

テキスト情報を対象とした ハイブリッド型情報推薦システムにおける擬似投票方式

宇田 隆 幸^{†,†††} 藤井 敦^{††} 石川 徹 也^{††}

情報推薦システムは、ユーザの嗜好に関する情報を用いて新規情報に対する嗜好を予測し、嗜好に合うアイテムを提示する。協調フィルタリングに基づくシステムは、アイテムの内容解析をせず、他のユーザからの評価（アイテム-ユーザ評価マトリクス）を用いて推薦アイテムを決定する。対象ユーザと嗜好の似たユーザが好むアイテムは優先され、嗜好の異なるユーザが好むアイテムの優先度は下がる。ユーザがアイテムを選ぶたびに嗜好情報が更新されるため、検索キーワードの入力は必要ない。しかし、評価値疎ら問題や再生起問題により、未評価の（推薦対象にならない）アイテムが多く存在する。本論文では、ユーザの評価とアイテム間の類似度を併用して推薦対象アイテム数を増やす「擬似投票方式」を提案し、地域新聞の記事を配信するウェブサイトでは評価実験を行い、その有用性を示す。

A Pseudo-voting Method in the Hybrid Recommendation System for Text Information

TAKAYUKI UDA,^{†,†††} ATSUSHI FUJII^{††} and TETSUYA ISHIKAWA^{††}

Recommender system utilizes user profiles to predict user preference for unseen information items, and presents preferable items. The collaborative filtering (CF) method does not analyze the content of items, but utilizes user rating (an item-user matrix) to determine recommending items. CF-based systems (do not) favor items rated highly by the users whose preference is (not) similar to that of a target user. Users can update their profiles by selecting items and do not need to submit search keywords. However, due to the sparsity and recurring startup problems, many items are not rated and thus cannot be recommended to users. In this paper, we propose “pseudo-voting method”, which increases the number of rated items by integrating user rating and content-based item similarity. The paper also show the effectiveness of our method by means of experiments in a website serving an article of a local newspaper.

1. はじめに

ユーザの嗜好に関する情報を用いて新規情報に対する嗜好を予測し、嗜好に合うアイテムを提示する情報推薦システムが研究されてきた。情報推薦システム^{12),21)}は、キーワード入力が必要であり、欲しいアイテムを選ぶだけで推薦の正確性が向上するので、ユーザのキーワード入力における苦手意識克服効果を望める。現在運用中の情報推薦システムは、研究において評価実験が行われているシステム^{1),5),7),12),15)}や、

Amazon.com^{10),14)}などに代表される電子商取引サイトが中心であり、ユーザ向けの情報を対象としたシステムはまだない。また、WWWは世界中の情報が対象であり、ユーザの行動範囲に的を絞った情報を対象とした情報推薦システムは、著者らが知る限りない。

関連研究^{7),12)}では、情報推薦システムのアルゴリズムの1つである協調フィルタリングアルゴリズムが扱われてきた。協調フィルタリングは、推薦アイテムを決定するとき「アイテム-ユーザ」評価マトリクスを参照し、推薦対象ユーザと嗜好の類似するかまたは正反対のユーザからの評価を参考にして、推薦対象ユーザが評価していないアイテムの推薦予測値を計算する。そのため、協調フィルタリングでは、情報（アイテム）数およびユーザ数が大規模になるほど、ユーザの評価値を記録するための「アイテム-ユーザ」評価マトリクスの未評価セルが増加する。いわゆる「評価値疎ら問題（Sparsity problem）」と呼ばれる現象で、推薦

† 筑波大学大学院図書館情報メディア研究科博士後期課程
Doctoral Program in Library, Information and Media Studies, University of Tsukuba

†† 筑波大学図書館情報メディア研究科
Graduate School of Library, Information and Media Studies, University of Tsukuba

††† 株式会社ネオジェイエスケー
NEO JSK Corporation

表 1 情報推薦システム(広義)の分類
Table 1 Classification of recommender systems (wide sense).

方式	利用する情報	概要	特徴
チェックボックス方式	ユーザが登録した嗜好情報	ユーザは推薦アイテムを受け取るために、自身の嗜好を前もって登録する。システムは登録された嗜好情報に基づいてアイテムを推薦する。	システムの構築は容易である反面、ユーザに意外性のあるアイテムが推薦される頻度は低い。
ルールベース方式	サイト運営者が登録した推薦ルール	サイトの管理者が、ユーザの評価したアイテム(例、閲覧や購入)と推薦したいアイテムとの対応ルールを作成する。システムは登録されたルールとユーザの評価したアイテムを照合することでアイテムを推薦する。	システムはサイト運営者の意志を反映することができる。ユーザ数やアイテム数が増加すると、ルール作成や保守が困難になる。
狭義の情報推薦方式	ユーザの行動履歴	システムはユーザの行動履歴から個々のユーザの嗜好を推測し、推薦アイテムを自動生成する ^{1), 5), 8), 11), 12)} 。	ユーザの行動履歴やアイテム数が増加すると、推薦アイテムを生成するためのオンライン計算コストが増大する。

アイテムを提示できないユーザが増加する。また、新しいアイテムは、そのアイテムがユーザに評価されるまで時間を要し、いわゆる「再生起問題(Recurring startup problem)」が発生し、新しいアイテムは推薦対象とならない。たとえば地域情報を配信するサイトのように、アイテムの内容以外でユーザコミュニティが形成されている場合、ユーザは特定のコミュニティ向けに配信されたアイテムのみを評価するので、推薦対象ユーザと異なるコミュニティに属するユーザの評価値は、たとえ内容的に類似するアイテムであっても、その評価値を利用できない。それゆえ、評価値疎ら問題および再生起問題がよりいっそう深刻になる。

一方、ユーザの評価したアイテムと類似するアイテムにユーザの評価値を付与することでアイテムの内容解析と協調フィルタリングを組み合わせる方式が研究²³⁾されている。しかし、この研究で対象とするアイテムは画像であり、テキストを対象とする研究はまだない。

