

Personalizability を考慮した推薦システムの提案

吉川 大弘^{1,a)} 森 貴章² 古橋 武¹

受付日 2012年8月16日, 再受付日 2012年10月10日,
採録日 2012年11月5日

概要: 近年, インターネットの普及とそれともなう各種メディアの発達により, ユーザがアクセス可能な情報の量は膨大なものとなったが, 一方で, ユーザにとって価値のある情報を選択することは困難となっている. そのため, ユーザのアイテム選択を支援する“レコメンデーション”に注目が集まり, 様々な推薦システムが実用化され始めている. このような推薦システムの評価には, これまで主に“精度”が用いられてきたが, 近年, ユーザ満足度の観点から, 他にも様々な側面を評価する必要性が指摘され始めている. “Serendipity”はその中の1つであり, 推薦アイテムの目新しさ, 意外性を表す. 本論文では, Serendipityの一面として“Personalizability”を定義し, Personalizability を考慮した推薦手法を提案する. また, 提案手法をベンチマークデータに適用し, 推薦における精度と Personalizability のバランスが調節可能であること, 従来手法と比べ, Personalizability において優れていることを示す.

キーワード: 推薦システム, Serendipity, Personalizability, 評価指標, 定量化

A Proposal of Recommendation System Considering Personalizability

TOMOHIRO YOSHIKAWA^{1,a)} TAKAFUMI MORI² TAKESHI FURUHASHI¹

Received: August 16, 2012, Revised: October 10, 2012,
Accepted: November 5, 2012

Abstract: Recently, a user can access a huge amount of information with the popularization of Internet and the development of media associated with that. On the other hand, it becomes difficult for him/her to choose desirable or valuable ones. Then various recommendation systems have been studied and put into practical use to support users' selection. Though “accuracy” has been used as the evaluation of these systems so far, it is said that other evaluation factors are also needed from the aspect of the satisfaction of users. “Serendipity,” which includes novelty or unexpectedness, is one of the evaluation indexes for user satisfaction. This paper defines and quantifies “Personalizability” as a part of “Serendipity.” This paper proposes a recommendation system considering “Personalizability,” and applies the proposed method to benchmark data. It shows that this method can adjust the balance between “Accuracy” and “Personalizability” and it is superior to the conventional method in terms of “Personalizability.”

Keywords: recommendation system, Serendipity, Personalizability, evaluation index, quantification

1. はじめに

近年, インターネットの普及とそれともなう各種メディアの発達により, ユーザがアクセス可能な情報の量は膨大なものとなったが, 一方で, ユーザにとって価値の

ある情報を選択することは困難となっている. そのため, ユーザのアイテム選択を支援する“レコメンデーション”に注目が集まり, 様々な推薦システムが実用化され始めている [1], [2], [3]. たとえば, EC サイトの Amazon.com [4] では, ユーザの購入・閲覧履歴に基づき, ユーザが好むであろうアイテムを“おすすめ”として推薦している. また, 音楽配信サイトの Pandora [5] では, ユーザの評価履歴 (Like/Don't Like) に基づき楽曲の推薦を行っている.

このような推薦システムの評価には, これまで主に“精

¹ 名古屋大学
Nagoya University, Nagoya, Aichi 464-8603, Japan
² ブラザー工業株式会社
Brother Industries, Ltd., Nagoya, Aichi 467-8561, Japan
^{a)} yoshikawa@cse.nagoya-u.ac.jp

度”が用いられてきた [1]. 精度は、推薦アイテムのうちユーザが「Like」と答えたものの割合であり、手法の性能を測るうえできわめて重要な指標である。しかし、近年、ユーザ満足度の観点から、他にも様々な側面を評価する必要性が指摘され始めており、たとえば、Explainability [6] や Serendipity [1], [7] などがあげられている。Explainability は、推薦の根拠をユーザに理解可能な形で説明できるかどうかを表す。また Serendipity は、推薦アイテムが「Like」であることに加え、目新しさ、意外性があるかを表す。本論文では、この“Serendipity”に着目する。

ほとんどのユーザが好むアイテムのみを推薦すれば、一般的に精度は高くなるが、目新しさ、意外性は少ない [1]. よって、個人性の高いアイテムの方が Serendipity が高くなる可能性が高い、すなわち、個人性は Serendipity の一面であると考えられる。そこで本論文では、Serendipity の一面として“Personalizability”を定義し、その定量化を試みる。これは推薦アイテムの個人特化性であり、推薦アイテムが全体においてどれだけ好まれていない中でその個人に特化して好まれるかを表す指標である。多くの人には好まれていないアイテムが、個人的には好きであることは、その推薦アイテムに対する Serendipity の向上につながる期待できる。本論文では、推薦における Personalizability の向上が Serendipity の向上に結び付くとの仮定のもと、Personalizability を考慮した推薦手法を提案する。また、提案手法をベンチマークデータに適用し、推薦における精度と Personalizability のバランスが調節可能であること、従来手法と比べ、Personalizability において優れていることを示す。

