

価値観に基づく情報推薦のためのレビュー閲覧履歴を用いたモデリング手法の検討

Consideration on User Modeling Method from User's Browsing Histories of Customer Reviews toward Personal Value-based Recommender System

清水 涼人†
Suzuto Shimizu

服部 俊一‡
Shunichi Hattori

高間 康史‡
Yasufumi Takama

1. はじめに

本稿では価値観に基づく情報推薦のためのレビュー閲覧履歴を用いたモデリング手法について検討する。ユーザの嗜好や特性に合わせたアイテムを提示する情報推薦システムが注目されており、代表的な手法として情報フィルタリングや内容ベースフィルタリングが存在する[1]。協調フィルタリングは嗜好が似ているユーザの閲覧履歴や購買履歴を利用してアイテムを推薦する手法であり、処理の手軽さなどから広く活用されている。しかし、協調フィルタリングでは新規のユーザやアイテムに対して十分な情報を得られないことから推薦精度が低下してしまうという cold-start 問題が指摘されている[2]。内容ベースフィルタリングは、ユーザが好むアイテムの属性値を取得し、同じ値を含むアイテムを推薦する方法である。しかし、ショッピングサイトなどには多くのアイテムやその属性値が存在することから、新規ユーザに対して推薦に十分な情報を集めることは困難と考える。

一方、個人の嗜好や消費行動を指定するための概念として価値観がある。価値観を、ユーザがどの属性を重視してアイテムの評価を決定するか、いわば属性に対する「こだわり」を表すと定義し、情報推薦に利用する手法が提案されている[5,6]。既存手法では、レビュー記事から投稿者の価値観に基づくユーザモデルを求めているが、レビュー文をユーザが閲覧・評価した履歴を用いてユーザの価値観をモデリングできれば、より多数のユーザに対してモデリングが可能になると考える。本稿では、レビュー評価履歴に基づくユーザモデリング手法を提案し、レビュー評価アンケート作成に必要な評価情報数について考察する。

2. 関連研究

情報推薦システムは、ユーザにとって有用と思われる情報を選別し、ユーザの目的に合わせて提示するシステムである。その代表的な手法として、協調フィルタリングや内容ベースフィルタリングが存在する[1]。しかし、従来の方法では新規ユーザやアイテムに対して有用な推薦を行えない cold-start 問題が指摘されている[2]。

また、ユーザが商品についての印象や評価などを書いたレビューが多くのショッピングサイトやクチコミサイトに投稿されている。レビューはユーザやアイテムの特性を分析するにあたって重要な情報源であるとされ、レビューを対象とした研究が広く行われている。平山らはレビュー文の中に含まれる商品の機能や特徴に対する評価文を解析し、その商品に対する全てのレビューを要約することでユーザ

のレビュー内容把握を支援する手法を提案している[3]。岡田らは、ユーザが書いたレビューテキストには嗜好やアイテムの特徴に関する情報が表現されていると考え、レビューテキスト間での類似度を用いて、情報推薦を行う手法を提案している[4]。

3. 価値観に基づくユーザモデル

3.1 評価一致率を用いたユーザモデリング手法

本稿で構築対象とする、価値観に基づくユーザモデルとは、ユーザの属性に対するこだわりの強さ、すなわちアイテムの各属性が評価に与える影響度をモデル化したものであり、評価一致率[5]と呼ばれる指標で表される。ユーザ u がアイテム i に対して行った評価 $e_{ui} \in E_u$ において、あるアイテムの極性 $P_{item}(u, i)$ と属性 j の極性 $P_{attr}(u, i, j)$ が一致するかどうかを調べ、一致する評価の回数を $O(u, j)$ とし、一致しない評価の回数を $Q(u, j)$ とする。この時、ユーザ u における属性 j の評価一致率 $P(u, j)$ は式(1)で算出される。ユーザのこだわりを表すユーザモデルは属性数を m とすると m 次元のベクトルとして表され、評価一致率が高い属性は評価への影響度が高い、すなわち推薦時に重要な属性であると推論される。

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

3.2 レビュー評価履歴に基づくユーザモデリング手法

ショッピングサイトには多くのユーザがアイテムに対してレビューを投稿しており、投稿されたレビューにはそのユーザのアイテムに対するこだわりが反映されていると考える。また、サイト利用者は、レビュー記事の中で自身がこだわりを持つ属性についての言及や評価を確認し、意思決定に利用していると仮定する。この仮定に従えば、ユーザがどのレビューを参考にしてアイテムの評価を決定したか判断できれば、ユーザの価値観をモデリングできると考える。モデリングに必要なのはユーザのレビュー評価履歴のみであることから、ユーザに対して大きな負担をかけることなくユーザモデルを作成可能という利点も期待できる。

レビューには、ユーザがアイテムの評価を決める理由が書かれており、それは多くの場合アイテムの属性に対する評価という形で表される。ユーザがアイテム評価の参考になったと判断したレビューについて、そのアイテムおよび各属性に対する評価情報が得られれば、評価一致率を用いてユーザのこだわりをモデリングできると考える。

モデリングの例として表 1 に示すように、あるアイテムに対して 3 件のレビューが投稿されている場合を考える。ここで、ユーザがこれらの中でレビュー Z が参考になったと回答したとする。加えて、ユーザがその商品を好評と評

† 首都大学東京 システムデザイン学部

‡ 首都大学東京大学院 システムデザイン研究科

表1: あるアイテムに対するレビューの例

	レビュー文	属性別評価
レビューX	1日500枚取ってもバッテリーが持つので良い。またポケットに入れられるので持ち運びが簡単。	バッテリー: ★5 携帯性: ★4
レビューY	デザインさ高級感溢れてて素晴らしいが、画質が思ったよりも粗い。	デザイン: ★5 画質: ★2
レビューZ	ボタンがでかくて操作しやすく、手振れ補正など機能が充実している。しかし、液晶は汚すぎる。	操作性: ★5 機能性: ★4 液晶: ★1