このような背景から、本研究では協調フィルタリングのかかえる評価値疎ら問題と再生起問題を解消することを目的にテキストの内容解析をともなう擬似投票方式を提案する。そして、地域情報が配信されている実際の Web ページのデータを用い、評価実験を行う。

以下、2章で関連研究について整理し、3章で本研究において提案する擬似投票アルゴリズムを提示し、4章で評価実験および考察を行う。最後に、5章でまとめを述べる。

2. 関連研究の検討

2.1 利用する情報による分類

情報推薦システムは、一般のユーザを対象とし、Web サイトですでに稼働していて、電子商取引では売上促進に効果が出ている¹⁰⁾。推薦情報を生成するときを利用する情報に着目すると表 1 に分類できる²²⁾。各方

式は、単独で実装されることや、各々の長所を生かし短所を補う目的で複数の方式を組み合わせられて実装されることがある。

表 1 の各方式を検討すると、狭義の情報推薦方式は、以下の点で他の方式より優れている。

- マーケティング専門家によって手動で行われる準備作業が不用である(最小限度人的資源対応)。
- ユーザ数やアイテム数の増加に追従できる(拡張性)。
- 意外性のあるアイテムを推薦できる(有効情報推薦対応)。
- 推薦アイテムの予測計算を自動で行える(運用コスト削減効果)。

以降は、狭義の情報推薦方式を情報推薦システムと同義として扱う。

2.2 アルゴリズムによる分類

情報推薦システムはフィルタリングアルゴリズムを利用する。推薦アイテムを提示するとき利用する情報に着目すると、フィルタリングアルゴリズムは表 2 のように分類できる。

表 2 の各方式を検討すると、相関係数法を用いた協調フィルタリング方式は、以下の点で他の方式より優れている。

- システム側の事由で(例、計算量過多の防止)予測精度を低下させない^{1), 3)}(ユーザ指向)。
- 意外性のあるアイテムを推薦できる(有効情報の推薦)。
- ユーザはマウスクリックするだけでよくキー入力が必要でない(簡易操作性)。

以下、相関係数法を用いた協調フィルタリング方式(協調フィルタリングアルゴリズム)について整理する。

2.3 協調フィルタリングアルゴリズム

相関係数法による協調フィルタリング方式は、嗜好の類似するまたは嗜好が逆である他のユーザの評価投

表 2 情報推薦システム(狭義)の分類
Table 2 Classification of recommender system methods.

方式	概要	特徴
内容ベース方式 ¹⁵⁾	ユーザの評価したアイテムや閲覧したアイテムが与えられると、それと類似する検索式を自動生成することにより推薦対象アイテムを探し出す。このため、検索ベースアルゴリズムとも呼ばれる。	評価履歴が少ない間は容易に検索式を生成できる反面、評価履歴数が増加すると検索式の生成が困難となる。他のユーザの意見を参考にしないので、意外性のある情報の推薦は難しい。ただし、意味的学習辞書を構築することで意外性のある情報を推薦するアルゴリズムも研究されている ¹¹⁾ 。
協調フィルタリング方式(CF方式)	推薦対象アイテムを提示するとき、嗜好が類似するまたは正反対である他のユーザによる評価履歴を参考にする。協調フィルタリングを分類すると、(a)相関係数法、(b)クラスタモデル、(a)(b)の組み合わせ型や(a)(b)の亜種 ^{11), 5), 13)} が存在する。	(a)(b)いずれの場合も、評価履歴の蓄積量が少ない間、ユーザによる評価履歴が疎らであるため、推薦可能なアイテムやユーザの網羅性に欠ける。または、推薦アイテムを提示できる場合でも、推薦予測の正確性は低い。
	(a)相関係数法 ^{12), 15)} N次元の「アイテム-ユーザ」評価マトリクスを用いてユーザ間の嗜好類似度に基づいて推薦対象アイテムを計算する。	計算量は、(a)に比べ(a)が優位である。 推薦予測の正確性は、(b)に比べ(a)が優位である ¹⁰⁾ 。
	(b)クラスタモデル ^{2), 16)} 類似するユーザを多数のセグメントに分類し、同じセグメントに属する他のユーザの意見を参考にして推薦アイテムを計算する。	(b)固有の問題として、ユーザの分類問題が研究されている。
ハイブリッド方式	内容ベース方式と協調フィルタリング方式の短所を補いかつ長所を生かす目的で、両方式を結合したタイプ ^{11), 5)} 。ユーザの評価履歴が少ない間は内容ベース方式を用いて推薦アイテムを提示し、評価履歴蓄積量が閾値を越えると協調フィルタリング方式で推薦アイテムを提示する。	推薦対象ユーザや推薦対象アイテムの網羅性が高い。長所短所は、内容ベースフィルタ方式と協調フィルタリング方式のいずれもあわせ持つ。そのため、両方式の結合技法、すなわち最適な切替タイミング(閾値)は、ユーザやアイテム固有の特性によって変化するので、閾値調整問題がある。

票値を参考にして推薦アイテムを予測計算するために「アイテム-ユーザ」評価マトリクスを持つ協調フィルタリングアルゴリズムを図1(上)に示す。図1において、 m はユーザ数、 n はシステムが蓄積しているアイテム数である。全ユーザを $U = \{u_1, u_2, \dots, u_m\}$ とし、全アイテムを $I = \{i_1, i_2, \dots, i_n\}$ とする。ユーザ u_j が評価投票したアイテム集合を $I_{u_j} \in I$ とする。ユーザ u_j の評価投票したアイテム k に対する評価投票値を $r_{j,k}$ とする。ユーザに対して推薦を行う場合、協調フィルタリングを行うとは、ユーザ $\{u_1, u_2, \dots, u_m\}$ のアイテム $\{i_1, i_2, \dots, i_n\}$ に対する評価 $r_{j,k}$ から、アイテム $i_j \notin I_{u_j}$ (ユーザ u_j が評価投票していないアイテム)の評価 $r_{j,k}$ の予測値 $P_{i,k}$ (例: 図1の X) を算出することである。ただし、ユーザ u_j に対して推薦されるアイテムは予測値の中で嗜好度の高い上位 N アイテムとする。代表的な $P_{i,k}$ の算出法は、下式である。

$$P_{i,k} = \bar{r}_i + \frac{\sum_m \text{sim}(u_i, u_m) \cdot (r_{m,k} - \bar{r}_m)}{\sum_m \text{sim}(u_i, u_m)} \quad (1)$$