2. 関連手法

2.1 想定する推薦システムと従来手法

本論文で着目する推薦システムを図 1 に示す。システムはユーザに 1 つのアイテムを推薦し、ユーザは推薦されたアイテムに対して評価 (Like/Don't Like) を行う。システムが持つデータベースにはユーザの評価履歴が蓄積されている。図 1 のデータベースにおいて、1 は“Like”，-1 は“Don't Like”とそれぞれ評価されたアイテムであることを示し、“0”は未評価のアイテムであることを表す。

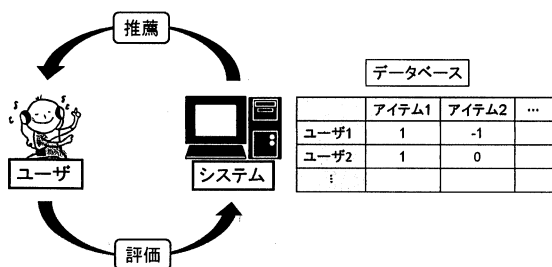


図 1 推薦システム

Fig. 1 Recommendation system.

本論文では、上述のようにして蓄積されたユーザの評価履歴情報を用いた、アイテムベース協調フィルタリングによる推薦システムを対象問題とする。アイテムのコンテンツ情報 (映画であればジャンル、監督、主演俳優など) をあわせて用いることで、より精度や Serendipity の高い推薦が可能となることも考えられるが、有効となる特徴量やその使い方は、アイテムの種類によって異なる。また、必ずしも、すべてのアイテムについて、それらのコンテンツ情報を持つことができるとは限らないため、ユーザの評価履歴情報のみに基づいた推薦システムを構築することは、実用上、有意義であると考えられる。

アイテムベース協調フィルタリングによる推薦システムのアルゴリズムには様々なものが用いられるが、代表的なものとして k -Nearest Neighbor 法 [8] があげられる。これは、各アイテムのスコアをその近傍 k 個のアイテムに対する評価履歴に基づいて算出し、最もスコアの高いアイテムを推薦する手法である。近傍アイテムは、各アイテムの評価履歴間の類似度を基に決定される。スコア算出式の例を以下に示す。

$$score(I_i) = \sum_{j=1}^k (s_{I_i I_j} \cdot e(I_j)) \quad (1)$$

ここで、 I はアイテムを表す。 I_j は I_i の近傍アイテムであり、 $s_{I_i I_j}$ は I_i - I_j 間の類似度を、 $e(I_j)$ は I_j に対する評価履歴 (1/-1) をそれぞれ表す。なお類似度 $s_{I_i I_j}$ は、アイテム間の類似度を表す値として任意の形で定義可能である。

また、最も単純な推薦アルゴリズムに Non-Personalized 法 [9] がある。これは、全ユーザに対し同じ推薦を行うものであり、本論文では、全ユーザ中における「Like」の割合が高い順にアイテムを推薦する手法を指す。

2.2 アソシエーションルールに基づく推薦

上述の k -Nearest Neighbor 法や Non-Personalized 法のほかに、アソシエーション分析におけるアソシエーションルールを用いた推薦方法がある。アソシエーション分析とは、データに内在する項目どうしの関連を、アソシエーションルールとして抽出する手法である [10]。ルールは $\{A \Rightarrow B\}$ の形式で表現され、 A を条件部、 B を結論部と呼ぶ。ルールには、その重要性を表す評価指標 [11] が存在し、閾値などの形で評価指標に関する条件を解析者が指定することによって、条件を満たすルールが抽出される。多く用いられる評価指標である *support*, *confidence*, *lift* を以下に示す。

$$support = P(A \cap B) = \frac{N(A \cap B)}{N_d} \quad (2)$$

$$confidence = P(B|A) = \frac{N(A \cap B)}{N(A)} \quad (3)$$

$$lift = \frac{P(B|A)}{P(B)} \quad (4)$$

ここで, N_d は全データ数, $N(A)$ は条件 A を満たすデータ数である.

アソシエーションルールに基づく推薦システムでは, ユーザの評価履歴をルールの条件部と見なし, 上述の評価指標 (*confidence*) を用いたスコアづけを未評価のアイテムに対して行い, スコアの高いアイテムを推薦する.

2.3 関連研究

Sarwar ら, Kim らは, 購入履歴データなどの 0/1 データを対象に, アソシエーションルールに基づく推薦手法を提案している [12], [13]. 両者とも, ユーザの評価履歴をルールの条件部と見なし, その評価指標を用いたスコアづけを各アイテムに対して行う点では, 提案手法と類似している. しかし, 用いる評価指標は *confidence* のみであり, また, アイテムのスコアを算出する際, 提案手法ではカルバック・ライブラ情報量上位 k_{kl} 個の和をとるが, Sarwar らは最大値を, Kim らは総和をとるなど, 推薦における *Personalizability* を考慮していないところに, 提案手法との大きな違いがみられる.

