

価値観アイテムモデリング手法を利用した 推薦理由提示手法についての考察

Consideration of Explanations for Recommender Systems with Personal-value-based Item Modeling

山口 貴之 服部俊一 高間康史*
Takayuki Yamaguchi Shunichi Hattori Yasufumi Takama

首都大学東京大学院システムデザイン研究科
Graduate School of System Design, Tokyo Metropolitan University

Abstract: 本稿では、価値観に基づくアイテムモデリング手法を提案し、これを用いた推薦理由の提示手法について考察する。近年、情報推薦では精度だけでなくその推薦過程をユーザに提示することで、システム全体に対する満足度の向上を意図した研究が行われている。提案手法ではユーザの価値観に着目し、相関ルールを用いて作成したアイテムモデルを推薦理由の説明に利用する。本稿ではアイテムモデルに関する予備実験結果について報告すると共に、推薦システムへの適用について考察する。

1 はじめに

近年、情報化技術の発展により、ユーザが膨大な情報の中から自分のニーズに合ったものを探すのが困難になるという問題が生じている。これに対する解決策として、ユーザの行動履歴から有用性の高い情報を推薦する情報推薦システムが注目されている。その中の一つに価値観に基づいたユーザモデルに関する研究があり、cold-start 問題 [2] や sparsity 問題 [3] に有用であることが示されている [1]。また、文献 [1] の研究では、ユーザモデリングを用いた手法を提案しているが、アイテムモデリングへの適用の可能性にも言及している。ここでのユーザモデリングとは特定ユーザのレビュー履歴からユーザのこだわりを求めるものであり、アイテムモデリングは特定アイテムに投稿されたレビューを収集し、そのアイテムがどの属性に着目して評価されているのかを求めることである。このことから一つのモデル構築に必要なレビュー数ではアイテムモデリングの方が集まりやすいという利点がある。さらにアイテムモデルを算出することで情報推薦システムの重要な要素技術の一つである、推薦理由の提示 [4] への応用が期待できる。

情報推薦システムにおいて、アイテムの推薦時にその推薦された理由を提示することは推薦アイテムに対して説得性を持たせることに繋がり、システム全体に

対する信用度を向上させる期待ができることから、近年、重要視されるようになってきている [4]。

そこで本稿では価値観に基づいたアイテムモデリングを元に、推薦理由の提示を行う手法を提案する。価値観に基づくアイテムモデリングでは、一般的に低評価でも、ある属性にこだわりの強い人は好む傾向にあることなどを推定可能であるため、低評価のアイテムに対しても推薦を行うことが可能になることが期待される。本稿ではモデリング結果を活かした推薦理由文について検討することを目的として、モデリング結果に対する解釈やその表現方法等をアンケートにより収集し、分析を行う。その結果に基づきアイテムモデリングに基づく推薦理由提示方法について考察する。

2 関連研究

2.1 協調フィルタリング

協調フィルタリングとは口コミによる推薦の過程を自動化したものであり、Amazon.com¹ などのショッピングサイトで幅広く利用されている推薦手法である。協調フィルタリングの代表的なものには、蓄積されたユーザの嗜好データから類似するユーザを予測し推薦を行うメモリベースと呼ばれる手法が存在する。メモリベースはユーザベース [6] とアイテムベース [7] の2つに大きく分類することが可能である。ユーザベース

*連絡先：首都大学東京大学院システムデザイン研究科
〒191-0065 東京都日野市旭が丘6-6
E-mail: ytakama@tmu.ac.jp

¹<http://www.amazon.com/>

では、類似した嗜好パターンを持つユーザを探し、そのユーザが好むアイテムを推薦するのに対し、アイテムベースは、ユーザの好むアイテムと類似したアイテムを推薦する。アイテムベースには、アイテムの類似度を事前に計算しておくことでユーザベースに比べ計算量を減らせる利点が存在する。