ここで、 $P_{i,k}$ はユーザ u_i に対するアイテム k への推

評価投票値は、ユーザが閲覧したアイテムに対する嗜好度を数値化したものである。本研究では、嗜好度が高いアイテムに投票値3、中程度のアイテムに2、好まないアイテムに1を割り当てた。

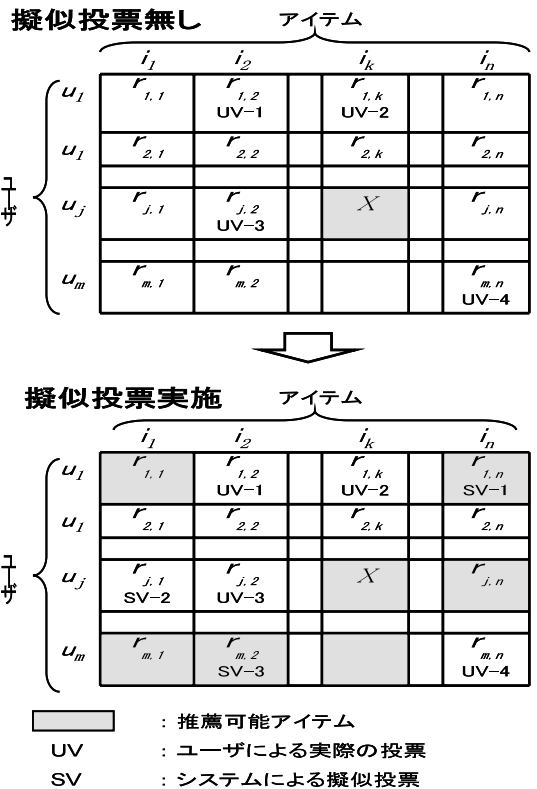


図1 擬似投票の効果
Fig. 1 Effectiveness of pseudo voting method.

薦スコアの予測値, \bar{r}_i はユーザ u_i が行ったすべての評価投票の平均投票値, Σ_m はアイテム k へ評価投票を行ったユーザ集合 $sim(u_i, u_m)$ はユーザ u_i とユーザ u_m の嗜好の類似度, $r_{m,k}$ はユーザ u_m がアイテム k へ行った評価投票値, \bar{r}_m はユーザ u_m が行ったすべての評価投票の平均投票値とする.

ユーザ間の類似度 $sim(u_i, u_m)$ は下式を用いる.

$$sim(u_i, u_m) = \frac{\sum_j (r_{i,j} - \bar{r}_i) \cdot (r_{m,j} - \bar{r}_m)}{\sqrt{\sum_j (r_{i,j} - \bar{r}_i)^2 \cdot (r_{m,j} - \bar{r}_m)^2}} \quad (2)$$

ここで, Σ_j はユーザ u_i とユーザ u_m の双方が評価投票を行っているアイテム集合, $r_{i,j}$ はユーザ u_i がアイテム j へ行った評価投票値, $r_{m,j}$ はユーザ u_m がアイテム j へ行った評価投票値である.

2.4 協調フィルタリングアルゴリズムの問題点

相関係数法による協調フィルタリングアルゴリズムは、「アイテム-ユーザ」評価マトリクスを維持することから、関連研究において以下の問題点が指摘されている.

多くの情報領域において、アイテム数は利用頻度の高いユーザにおいても吸収できなほど存在しており、すべての情報の評価を含む「アイテム-ユーザ」評価マトリクスは巨大になる反面、その評価値のセルはスパースになる^{5),9)}。相対的に密集したフィルタリング領域でも、98~99%が抜け¹⁰⁾、協調フィルタリング予測の基礎を形成するために必要となる十分な情報を見つけないことが難しくなる (Sparsity Problem: 評価値疎ら問題)。

また、新聞記事のように毎日発行される新着アイテムは、ユーザに評価されるまでに時間を要する。このため、推薦対象とならないアイテムが多出し「アイテム-ユーザ」評価マトリクスのアイテム列を占有する⁹⁾。評価値を与えられていないアイテムは、推薦スコアを計算できないので推薦対象アイテムからはずれる (Recurring Startup Problem: 再生起問題)。

評価値疎ら問題と再生起問題は、推薦アイテムの予測計算で使用する式 (1) と式 (2) に現れる集合 (Σ_m や Σ_j) を狭める点で問題が指摘されている。これら問題の悪影響として、評価実験 (後述) で示したとおり、推薦対象ユーザと推薦対象アイテムの網羅性が低下する。

関連研究において、評価値疎ら問題と再生起問題とを解消する目的で協調フィルタリングと内容ベースフィルタリングを組み合わせたハイブリッド方式が研

究されてきた。アイテムへの評価値が少ない間は内容ベースフィルタリングを用いてアイテムの推薦スコアを計算し、その後アイテムへの評価値が多くなったとき、協調フィルタリングを用いてアイテムの推薦スコアを計算する。しかし、ハイブリッド方式を用いた問題解消方式では、各方式の組合せ問題や方式の切替えタイミング問題が発生する。また、内容ベースフィルタリングは検索ベースで推薦情報を生成するため、意外性のあるアイテムを推薦することができない。そのうえ、推薦スコアの計算量が增大するため即時応答性が難しくなる。ただし、推薦スコアの計算量が增大する問題については、協調フィルタリングにクラスタモデルを採用することで計算量問題を解消する方式が提案されている^{2),10),16)}。