価した場合、属性「操作性」「機能性」に対する評価は好評でアイテムに対する評価と一致しており、「操作性」「機能性」に着目してアイテムの評価を決めたと推論できる。一方、属性「液晶」は不評でアイテムに対する評価と一致していないことから、「液晶」アイテムの評価にあまり影響していないことが推論される。以上の例から評価一致率を求めた結果を表2に示す。

表2: モデリングした価値観の例

属性	評価一致率	こだわり
操作性	1.0	強い
機能性	1.0	強い
液晶	0.0	弱い

4. アンケートによるレビュー評価履歴収集

4.1 概要

レビュー閲覧履歴から価値観に基づくユーザモデルの構築可能性について検討するため、アンケート形式でユーザの評価履歴を収集する。アンケートに掲載するレビューは、レビュー文に加えて属性別評価が5段階形式で投稿されている価格.com¹を対象とし、ジャンルは「デジタルカメラ」「デジタル一眼カメラ」を対象とする。この時、ユーザに提示するレビューの質が、結果として得られるモデリングの質に影響すると考える。本稿では、モデリングに用いるレビュー文の選択基準として、以下の条件に合致するものを収集して利用する。

- (1) 全属性が高評価(低評価)ではない
- (2) 特定の属性に対して根拠を示し言及している
- (3) レビュー文の評価と属性別評価(5段階形式)が一致する

全属性が同じ評価である場合、どの属性を重視して評価を決定したか判断できないことから、条件(1)に合致するレビューを集める必要があると考える。また、ユーザはこだわりのある属性に着目してレビューを評価するとの仮定から、条件(2)に示したように、特定の属性に対して根拠を示した上で言及している必要がある。加えて、ユーザがレビュー文を評価した際に、提案手法ではレビューに付属する属性別評価(5段階形式)を用いてモデリングすることから、条件(3)で示したレビュー文の評価と属性別評価が一致している必要がある。また、1つのレビューで常に全属性に対して言及されているとは限らないため、1つのアイテムに対して言及されている属性がそれぞれ異なるレビューを収集し、全ての属性を網羅できるようなレビューの組を1つ

表3: レビュー60件における属性別評価

属性	被言及数	属性	被言及数
デザイン	27回	携帯性	30回
画質	48回	機能性	45回
操作性	45回	液晶	29回
バッテリー	35回	ホールド感	32回

の設問とする。ユーザは各設問において1組のレビューを閲覧し、最も参考になったと感じたレビューを回答する。加えて、そのレビューを踏まえたアイテムへの評価についても回答してもらうことで、ユーザがどの属性を重視して評価を決定したか判断できる。

4.2 抽出結果と必要な評価情報数について

実際のレビュー記事を収集して調査した結果、1つのアイテムに対して3件のレビューがあれば全ての属性を網羅可能との見通しが得られた。また、ユーザの負担を考慮して設問数は20問以下とする。4.1節で述べた条件に基づいて手動で抽出したレビュー60件において、8個の属性に対してそれぞれ言及されている頻度を表3に示す。3件のレビューの中から参考になったと感じた1件を選択することから、収集できる評価情報の期待値は属性に対する被言及数の1/3となる。また、属性あたり7件の評価情報が得られれば評価一致率を概ね正確に計算できるとされていることから[6]、レビュー60件中21回以上言及されている属性はモデリングに必要な量の評価情報が得られることが期待できる。表3から全ての属性において必要な被言及数21回を上回っていることから、本稿で作成する20問からなるアンケートを用いて、モデリングに十分な評価情報を得られることが期待できる。ユーザの回答によっては評価情報が7件に満たさない場合も起こり得るが、その場合はユーザの関心が低い属性としてこだわりが弱い属性と判断できる。

5. おわりに

本稿では価値観に基づく情報推薦のためのレビュー閲覧履歴を用いたモデリング手法を提案した。また、モデリングのためのアンケートで用いるレビューを収集し、モデリングに必要な評価情報が含まれることを示した。今後は提案手法を用いたアンケートを実施し、作成したユーザモデルからアイテムの推薦を行うことで提案手法の有用性および課題について考察する。

参考文献

- [1] 神嶋 敏弘, 推薦システムのアルゴリズム(1), 人工知能学会誌 22巻6号, pp.826-837, 2007.
- [2] A. I. Schein, A. Popescul, L. H. Ungar and D.M. Pennock, "Methods and metrics for cold-start recommendations," Proc. of SIGIR, pp.253-260, 2002.
- [3] 平山 拓央, 湯本 高行, 新居 学, 高橋 豊, 属性評価モデルに基づく商品評価の抽出と提示, DEIM 2011, F2-5, pp.1-7, 2011.
- [4] 岡田 瑞穂, 藤井 敦, レビューテキスト間の類似度を用いた協調フィルタリング, NLP2012, F3, 2012.
- [5] 服部 俊一, 高間 康史, 価値観に基づくユーザモデルを用いた情報推薦手法に関する検討, 第3回インタラクティブ情報アクセスと可視化マイニング研究会, pp.1-6, 2013.
- [6] Shunichi Hattori and Yasufumi Takama, "Proposal of User Modeling Method Employing Reputation Analysis on User Reviews Based on Personal Values," JSAI2013, IA3-IOS-3a-4, 2013.

¹ <http://kakaku.com/>