2.2 価値観に基づいたユーザモデル

価値観とはユーザがアイテムのどの要素を重視するかというものであり、消費行動に影響を与える要素であることからマーケティング等に利用されている。この価値観を用いることで、より少ない情報からユーザの嗜好や特性を推論できると考えられている。

ユーザベース協調フィルタリング [6] では他のユーザと共通に評価を行っているアイテムの評価値を元に Pearson 相関を用いて類似度を計算するのに対し、価値観に基づくユーザモデルを用いて協調フィルタリングを拡張した手法が提案されている [1]。拡張手法 [1] ではユーザの価値観 (Personal Values) に着目し、評価一致率と呼ばれる指標を用いて類似度を計算する。評価一致率の P_{uj} は、 u のアイテム i に対する評価極性 p_{ui} と、 i の属性 j に対する評価極性 p_{uij} から求められる。ユーザ u が評価した全アイテムについて、属性 j に対する評価とアイテムに対する評価の極性が一致した回数を $O(u, j)$ 、不一致であった回数を $Q(u, j)$ とした場合、評価一致率 P_{uj} は式 (1) のように計算される。

$$P_{uj} = \frac{O(u, j)}{O(u, j) + Q(u, j)} \quad (1)$$

これによりユーザモデルは属性数を n とした n 次元ベクトルで表され、Pearson 相関を用いて他のユーザとの類似度が計算される。レビューサイトのデータセットを用いた評価実験により、低評価のアイテムにおける MAE が低減する結果などが得られている [9]。

2.3 推薦理由に関する研究

情報推薦において、推薦アイテムと共にその推薦理由を提示することで、以下の様な効果が期待されており、近年その必要性が高まってきている [4]。

- ユーザのシステムに対する親密性や信頼性を高める。
- ユーザの満足度を上げる。
- ユーザが目的のアイテムを早く探せるようになる。
- 説得性を持たせユーザにアイテムを購入させやすくする。

特に協調フィルタリングにおいては、推薦の仕組みや過程がユーザに伝わらない状態でアイテムが推薦されるため、推薦理由を提示することが効果的であると指摘されている [5]。協調フィルタリングに関する推薦理由提示手法として、Amazon.com では推薦時に「この商品を買った人はこんな商品も買っています」という説明文を提示し、推薦の仕組みをユーザに示している。また文献 [5] では、対象ユーザと類似度の高いユーザに付与した評価値を、ヒストグラムや表、説明文などの形式で提示することを提案し、有用性を示している。

類似ユーザが対象アイテムに下した評価を、アイテムの推薦理由として用いる手法に関しては欠点も指摘されている。文献 [8] では、推薦アイテムに対するユーザの評価を推薦時とアイテム使用後で比較した結果、評価が大きく変動することを報告している。この問題の解決策として、ユーザの過去の評価アイテムが推薦に与えている影響力を数値化し提示する手法を提案し、類似ユーザの評価値を利用した手法よりも正確に評価を推定できることを示している。

3 価値観アイテムモデリング手法

3.1 アイテムモデリングへの適用

価値観に基づいたユーザモデリング手法 [1] では、あるユーザが評価した全アイテムのレビューよりユーザモデルを構築したのに対し、本稿で提案する価値観アイテムモデリング手法では、あるアイテムに評価されている全ユーザのレビューを元にアイテムモデルを作成する。提案手法ではどの属性に対する評価がアイテムの評価に影響を与えるのかを推論し、アイテムを推薦する。

3.2 評価一致率の拡張

2.2 節で述べたとおり、価値観ユーザモデリング手法 [1] では、アイテムに対する評価と属性に対する評価の極性が一致した回数に着目している。すなわち、好評で一致した場合と不評で一致した場合を区別せずに扱っている。提案手法では、アイテムの特性をより詳しく分析することを試みる。具体的には、アイテムに対する好評・不評、属性に対する好評・不評の組み合わせ 4 種類に分類してモデル化を行う。また、評価一致率の代わりにリフト値 [10] によって検出を行う。リフト値は、相関ルールの評価指標の一つであり、式 (2) で定義される。ここで、 $P(X)$ は事象 X の生起確率である。式よりリフト値は、 Y の生起確率が条件 X により何倍

