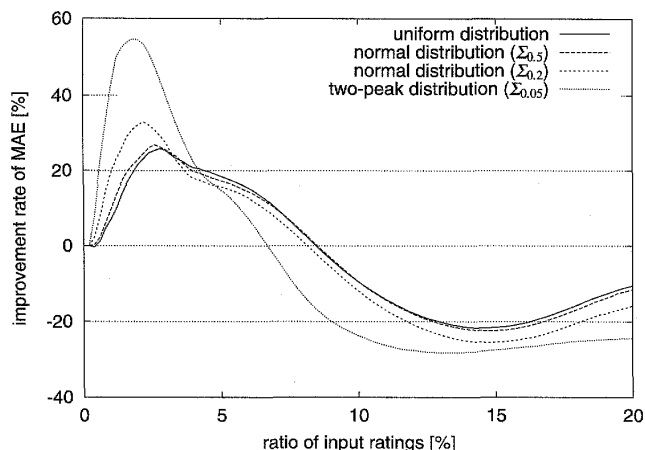


図4 エージェントモデルを用い、 $T = 0.1$ 、 $\mathbf{p}_u$ を各分布で与えたときのランダム選択に対する MAE の改善率

Fig. 4 Improvement rate of MAE over random selection by using agent model in setting  $T = 0.1$  and  $\mathbf{p}_u$  is generated based on each distribution.

次に、嗜好ベクトル  $\mathbf{p}_u$  を各分布で与えた場合の実験結果を図4に示す。ただし、ボルツマン選択の温度パラメータ  $T = 0.1$  とした。

図4より、 $\mathbf{p}_u$  をいずれの分布で生成しても、格付けの収集率が低ければ、MAEはランダム推薦より改善されている。また、嗜好ベクトル  $\mathbf{p}_u$  を一様分布で与えた場合と比較して、正規分布など偏りが存在する方が、格付けの収集率が低い段階での改善率が高くなっている。特に、二つの頂点をもつ分布では、その他の分布と比較して改善率が高い。これは、エージェントの嗜好が二つに分かれていることで、中心的なユーザの格付け推定への影響力が増し、格付けの寄与度の効果が大きくなったためと考えられる。

一方で、実運用を考慮した場合、格付け可能なアイテム数がユーザごとに限られている状況も想定される。そこで、各ユーザが格付けできるアイテム数に制限を設けた状況での実験を行った。詳細は割愛するが、格付けできるアイテム数の制約によって、中心的なユーザの影響が弱くなり、改善率は全体的に低下したものの、格付け寄与度に基づく収集がランダムな収集と比較して有効であることには変わりなかった。

### 5.3 実データに基づく実験結果

次に、実データに基づく検証実験の結果について述べる。実験で用いたデータは、MovieLensによるテストデータ<sup>(注4)</sup>であり、943名のユーザが1682タイトルの映画を5段階で格付けした10万個の格付けサンプルである。全く格付けが入力されていない状態から、

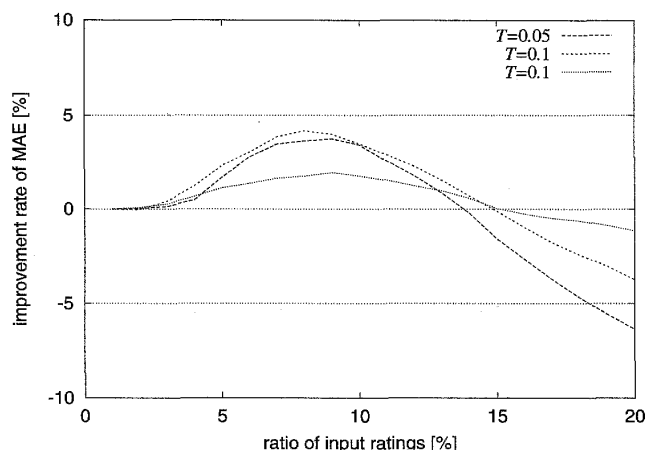


図5 実データに基づき、 $T = \{0.05, 0.1, 0.2\}$  としたときのランダム選択に対する MAE の改善率

Fig. 5 Improvement rate of MAE over random selection based on real data in setting  $T = \{0.05, 0.1, 0.2\}$ .