館村らは、ユーザの評価したアイテムと類似するアイテムにユーザの評価値を付与することでアイテムの内容解析と協調フィルタリングを組み合わせることで評価値疎ら問題と再生起問題とを解消するハイブリッド方式を提案している²³⁾。しかし、この研究は画像を対象としているため、アイテムの内容解析では、色・形・テクスチャの類似度を比較している。このため、アイテムの内容解析手法はテキストアイテムに適用することができない。

一方、Sarwar らは、評価値疎ら問題を解消させるためにアイテムに基づく協調フィルタリングを提案している¹³⁾。この方式は、ユーザ投票値を利用してアイテム間の類似度を計算している。そしてこの結果を推薦アイテム予測計算に利用しているため、アイテムの内容の類似度は対象にしていない。この結果、人気のあるアイテムが優先され、意外性のあるアイテムは推薦されない。

3. 本 研 究

問題点の項で示したとおり、評価値が疎らであると相関係数を用いて嗜好の類似度を計算することができないので、評価値疎ら問題が発生する。また、新規登録された直後のアイテムへは、どのユーザも評価投票を行っていないので、再生起問題が発生する。本研究では、伝統的な協調フィルタリングの推薦予測精度を低くおさえたうえで、これら2つの問題点が解消する

アイテムに基づく協調フィルタリングとは、図1(上)のマトリクスにおいて、列方向に着目した投票値の相関からアイテム間の類似度を計算し、推薦アイテムを提示する手法である。これに対して、行方向すなわちユーザ間の投票値の相関に着目する方式をユーザに基づく協調フィルタリングという。一般に、評価値が疎らである場合、推薦対象アイテムの網羅性は、アイテムに基づく CF の方が優位である。

方式を提案する．

3.1 擬似投票方式の提案

評価値疎ら問題と再生起問題とを同時に解消する方式として、我々は、「アイテム-ユーザ」評価マトリクスでユーザ評価値が未設定のセルに対し、システムが自動的に評価予測値を設定する方式を提案する．同提案方式を 図 1 (下) に示す．

(1) 評価値疎ら問題の解消方式

評価値が疎らであると相関係数を用いて嗜好の類似度を計算することができない．この問題を解消するために、ユーザが評価投票を行ったとき、システムは、①投票されたアイテムと類似するアイテムを見つけだし、②見つかったセル (図 1 (下) SV) がユーザによって評価投票されていない場合、そのセルにユーザ投票値と同じ値を投票する方式とする．

(2) 再生起問題の解消方式

新しいアイテムが登録されたとき、そのアイテムに対するユーザによる評価値は記録されていない．この問題を解消するために、新しいアイテムが登録されたとき、システムは、①新しく登録されたアイテムと類似するアイテムを見つけだし、②見つかったアイテムに対してユーザによる評価投票がされていた場合、そのセルにユーザ投票値と同じ値を投票する方式とする．

これらの処理方式を「擬似投票」と名称する．すなわち、擬似投票は、アイテムの内容に基づく類似性を利用した欠損値補足を目的としている．

擬似投票は、「あるアイテムを嗜好するユーザは、内容的に類似する他のアイテムも同じように嗜好する」という仮説に基づく解消方式である．

擬似投票アルゴリズムを以下に示す．

```

擬似投票 (UserId, ItemId_U, RealVoteVal) {
  For each (システムに蓄積されている全アイテム: ItemId_S) {
    If (類似判定 (ItemId_U, ItemId_S, 類似閾値) == 類似
      && ユーザが過去に評価していない (ItemId_S) == YES) {
      擬似投票値記録 (ItemId_S, RealVoteVal);
    }
  }
}

```

擬似投票を行うことにより、推薦スコア予測計算時にノイズ、すなわち、そのアイテムに対するユーザ評価値と、他の内容的に類似するアイテムにユーザが付けた評価値 (擬似投票値) との差異が発生する．このノイズ除去を行う機構として、推薦アイテムの計算処理後、上位 N アイテムを推薦するためにアイテムを並べ替える．並べ替え処理では、擬似投票履歴も併用参照して推薦されたアイテムの優先度を下げる．

3.2 擬似投票における類似判定基準

擬似投票の実施有無は、アイテム間の内容に基づく類似度を判定基準とする．本研究に先立ち、最適なアイテム間の類似度計算方式を決定するために、様々な類似度計算式を予備的実験²⁰⁾ で評価した．予備的実験では、類似度計算式の優劣比較の指標として、下式に示す MAE (mean absolute error) および計算処理時間の両方を評価した．

この節において、下式は、予備的実験で用いた MAE であり、ユーザによる評価投票値と擬似投票値間の平均絶対値エラーを意味する．

$$MAE = \sum_{j \in U} \frac{\sum_{k \in T_{u_j, s_j}} \frac{|r_{j,k} - \overline{sv}_{j,k}|}{N_k}}{N_j} \quad (3)$$

上式において、 j はアイテムへの評価投票を行ったユーザの集合で、 N_j は集合 j の要素数 (異なりユーザ数)、 k はユーザ j が評価したアイテムとシステムがユーザ j のために擬似投票したアイテムの積集合、 $r_{j,k}$ はユーザ j がアイテム k に与えた評価値で、 $\overline{sv}_{j,k}$ はシステムがユーザ j のアイテム k に対して行った擬似投票値の平均値、 N_k は積集 k の要素数 (アイテム数) である．なお、計算に先立ち、 $r_{j,k}$ と $\overline{sv}_{j,k}$ は、 $-1.0 \sim +1.0$ の値域に基準化した． MAE がゼロに近づくほど、ユーザによるアイテムへの評価値とシステムによる擬似投票値の嗜好パターンが類似していることを意味する．

予備的実験の結果から、Dice 係数方式が最も好成績であった．また、索引語の重み付けは IDF を用いないで TF のみで求める．これは、TF-IDF で計算する場合と比べ、推薦予測精度に大きな差がなく、かつ、TF のみで計算した方が計算量を少なくできるからである．