に増加するかを表している。

$$lift(X \Rightarrow Y) = \frac{P(X \cap Y)}{P(X)P(Y)} \quad (2)$$

相関ルールでは、 X, Y はあるアイテム集合がトランザクションに含まれる事象を表すが、提案手法では属性に対する評価、アイテムに対する評価が X, Y にそれぞれ対応する。アイテム i に対するレビューの評価極性が $p_t (\in \{ \text{好評}, \text{不評} \})$ となる事象 X_i^t と、アイテム i に対するレビューにおいて属性 j の評価極性が $p_a (\in \{ \text{好評}, \text{不評} \})$ となる事象 X_{ij}^a の間のリフト値 $lift_{ij}(p_a \Rightarrow p_t)$ は式 (3) で定義される。

$$lift_{ij}(p_a \Rightarrow p_t) = \frac{P(X_{ij}^a \cap X_i^t)}{P(X_{ij}^a)P(X_i^t)} \quad (3)$$

例えば $lift_{ij}(\text{好評} \Rightarrow \text{好評})$ の値は、 i が好評となる確率が、属性 j が好評の場合にどのくらい高くなるかを意味している。

式 (3) を用いて、 p_a, p_t の組み合わせにより 4 通りの値が求められる。例として、ある映画に対して 4 人のユーザが評価を行った結果を表 1 に示す。また、表 1 の評価に基づいて属性ごとに 4 通りのリフト値を計算した例を表 2 に示す。ここで、「好」、「不」はそれぞれ好評、不評を表し、例えば 2 列目は属性に対する評価が好評の時に総合評価も好評となる場合のリフト値を意味する。この例では、属性「物語」が好評だとアイテム

表 1: アイテムにおける評価例

属性	ユーザ1	ユーザ2	ユーザ3	ユーザ4
総合評価	不評	不評	好評	不評
物語	不評	不評	好評	好評
映像	好評	好評	不評	不評

表 2: リフト値の計算例

属性	好⇒好	好⇒不	不⇒好	不⇒不
物語	2.00	0.67	0	1.33
映像	0	1.33	2.00	1.33

も好評になる傾向にあるといえる。また、属性「映像」に関しては不評の場合にアイテムは好評になると言える。このようにリフト値を計算することで、アイテムの評価に影響を与える「推薦時に重要度の高い属性の極性」を推論することができる。

3.3 推薦理由のための説明文作成

推薦理由の説明文を、アイテムモデルを元に作成する。前述のとおり、表 2 の場合では属性「物語」を好む

人はアイテムに対して高評価を付ける傾向にあることが推論できる。これより「物語を気に入った人は、この映画を好む傾向があります。」のような説明文の提示が考えられる。また属性「映像」については「映像を気に入らなかった人でも、映画自体には満足する傾向があります。」のような説明文を提示可能である。

4 予備実験

アイテムモデリングに基づく推薦理由提示方法について考察を行うため、予備実験を実施した。4.1 節に実験の概要を示し、結果を 4.2 節に示す。

4.1 実験概要

本実験では、同じ研究室に所属する 20 代の工学系大学院生 14 名に、提案するアイテムモデリングの結果を提示し、結果に対する解釈やこれを利用したアイテムの説明文等をアンケート形式で収集した。アンケートでは、アイテムの基本情報 (属性値やサムネイル画像等) とモデリング結果を提示し、実験協力者に「売り手として、アイテムの魅力を買い手に宣伝すること」を想定し、提示された情報を元にそのアイテムの魅力や特徴を説明する文を記述してもらった。アイテムモデリングの結果として、各属性における $(X_{ij}^a \cap X_i^t)$ を満たすレビューの数と、4 種類のリフト値の値を表及びレーダグラフにしたものを提示した。