また, 1章で述べた *Serendipity* は, 感情的な応答をとまなうものであり, 定量的に評価することは難しいが, これまでいくつかの点から定量化の試みがなされている. Ziegler らは, 推薦するアイテムの多様性を重視することで *Serendipity* を高めようとした [14]. しかし, これにはアイテムの特徴 (メタ情報) が必要となる. Herlocker らは, グループ, もしくは全体向けの (Non-Personalized な) 推薦には現れないが, 個人向けの推薦には現れるアイテムを *Serendipity* が高いと見なしている [7]. またこの Non-Personalized な推薦との違いを, Chandrashekar ら, Schickel-Zuber らはそれぞれ “*Novelty*” と呼んでいる [15], [16].

奥らは, ユーザが2つのアイテムを混ぜ合わせて新たな推薦アイテムを生成するフュージョンベース推薦システムを提案し, 外発的/内発的偶然性, およびユーザの価値への結び付けを高めることを試みている [17]. 文献 [17] では, アイテムのコンテンツ情報を積極的に用いている点や, 複数のアイテムを提示し, ユーザがインタラクティブに推薦アイテムを生成している点など, 本論文で想定している推薦システムとは異なる点が多い. また Murakami らは, 評価対象の予測方法による結果と, プリミティブな予測方法による結果との差異を用いて “*unexpectedness*” という指標を定義し, アイテムの意外性を定量的に評価している [18]. この手法では, ユーザの嗜好モデルやアイテムのジャンル情報などを必要とするという点で, 提案手法とは異なる. また本論文で提案する *Personalizability* と *Unexpectedness* とでは, どちらが優れている (*Serendipity* に近い) という評価は難しく, また互いに排他的な関係にあるものではないため, 統合して用いることなども可能であると考えられる. Ge らは, 推薦システムにおけるユーザ

満足度向上のために “*coverage*” と “*serendipity*” の必要性, およびそれらのトレードオフについて述べている [19]. しかし文献 [19] では, アプローチは Murakami らの手法 [18] に従うとされており, また具体的なシステムや評価指標は示されていない.

2.4 推薦システムの評価指標

以下に, 精度, および *Novelty* [15], [16] に対する評価指標を示す.

精度

推薦回数を N , 推薦アイテムの集合を $I = I_1, I_2, \dots, I_N$, I_i に対する評価履歴を $e(I_i) = 1/-1$ とすると, 以下の式で表される.

$$\frac{1}{N} \sum_{i=1}^N t_i \quad (5)$$

$$t_i = \begin{cases} 1 & \text{if } e(I_i) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Novelty

式 (5) における t_i を以下のものとして表される.

$$t_i = \begin{cases} 1 & \text{if } e(I_i) = 1 \text{ and } I_i \notin \mathbf{I}_{NP} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

\mathbf{I}_{NP} は Non-Personalized 法における推薦アイテムの集合を表す. これは, 推薦アイテムが 「Like」 であり, かつ Non-Personalized な推薦には現れない割合である [15], [16].

3. 提案手法

3.1 *Personalizability* を考慮した推薦システム

提案手法は, “*Personalizability*” を考慮した, アソシエーションルールに基づく推薦手法である. 提案手法では, 2.2 節で述べた方法と同様に, 全ユーザの評価履歴をアソシエーションルールの条件部と見なし, 結論部にアイテムに対する評価 「Like」 をあてはめる. そして, そのルールの評価指標 *confidence* と *lift* を用いたスコアづけを未評価のアイテムに対して行い, 最もスコアの高いアイテムを推薦する.

confidence は, “同じ評価履歴を持つ人が, 対象となるアイテムをどれだけ好きであるか” を表し, 評価指標における “精度” の期待値に相当すると考えられる. しかし, たとえば, ある評価履歴を持つ人の 90% がアイテム a を好きであるが, そもそも全ユーザの 90% がアイテム a を好きである, といった場合のように, *confidence* による指標だけでは, Non-Personalized 法に近い, 個人性の低いアイテムが多く推薦されうる. そこで提案手法では, “ある評価履歴を持つ人があるアイテムをどれだけ好きであるか” と, “皆があるアイテムをどれだけ好きであるか” の比の値であ

表 1 アソシエーション行列
Table 1 Association matrix.

	アイテム 1 =Like	アイテム 2 =Like	アイテム 3 =Like
アイテム 1 =Like		0.88	0.60
アイテム 1 =Don't Like		0.18	0.90
アイテム 2 =Like	0.50		0.88
アイテム 2 =Don't Like	0.78		0.94

る lift をあわせて用いる。lift は, confidence が大きく, 全ユーザ中における「Like」の割合が低くなるほど大きな値となるため, 次節で示す評価指標 Personalizability の期待値に相当すると考えられる。以下, 提案手法の詳細について述べる。

提案手法では初めに, 全ユーザの評価履歴データから, アソシエーション行列を作成する。アソシエーション行列の例を表 1 に示す。行頭は各アイテムに対する評価「Like」と「Don't Like」, 列頭は「Like」のみである。各要素 a_{ij} は, 行列の行頭 A_i , 列頭 B_j を用いて以下のように算出される。

$$a_{ij} = confidence_{(A_i \Rightarrow B_j)} \cdot lift_{(A_i \Rightarrow B_j)}^\alpha \quad (8)$$

