

擬似投票方式に基づくハイブリッドフィルタリングシステムにおける推薦予測精度の改良

宇田 隆幸^{†1,†2} 木下 哲男^{†1,†3}

フィルタリングシステムは、電子商取引や広告表示などにおける情報推薦サービスのエンジンとして広く用いられているが、ユーザ数と比べてアイテム数が多い場合には、推薦されないアイテムや推薦アイテムを受け取れないユーザも増大する。これは網羅性問題と呼ばれる推薦システムにおける1つの問題であるが、これに対して、筆者らは、擬似投票方式を提案し、網羅性が改善できることを示した。一般に、網羅性向上と推薦予測精度向上はトレードオフの関係にあり、両者を同時に満足させる手法は提案されていない。加えて、推薦システム稼働中に発生するユーザの嗜好の変化によって推薦予測精度の低下が増大するため、これに対処する手法の開発が急務となっている。そこで、本論文では、擬似投票方式に基づくハイブリッドフィルタリングシステムにおいて、網羅性を維持したまま推薦予測精度を向上させる新たな手法を提案し、実験システムの試作と評価により提案手法の有効性を示す。

Improvement of Recommendation Prediction Accuracy for a Hybrid Filtering System Based on the Pseudo-voting Method

TAKAYUKI UDA^{†1,†2} and TETSUO KINOSHITA^{†1,†3}

Although the filtering systems are widely used as the engines of recommender services of electric commerce and advertisement systems, the non-recommend items and the users who cannot receive the recommended items are increased when the number of items is greater than the number of users. This is known as the coverage problem, one of problems of the recommender system. To solve this problem, the authors proposed the pseudo-voting method and showed that the coverage could be improved effectively. Generally, there exists a trade-off between the coverage and the predicted accuracy of the recommended items, and no method have been proposed to improves both coverage and accuracy simultaneously. Moreover, it becomes an important problem to establish an effective method to improve the predicted accuracy which degrades due to the changes of the users' preference at the runtime of recommender systems. Hence,

in this paper, a new method is proposed to improve the predicted accuracy without the degradation of coverage of the recommended items in the hybrid filtering system based on the pseudo-voting method. Furthermore, the effectiveness of the proposed method is demonstrated through the implementation and evaluation of the experimental recommender system.

1. はじめに

近年、Web上を流通する情報は爆発的に増加している。BlogやSNSは、ユーザ層の拡大や、情報発信の容易性向上に貢献している。情報発信の容易性にささえられ、特に、テキスト情報の流通件数が増加している。しかし、ユーザが情報にアクセスする手段は、BlogやSNS出現前と比べ方式に変化がない。また、ユーザには、ぼんやりとした好みの情報があり、それを情報検索式によって表現することは難しいので「ユーザ個人の嗜好に適合するアイテムの一覧を提示するシステム」があれば便利である。しかし、ECサイト¹⁾・広告推薦サイト・研究用のサイトを除くと、このようなシステムは存在しないので、現在もユーザは検索エンジンやネットサーフィンに頼っている。このため、情報の取りこぼしや、好まない情報の閲覧による時間の浪費が発生している。

問題解決策の1つとして情報推薦システムの利用がある。情報推薦システムとは、利用者の好みそうなアイテムを見つけ出し、見つかったアイテムを提示するシステムである。情報推薦システムには、利用者の好みそうなアイテムを見つけ出す方式、すなわち推薦アルゴリズムには様々な方式がある。各推薦アルゴリズムには長所や短所があり、情報提供者側の目的・利用者側の目的・提示するアイテムの特性・利用者の特性などを総合的に考慮し、情報推薦システムのコアエンジンとして最適な推薦アルゴリズムを決める。

情報推薦システムのコアエンジンとして、協調フィルタリング(CF: collaborative filtering)は、多くの先行研究があり^{2),3)}、商用システムでも採用されている。CFアルゴリズムは、アイテムの内容を解析しないので、どのようなアイテムでも推薦できるという特長を持つ。たとえば、書籍・映画・音楽などのタイトル、タイトルのない画像や映像な

†1 東北大学大学院情報科学研究科
Graduate School of Information Science, TOHOKU University

†2 情報セキュリティ大学院大学事務局
Secretariat, INSTITUTE of INFORMATION SECURITY

†3 東北大学サイバーサイエンスセンター
Cyberscience Center, Tohoku University

ども推薦対象となり、マルチメディア情報が多数含まれる Web ページを対象とする推薦では、特に有用である。

CF アルゴリズムにおける推薦予測アルゴリズムは、被推薦ユーザ（推薦アイテムを受け取るユーザ）の嗜好情報と類似または正反対の嗜好情報を持つ標本ユーザ（被推薦ユーザ以外のユーザ）の嗜好情報を利用して、被推薦ユーザの見ていないアイテムの推薦スコアを予測して、スコアの高いアイテムを推薦する。被推薦ユーザと嗜好対象の似た標本ユーザが好むアイテム（または、被推薦ユーザと嗜好対象が正反対の標本ユーザが嫌うアイテム）は、優先され（推薦スコアが上がる）、嗜好対象の似た標本ユーザが嫌うアイテムの推薦スコアは下がる。ユーザがアイテムを選ぶたびに嗜好情報が更新され、更新された情報はその後の推薦スコア計算で利用される。

多くの CF アルゴリズムは、推薦スコア予測計算の基となる嗜好情報として、多くのユーザの行動履歴を必要とする。ここで、ユーザの行動履歴とは、ユーザによるアイテムへの評価投票や、Web アクセスログから得られるアイテムの閲覧に関する情報などである。多くのユーザの行動履歴を必要とする理由は、被推薦ユーザと相関が類似するかまたは正反対の標本ユーザを探すためである。この点、CF を少ないユーザの行動履歴で稼働させるために、アイテムのクラスタ化やユーザのクラスタ化を実装するアルゴリズムもある⁴⁾。クラスタ化を行うことで推薦対象となるアイテムやユーザの網羅性が高まる反面、クラスタとクラスタの間に位置するアイテムやユーザに対して、推薦スコア予測の正確性は低下する。一般に、情報推薦システムにおいて、推薦対象アイテムやユーザの網羅性向上と推薦スコア予測の正確性向上は、トレードオフの関係にある。

筆者らは、ハイブリッドフィルタ^{*1}において、アイテムの内容解析を行い、嗜好情報履歴の自動追加を行う擬似投票方式を提案し評価した⁵⁾。この結果、情報推薦システムにおける網羅性問題は解消した。本研究では、被推薦対象の網羅性を維持したまま、「アイテムの内容の特徴・ユーザ評価の特徴・ユーザ評価時と推薦予測時の時間差」に着目してアイテムの推薦スコア予測アルゴリズムを改良することによって、推薦予測精度の正確性を高める方式を提案し評価する。

*1 ハイブリッドフィルタとは、内容ベース方式と CF 方式を組み合わせた情報推薦方式である^{6),7)}。推薦するアイテムを決めるとき、内容ベース方式または CF 方式のいずれかを利用する手法や、アイテムの内容解析を行って CF が保持する嗜好情報履歴の欠損値を補完する手法などがある。本研究におけるハイブリッドフィルタの構成は、CF 方式と擬似投票方式（後述）とを組み合わせている。なお、Burke は、ハイブリッドフィルタに関するサーベイと評価を行い、各フィルタの結合に応じた分類を行っている¹⁶⁾。