アイテム間の類似度計算方式は、予備的実験の結果最も好成績 (優劣比較方法は後述) を残した下式を用いる．

$$\text{Dice 係数方式 (Dice formulation)}^{18)} \\ S_{u,s} = \frac{2 \sum_{t \in T_{u,s}} (w_{u,t} \cdot w_{s,t})}{W_u^2 + W_s^2} \quad (4)$$

ここで、 $S_{u,s}$ はアイテム u とアイテム s の類似度、 T は異なり語の集合、 $t \in T_{u,t}$ における t はアイテム u とアイテム s の両方に含まれる索引語の集合である．

アイテム中から「日付情報、地域を特定する情報、電話やその他のアクセス情報」に関するフレーズを除去後、形態素解析を行った．索引語は、[名詞]、[名詞-接尾] が連続する場合、名詞を結合する処理を行うことで特徴語に近づくような配慮を施した．

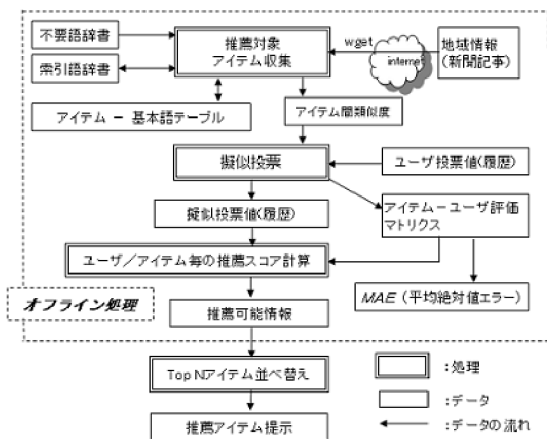


図2 システム全体図
Fig. 2 System overview.

$w_{u,t}$ と $w_{s,t}$ はアイテム u またはアイテム s に含まれる語 t の重みとする。ここで、 $w_{u,t} = 1 + \ln(f_{u,t})$ 、 $w_{s,t} = 1 + \ln(f_{s,t})$ とし、 $f_{u,t}$ と $f_{s,t}$ は、アイテム u またはアイテム s における語 t の出現頻度である。アイテムの長さは、 $W_u = \sqrt{\sum_{t=T_u} w_{u,t}^2}$ 、 $W_s = \sqrt{\sum_{t=T_s} w_{s,t}^2}$ とする。

類似判定基準は、 $S_{u,s}$ の値で判定する。類似性判定基準の閾値緩和（擬似投票件数増加）と正確性向上は、トレードオフの関係にある。そこで、類似性判定基準（閾値）を 0（アイテムは類似しない）から 1.0（アイテムは完全一致）に向けて変化させ、擬似投票件数と正確性とを観測した。この結果、擬似投票件数が急増するポイントと正確性が急低下するポイントを観測した。本研究では、同ポイントを類似度判定基準閾値とした²⁰⁾。

3.3 実装

処理概要を図2に示す。擬似投票は、3.1節で示した「ユーザが評価投票を行ったとき」と「新しいアイテムが登録されたとき」の両タイミングで実行する。ユーザ/アイテムごとの推薦スコア計算処理は、式(1)で示したアルゴリズムを用いて計算する。

一方、情報推薦システムはユーザに対して、短時間に多くの推薦アイテムを提示しなければならない。

相関係数法に基づく協調フィルタリングでは、推薦アイテムの予測計算時に完全な「アイテム-ユーザ」評価マトリクスを使用する。そのため、その計算量は $O(User_{max} \times Item_{max})$ の空間サイズに比例して増大する。しかし実際は、評価マトリクスは非常にスパースであるため $O(User_{max} + Item_{max})$ に近似する。ところが、擬似投票を行うため、本研究における空間

サイズは、 $O(User_{max} \times Item_{max})$ に近づく。そこで、擬似投票およびユーザ u_i に対するアイテム k への推薦スコア予測値 $P_{i,k}$ の計算と、上位 N アイテムの選択処理を分離することにした。具体的には前者をオフラインで実行し、後者をオンラインで実行する⁴⁾。この結果、特に計算処理量が問題となる前者をオンライン計算（メモリベース計算）から切り離せるので、応答速度の高速化とスケーラビリティを確保できる方式にした。

4. 評価実験

4.1 評価実験の方針

評価実験では、前もって①擬似投票の実施有無の判定基準を設定した^{20),21)}。この判定基準は、3.2節の方法で求めた類似性判定基準閾値(0.8)を用い、同値を超える場合に擬似投票を実施する。次に、実験データ収集後、②擬似投票を実施し¹⁹⁾、③協調フィルタリングに基づいてユーザごとアイテムごとの推薦予測スコアを計算し、④ユーザごとに上位 N 件のアイテムを決定する。ただし、実験対象のデータが地域性を強く含んだ情報であるため、ユーザ居住地に基づいて、居住地外のアイテムを除去する処理を行った。

評価用の測定値として、擬似投票を実施した件数を記録し、「アイテム-ユーザ」評価マトリクスの空塞率を計算し、ユーザ評価とシステム評価の絶対誤差を計算する。

4.2 実験データ（評価実験の準備）

(1) アイテムの標本

評価実験用データは、図3で示すサイトに蓄積されているアイテムのうち、2002年10月27日～2003年2月15日（12週間）の間、1つ以上の地域に配信されていて、各ユーザに閲覧される可能性があるアイテムを対象とし、そのうちの6%を標本抽出した。アイテムの標本数は2,424である。

(2) ユーザ評価履歴の標本

研究用ユーザ評価値投票履歴データは、図3で示すページに対し、ユーザによって直接評価投票された履歴のうち、2002年10月27日～2003年2月15日

サンケイリビング新聞社が提供している地域情報サイト「えるこみ」。同サイトでは、主に女性のために生活情報が提供されている。評価実験で使用したアイテムの種類は、イベント・教養&カルチャー・生涯学習・グルメなどの地域情報である。地域新聞紙と同様に、それぞれのアイテムは、そのアイテムの発生地と広域性に依りて、1つ以上の地域に配信されている。同サイトにおけるアイテムの配信地域は、首都圏エリアが18地区、大阪神戸エリアが13地区、国内のその他地域を42地区にそれぞれセグメント化している。http://www.lcomi.ne.jp/