ここで, α は lift の重みに関するパラメータであり, α を小さくするほど lift の影響は小さくなる。 $\alpha \neq 0$ のとき, a_{ij} は lift が 1 より大きい場合には confidence より大きくなり, lift が 1 より小さい場合には confidence より小さい値となる。なお, アソシエーション行列の作成はオフライン (ユーザとのコネクションを行っていない状態) で行う。

ただし, 式 (8) の計算に用いられる $P(B|A)$, $P(B)$ には, Laplace expected error estimate [11], [20] に基づくスムージングを行い, その値はそれぞれ $(N(A \cap B) + 1) / (N(A) + 2)$, $(N(B) + 1) / (N_d + 2)$ となる。これは, データ件数が少数の場合の $P(B|A)$, $P(B)$ に補正を行うためであり, あてはまるデータ数が 0 のとき, $P(B|A)$, $P(B)$ はそれぞれ $1/2$ となる。4 章で示す実験における Non-Personalized 法においても, 同様のスムージングを行う。

ユーザに対するアイテムの推薦は, そのユーザの評価履歴とアソシエーション行列を用いて以下のように行う (図 2)。

- (a) アソシエーション行列から, 行頭がユーザの評価履歴に該当する行を抽出する。ただし, ユーザが評価済みアイテムに関する列は除く。
- (b) (a) で抽出した行群において, 列ごとに要素の値の和を取り, その値を各アイテムのスコアとする。

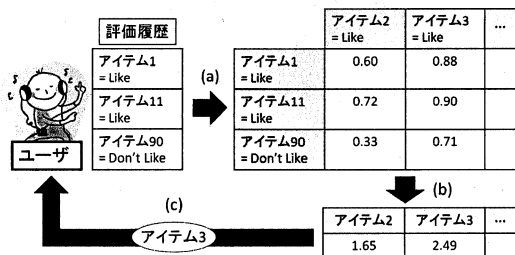


図 2 アイテムの推薦

Fig. 2 Recommendation of items.

- (c) 最もスコアの高いアイテムを推薦する。

ここで, (b) のスコア算出の際, アイテムの評価と関連がない評価履歴が多く用いられると, Non-Personalized 法による推薦と類似したものとなり, Personalizability が失われてしまう。そこで提案手法では, (b) のスコア算出において, 以下に示すカルバック・ライブラ情報量 [21] に基づき, 用いる評価履歴の選択を行う。

$$\begin{aligned}
 D_{KL} &= P(B|A) \log_2 \frac{P(B|A)}{P(B)} \\
 &\quad + P(\bar{B}|A) \log_2 \frac{P(\bar{B}|A)}{P(\bar{B})} \quad (9) \\
 &= confidence_{(A \Rightarrow B)} \log_2 lift_{(A \Rightarrow B)} \\
 &\quad + confidence_{(A \Rightarrow \bar{B})} \log_2 lift_{(A \Rightarrow \bar{B})} \quad (10)
 \end{aligned}$$

カルバック・ライブラ情報量は情報利得とも呼ばれる。ここでは, 評価履歴 (A) から得られる, アイテムの評価 (B/ \bar{B}) に関する情報量の期待値となり, “ある評価履歴が, あるアイテムの評価とどれだけ関連があるか” を表している。提案手法では, スコア算出に用いる評価履歴は, カルバック・ライブラ情報量が上位 k_{kl} 個のものとする。なお, ユーザに対するアイテムの推薦はオンライン (ユーザとのコネクションを行っている状態) で行う。

さらに提案手法では, アソシエーションルールに基づく推薦を行うことで, アイテムのスコア算出に用いた評価履歴を推薦の根拠として提示可能であり, Explainability においても優れていると考えられる。

3.2 推薦システムの評価指標

本論文では, 推薦アイテムが全体においてどれだけ好まれていない中で, その個人に特化して好まれるかを表す指標として Personalizability を定義し, その定量化を試みる。具体的には, 推薦したアイテムが「Like」であることに加え, 推薦アイテムの全ユーザ中における「Like」の割合の低さを情報量 [21] により表す。

Personalizability

式 (5) における t_i を以下のものとして表される。

$$t_i = \begin{cases} \log_2 \frac{1}{P(e(I_i)=1)} & \text{if } e(I_i) = 1 \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

式 (11) の上段は、推薦アイテムの全ユーザ中における「Like」の割合の低さを情報量 [21] により表したものである。ここでは対数の底として“2”を選んでいるため、「Like」の割合が 1/2 より小さい、すなわちマイノリティに好まれるアイテムにおいて、情報量は 1 より大きくなる。また、式 (8) と同様、 $P(e(I_i) = 1)$ の計算には、Laplace expected error estimate に基づくスムージングを行う。

4. 実験

4.1 使用データ

本実験では、ベンチマークデータである MovieLens データ [22] と Jester Joke データ [23] を用いた。MovieLens データは、映画に対する評価履歴データであり、ユーザ数は 943、アイテム数は 422 である。また、1 ユーザあたりの平均評価アイテム数は 76.6 である。評価は 1~5 の 5 段階でつけられているが、Like/Don't Like のバランスを考慮し、1~3 の評点を「Don't Like」、4、5 を「Like」に変換して実験を行った。また、Jester Joke データは、ジョークに対する評価履歴データであり、ユーザ数は 24,983、アイテム数は 100、平均評価アイテム数は 72.5 である。評価は -10.00~10.00 であるが、MovieLens と同様に、2.00 未満の評点を「Don't Like」、2.00 以上を「Like」に変換して実験を行った。