以下、2 章では関連研究で指摘されている問題点について触れ、3 章では提案方式の設計内容と、本研究の特徴である「アイテムの内容の特徴・ユーザ評価の特徴・ユーザ評価時と推薦予測時の忘却性」を推薦スコア予測計算アルゴリズムに反映させるための予備実験の方法と予備実験の結果について述べる。4 章では提案方式に基づく情報推薦システムの試作と評価実験結果について述べ、5 章で結論と今後の課題をまとめる。

2. 関連研究

2.1 CF と擬似投票方式

相関係数法²⁾による CF は、図 1 に示す「アイテム-ユーザ」評価マトリクスを持つ。このマトリクスは、ユーザが閲覧したアイテムを評価したときに記録される。そして、ユーザがまだ見ていないアイテムの推薦スコアを予測するときに参照される。

たとえば、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 ）は、式 (1) で求める。

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

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。

		アイテム			
		i_1	i_2	i_k	i_n
ユーザ	u_1	$r_{1,1}$	$r_{1,2} (V)$	$r_{1,k} (V)$	$r_{1,n}$
	u_2	$r_{2,1}$	$r_{2,2}$	$r_{2,k}$	$r_{2,n}$
	u_j	$r_{j,1}$	$r_{j,2} (V)$	X	$r_{j,n}$
	u_m	$r_{m,1}$	$r_{m,2}$		$r_{m,n}$

図 1 「アイテム-ユーザ」評価マトリクス
Fig.1 Evaluation matrix for items and users.

ユーザ間の類似度 $sim(u_i, u_m)$ は式 (2) 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 の両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 .

CF は、ユーザ間の相関を利用して嗜好の類似するユーザを探し、それらユーザの評価値を参考にすることで推薦スコア予測を計算する . このため、たくさんのユーザ評価値が必要になる .

擬似投票方式⁵⁾ は、関連研究において指摘されていた 2 つの問題を解決した . 1 つ目は、評価値疎ら問題 (Sparsity Problem) である . この問題の原因は、「アイテム-ユーザ」評価マトリクスがスパースであるために発生する . スパースであるため、ユーザ間の相関を計算できず、推薦アイテムを受け取れるユーザの割合は、20%程度になる (評価実験による) . 2 つ目は、再生起問題 (Recurring Startup Problem) である . この問題は、新聞記事のように毎日発生する新しいアイテムは、どのユーザも評価していないので、推薦対象から外れるということである . これら 2 点の問題を解消することで、推薦アイテムを受け取れるユーザ割合や推薦されるアイテムの割合が高まる .

擬似投票方式は、「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 . 評価予測値の自動的な設定は、「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 .

評価値疎ら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出し、② 見つかったセル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 .

再生起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出し、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 .

2.2 擬似投票方式の有用性

伝統的 CF 方式と擬似投票方式の性能を比較するために、それぞれを実装し、図 2 に示



図 2 アイテム詳細ページの例

Fig. 2 Example of the detailed information of an item.

すサイト^{*1}で収集したデータを使って評価実験を行った . 実験の目的は、本研究の提案方式検討に先立ち、関連研究が示す CF アルゴリズムの問題点を確認することである . 実験結果から、地域情報を対象とした推薦アイテムの場合、ユーザが閲覧するアイテムは、居住地や勤務先近辺などに集中する傾向が強く、他の地域のアイテムはほとんど閲覧されないことが分かった . これは、複数のユーザコミュニティで構成される Blog や SNS において、ユー

*1 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、アート&シアター/イベント/グルメ/ショッピング/レジャー/教養&カルチャー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応じて、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/>) .

は自身の属するコミュニティ内で活動し、他のコミュニティへは参加しないという行動パターンと似ている。Web 全体のユーザ数は増加しても、それ以上にコミュニティの数が増加している。このため、各コミュニティのユーザ数は少なくなる。このような背景から、大多数のユーザコミュニティは少人数で構成されるので、ユーザに評価されないアイテムが、多数存在することになる（評価値疎ら問題）。したがって、相関係数法による伝統的 CF をコアエンジンに採用する情報推薦システムは、被推薦対象が狭くなるので実用的でない。

一方、擬似投票方式は、異なるユーザコミュニティ間でも、アイテムのテキスト情報で媒介され、嗜好情報の共有が可能になる。このため、複数のユーザコミュニティを結合して作った巨大な「アイテム-ユーザ」評価マトリクスでも、嗜好情報の共有によって、スパース性が大きく改善される。この結果、推薦可能なユーザの網羅性や推薦可能なアイテムの網羅性は、商用環境でも稼働できるほど実用的である。

2.3 擬似投票方式の問題点

情報推薦システムは、できる限り、ユーザの嗜好するアイテムを推薦すべきである。また、ユーザはそれを期待しているので、推薦予測精度は、重視されるべき性能評価基準である⁸⁾。当然、推薦予測精度の正確性や最新的话题を提供できる推薦機能は、理想的な情報推薦システムにおいて欠かせないものである。ところが、CF と擬似投票方式を組み合わせた情報推薦システムを含む情報推薦システム全般において、推薦対象アイテムやそこに集うユーザの嗜好を組織的に系統立てて分析し、分析結果を推薦予測精度に反映させる研究は、始まったばかりである⁹⁾⁻¹¹⁾。

単一のユーザコミュニティを対象として、伝統的 CF 方式の推薦予測精度を改良した関連研究がある。しかし、情報爆発時代では、複数のユーザコミュニティを横断した嗜好情報の共有や、組織的・系統的に知識共有する仕組みの提案が必要である。しかし、それらの提案や、それらを応用したシステムの実装は、まだない。

筆者らは、評価値疎ら問題や再生起問題を解消する目的で、システムがユーザに代わってアイテムを評価投票する擬似投票方式⁵⁾を提案した。擬似投票方式を実装することで網羅性問題および複数のユーザコミュニティを横断した嗜好情報の共有機能は解消した。しかし、推薦予測精度の正確性は、伝統的な CF アルゴリズム²⁾と同程度であり、正確性の向上問題が未解決である。加えて、システムの稼働期間が長くなるにつれ、嗜好対象の変化するユーザを数多く観測している（図 2 の商用サイト）。伝統的な CF は、ユーザによるアイテムの評価を不変の嗜好として扱うので、「古い評価履歴がノイズになっていく」という問題も発生している。これらのことから、CF と擬似投票方式を組み合わせた情報推薦システムの場合、推薦予測精度の改良が必要である。

3. 提案方式の設計

本研究では、擬似投票方式の実装によって獲得した被推薦対象アイテムやユーザの網羅性および複数のユーザコミュニティを横断した嗜好情報の共有機能を維持したまま、推薦予測精度を高める方式を提案し実装する。これにより、情報爆発時代において複数のユーザコミュニティをまたいで存在する Blog や SNS におけるアイテムも推薦対象にできる。