旅行サイト 4travel² より宿泊施設 3 件 (アイテム 1, 2, 3)、映画情報サイト Yahoo!映画³ より映画 2 件 (アイテム 4, 5) の合計 5 件のアイテムを提示した。図 1 に、各アイテムについて提示したレーダーチャートを示す。

回答項目には文章の記述以外に、提示されたアイテムに対する認知度を 5 段階で評価してもらった。また説明文の記述に関しては、以下の 4 点に従ってもらった。

- 売り手として、買い手にアイテムの魅力を伝える表現を考えること。
- 売り手側としての評価や信頼を失わない表現にすること。
- 説明文に加え、提示情報内でその根拠としたものについても記述すること (理由文は買い手を想定した表現である必要はない)。
- リフト値を参考にした説明文を必ず 3 つ以上記述すること。

²<http://4travel.jp/>

³<http://movies.yahoo.co.jp/>

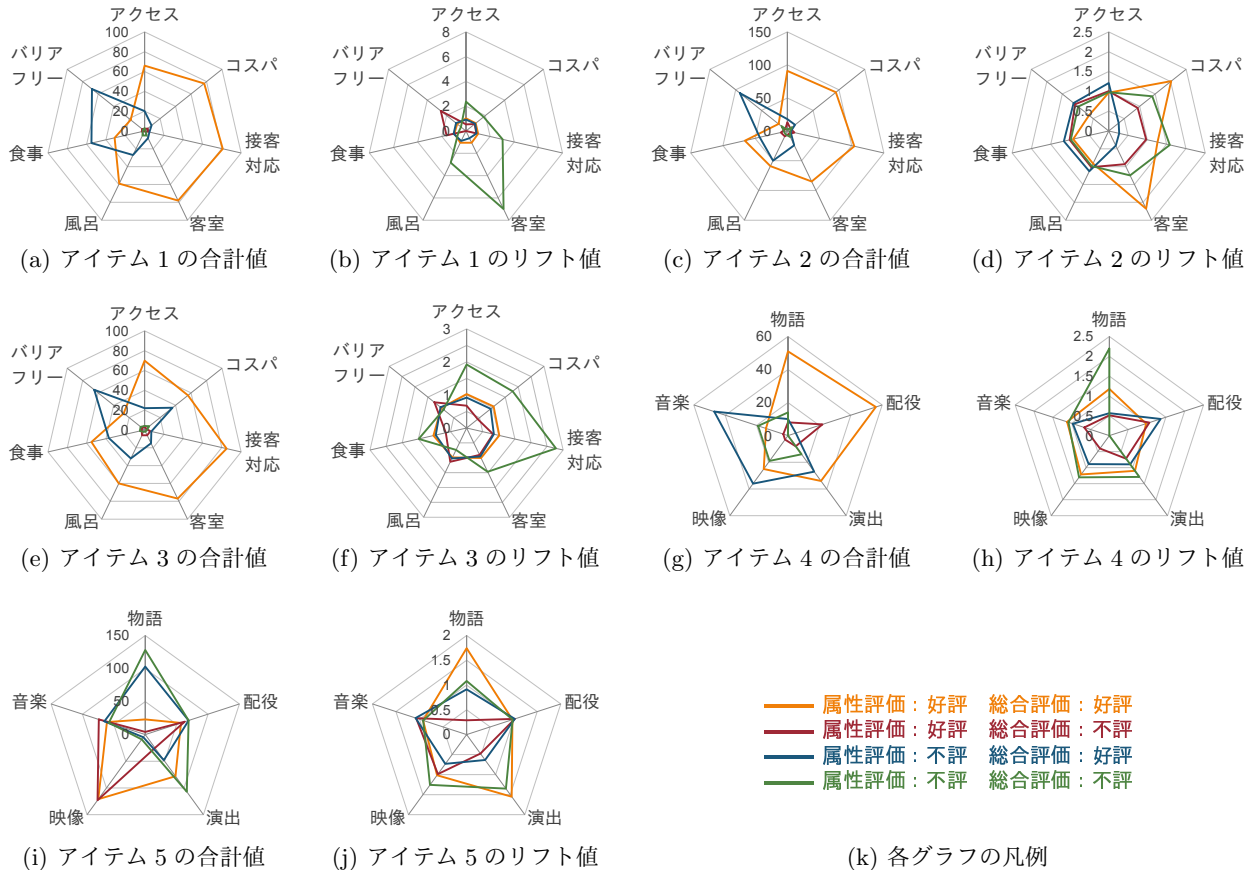


図 1: 実験にて提示したアイテム情報 (レーダチャート)

協力者にはアンケート前に、アイテムモデリングの作成方法やリフト値について説明を行った上で、2 週間の期限を設けアンケートに回答してもらった。

4.2 実験結果

アンケート結果より、アイテムに対する認知度をまとめたものを図 2 に示す。図 2 より、対象のアイテムの利用経験による差は小さく、協力者はほぼアンケート内の提示内容のみからアイテムの特徴を捉え、説明文を作成していると考えられる。

アイテムモデリングに対する説明文は、協力者 14 名より 221 個の文章を得られた。得られた文章について、「リフト値が大きいこと」を根拠として作成された説明文は 190 個存在し、その中には「『属性』がおすすめです」、「『属性』が気に入られています」等の共通した説明文が見られた。そこで、回答された説明文を以下の 13 種類に分類し、それぞれの頻度を集計した結果を表 3 に示す。