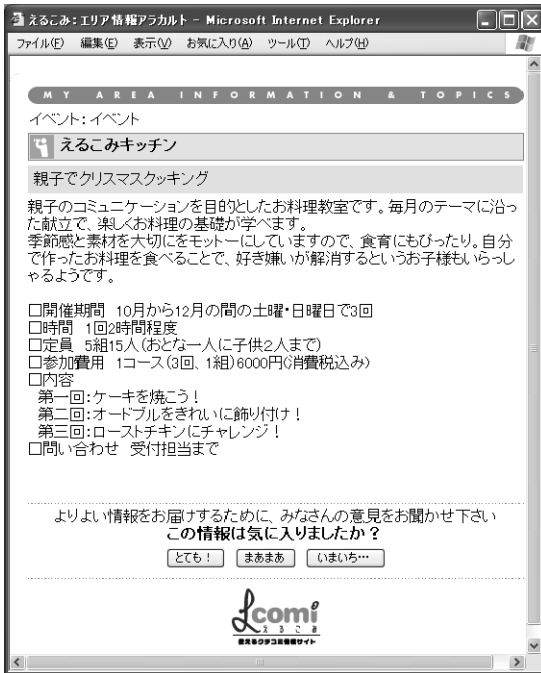


図 3 アイテムページの例
Fig. 3 Example item page.

(12 週間) に評価投票のあった 17,869 である。その後、アイテムの標本抽出にともない、評価実験対象外となったアイテムへの投票も除外したので、アイテム標本に存在するアイテムに対するユーザ評価投票履歴の標本数は 2,197 件である。なお、標本抽出前の投票者異なり数は 1,001 人で、ユーザ評価投票履歴標本における投票者の異なり数は 756 人である。

(3) 実験データ観察期間

ユーザ評価投票履歴の標本数は、時間経過とともに増加する。標本数が増加すると、嗜好度の相関係数が計算可能になるユーザペア数が増加する。また、ユーザペア内でどちらも評価しているアイテム数も増加する。同時に、情報推薦システムで利用されている性能評価指標の MAE 値も変化する。そこで、 MAE と「アイテム-ユーザ」評価マトリクスのスパース率の変化を観察するために、観察期間を分割した(表 3)。

4.3 実験結果

(1) 問題点の解消度合い

擬似投票の効果を示すため、ユーザによる現実の投票数と擬似投票数との比較を表 3 の「スパース性」欄に示した。擬似投票によって「アイテム-ユーザ」評価マトリクスの欠損値が設定され、その結果、表 3 の「スパース性」欄から、同マトリクスのスパース性が解消されていることが確認できる。同欄の「ユーザ投

票」および「擬似投票」列は、「アイテム-ユーザ」評価マトリクスのセルに投票値が設定されていない割合を示している。また、同欄の「解消率」は、「擬似投票」から「ユーザ投票」を減じた値を示しており、この値が擬似投票処理による効果である。

式 (2) の Σ_j は、ユーザ u_i とユーザ u_m の双方が評価投票を行っているアイテム集合を指す。このアイテム集合数(ユーザペア数)が増加するほど、推薦可能なユーザ数が増加する。このアイテム集合の平均要素数が増加すると、ユーザ間の嗜好の類似度計算結果において、その正確性が向上した。

実験開始から 12 週間経過後の推薦可能ユーザ数を測定した。擬似投票を実施後に協調フィルタリングを行った場合、665 人に何らかのアイテムを推薦可能であったのに対し、伝統的な協調フィルタリングの場合、73 人であった。

(2) 推薦予測精度

表 3 の「 MAE 」欄に、ユーザによる評価投票値のみの場合と、擬似投票も行った場合との MAE 比較を示した。表 3 の「 MAE 」欄から、システムによる擬似投票処理を行っても、情報推薦システムの最終評価、すなわち推薦アイテムの予測の正確性は、大きく低下しないことを確認できた。 MAE は下式によって算出した。

$$MAE = \frac{\sum_{i=1}^N |rating_i - prediction_i|}{N} \quad (5)$$

ここで、 $rating$ はユーザによる評価値を示し、 $prediction$ はシステムによる推薦スコアを示し、 N が $rating$ と $prediction$ のペア数を指す。なお、 MAE の計算対象とした評価投票ペアは表 4 に示した。本研究では、ユーザに対して推薦しないことをよしとするアイテムの予測は研究対象外であるため、システムが推薦したアイテムと、ユーザが「とても良い」または「いまいち」と評価したアイテムペアのみを対象とした。

(3) 考察

表 3 の「投票件数-増加率」欄に示したとおり、システムによる擬似投票処理によって、その処理を行わない場合と比較して、評価値が約 250% 増加した。この結果、式 (2) の Σ_j の集合要素数が増加した。このことから、嗜好の類似度を計算可能となるユーザ数が増加し、推薦可能なユーザの網羅性と推薦可能なアイテムの網羅性が高まることを確認できた。

また、擬似投票処理を実行することで「アイテム-ユーザ」評価マトリクスのスパース性は、約 27% 解消した。表 3 に示したとおり、この解消率はシステム

表 3 実験結果

Table 3 Experimental results.

期間 ID	観察期間	投票件数			スパース性			MAE		
		ユーザ	擬似	増加率	ユーザ投票	擬似投票	解消率	ユーザ投票	擬似投票	差分
1	～ 1 週間	91	24,765	272.1	99.9506%	98.7133%	1.2373	0.728	0.760	-0.032
2	～ 2 週間	251	68,622	273.4	99.9499%	96.4722%	3.4777	0.609	0.650	-0.041
4	～ 4 週間	602	161,020	267.5	99.9479%	91.8099%	8.1380	0.494	0.526	-0.032
8	～ 8 週間	1,256	323,529	257.6	99.9487%	83.8972%	16.0515	0.438	0.494	-0.056
12	～12 週間	2,197	550,914	250.8	99.9486%	73.1810%	26.7676	0.397	0.468	-0.071

表 4 表 3 における MAE の計算方法

Table 4 MAE calculation rule in Table 3.