3.1 推薦予測精度を高める方式の検討

この節では、推薦予測精度を高めるために、利用可能な観測値（説明変数）を検討する。具体的な方法は、図 2 のサイトで収集したユーザ評価値および推薦アイテムについて、以下の 3 点に着目して、ユーザの評価履歴などの観測データを調べながら利用可能な観測値について検討する。

- アイテムの内容
- ユーザ評価の傾向
- ユーザ評価と推薦予測の時間のずれ

3.1.1 アイテムの内容

本研究が対象とするアイテムはテキスト情報を含むので、ユーザ評価の高低は、テキストに含まれる語、すなわち内容と何らかの関係があると考えた。

そこで、ユーザ評価が高いアイテムとそうではないアイテム間において、

- アイテムを特徴づける語
- 特徴語の出現位置（見出しまたはリード文/本文）

の 2 点に、統計上の差があるか否かを検討する。

まず「特徴語ペア」の意義について説明する。ユーザはアイテムを評価するとき、その内容に着目する。テキストを含むアイテムの場合、アイテムの内容を決定づける要素は、アイテムに出現する語が影響する¹²⁾。たとえば、多くのユーザが高く評価したアイテムに含まれる語のいずれかは、アイテムが高く評価された要因であろう。そこで、特徴語の語数（＝形態素数）を 1 語および 2 語として、ユーザが高く評価したアイテムから特徴語を発見する実験を行った。

ただし、2 語からなる特徴語は（以降「特徴語ペア」）、それぞれの語が 10 語以内に出現する組合せに絞った。これは、アイテムの見出しが 10 語程度で記されていることと、仮に 2 語が離れているなら、同時に視界に入ることなく、その印象は薄いと推測できるからである。

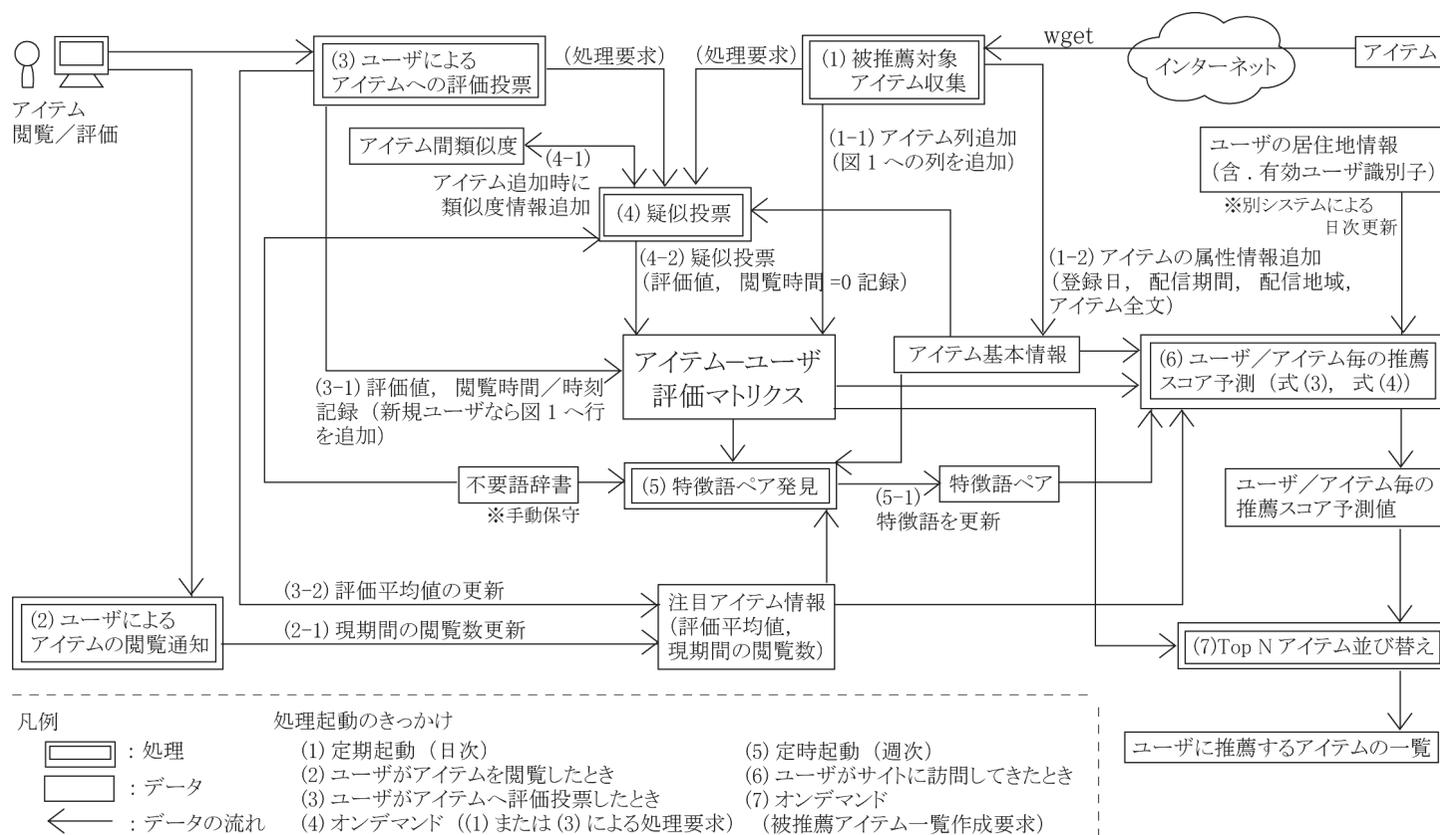


図3 システム全体図
Fig. 3 System overview.

さらに、特徴語の出現位置も推薦スコア調整係数算出の基に加えた。これは、アイテムが新聞記事と同様、図2に示すように、「見出し、リード文、本文」の順で出現するので、特徴語の出現位置がアイテムの先頭に近いほど、ユーザは強い印象を受けるからである。また、ユーザはアイテム一覧ページからアイテムを選ぶとき、「見出し」を参考にするので、見出しから得る印象は特に強いと考える。

特徴語を発見するために、次の実験を行った。最初に、① アイテムの母集団から、2つ

の部分集合を切り出す。いずれの集合も、図3のアイテム基本情報に記録している「アイテムの配信地域」と「アイテムの配信時期」は同じとする。ただし、一方は、図3の注目アイテム情報における「評価平均値」が2.8以上を満たし(約90%のユーザが満足するアイテム)、他方は2.8未満とする。次に、② アイテムを形態素解析し、アイテム本体から不要語を除去し、[名詞]、[名詞-接尾]が連続する場合は、名詞を結合する。特徴語ペアについては、③ 各アイテムで10語以内に出現する「語のペア」をつくり、集合内における各語の

ペアの出現回数を記録（特徴語を1語とする条件下では、名詞ごとの出現回数を記録）する。最後に、④ 特徴語の出現割合が2つの部分集合間で差があるか、有意水準5%で χ^2 仮説検定を行う¹³⁾。そして、有意差がある特徴語を、アイテムの配信地域、アイテムの配信時期とともに記録する。