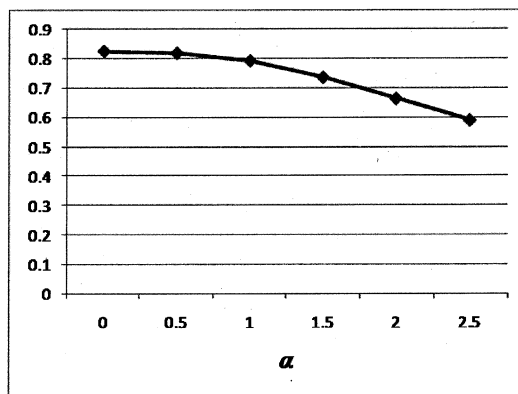
4.2 実験方法

実験手順は次のとおりである。初めに、4.1 節で述べた評価履歴データにおいて、「Like」と 21 回以上答え、「Don't Like」と 20 回以上答えたユーザを実験の対象とした（以降、単にユーザと呼ぶ）。ユーザ数は、MovieLens データにおいて 416、Jester Joke データにおいて 13,801 であった。次に、ユーザ群に対する 10-fold cross-validation [24] により、2.4 節で示した評価指標に基づき評価を行った。すなわち、ユーザ群をランダムに 10 分割し、その 9/10 でトレーニング（アソシエーション行列の作成）、1/10 でテスト（評価）を行うことを、分割されたすべてのグループについて繰り返した。

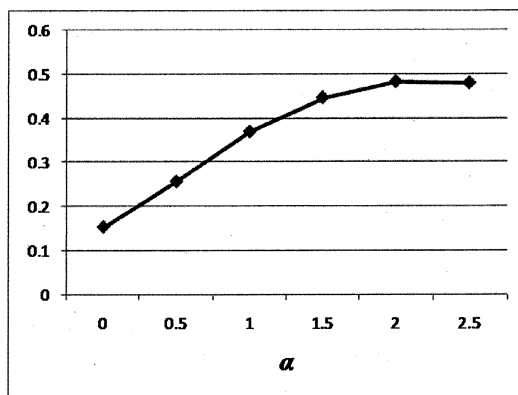
実験の際、テストユーザは新規ユーザ（全アイテムに対し未評価）として扱うが、初期の評価履歴として、「Like」と答えたアイテムの中からランダムに選ばれたものを 1 つ用いることとした。これは、新規ユーザには、最初に好きなアイテムを 1 つ入力してもらうことを想定している。評価履歴データで「未評価」となっているアイテムが推薦された場合は、カウントせず次のアイテム推薦（次にスコアの高いアイテム）を行い、評価済みアイテムが 10 個推薦されるまで推薦を行った。

4.3 実験 1

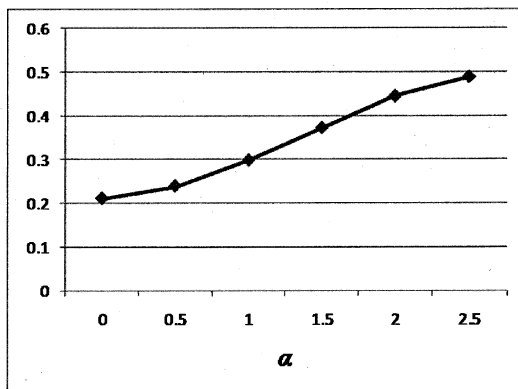
提案手法において、式 (8) に示す重みパラメータ α に対



(a) 精度



(b) Novelty



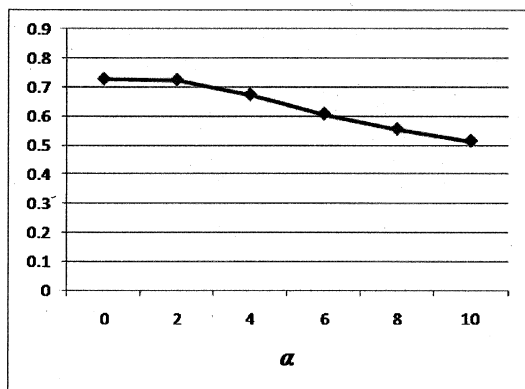
(c) Personalizability

図 3 重みパラメータ α に対する性能の変化 (MovieLens データ)

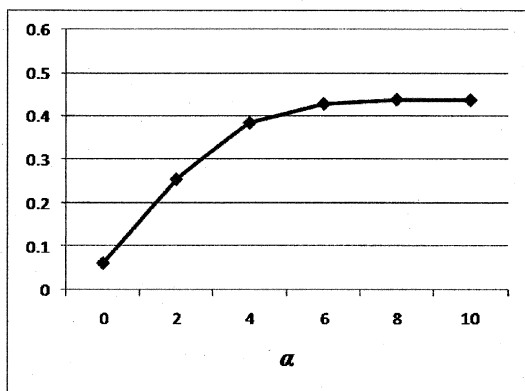
Fig. 3 Performance for difference of α (MovieLens data).