1. 「属性」が好評 (おすすめ) である。
2. 「属性」を好評とした人は、アイテム自体には満足する傾向がある。

3. 「属性」に興味やこだわりを持つ人には、アイテムをおすすめできる。
4. 「属性」が不評である。
5. 「属性」を不評とした人は、アイテム自体には満足しない傾向にある。
6. 「属性」に興味やこだわりを持つ人には、アイテムをおすすめできない。
7. 「属性」がアイテムの評価に影響を与えている。
8. 「属性」を好評とした人は、アイテム自体には満足しない傾向にある。
9. 「属性」を不評とした人は、アイテム自体には満足する傾向がある。
10. 「属性」に興味やこだわりを持たない人でも、アイテムをおすすめできる。
11. 「属性」に興味やこだわりを持たない人には、アイテムをおすすめできない。
12. 「属性」がアイテムの評価に影響を与えていない。
13. その他

ここで 1 行目の文章の種類 (Index) は、分類した 13 種類の文章の番号に対応し、例えば一行一列目は「 $lift_{ij}$ (好評 \Rightarrow 好評) の値が大きい」という理由で、「『属性』はおすすめです」と表現した説明文が 35 件あったことを示している。各リフト値の中で高頻度の 3 種類を赤字で示している。

表 3: リフト値が大きいことを根拠とした、説明文の分類結果

説明文を記述した理由	文章の種類 (Index)												
	1	2	3	4	5	6	7	8	9	10	11	12	13
$lift_{ij}$ (好評⇒好評)が大きい	35	7	19	0	0	2	1	0	0	2	0	0	9
$lift_{ij}$ (好評⇒不評)が大きい	3	1	2	0	1	4	0	1	0	11	0	7	2
$lift_{ij}$ (不評⇒好評)が大きい	2	1	0	1	0	3	0	0	0	15	0	5	2
$lift_{ij}$ (不評⇒不評)が大きい	6	0	1	18	1	21	2	1	0	2	0	0	8