なお、アイテムの配信地域、配信時期が同じ特徴語は、翌年の配信時期において、同じ特徴語が含まれる傾向が高い。たとえば、特徴語ペアの場合、8月は「親子-体験」「親子-学習」、10月は「紅葉-旅行」など。この傾向を利用して、既登録の特徴語を翌年の同時期の特徴語として利用する。

なお、有意差のあった1語からなる特徴語は、固有名詞の割合が多く、翌年の配信時期において利用価値が少ない^{*1}ので、特徴語は2語からなる特徴語ペアを対象にすべきであることが分かった。

3.1.2 ユーザ評価の傾向

複数のユーザによる評価の集計や閲覧数の集計は、そのアイテムの推薦スコア調整に利用できると思った。そこで、ユーザ評価の傾向に関して、

- ユーザ評価値の高低（満足度）
- アイテムの閲覧数の増減（注目度）

を検討する。これらの2項目を推薦スコア調整係数算出の基となる要因に選んだのは、今まで商用の推薦システムを開発・運用してきた経験からである。ここにいう経験とは、アイテムを先に評価したユーザの評価値と、後から評価したユーザの評価値には、相関関係があるということを目指す。また、アイテムの閲覧数についても同様である。

ユーザ評価値の高低については、図3の注目アイテム情報における「評価平均値」が2.8以上（約90%のユーザが満足するアイテム）のアイテムを推薦スコア調整の対象にした。このような人気の高いアイテムを推薦すると、高い確率で、推薦アイテムを受け取ったユーザも満足する。しかし、① 評価平均値を算出したときの母数が少ないアイテム、② 配信開始から時間の経過したアイテム、③ 配信終了時期が近づいているアイテムの場合は、それらの推薦アイテムを受け取ったユーザの評価値が低いこともある。評価値の低い事例を追跡し、その要因を調査した。①の場合、その後の評価平均値が低くなることや、分散値が大

きくなることがある。この要因は、母数の少なさによる評価のゆれである。加えて、アイテムの内容がすでに古くなったことから評価が低くなる事例も観測した。②および③の場合、その後の評価平均値が低くなる主な要因は、アイテムの内容が古くなったためである。結局、ユーザ評価値の高低は、①、②、③の問題があるものの、それらは一部のアイテムに限られるので、推薦スコア予測精度の向上において、利用価値が高いことが分かった。

アイテムの閲覧数の増減については、図3の注目アイテム情報における「現期間の閲覧数」が直前の期間の閲覧数よりも多い場合を推薦スコア調整の対象にした。ユーザ評価値の高低と同様、観測数の少なさやアイテムの内容が古くなったことを要因とする評価のぶれを観測した。アイテムの閲覧数が増加しているときは、満足度が高くなる傾向があるので、推薦スコア予測精度の向上において、利用価値があることが分かった。

3.1.3 ユーザ評価と推薦予測の時間のずれ

ユーザの嗜好は時間とともに変化するので、ユーザ評価は、時間がたつことでその価値が低下していくと考えた。ただし、ユーザ嗜好の変化は、個人差が大きいので、それを吸収する仕組みが必要だと考えている。この項では、ユーザ評価のときと推薦予測において評価値を参照するときの時間のずれに着目すれば、推薦予測精度の向上に有意であるかどうかを検討する。

ユーザによるアイテムの評価時期と推薦予測を行う時期の時間差が、推薦予測精度に影響をもたらすか否かを、以下の抽出条件を満たすユーザに対して、Aグループ（分類条件＝興味対象の変化あり）とBグループ（分類条件＝興味対象の変化なし）それぞれを観測した。

（抽出条件）. 5年以上のアクティブユーザで、アイテムの評価数が年間120を超えていること（絶対条件）

（分類条件）. ユーザの評価したアイテムのジャンルに変化があるか/変化がないか。なお、アイテムのジャンルとして、実験データ貸与元の分類（アート&シアター/イベント/グルメ/ショッピング/レジャー/教養&カルチャー）を利用した。

具体的なグループ分けの方法は、最初に、ユーザ全体を対象としたジャンルごとの閲覧アイテムカウンタと、評価投票平均値を記録する。次に、ユーザごとジャンルごとの閲覧アイテムカウンタと評価投票平均値を記録する。続いて、ユーザ全体と各ユーザに対する閲覧アイテム数と評価投票平均値について、90%の信頼区間で χ^2 による有意差検定を行う。有意差なしなら平均的（値1）、有意差ありの場合、平均値以上（値2）または平均値以下（値0）について、ユーザごと/アイテムのジャンルごとの閲覧アイテムカウンタと評価投票平均値とを記録する。最後に、測定開始年とその3年後・測定開始

*1 交差確認法による実験を実施したところ、翌年度の同時期のアイテム集合において、1語からなる特徴語の場合、アイテムを特徴付けるほどの有意差が見られなかった。また、実験で観測した固有名詞は、大量の広告によってユーザの記録に残った商品名や施設名、翌年に実施されないイベント名などが多く、特徴語として記録する価値は低いと判断した。

年とその4年後・測定開始年とその5年後のいずれかにおいて、記録値の絶対誤差の合計が、8ポイント*1を超えるユーザをAグループとし、それ以外をBグループとした。推薦スコア予測で参照するユーザ-アイテム評価マトリクスデータを2種類用意した。1つは評価値の記録期間を1年以内とし、もう一方は5年間とした。

式(1)・式(2)を使って推薦予測の正確性を観測した。記録期間が1年以内の場合、A・Bのグループ間で正確性の有意差はない。しかし、記録期間が5年間の場合、AグループはBグループより正確性が低くなった。Aグループは興味対象が変化したことにより、時間の経った評価データの価値が低下したためだと考えられる。このことから、アイテムの評価から経過時間に応じた重み付け調整が必要になることが分かった。

なお、本研究では、アイテムの評価から時間が経過するほど、価値が低下するユーザ投票値について検討した。これとは逆に、価値の上昇するユーザ投票値の存在も否定できないので、今後も注意深くユーザ投票値の価値(推薦予測精度との相関)について観察していく。

3.1.4 アイテムの閲覧時間とユーザ興味の関係

アイテムの閲覧時間の長さやユーザの興味の深さには、正の相関関係があるといわれている¹⁴⁾。この項では、アイテムの閲覧時間に注目すれば、推薦予測精度の向上に結び付くかどうかを検討する。

この点を確認する目的で、ランダムに100人のユーザを選び、アイテムが満足(「とても!」)と評価されたアイテムの平均閲覧時間と、アイテムが「まあまあ」と評価されたアイテムの平均閲覧時間を比較した。「とても!」と評価したグループの平均閲覧時間は「まあまあ」よりも約8秒長いので、文献14)の主張を確認できた。提案システムへも、この点を組み入れ可能かどうか確かめるための実験を行った。実験では、閲覧時間の長さやユーザの興味の関係を利用して、「とても!」と評価し、かつ閲覧時間が「平均閲覧時間+標準偏差」よりも長い場合は、評価値に0.5を加算し、推薦予測精度の変化を観察した。この結果、わずかに推薦予測精度は改善されたが、これは誤差の範囲であった。