する性能の変化を調べた。MovieLens データ、Jester Joke データにおける実験結果をそれぞれ図 3、図 4 に示す。なお、提案手法における k_{kl} の値は、 $\alpha = 0$ の場合において最も精度が高くなるものを選び、MovieLens データ、Jester Joke データにおいてそれぞれ $k_{kl} = 5$ 、 $k_{kl} = 4$ となっている。

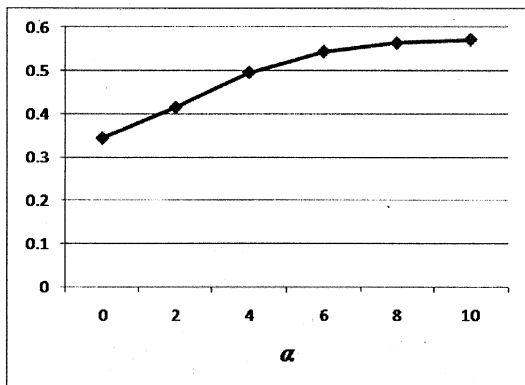
図 3、図 4 から、 α を大きくするほど、精度は低くなるものの、Novelty、Personalizability は高くなっていくことが分かる。これより、重みパラメータ α により、推薦における“精度”と“Personalizability”のバランスが調節可能



(a) 精度



(b) Novelty



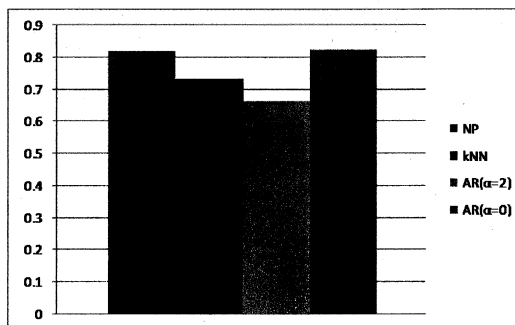
(c) Personalizability

図 4 重みパラメータ α に対する性能の変化 (Jester Joke データ)

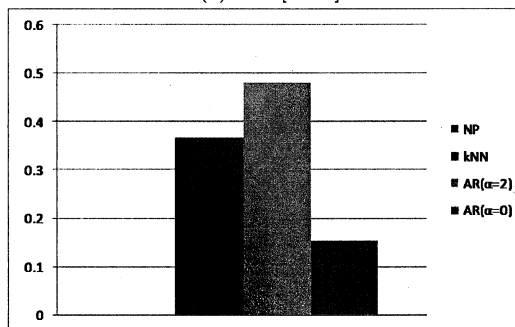
Fig. 4 Performance for difference of α (Jester joke data).

であると考えられる。実際の運用のうえでは、初期のうちには $\alpha = 0$ として精度を重視した推薦を行い、徐々に α を大きくして Personalizability 重視の推薦を行うことや、 α を変えたときの「Like」/「Don't Like」の評価状況の変化に応じて α を調整するなどの運用が考えられる。

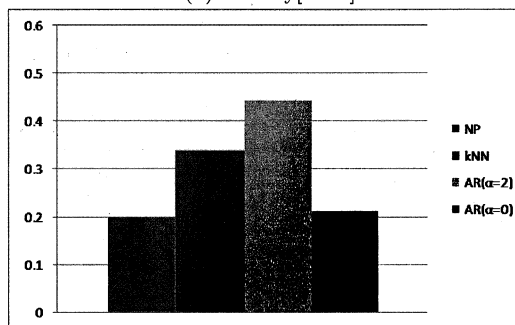
なお、図 3 と図 4 の α の値から見て取れるとおり、重みパラメータ α に対する性能の変化はデータによって異なる。式 (8) における *lift* は、アイテムの評価履歴間の関連性 (あるアイテムを Like/Don't Like であることと、別のアイテムを Like であること) の強さに関係している



(a) 精度 [2mm]



(b) Novelty [2mm]



(c) Personalizability

図 5 従来手法との性能比較 (MovieLens データ)

Fig. 5 Comparison of performance with conventional method (MovieLens data).

ため、全体的に関連が高いデータの方が α の変化に対して敏感になると考えられる。そこで、カルバック・ライブラ情報量の期待値である相互情報量 [21] を用いて、適用したデータにおける評価履歴間の関連性を調べた。その結果、MovieLens データ、Jester Joke データにおけるアイテム間の相互情報量の平均値は、それぞれ 0.0385, 0.0080 であった。このことから、MovieLens データの方が、評価履歴間の関連性が高く、それにより α の変化に対して各評価指標の変化が敏感になっていたと考えられる。

4.4 実験 2

提案手法 (AR (Association Rule)), Non-Personalized 法 (NP), *k*-Nearest Neighbor 法 (kNN) における性能比較を行った。MovieLens データ、Jester Joke データにおける結果をそれぞれ図 5, 図 6 に示す。ただし、図 5(b) お