2万個の格付け（収集率20%）が入力されるまで推薦と格付けを繰り返し、MAEの推移を検証した。ボルツマン選択の温度パラメータは  $T = \{0.05, 0.1, 0.2\}$  として実験を行った。シミュレーション結果を図5に示す。

実データに基づく結果からも、格付けの収集率が一定以下であれば、格付け寄与度に基づく格付けの収集によって推定精度が改善することが明らかとなった。エージェントモデルに基づく検証結果と比べ、全体的な改善率は低くなっているが、これは理想化されたエージェントモデルと比較して、そもそも全体的な推定精度が低く、相対的に改善率も下がるためである。したがって、前述のとおり、実運用では低い収集率で高い推定精度を実現することが重要であり、格付けの寄与度に基づいた格付け収集の有効性が示された。

## 6. その他のCFへの応用

CFに基づく推薦アルゴリズムは多様化している。アイテム間CFは、格付けをもとにアイテム間の類似性を計算し、推薦する代表的な手法である[17]。基本的な考え方はユーザ間CFと同じであるが、ユーザ間類似度の代わりに、アイテム間類似度を用いる点異なる。類似度の計算方法は、ユーザ間類似度と同様、ピアソン相関係数やコサイン距離が用いられる。アイテム間類似度に基づき、未入力格付けを推定し、推定値が最も高いアイテムを推薦する。つまり、ユーザ

(注4) : <http://www.grouplens.org/node/73#attachments>

間 CF との本質的な違いは、格付け行列の行（ユーザ）に着目して推薦するか、列（アイテム）に着目して推薦するかの違いである。したがって、アイテム間 CF での格付け寄与度も、本論文で提案した方法と同様に定義でき、格付け寄与度に基づいた収集により、推定精度の効率的な向上が実現できると考えられる。

また、ユーザ間 CF とアイテム間 CF を統合する手法も提案されている [18]。基本的には、ユーザ間 CF に基づく推定値と、アイテム間 CF に基づく推定値の加重和を推定値とする方法である。そのため、両者を融合する場合も、本論文と同様に定義可能であると考えられる。今後は、より統合的な手法にも適用するための、格付け寄与度の拡張が必要である。

## 7. お 結 語

協調フィルタリングに基づく推薦システムは、ユーザの格付けをもとに推薦結果を決定し、ユーザはその推薦結果に影響を受けながら次に格付けするアイテムを選択するため、相互作用的である。そのため、推薦システムは、推薦の結果得られる格付けをフィードバックとしてその有効性を高めていく。一般的に、ユーザから得た格付けの数が多いほど推薦精度は向上するため、実際のサイトでは、メールで格付けを促したり、特典を設けるなどの工夫をしている。しかし、それにはコストが必要であり、推定精度向上に寄与する格付けを重点的に収集した方が効率的である。

本研究では、未入力の格付けが仮に得られた場合、システム全体の推定精度に与える寄与の割合を、格付け寄与度と定義し、定式化を行った。また、格付け寄与度に基づいた格付けの収集により、効率的に推定精度を向上させる方法を提案し、エージェントモデルと実データの双方でその有効性を検証した。実データは、様々な要因が複雑に関連した結果であり、それだけでは一般的とはいえないが、本論文は、単純化したエージェントモデルによる検証実験も含めて格付け寄与度の有効性を検証した結果として価値ある知見を与えるものである。