予測精度が改善されないのは以下の理由による。擬似投票には閲覧時間を持たないので、推薦スコア予測計算で閲覧時間を利用した評価値の加算は行わない。推薦スコア予測計算では、ユーザ投票値と擬似投票値の両方を参照する。しかし、ユーザ投票数は擬似投票数より

*1「アート&シアター/イベント/グルメ/ショッピング/レジャー/教養&カルチャー」のほか、複数のアイテムジャンルを含む広告などについては計測対象外にしたので、ユーザに提供しているアイテムジャンル数は8種類になる。3年間に8ポイント以上を獲得したユーザは、嗜好変化があると判定した(この場合、2ジャンルの嗜好変化/提供アイテムのジャンルの約25%の嗜好変化)。

も極端に少ないので、ユーザ投票値に加算しても擬似投票値の中に埋もれてしまい性能改善には寄与しないからである。

本研究では、擬似投票方式の実装によって獲得した被推薦対象アイテムやユーザの網羅性を維持したまま推薦予測精度を向上させることを目的にしている。このため、擬似投票方式を取り除くことはできないので、アイテムの閲覧時間に着目した手法による推薦予測精度の改良は、難しいと判断した。

3.2 推薦予測精度を高める方式

前節までの検討結果をふまえ、式(1)・式(2)を改良して調整係数や重み付けを実施する式(3)・式(4)を提案する。これらの提案式を用いることで、推薦スコアの調整が実施され、ユーザに提示する推薦スコアTop Nのアイテムが入れ替わることで、ユーザが閲覧するアイテムも入れ替わり、その結果、推薦予測精度が改良される。

ユーザ・アイテムごとの推薦スコアの計算は、式(1)を改良した式(3)を用いる。

$$P_{i,k,t} = \left(\bar{r}_i + \frac{\sum_m \text{sim}(u_i, u_m) \cdot (r_{m,k} - \bar{r}_m)}{\sum_m \text{sim}(u_i, u_m)} \right) \times (\alpha_{k,t} + 1.0) \quad (3)$$

ここで $P_{i,k,t}$ は、時刻 t におけるユーザ u_i に対するアイテム k への推薦スコアの予測値である。 $\alpha_{k,t}$ は、時刻 t におけるアイテム k への推薦スコア調整係数である。詳細は後述する(4.1節)。

ユーザ間の類似度の計算は、式(2)を改良した式(4)を用いる。

$$\text{sim}(u_i, u_m) = \frac{\sum_j (r_{i,j} - \bar{r}_i) \cdot (r_{m,j} - \bar{r}_m) \cdot w_{tl}}{\sqrt{\sum_j (r_{i,j} - \bar{r}_i)^2 \cdot (r_{m,j} - \bar{r}_m)^2 \cdot w_{tl}^2}} \quad (4)$$

ここで w_{tl} は、ユーザ u_i がアイテム j への評価投票($r_{i,j}$)を行った時刻と現在時刻との差(時間差)に応じた重み付け値である。時間差が小さいほどユーザの嗜好変化は小さいので、重み付け値を大きくすることで、その投票値に価値を与える。逆に、時間差が大きいほど重み付け値を小さくする。なお、評価値がシステムによる擬似投票の場合は、重み付け値を最小にした。擬似投票は、同じユーザ/同じアイテムに対して複数の評価値を付与することがあるので、時間差の概念になじまないからである。推薦スコア予測処理結果は、「ユーザ/アイテムごとの推薦スコア予測値」領域に記録する。

3.3 提案システムの全体図

システム全体図を図3に示した。図3の二重線で囲んだ長方形は処理ブロックを表している(凡例参照)。提案システムは7個の処理ブロックを持つ。各処理ブロックには、図3下

部で示した処理起動のきっかけがあり(たとえば, 定時起動・オンデマンド起動など), 起動パラメータと図中の実線で囲んだデータ領域を読み込むことで処理が始まる. 予定する処理が終わると, 結果を矢印で示したデータ領域に書き出す.

情報推薦システムは, (1) から (7) の処理が繰り返されることで, 新しいアイテムをユーザに推薦できる. 以降, 図 3 に記した (1) から (7) の処理ブロックについて, それぞれ説明する.

(1) 被推薦対象アイテム収集処理

この処理の起動契機は, 定時起動である. 提案システムが取り込んでいないアイテムを収集し, 取り込んだアイテムに向けたユーザの評価値を記録するための領域を「アイテム-ユーザ評価マトリクス(図 1 参照)」に追加し, 取り込んだアイテムの情報を「アイテム基本情報」領域に記録し, 擬似投票処理へ処理要求を送る. なお, アイテム基本情報を記憶する領域には, アイテム本体(見出し・リード文・本文の区別付き/図 2 参照)・アイテムの登録日・ユーザへの配信期間・配信対象地域(行政区区町村に相当する地域)を記録する.

(2) ユーザによるアイテムの閲覧通知処理

この処理の起動契機は, ユーザがアイテム詳細ページ(図 2)を閲覧したときである. 1 週間単位/アイテム単位で閲覧数を記録する. この情報は後に, 推薦スコア予測精度を高める目的で使用される.

(3) ユーザによるアイテムへの評価投票処理

この処理の起動契機は, ユーザがアイテム詳細ページ下部に付けた評価投票ボタン(図 2 下部にある 3 個のボタン「とても!」「まあまあ」「いまいち...」)がクリックされたときである. アイテム-ユーザ評価マトリクスの該当セルに, 評価値・閲覧時間を記録し, 擬似投票処理へ処理要求を送る. クリックされた評価投票ボタンが「とても!」の場合は該当セルの評価値らんに 3 を設定し, 同様に「まあまあ」なら 2 を, 「いまいち...」なら 1 を設定する. なお, 閲覧時間の長さともユーザの興味の深さには正の相関を観測している(閲覧時間が長いとき).

(4) 擬似投票処理⁵⁾

この処理の起動契機は, 「(1) 被推薦対象アイテム収集処理」または「(3) ユーザによるアイテムへの評価投票処理」から処理要求を受け取ったときである.

擬似投票はユーザの評価したアイテムと類似するアイテムに, システムが代理投票する仕組みである. ユーザによるアイテム評価には「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」という特性があるため⁸⁾, 擬似投票が成立す

る. 擬似投票の効果は, アイテム-ユーザ評価マトリクスのスパース性が大幅に改善される点である. この結果, 推薦対象になるアイテムやユーザの網羅性が向上する.

起動契機が「(1) 被推薦対象アイテム収集処理」の場合は, 新しく登録したアイテムと類似するアイテムを見つけ出し, 見つかったアイテム集合に対するユーザ個々の評価値平均を新しく登録したアイテムの評価値として記録する.

起動契機が「(3) ユーザによるアイテムへの評価投票処理」の場合は, ユーザが評価したアイテムと類似するアイテムを見つけ出し, 見つかったアイテムにユーザの評価値を記録する.

なお, 類似アイテムか否かの判定は, Dice 係数を用い, 演算結果が 0.8 を超えたときに両アイテムは類似すると判定する^{5), 12)}.