No	ユーザ評価	システム評価	妥当性	計算対象
1	1:とても好き	1:上位 100 件	正解	対象
2	1:とても好き	上位 100 外	不正解	対象外
3	0:まあまあ	1:上位 100 件	—	対象外
4	0:まあまあ	上位 100 外	—	対象外
5	-1:いまいち	1:上位 100 件	不正解	対象
6	-1:いまいち	上位 100 外	正解	対象外

稼働期間が長くなるほど改善している。ところで、アイテムには有効期間がある。しかし、ユーザの嗜好は急激に変化しないと予想されるので、ユーザが過去に行った評価投票も有効であると考えられる。

一方、システムによる擬似投票処理を実施した場合、表 3 に示したとおり、情報推薦システムの性能評価指標として広く用いられている MAE が劣化する。すなわち、擬似投票を実施することによる「評価値疎ら問題と再生起問題の解消」と「MAE の劣化」は、トレードオフの関係にある。表 3 の期間 ID=12 の実験結果の場合、MAE の劣化は -0.071 である。-0.071 とは、ユーザにアイテムを 100 件推薦した場合、不正解アイテムがおよそ 3 件増加することに相応する。したがって、MAE の劣化はわずかであると考えられる。この劣化に対し、問題点が解消したことによって得られた効果、すなわち、推薦可能なユーザやアイテムの網羅性向上は顕著であり、この効果によって擬似投票の有効性が示された。

推薦スコア上位 100 件をスコア降順に並べてユーザに提示した場合、MAE 値に対する正解と不正解の割合は、MAE = 0.4 なら 80:20、MAE = 0.46 なら 77:23、MAE = 0.5 なら 75:25 である。ここで、正解とはユーザが「とても! (興味あり)」を示した場合を指し、不正解とは「いまいち... (興味なし)」を示した場合を指す。表 3 の期間 ID=12 で比較すると、従来の協調フィルタリングの場合 MAE は 0.397 であり、擬似投票を行うと 0.468 となり、MAE の差は -0.071 である。-0.071 とは、ユーザにアイテム 100 件を推薦した場合、不正解数がおよそ 3 アイテム増加することに相応する。

5. おわりに

「あるアイテムを嗜好するユーザは、内容的に類似する他のアイテムも同じように嗜好する」という仮説の妥当性が、評価実験により証明された。この仮説成立を前提条件とした「システムによる擬似投票アルゴリズム」は、協調フィルタリングにおける推薦可能アイテムと推薦可能ユーザの両方に対して、推薦予測の網羅性を高めることができた。そして、同アルゴリズムを搭載することで表面化する大規模化の困難性に関して、大部分の処理をオフラインで処理できるので、Web ページ閲覧時の応答速度において問題は発生しない。

今後は、協調フィルタリングにおいて推薦スコア予測計算を行うとき、ユーザの評価値以外でユーザの嗜好を予測する方式を研究する。また、ユーザごとの推薦アイテム数平均や分布を利用した新たな推薦スコア予測方式や、評価実験期間を長くとり、時間とユーザ嗜好の変化をも考慮した推薦スコア予測方式を研究する。

受託業務でサンケイリビング社の地域情報を配信する Web サイト開発にたずさわって、ユーザにとってより精緻な情報を積極的に推薦するシステムに変更することによってユーザの利便性を高めることを考えてきた。この考えは、Amazon.com で行われている。しかし、実際にこのようなシステムを構築することになると、評価値疎ら問題と再生起問題を解決するために、ユーザの評価したアイテムとユーザの評価していないアイテムとの類似度をいかに判定するかという課題を解決する研究が必要であると考え、情報検索問題について活発に研究されている石川教授の指導を仰ぐことにし、博士前期課程に入学した (2002 年 4 月)。大学院での研究成果は、本論文に示したとおりで、協調フィルタリングに擬似投票方式を組み入れることで、推薦情報を受け取れるユーザ数のが飛躍的に向上した。この研究によって基礎研究は完了したので、今後は実用システムに適用することを計画する。