検証実験の結果、格付け寄与度の利用により、効率的に推定精度を向上できることが明らかとなった。

## 文 献

- [1] P. Resnick, and H. Varian, "Recommender systems," *Commun. ACM*, vol.40, no.3, pp.56-58 1997.
- [2] J.B. Schafer, J.A. Konstan, and J. Riedl, "E-Commerce recommendation applications," *Data Mining and Knowledge Discovery*, vol.5, pp.115-153 2001.
- [3] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: Item-to-item collaborative filtering," *IEEE Internet Computing*, Jan. / Feb. 2003.
- [4] 山下晃弘, 川村秀憲, 飯塚博幸, 大内 東, "推薦と格付けの相互作用に基づく推薦アルゴリズムの性能評価," *情処学論, 数理モデル化と応用 (TOM)* vol.2, no.2, pp.47-57, 2009.
- [5] J.S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," *Proc. 14th Conference on Uncertainty in Artificial Intelligence (UAI'98)*, pp.43-52, 1998.
- [6] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: An open architecture for collaborative filtering of netnews," *Proc. Conf. on Computer Supported Cooperative Work*, pp.175-186, 1994.
- [7] L.H. Ungar and D.P. Foster, "Clustering methods for collaborative filtering," *Proc. Workshop on Recommendation System at the 15th National Conf. on Artificial Intelligence*, 1998.
- [8] D.M. Pennock, E. Horvitz, S. Lawrence, and C.L. Giles, "Collaborative filtering by personality diagnosis: A hybrid memory-and model-based approach," *Proc. Sixteenth Conference on Uncertainty in Artificial Intelligence (UAI)*, 2000.
- [9] D. Goldberg, D. Nichols, B.M. Oki, and D.B. Terry, "Using collaborative filtering to weave an information tapestry," *Commun. ACM*, vol.35, no.12, pp.61-70, 1992.
- [10] J.L. Herlocker, J.A. Konstan, L.G. Terveen, and J.T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Trans. Inf. Syst.*, vol.22, no.1, pp.5-53, 2004.
- [11] J.L. Herlocker, J.A. Konstan, A. Borchers, and J. Riedl, "An algorithmic framework for performing collaborative filtering," *Proc. 22nd annual international ACM SIGIR Conference on Research and Development in Information Retrieval*, pp.230-237, Aug. 1999.
- [12] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowledge and Data Engineering*, vol.17, no.6, pp.734-749, 2005.
- [13] 平山巧馬, 小柳 滋, "協調フィルタリングにおける相関係数法の予測性能向上," *信学論 (D)*, vol.J90-D, no.2, pp.223-232, Feb. 2007.
- [14] J.L. Herlocker, J.A. Konstan, and J. Riedl, "An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms," *Information Retrieval*, vol.5 no.4, pp.287-310, 2002.
- [15] H. Ma, I. King, and M.R. Lyu, "Effective missing data prediction for collaborative filtering," *Proc. 30th Annual International ACM SIGIR Conference*



on Information Retrieval, pp.39–46, 2007.

- [16] 神脇敏弘, “推薦システムのアルゴリズム (1)~(3),” 人工知能誌, vol.22, no.6~vol.23, no.2, 2007, 2008.
- [17] B. Sarwar, G. Karypis, J.A. Konstan, and J. Riedl, “Item-based collaborative filtering recommendation algorithms,” Proc. WWW’ 01, pp.285–295, 2001.
- [18] J. Wang, A.P. Vries, and M.J.T. Reinders, “Unifying user-based and item-based collaborative filtering approaches by similarity fusion,” Proc. 29th ACM SIGIR Conference on Information Retrieval, pp.501–508, 2006.

## 付 録

### $\hat{r}_{a,i} - \hat{r}'_{a,i}$ の導出

格付けの推定値  $\hat{r}_{a,i}$  は, 式 (2) で与えられる. ここで, ユーザ  $u'$  がアイテム  $i$  に対して格付け  $r_{u',i}$  を新たに入力したと仮定すると, 格付けの推定値  $\hat{r}'_{a,i}$  は次の式で計算される.

$$\hat{r}'_{a,i} = \bar{r}_a + \frac{\sum_{u \in U_i \cap S(a)} \text{sim}(a,u)(r_{u,i} - \bar{r}_u) + \text{sim}(a,u')(r_{u',i} - \bar{r}_{u'})}{\sum_{u \in U_i \cap S(a)} |\text{sim}(a,u)| + |\text{sim}(a,u')|}$$

また, 式 (2) は以下のように変形できる.

$$\hat{r}_{a,i} = \bar{r}_a + \frac{\sum_{u \in U_i \cap S(a)} \text{sim}(a,u)(r_{u,i} - \bar{r}_u)}{\sum_{u \in U_i \cap S(a)} |\text{sim}(a,u)|}$$

$= \bar{r}_a$

$$\begin{aligned} & \sum_{u \in U_i \cap S(a)} \text{sim}(a,u)(r_{u,i} - \bar{r}_u) \times \left( 1 + \frac{|\text{sim}(a,u')|}{\sum_{u \in U_i \cap S(a)} |\text{sim}(a,u)|} \right) \\ & + \frac{\sum_{u \in U_i \cap S(a)} |\text{sim}(a,u)| + |\text{sim}(a,u')|}{\sum_{u \in U_i \cap S(a)} |\text{sim}(a,u)| + |\text{sim}(a,u')|} \\ & = \bar{r}_a + \frac{\sum_{u \in U_i \cap S(a)} \text{sim}(a,u)(r_{u,i} - \bar{r}_u) + (\hat{r}_{a,i} - \bar{r}_a) |\text{sim}(a,u')|}{\sum_{u \in U_i \cap S(a)} |\text{sim}(a,u)| + |\text{sim}(a,u')|} \end{aligned}$$