(5) 特徴語ペア発見処理

この処理の起動契機は, 週次の定時起動である. ユーザの満足度が高いアイテムに含まれる特徴語のペアを発見して記録する. 特徴語ペアを発見する手法は, 後述する. なお, この情報は, 推薦スコア予測精度を高める目的で使用される.

(6) ユーザ/アイテムごとの推薦スコア予測処理

この処理の起動契機は, ユーザがサイトに訪問してきたときである. 推薦スコア予測とは, ユーザが未評価のアイテムに対する推薦スコアをそれぞれ計算し, 推薦スコア上位 N 件のアイテムを提示するための準備処理である. 推薦対象になるアイテムの条件は, ユーザが未評価であるほか, ユーザの居住地とアイテムの配信値が一致していて, かつ, 配信期間中のアイテムに限る.

ユーザ・アイテムごとの推薦スコアの計算は, 式 (1) を改良した式 (3) を用いる. ユーザ間の類似度の計算は, 式 (2) を改良した式 (4) を用いる.

(7) Top N アイテム並べ替え処理

この処理の起動契機は, 推薦アイテム一覧の作成要求があったときである. この処理は, 「ユーザ/アイテムごとの推薦スコア予測値」領域に記録されたアイテムのうち, ユーザが未評価のアイテムを抜き取り, 推薦スコアの高い順に並べ, 上位 N 個のアイテムを「ユーザに推薦するアイテムの一覧」領域に記録する.

次に, 図 3 の「アイテム-ユーザ評価マトリクス」領域への評価情報記録と, 同領域からの評価情報読み出しの方法について説明する. 評価情報の記録は, アイテム-ユーザ評価マトリクスのセルに行く. 同じセルに複数の評価情報を記録できる構造である. たとえば, あるユーザがあるアイテムを 2009 年 1 月 1 日と 2009 年 2 月 2 日と 2 回評価した場合, 2 回

とも評価値が記録される。また、擬似投票処理によって何回か当該アイテムが評価されたときも、すべての評価情報が記録される。

セルに記録された評価情報の読み込みは、① 擬似投票を行うとき、② ユーザの満足度が高いアイテムを見つけるとき、③ 他のアイテムの推薦スコアを予測するときに発生する。① から ③ における評価情報の参照値は、そのセルにユーザによる評価投票がある場合、最新のユーザ評価投票値が評価値になる。そのセルにユーザによる評価投票がない場合、擬似投票値の平均値が評価値になる。

4. 情報推薦システムの試作と評価実験

3章で提案した方式の実装と実験に先立って、アイテムの内容や係数の設定について、式(3)および式(4)における最適な調整係数と重み付けの調整を検討する(予備実験)。その後、予備実験によって得られた調整係数や重み付け値を用いて、システム全体の評価実験を行う。

4.1 予備実験(推薦スコア調整係数および重み付けの調整)

3.1節で、「アイテムの内容・ユーザ評価の傾向・ユーザ評価と推薦予測の時間のずれ」に着目して推薦スコア予測の正確性に結び付く観測値(説明変数)について検討した。この節では、推薦スコア予測の正確性に結び付く観測値について、推薦スコア予測(式(3)および式(4))でそれらの観測値を推薦予測精度向上に利用するときの最適な調整係数や重み付け値を求める目的で予備実験を行う。

4.1.1 推薦スコア調整係数

3.1.1項で「アイテムの内容」、3.1.2項では「ユーザ評価の傾向」が、推薦スコア予測の正確性向上に寄与することを検討した。いずれの観測値(表1)も、アイテムの特長を反映できるので、式(3)における推薦スコア調整係数とした。

式(3)における推薦スコア調整係数の計算は、調整係数の基となる情報に対する係数の総和とした。まず、各係数の初期値を0.1に設定して、推薦予測スコアを計算し、式(5)に示す平均絶対値誤差(MAE)を記録する。次に、各係数それぞれについて値を単独で増減させ、成績の良いMAEを得たときの調整係数を記録する。

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

ここで、*rating* はユーザによる評価値を示し、*prediction* はシステムによる評価予測値を示し、*N* が *rating* と *prediction* のペア数を指す。ここでは、システムによる評価予測値は、「満足」に相当するアイテムのみを対象にする。すなわち、各ユーザに対して推薦できるアイ

表1 調整係数

Table 1 Adjustment coefficients.

調整係数が機能する条件	調整係数
特徴語ペアがアイテムの見出し部に出現した場合	+0.10
特徴語ペアがアイテムの本文に出現した場合	+0.05
ユーザ評価投票数の90%以上が満足	+0.15
アイテム参照数が前週より多い	+0.05

テムのうち、式(3)の推薦スコアの上位20アイテムをシステムの推薦アイテム(満足=3)とした。

また、推薦スコア調整係数の個々の調整係数(表1)の最適値を求める実験では、評価時期を2005年12月・2006年12月・2007年12月の最終週とし、配信地域は首都圏のみ、被推薦対象ユーザは無作為に抽出した50人に絞った。

最適値を求める実験の結果、調整係数は、表1の値を採用することにした。

最適値を求める実験中の考察は次のとおりである。

まず、特徴語ペアは、アイテムの見出し部分に出現した方が推薦精度の正確性が高まる。これは、ユーザが嗜好に適合するアイテムを見つけるとき、アイテム一覧に表示された「見出し」を見ているためだと考えられる。しかし、見出し部分は語数が少ないので、アイテム本文に出現した場合も、係数調整の範囲に加えた方が正確性向上に結び付くアイテム数が増加することを観測した。

次に、ユーザ評価の満足度割合が高いアイテムを推薦することは、当然、推薦精度の正確性向上に好影響を与える。しかし、アイテムが配信されてから時間が経っている場合、新しいアイテムを積極的に推薦した方が良い成績に結び付くこともある。これは、ある一時点で、1人のユーザに推薦するアイテム数が最大20のため、推薦すべき意外性のあるアイテムが推薦スコア上位20から外れると考えられることによる。そこで、アイテム参照数が増加(着目度の高いアイテム)しているか否かも調整係数に加えた。これによって、新しいアイテムの推薦スコアが上昇する。

4.1.2 重み付けの調整

3.1.3項では、「ユーザ評価と推薦予測の時間のずれ」が、推薦スコア予測の正確性向上に寄与することを検討した。アイテムの評価時期ごとに推薦予測計算時期との時間差が異なるので、時間差特性を反映できるのは、式(4)におけるユーザの評価値をどの程度取り入れるかを示す重み付け調整である。ユーザが評価投票を行った時刻と現在時刻との時間のずれ

表 2 重み付けの調整

Table 2 Weighted values for adjustment.