謝辞 評価実験用データを貸与していただきました

サンケイリビング新聞社の皆様に感謝します。

参 考 文 献

- 1) Balabanovic, M. and Shoham, Y.: Fab: Content-Based, Collaborative Recommendation, *CACM*, Vol.40, No.3, pp.66–72 (1997).
- 2) Basu, C., Hirsh, H. and Cohen, W.W.: Recommendation as Classification: Using Social and Content-based Information in Recommendation, *Proc. Recommender System Workshop 1998*, pp.11–15 (1998).
- 3) Breese, J.S., Heckerman, D. and Kadie, C.M.: Empirical Analysis of Predictive Algorithms for Collaborative Filtering, *Proc. 14th Conference on Uncertainty in Artificial Intelligence*, Madison, WI, July 1998, pp.43–52, Morgan Kaufmann Publisher (1998).
- 4) Buono, P., Costabile, M.F., Guida S., Piccinno A. and Tesoro G.: Integrating User Data and Collaborative Filtering in a Web Recommendation System, *Proc. 3rd Workshop on Adaptive Hypertext and Hypermedia* (2001). <http://www.wis.win.tue.nl/ah2001/proceedings.html>
- 5) Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D. and Sartin, M.: Combining Content-Based and Collaborative Filters in an Online Newspaper, *Proc. Recommender Systems Workshop at ACM SIGIR* (1999).
- 6) Gokhale, A. and Claypool, M.: Thresholds for More Accurate Collaborative Filtering, *Proc. IASTED International Conference on Artificial Intelligence and Soft Computing*, Honolulu, Hawaii, USA, Aug. 9–12 (1999).
- 7) Goldberg, D., Nichols, D., Oki, B.M. and Terry, D.: Using Collaborative Filtering to Weave an Information Tapestry. *CACM*, Vol.35, No.12 (1992).
- 8) Hirooka, Y., Terano, T. and Otsuka, Y.: Extending Content-Based Recommendation by Order-Matching and Cross-Matching Methods, *Electronic Commerce and Web Technologies — 1st Int. Conf. EC-Web 2000*, Bauknecht, K. et al. (Eds.), LNCS 1875, pp.177–190, Springer-Verlag (2000).
- 9) Lee, W.S.: Collaborative learning for recommender systems, *Proc. 18th International Conf. on Machine Learning*, pp.314–321. Morgan Kaufmann, San Francisco, CA (2001).
- 10) Linden, G., Smith, B. and York, J.: Amazon.com Recommendations: Item-to-Item Collaborative Filtering, *Proc. Feb. 2003 issue of IEEE Internet Computing* (2003). <http://www.cs.helsinki.fi/u/gionis/linden03amazon.pdf>
- 11) Polcicova, G. and Navrat, P.: Semantic Similarity in Content-based Filtering, *Proc. Advances in Databases and Information Systems, 6th East European Conference, ADBIS 2002*, Bratislava, Slovakia, Sept. 8–11, 2002, Lecture Notes in Computer Science 2435, pp.80–85, Springer 2002, ISBN 3-540-44138-7, (2002).
- 12) Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl, J.: GroupLens: an open architecture for collaborative filtering of netnews, *Proc. CSCW 1994*, New York, pp.175–186, ACM (1994).
- 13) Sarwar, B., Karypis G., Konstan, J. and Riedl, J.: Item-Based Collaborative Filtering Recommendation Algorithms, *Proc. 10th International World Wide Web Conference (WWW10)*, Hong Kong, May 1–5, pp.285–295 (2001).
- 14) Schafer, J.B., Konstan, J. and Riedl, J.: E-Commerce Recommendation Applications, *Proc. Data Mining and Knowledge Discovery*, pp.115–153, Kluwer Academic Publishers (2001).
- 15) Shardanand, U. and Maes, P.: Social Information Filtering: Algorithm for Automating “Word of Mouth”, *Proc. CHI '95 Conference*, pp.210–217 (1995).
- 16) Ungar, L.H. and Foster, D.P.: Clustering Methods for Collaborative Filtering, *Workshop on Recommender Systems at the 15th National Conference on Artificial Intelligence 1998* (1998).
- 17) Viappiani, P., Pu, P. and Faltings, B.: Acquiring User Preferences for Personal Agents, *Proc. AAAI Fall Symposium 2002*, North Falmouth, MA (2002). <http://hci.epfl.ch/publications/2002/a3iSymposium.pdf>
- 18) Zobel, J. and Moffat, A.: Exploring the Similarity Space, *ACM SIGIR FORUM*, Vol.32, No.1, Spring (1998).
- 19) 宇田隆幸, 藤井 敦, 石川徹也: ユーザ投票と情報アイテム間類似度を併用した情報推薦システム: 擬似投票方式の提案と評価, 情報処理学会研究報告, 2004-DBS-132, pp.105–112 (2004).
- 20) 宇田隆幸, 藤井 敦, 石川徹也: 協調フィルタリングに基づく情報推薦システムにおける擬似投票方式, 言語処理学会第 10 回年次大会発表論文集, pp.381–384 (2004.3).
- 21) 宇田隆幸, 藤井 敦, 石川徹也: アイテムの主題分析に基づく協調フィルタリングシステムの推薦予測方式, 情報処理学会研究報告, 2004-NL-163,

pp.67-74 (2004).

- 22) 寺野隆雄：e ビジネスの理論と応用，情報推薦システム，第3章，北村泰彦，山田誠二（編），pp.59-87，東京電機大学出版局，東京（2003）。
- 23) 館村純一，鈴木啓介，Santini, S. and Jain, R.：社会的フィルタリングと内容に基づくフィルタリングを融合した WWW デザイン画像推薦システム：インタラクティブシステムとソフトウェア VII(日本ソフトウェア学会 WISS'99), pp.153-162, 近代科学社 (1999.12).
- 24) 山西健司：Web マイニングと情報論的学習理論，電子情報通信学会情報論的学習理論時限研究専門委員会第5回情報論的学習理論ワークショップ招待講演，Information-Based Induction Science: ISBS2002 (2002).
<http://www.ieice.org/~ibis-tg/>

(平成 16 年 9 月 2 日受付)

(平成 17 年 2 月 1 日採録)



宇田 隆幸（学生会員）

1961 年生．1984 年図書館情報大学図書館情報学部卒業．同年日本情報サービス（株）（現，日本総合研究所）入社．2001 年より（株）ネオジェイエスケーにてコンピュータネットワークと DB に関する研究開発に従事．この間，2004 年図書館情報大学情報メディア研究科博士前期課程修了．引き続き，筑波大学大学院図書館情報メディア研究科博士後期課程に進学後，同年 9 月同中退．2004 年 10 月より，東北大学大学院情報科学研究科博士後期課程に在学．電子情報通信学会，ACM 等各学生会員．



藤井 敦（正会員）

1993 年 3 月東京工業大学工学部情報工学科卒業．1998 年 3 月同大学大学院博士課程修了．図書館情報大学助手を経て，現在，筑波大学大学院図書館情報メディア研究科助教授，博士（工学）．自然言語処理，情報検索，音声言語処理，Web マイニングの研究に従事．電子情報通信学会，人工知能学会，言語処理学会，Association for Computational Linguistics 各会員．



石川 徹也（正会員）

1977 年 3 月慶應義塾大学大学院修士課程（図書館情報学専攻）修了．富士フイルム（株）足柄研究所，図書館短期大学，図書館情報大学を経て，現在，筑波大学大学院図書館情報メディア研究科教授．工学博士．情報管理システムの高度化に関する研究に従事．人工知能学会，言語処理学会，ACM 等各会員．