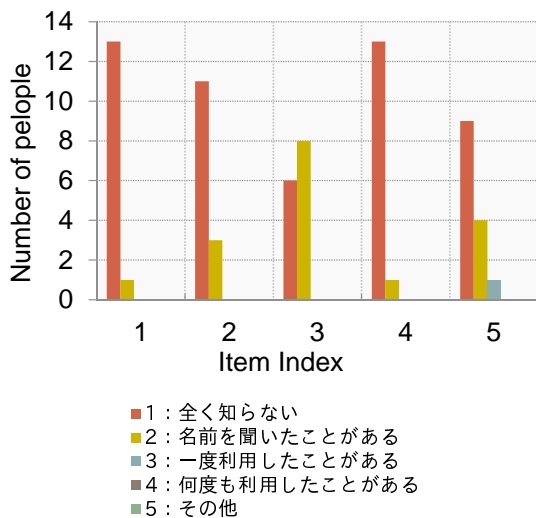


図 2: アイテムに対する認知度のグラフ

表 3 より, $lift_{ij}$ (好評 ⇒ 好評) の値が大きいという理由により, (1)「『属性』が好評である」, (3)「『属性』に対してこだわりを持つ人にはおすすめできる」という表現が多く作成されたことがわかる. また $lift_{ij}$ (不評 ⇒ 不評) の値が大きいとき, (4)「『属性』が不評である」, (6)「『属性』に対してこだわりを持つ人にはおすすめできない」という解釈のされ方が多いことがわかる.

複数のリフト値を考慮した説明文も存在した. 例えば, $lift_{ij}$ (好評 ⇒ 好評) と $lift_{ij}$ (不評 ⇒ 不評) の値が両方とも大きい時,「このアイテムは『属性』によって評価が分かれる, 好き嫌いの大きいアイテムである」等の説明文が多く見られた. これらについては, その他に分類した. $lift_{ij}$ (好評 ⇒ 好評) と $lift_{ij}$ (不評 ⇒ 不評) において (13)「その他」に分類されている説明文が多いのはそのためである.

$lift_{ij}$ (好評 ⇒ 不評) や $lift_{ij}$ (不評 ⇒ 好評) の値が大きいという理由により, (10)「『属性』に対してこだわりを持っていないくても, アイテムをおすすめできる」という表現が多くされている. これらは, 属性の評価がアイテムの評価に与えている影響は小さいという解釈に基づくものと言える.

回答された説明文の文中には,「リフト値」や「確率」

といった用語を使用した表現は一切見られなかった. また $lift_{ij}$ (好評 ⇒ 好評) に関する説明文以外では「『属性』を好評 (不評) とした人はアイテム自体を満足とする (しない) 傾向にある」といった表現はあまり見られなかった. これらの結果は協力者が売り手側としてアイテムの宣伝を考慮し, 直接的な表現を控えたためと考える.

回答された説明文の中には, 推薦システムで利用可能な有用な表現も存在した. 例えば $lift_{ij}$ (不評 ⇒ 不評) の値が大きいとき,「この『属性』にこだわりがある方は, レビューなどを確認の上ご利用下さい」,「この『属性』の評価はアイテムの評価に大きな影響を与えています」等の表現が見られた. また, $lift_{ij}$ (好評 ⇒ 不評) のリフト値が大きいとき,「この『属性』を重視する方は, 他の要素も気にして判断していただくと良いです」という表現が見られた. このような表現は「不評, おすすめできない」といった消極的な表現を使わないため, アイテムに対するイメージを下げずにユーザに注意を促すことが期待できる.

また各リフト値がとる値にあまり差が見られない場合,「バランスのとれたアイテムです」,「各要素よりもアイテム全体を見て評価されている」等の表現がされていた. 一方, リフト値に大きな特徴が見られなくても $(X_{ij}^a \cap X_i^t)$ を満たすレビューの数から判断している回答も存在した.

4.3 推薦システムでの利用に関する考察

実験結果を踏まえ, 4 種類のリフト値それぞれを根拠とした説明文として, 表 4 に挙げるものが利用可能と考える. 推薦システムでの利用方法としては, 各リフト値が設定した閾値を超えた場合に該当する説明文をテンプレートとして生成することが考えられる.

またリフト値に大きな特徴が見られない場合は, $(X_{ij}^a \cap X_i^t)$ を満たすレビューの数による判断を行うか,「バランスのとれた (平均的に見られている) アイテムです」等の説明を行うことが考えられる.