調整方法	時間差					MAE
	1年以下	2年以下	3年以下	4年以下	4年超	
調整なし	1.0	1.0	1.0	1.0	1.0	0.387
パターン1	5.0	4.0	3.0	2.0	1.0	0.310
パターン2	8.0	4.0	3.0	2.0	1.5	0.282
パターン3	17.0	9.0	5.0	3.0	2.0	0.295

擬似投票値との比較における重みは、常に1とした

(時間差)に応じた重み付けは、ずれの大きさとユーザ嗜好の変化に要する時間には、正の相関がある。時間のずれが大きいと、嗜好の変化する可能性が高まるので、参照した評価投票値の値は下がるので、重みは減らすという考え方である。

式(4)の重み付けの調整は、時間差だけを基準に重みの調整をする方式と、時間差に加えユーザの嗜好対象特性の両方を基準に重みの調整をする方式とがある。本研究では、後者の方式を採用しない。この理由は、3.1.3項の観測結果から、ユーザの嗜好対象特性と時間差の関係は、個人ごと/時間ごとに変化していて、自動的に重みを調整する方式を見つけれなかったことによる。

ここでは、重み付け調整として表2の各パターンについて、式(1)と式(4)を用いて推薦スコアの予測計算を実施し、推薦スコアの予測とユーザの評価投票値の平均絶対値誤差(MAE)を式(5)に示す方法で求める。

実験を行った調整方法では、パターン2が最も良い成果を出した。時間差が小さいときほど、重みの調整を大きくすることで良い結果が得られた。

4.2 評価実験

4.2.1 実験データと前提

評価実験用アイテムは、図2で示すサイトに蓄積されているアイテムのうち、2004年1月1日～2008年12月31日(5年間)の間に、首都圏または関西圏に配信実績のあるアイテムのうち、50,000アイテムを無作為に抽出した。アイテムの配信地域に首都圏と関西圏を選んだ理由は、「アイテム-ユーザ」評価マトリクスの評価値が、他の地域と比べ密集しているからである。ここで、首都圏とは「東京都・神奈川県・千葉県・埼玉県および茨城県南部」を指し、関西圏とは「大阪府と兵庫県」である。

ユーザの評価投票データは、2004年1月1日～2008年12月31日の間に、上50,000アイテムのいずれかに評価投票のあった(約310,000件)を対象とした。異なるユーザ数は、

表 3 伝統的 CF と提案方式の比較

Table 3 Comparison of the traditional CF and the proposed method.

評価 時期	推薦可能ユーザ割合			MAE		
	伝統的 CF	擬似投票のみ	提案方式	伝統的 CF	擬似投票のみ	提案方式
2004.12	8.3%	95.5%	95.5%	0.375	0.406	(0.239)
2005.12	10.1%	97.1%	97.1%	0.366	0.387	0.234
2006.12	12.5%	97.8%	97.8%	0.351	0.361	0.231
2007.12	13.6%	97.6%	97.6%	0.339	0.343	0.227
2008.12	18.1%	98.8%	98.8%	0.340	0.351	0.226

およそ38,500人である。なお、同じユーザが同じアイテムに複数回評価投票を実施した場合、投票回数は1回とカウントした。

アイテムの評価投票数が3件未満、つまり評価投票数が少ないと、正確な推薦スコア予測を行えないので、システムは推薦スコア予測計算の対象ユーザから除外した。

評価実験は、図3に示した各処理用を起動するためのテストドライバを作り、7日間単位に、①新規アイテムの登録・②ユーザの閲覧情報登録・③ユーザの評価投票の登録を行い、④翌7日間に評価投票したユーザについて推薦アイテム上位20件のリストを作成し、⑤翌7日間のユーザ評価値と④のリストからMAEを観測した(測定方法は3.3節と同様)。

4.2.2 実験結果

表3は、実験を行った2004年1月1日～2008年12月31日における12月の最終週の測定値を表にしたものである。

先に、表3の列見出しの各要素の意味について説明する。伝統的CFとは、擬似投票を行わず、推薦スコア予測計算は式(1)と式(2)を利用した方式である。また、擬似投票のみとは、擬似投票は実行するが、推薦スコア予測計算は式(1)と式(2)を利用し、推薦予測の正確性を改良しない方式である。さらに、提案方式とは、擬似投票を行い、推薦スコア予測計算は式(3)と式(4)を利用した方式である。続けて、推薦可能ユーザ割合とは、各評価時期において推薦アイテムを受け取れるユーザの割合で、分母は図3のユーザの居住地情報領域に登録されている全ユーザ数で、分子は推薦アイテムを1つ以上受け取れるユーザ数である。また、MAE(絶対誤差)とは、式(5)を用いて算出する。算出方法は、4.1.1項で示した平均絶対誤差の算出と同じ方法である。

2008年の最後の評価期間において、推薦可能ユーザの割合は、伝統的CFが約18.1%、提

案方式が 98.8%である。表 3 のとおり、網羅性は、どの評価時期においても、提案方式が優れている。なお、各評価時期において、提案方式による推薦可能アイテムの割合は、100%である（表 3 の記載外事項）。

2008 年の最後の評価期間において、絶対誤差（MAE）は、伝統的 CF が 0.340、提案方式が 0.226 である。

4.2.3 評価

表 3 に示した実験結果を参照しながら、以下、高い網羅性の維持と推薦予測精度の向上の 2 点について、評価する。

高い網羅性の達成は、先行研究⁵⁾においてその方式を確立していた。このため、本研究で心がけたのは、擬似投票処理部分を変更せずに、網羅率の高さを維持することである*1。評価実験で対象としたアイテムは、ユーザによる評価投票が比較的密集している地域（首都圏と関西圏への配信アイテム）を選んだ。しかし、ユーザ-アイテム評価マトリクスの評価投票が比較的密集している地域でさえも、評価値疎ら問題の悪影響を受け、推薦アイテムを受け取れるユーザは 20%未満で、情報を推薦できない。一方、提案方式では、98.8%のユーザが何らかの推薦アイテムを受け取る。これは、商用環境においても十分な性能である。ただし、およそ 1%のユーザは推薦アイテムを受け取ることができない。これは、システムの前条件で、評価投票数が少ないと、正確な推薦スコア予測を行えないので、システムの推薦対象ユーザから外したためである。なお、推薦対象となりうるアイテムの網羅性は、提案方式の場合 100%を達成している。したがって、擬似投票方式は最適な方式で、これ以上改善の余地はない。

推薦予測精度の正確性向上の達成は、本研究の中心をなすテーマである。かつて擬似投票を行う段階で推薦予測性を向上させる手法を試みた。しかし、計算コストに見合う効果を見い出せなかった¹⁵⁾など、性能が悪かった。加えて、システムの稼働期間が長くなると、ユーザの嗜好変化を原因とする予測精度の低下が観測されるようになり、正確性を向上させるため、試行錯誤を繰り返していた。

伝統的 CF と擬似投票のみを実施したときの推薦予測精度の正確性は、伝統的 CF がわずかに良い。これは、擬似投票による投票ノイズが推薦予測精度を低下させているため

*1 提案方式は、推薦予測精度改良のために、推薦予測スコアを加点することでその調整を行う。すなわち、スコア Top N アイテムを推薦するとき、推薦の優先度（スコア）によって、推薦アイテムの順序は変更になるが、推薦可能ユーザ割合への影響はない。したがって、擬似投票のみの場合と提案方式の場合における推薦可能ユーザ割合は同値になる。