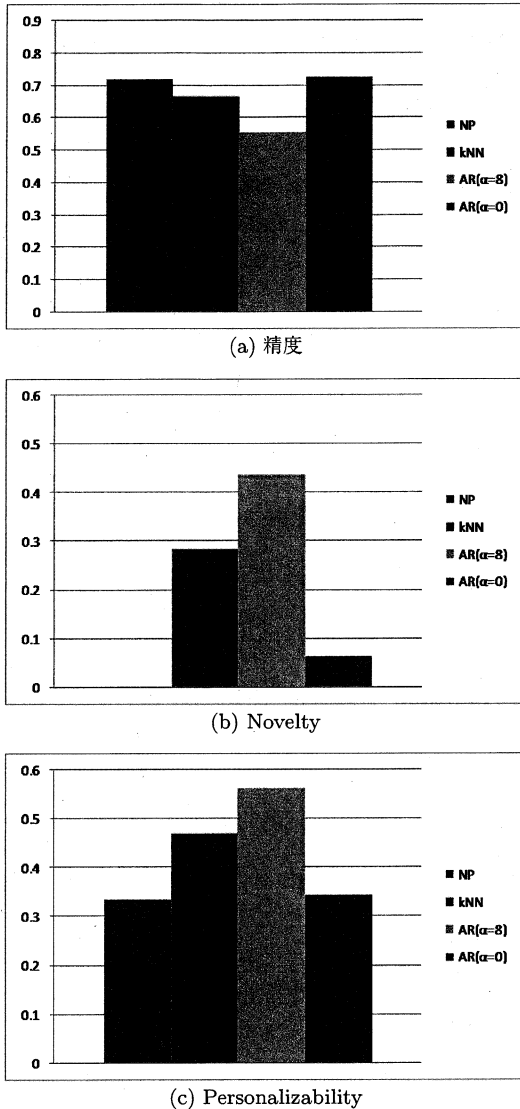


図 6 従来手法との性能比較 (Jester Joke データ)

Fig. 6 Comparison of performance with conventional method (Jester joke data).

よび図 6(b)において、その定義上、Non-Personalized 法による Novelty の値は 0 となる。なお、kNN におけるアイテム間類似度には、各アイテムに対する全ユーザの評価履歴をそのままベクトルの要素としたときのコサイン類似度 [8] を用いた。また、kNN における k の値は最も精度が高くなるものを選び、MovieLens データ、Jester Joke データにおいてともに $k = 10$ とした。

図 5, 図 6 から、提案手法において Personalizability を重視した場合、Novelty, Personalizability の基準においてともに最も高くなる事が分かる。また図 5(b), 図 6(b) から、提案手法 (Personalizability を重視) では Novelty のあるアイテムを、MovieLens データにおいて 48.1%, Jester Joke データにおいて 43.7% の割合で推薦していることが分かる。さらに図 5(c), 図 6(c)において、提案手法によ

る推薦において受け取る情報量は、Non-Personalized 法と比べ、MovieLens データにおいて 2.24 倍、Jester Joke データにおいて 1.67 倍であった。精度については、提案手法において精度を重視 ($\alpha = 0$) した場合、その性能を Non-Personalized 法と同程度にまで高くすることが可能であり、これは、 k -Nearest Neighbor 法に勝るものであることが分かる。

5. まとめ

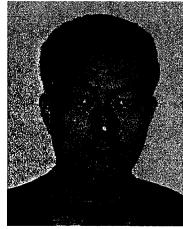
本論文では、推薦システムにおける Serendipity に着目し、Serendipity の一面として “Personalizability” を定義するとともに、その定量化を試みた。推薦における Personalizability の向上が Serendipity の向上に結び付くとの仮定のもと、Personalizability を考慮したアソシエーションルールに基づく推薦手法を提案した。提案手法では、ルールの条件部にユーザの評価履歴、結論部にアイテムに対する評価「Like」をあてはめ、その評価指標 *confidence* と *lift* を用いたスコアづけを未評価のアイテムに対し行い、最もスコアの高いアイテムを推薦する。さらにカルバック・ライブラ情報量に基づき、用いる評価履歴の選択を行う。提案手法を映画およびジョークの評価に関するベンチマークデータに適用し、*lift* に課した重みパラメータにより、推薦における “精度” と “Personalizability” とのバランスが調節可能であること、また、従来手法と比べ、Personalizability において優れていることを示した。

今後の課題として、ユーザベース協調フィルタリングにおける Personalizability の考慮方法に対する検討や、より大規模なデータに適用し、多くの推薦回数を設定した実験による、提案手法の有効性の検証を行うことがあげられる。またあわせて、提案手法による推薦アイテムについて、Serendipity としての妥当性を、ユーザ評価により検証していく必要がある。