表 4: 閾値を超えたリフト値に応じて提示する説明文の例

リフト値の種類	提示する文章の例
$lift_{ij}$ (好評⇒好評)	<ul style="list-style-type: none"> この「属性」に興味(こだわり)を持つ方にはおすすめてできます。 この「属性」の評価は、アイテム全体の評価に大きな影響を与えています。 この「属性」の評判は良いです。
$lift_{ij}$ (好評⇒不評)	<ul style="list-style-type: none"> この「属性」にこだわりを持たない方は、他の要素で判断するほうが良いかもしれません この「属性」の評価が、アイテムの評判に繋がるとはあまりありません。 この「属性」に興味(こだわり)を持たない方にもおすすめてできます。
$lift_{ij}$ (不評⇒好評)	<ul style="list-style-type: none"> この「属性」の評価が、アイテムの評判に繋がるとはあまりありません。 この「属性」に興味(こだわり)を持たない方にもおすすめてできます。
$lift_{ij}$ (不評⇒不評)	<ul style="list-style-type: none"> この「属性」に興味(こだわり)を持つ方は、事前にそれに関するレビューに目を通すことをおすすめてします。 この「属性」の評価は、アイテム全体の評価に大きな影響を与えています

5 おわりに

本稿では推薦システムにおいて、価値観アイテムモデリング手法を提案した。提案手法では、価値観に基づいたユーザモデルに関する先行研究を拡張し、アイテムモデルに適用した。評価一致率に代わる指標として、リフト値を用いてモデリングを行った。

モデリング結果に対する協力者の意見を予備実験にて収集し分析することで、アイテムモデリングに基づく推薦理由の説明文生成の可能性について考察し、推薦システムでの利用について検討した。

今後は、提案手法を用いた推薦システムを構築する予定である。そのために、各リフト値を根拠とした説明文を生成する際の閾値の設定、およびテンプレート文章の増加について検討を行う。

参考文献

[1] S. Hattori and Y. Takama, “Consideration about Applicability of Recommender System Employing Personal-Value-Based User Model,” TAAI2013, pp.282-287, 2013.

[2] A. I. Schein et al, “Methods and metrics for cold-start recommendations,” Proceedings of the 25th annual international ACM SIGIR confer-

ence on Research and development in information retrieval, pp. 253-260, 2002.

[3] S. Lee, J. Yang and S. Park, “Discovery of hidden similarity on collaborative filtering to overcome sparsity problem,” Discovery Science. Springer Berlin Heidelberg, pp. 396-402, 2004.

[4] N. Tintarev and J. Masthoff, “A survey of explanations in recommender systems,” IEEE 23rd International Conference on Data Engineering Workshop, pp. 801-810, 2007.

[5] J. L. Herlocker, J. A. Konstan and J. Riedl, “Explaining collaborative filtering recommendations,” 2000 ACM conference on Computer supported cooperative work, pp. 241-250, 2000.

[6] P. Resnick et al, “GroupLens: an open architecture for collaborative filtering of netnews,” 1994 ACM conference on Computer supported cooperative work, pp. 175-186, 1994.

[7] G. Linden, B. Smith and J. York, “Amazon.com recommendations: Item-to-item collaborative filtering,” IEEE Internet Computing, Vol. 7, No. 1, pp. 76-80, 2003.

[8] M. Bilgic and R. J. Mooney, “Explaining recommendations: Satisfaction vs. promotion,” Beyond Personalization 2005, pp. 13-18, 2005.

[9] 三澤 遼理, 服部 俊一, 高間 康史, “価値観に基づくユーザモデルによる協調フィルタリングの拡張手法の提案,” 第 27 回人工知能学会全国大会 (JSAI2014), 1H4-NFC-01a-5, 2014.

[10] R. J. Bayardo Jr and R. Agrawal “Mining the most interesting rules,” 5th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 145-154, 1999.