ある。一方、伝統的 CF と提案方式の推薦予測精度の正確性は、提案方式が良い。これは、擬似投票による投票ノイズを克服してさらに余ほどの性能が改善されたからである。評価実験では、季節の特長を反映した特長語ペアや、固有名詞と関連語の組合せによる特長語ペアを観測しており、それらは推薦予測精度の向上に貢献していた。

ユーザの嗜好変化による予測精度の低下を食い止めるため、「ユーザ評価と推薦予測の時間のずれ」に着目した性能改善を行った。提案方式では、評価投票が行われたときと、推薦予測を行うときの時間差に応じた重み付けを行うことで、絶対平均誤差が減少したので、ユーザの嗜好の変化にシステムの予測値が近づいたといえる。ここでは、時間差に応じた重み付けを行い、すべての評価投票値を利用しているので、網羅性への影響は出ていない。ただし、ユーザの嗜好の変化への追従は、時間差のほかにも精度を高める要因があると考えられるので、今後の性能改善のポイントでもある。

評価実験用のデータを貸与いただいた商用サイトは、システム稼働後 6 年 5 カ月が経過している。本研究の成果を得るまで、ユーザによる評価ランクが 3 段階（とても良い・まあまあ・いまいち）にしたことを後悔していた。それは、評価ランク数が多いと推薦予測精度が向上との憶測があったからである。しかし、提案方式のようにユーザによる評価値以外の情報、たとえば、アイテムの特徴・ユーザ評価の特徴・時間差による調整を利用することで、推薦予測精度が向上することを確認できた。

5. おわりに

本研究の目的は、擬似投票方式に基づくハイブリッドフィルタを用いて、網羅性を維持したまま、推薦予測精度を向上させる改良方式を提案することである。提案方式は「アイテムの内容の特徴・ユーザ評価の特徴・ユーザ評価時と推薦予測時の時間差」に着目して、2 点のアルゴリズムを改良した。1 点目は推薦スコア予測計算方式の改良で（式 (3)）、2 点目は被推薦ユーザと標本ユーザ間の嗜好の類似性計算方式の改良（式 (4)）である。

上 2 点を改良した提案方式と、伝統的 CF 方式を実装し、商用サイトにおける 5 年間にわたるユーザの評価履歴・Web サーバのログ・アイテムデータを使い評価実験を行い両者の性能比較を行った。この結果、研究目的「網羅性を維持したまま推薦予測精度を向上させる新たな手法の提案」の達成を確認できた。

本研究の評価実験では、複数のユーザコミュニティの存在する地域情報を推薦対象にした。これは、情報爆発時代における Blog や SNS に多くのユーザコミュニティがあるという点に似ている。つまり、あるユーザコミュニティから推薦予測精度向上に関する知識を発

見し、別のユーザコミュニティでその知識を活用できるので、Blog や SNS インフラを改造することなく、本方式を情報閲覧ツールとして適用できる。

現在、システムの稼働時間のさらなる長期化にともない、ユーザ評価履歴の蓄積量が増え続け、推薦スコアの予測計算量も増加している。今後は、現在の網羅性と予測精度に関する性能を維持したまま、推薦スコア予測計算量と保存データ量を減らす手法を確立したい。これが次の研究テーマである。

謝辞 評価実験用データを貸与していただきました株式会社サンケイリビング新聞社の皆様、評価実験用試験環境を借用させていただきました株式会社ネオジェイエスケーの皆様に感謝します。

参 考 文 献

- 1) Linden, G., Smith, B. and York, J.: Amazon.com Recommendations: Item-to-Item Collaborative Filtering, *Proc. February 2003 Issue of IEEE Internet Computing* (2003).
- 2) 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).
- 3) 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, pp.285–295 (May 2001).
- 4) Ungar, L.H. and Foster, D.P.: Clustering Methods for Collaborative Filtering, *Proc. Workshop on Recommender Systems at the 15th National Conference on Artificial Intelligence 1998* (1998).
- 5) 宇田隆幸, 藤井 敦, 石川徹也: 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, *情報処理学会論文誌*, Vol.45, No.5, pp.1246–1255 (2005).
- 6) Balabanovic, M. and Shoham, Y.: Fab: Content-Based Collaborative Recommendation, *Comm. ACM*, Vol.40, No.3, pp.66–72 (1997).
- 7) 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).
- 8) Sinha, S., Rashmi, K.S. and Sinha, R.: Beyond Algorithms: An HCI Perspective on Recommender Systems, *Proc. ACM SIGIR 2001 Workshop on Recommender Systems* (2001).
- 9) Cosley, D., Lam, S.K., Albert, I., Konstan, J.A. and Riedl, J.: Is seeing believing?: How recommender system interfaces affect users' opinions, *Proc. ACM SIGCHI*

2003, pp.585–592 (2003).

- 10) McNee, S.M., Riedl, J. and Konstan, J.A.: Accurate is not always good: How Accuracy Metrics have hurt Recommender Systems, *Proc. ACM SIGCHI*, pp.1097–1101 (2006).
- 11) McNee, S.M., Riedl, J. and Konstan, J.A.: Making Recommendations Better: An Analytic Model for Human-Recommender Interaction, *Proc. ACM SIGCHI*, pp.1103–1108 (2006).
- 12) 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).
- 13) 長尾 真, 水谷幹男, 池田浩之: 日本語文献における重要語の自動抽出, *情報処理学会論文誌*, Vol.17, No.2, pp.110–117 (1976).
- 14) Morita, M. and Shinoda, Y.: Information Filtering Based on User Behavior Analysis and Best Match Text Retrieval, *Proc. 17th Annual International ACM SIGIR Conference*, pp.272–281 (1994).
- 15) 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, *情報処理学会第 163 回自然言語処理研究発表会論文集*, pp.67–74 (2004).
- 16) Burke, R.: Hybrid Recommender Systems: Survey and Experiments, *User Modeling and User-Adapted Interaction*, Vol.12, Issue 4, pp.331–370 (2002).

(平成 21 年 5 月 18 日受付)

(平成 21 年 10 月 2 日採録)



宇田 隆幸 (正会員)

昭和 59 年図書館情報大学卒業。同年日本情報サービス (現、日本総合研究所) 入社。その後、ネオジェイエスケー代表取締役、アラン取締役 CTO、ネオジェイエスケー取締役等を歴任後、平成 19 年学校法人岩崎学園入職。現在、情報セキュリティ大学院大学に勤務。知識処理、Web2.0、情報共有環境のセキュリティに興味を持つ。この間、平成 21 年東北大学大学院博士後期課程修了。博士 (情報科学)。電子情報通信学会、人工知能学会、ACM 各会員。



木下 哲男 (フェロー)

昭和 54 年東北大学大学院修士課程修了。同年沖電気工業 (株) 入社。平成 8 年東北大学電気通信研究所助教授，平成 13 年同大学情報シナジーセンター教授，現在，サイバーサイエンスセンター教授。知識工学，エージェント工学，エージェント応用システム等の研究開発に従事。情報処理学会平元年度研究賞および平成 8 年度論文賞，電子情報通信学会平成 13 年度業績賞等。工学博士。電子情報通信学会，人工知能学会，日本認知科学会，IEEE，ACM，AAAI 各会員。