得られた式より,  $\hat{r}_{a,i} - \hat{r}'_{a,i}$  を計算する.

$$\hat{r}_{a,i} - \hat{r}'_{a,i} = \frac{(\hat{r}_{a,i} - \bar{r}_a) |\text{sim}(a,u')| - \text{sim}(a,u')(r_{u',i} - \bar{r}_{u'})}{\sum_{u \in U_i \cap S(a)} |\text{sim}(a,u)| + |\text{sim}(a,u')|}$$

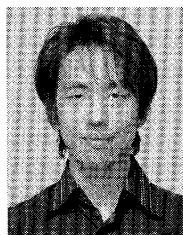
$$= \begin{cases} \frac{\text{sim}(a,u') \{ (\hat{r}_{a,i} - \bar{r}_a) - (r_{u',i} - \bar{r}_{u'}) \}}{\sum_{u \in U_i \cap S(a)} |\text{sim}(a,u)| + |\text{sim}(a,u')|} & (\text{sim}(a,u') \geq 0) \\ \frac{\text{sim}(a,u') \{ -(\hat{r}_{a,i} - \bar{r}_a) - (r_{u',i} - \bar{r}_{u'}) \}}{\sum_{u \in U_i \cap S(a)} |\text{sim}(a,u)| + |\text{sim}(a,u')|} & (\text{sim}(a,u') < 0) \end{cases}$$

ここで, 式 (3) により,

$$\hat{r}_{a,i} - \hat{r}'_{a,i} = \begin{cases} c_{u',i}(a,i) \{ (\hat{r}_{a,i} - \bar{r}_a) - (r_{u',i} - \bar{r}_{u'}) \} & (\text{sim}(a,u') \geq 0) \\ c_{u',i}(a,i) \{ -(\hat{r}_{a,i} - \bar{r}_a) - (r_{u',i} - \bar{r}_{u'}) \} & (\text{sim}(a,u') < 0) \end{cases}$$

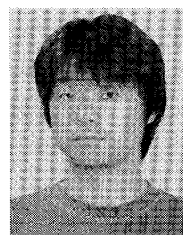
よって導出された.

(平成 21 年 1 月 14 日受付, 5 月 11 日再受付)



山下 晃弘

2006 北大・工・情報工学卒, 2008 同大学院情報科学研究科修士課程了. 同年同大学院情報科学研究科博士課程入学, 同年より日本学術振興会特別研究員 (DC1), 現在に至る. 推薦システム, マルチエージェントシステム, 観光情報学に関する研究に興味をもつ. 情報処理学会, 人工知能学会, 観光情報学会各学生会員.



川村 秀憲 (正員)

1996 北大・工・情報工学卒, 2000 同大学院工学研究科博士後期課程期間短縮了. 同年同大学院工学研究科助手. 2004 同大学院情報科学研究科助手, 2006 同助教授, 現在に至る. 飛行船ロボット, マルチエージェントシステム, 複雑系工学, 観光情報学等の研究に従事. 博士 (工学). 情報処理学会, 人工知能学会, 日本オペレーションズ・リサーチ学会, 観光情報学会各会員.



鈴木 恵二

1988 北大・工・精密工学卒, 1993 同大学院工学研究科博士後期課程了. 同年同大学工学部助手, 助教授, 2000 公立はこだて未来大学助教授. 2004 同大学教授. 2008 北海道大学大学院情報科学研究科教授, 現在に至る. マルチエージェントシステム, 複雑系工学, 観光情報学等の研究に従事. 博士(工学). 情報処理学会, 人工知能学会, 日本オペレーションズ・リサーチ学会, 観光情報学会各会員.



大内 東

1968 北大・工・応用物理卒, 1974 同大学院工学研究科博士後期課程了. 同年同大学工学部助手, 助教授, 1989 同大学院工学研究科教授. 2004 同大学院情報科学研究科教授, 現在に至る. 飛行船ロボット, DNA コンピューティング, マルチエージェントシステム, 医療システムなどの研究に従事. 博士(工学). 情報処理学会, 日本オペレーションズ・リサーチ学会, 医療情報学会, 観光情報学会各会員.