参考文献

- [1] 神宮敏弘：推薦システムのアルゴリズム (1), 人工知能学会誌, Vol.22, No.6, pp.826-837 (2007).
- [2] 神宮敏弘：推薦システムのアルゴリズム (2), 人工知能学会誌, Vol.23, No.1, pp.89-103 (2008).
- [3] 神宮敏弘：推薦システムのアルゴリズム (3), 人工知能学会誌, Vol.23, No.2, pp.248-263 (2008).
- [4] Amazon.com, available from (<http://www.amazon.com>).
- [5] Pandora, available from (<http://www.pandora.com>).
- [6] Koren, Y.: Tutorial on recent progress in collaborative filtering, *Proc. 2008 ACM conference on Recommender Systems*, pp.333-334 (2008).
- [7] Herlocker, J.L., Konstan, J.A., Terveen, L.G. and Riedl, J.T.: Evaluating collaborative filtering recommender systems, *ACM Trans. Info. Syst.*, Vol.22 (2004).
- [8] Sarwar, B., Karypis, G., Konstan, J. and Reidl, J.: Item-based collaborative filtering recommendation algorithms, *Proc. 10th international conference on World Wide Web*, pp.285-295 (2001).

- [9] Lee, W.S.: Collaborative learning for recommender systems, *Proc. 18th international conference on Machine Learning*, pp.314-321 (2001).
- [10] Agrawal, R. and Srikant, R.: Fast algorithms for mining association rules, *Proc. 20th international conference on Very Large Data Bases*, pp.487-499 (1994).
- [11] Geng, L. and Hamilton, H.J.: Interestingness measures for data mining, A survey, *ACM Computing Surveys*, Vol.38 (2006).
- [12] Sarwar, B., Karypis, G., Konstan, J. and Riedl, J.: Analysis of recommendation algorithms for e-commerce, *Proc. 2nd ACM conference on Electronic Commerce*, pp.158-167 (2000).
- [13] Kim, C. and Kim, J.: A recommendation algorithm using multi-level association rules, *Proc. 2003 IEEE/WIC international conference on Web Intelligence*, pp.524-527 (2003).
- [14] Ziegler, C.N., McNee, S.M., Konstan, J.A. and Lausen, G.: Improving recommendation lists through topic diversification, *Proc. 14th international conference on World Wide Web*, pp.22-32 (2005).
- [15] Chandrashekar, H. and Bhasker, B.: Personalized recommender system using entropy based collaborative filtering technique, *Journal of Electronic Commerce Research*, Vol.12, No.3, pp.214-237 (2011).
- [16] Schickel-Zuber, V. and Faltings, B.: Inferring user's preferences using ontologies, *Proc. 21st national conference on Artificial Intelligence*, pp.1413-1418 (2006).
- [17] 奥 健太, 服部文夫: セレンディピティ指向情報推薦のためのフュージョンベースアプローチのユーザ評価, DEIM 2012 第4回データ工学と情報マネジメントに関するフォーラム, A1-3 (2012).
- [18] Murakami, T., Mori, K. and Orihara, R.: Metrics for evaluating the serendipity of recommendation lists, *JSAI 2007, Lecture Notes in Artificial Intelligence*, Vol.4914, pp.40-46 (2008).
- [19] Ge, M., Delgado-Battenfeld, C. and Jannach, D.: Beyond accuracy: evaluating recommender systems by coverage and serendipity, *Proc. 4th ACM conference on Recommender systems*, pp.257-260 (2010).
- [20] Yin, X. and Han, J.: CPAR: Classification based on predictive association rules, *Proc. 3rd SIAM international conference on Data Mining*, pp.331-335 (2003).
- [21] Kullback, S.: Information theory and statistics, Dover Publications, Mineola, N.Y. (1997).
- [22] Miller, B.N., Albert, I., Lam, S.K., Konstan, J.A. and Riedl, J.: Movielens unplugged: experiences with an occasionally connected recommender system, *Proc. 8th international conference on Intelligent User Interfaces*, pp.263-266 (2003).
- [23] Gupta, D., Digiovanni, M., Narita, H. and Goldberg, K.: Jester 2.0 (poster abstract): evaluation of a new linear time collaborative filtering algorithm, *Proc. 22nd annual international ACM SIGIR conference on Research and Development in Information Retrieval*, pp.291-292 (1999).
- [24] Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection, *Proc. International Joint Conference on Artificial Intelligence 1995*, pp.1137-1143 (1995).



吉川 大弘 (正会員)

1997年名古屋大学大学院博士課程修了。同年カリフォルニア大学バークレー校ソフトコンピューティング研究所客員研究員。1998年三重大学工学部助手。2005年名古屋大学大学院工学研究科 COE 特任准教授。

2006年10月同研究科准教授、現在に至る。主としてソフトコンピューティングとその応用に関する研究に従事。博士(工学)。IEEE, 電子情報通信学会, 日本知能情報フェジ学会, 進化計算学会, 計測自動制御学会各会員。



森 貴章

2008年3月名古屋大学大学院工学研究科計算理工学専攻博士課程前期課程修了。同4月ブラザー工業株式会社入社、現在に至る。



古橋 武

昭和60年名古屋大学大学院工学研究科博士後期課程電気系専攻修了。工学博士。平成16年名古屋大学大学院工学研究科計算理工学専攻教授、現在に至る。ソフトコンピューティング, 感性工学に関する研究に従事。平成8年

日本フェジ学会論文賞受賞。IEEE, 日本知能情報フェジ学会, 電気学会等の各会員。