

書評SNSにおけるレビューの観点の違いを考慮したフィードバック型 協調フィルタリング

南 大智[†] 牛尼 剛聡^{††}

[†]九州大学芸術工学部芸術情報設計学科 〒815-8540 福岡県福岡市南区塩原 4-9-1

^{††}九州大学大学院芸術工学研究院 〒815-8540 福岡県福岡市南区塩原 4-9-1

E-mail: [†]1DS12200T@s.kyushu-u.ac.jp, ^{††}ushiana@design.kyushu-u.ac.jp

あらまし 現在、書籍は莫大な数が存在しており、ユーザは自分の好みに合った書籍を発見することが困難である。このことから、書籍を推薦するシステムの必要性が増大している。例えば SNS 形式の書評サイト「読書メーター」では「この本を読んだ人はこんな本も読んでいます」という、協調フィルタリングを利用した推薦が行なわれている。協調フィルタリングでは、同一の書籍を読んだユーザはすべて同等に扱われるのが一般的である。しかし、同一の書籍に対する感想には多様性があるため、ユーザは同等と考えられない場合も多い。そこで本論文では、各ユーザが書籍を評価する際、その書籍をどの観点から評価するかの違いに着目し、よりユーザの興味に合った推薦を実現する手法を提案する。本論文では、レビューには、評価における観点の違いが表れると考え、各レビューに対する評価のフィードバックを入力とし、トピックモデルを用いることで対象ユーザの評価観点を推定し、類似した評価観点をもつレビューの重要度を上げることで、推薦の結果に影響させる手法を提案する。そして、提案手法の有効性を評価する。

キーワード 書籍推薦, レビュー解析, トピック抽出, LDA, 適合フィードバック

1. はじめに

1.1 背景

現在、世界中には莫大な数の書籍が存在し、Google Books によると 2010 年 8 月 5 日時点での全世界の書籍総数は 1 億 2986 万 4880 冊だといふ [1]。莫大な数の書籍の中から、ユーザが自分の好みにあった書籍を探すことは困難であり、その問題を解決するために書籍推薦の重要性が増大している。

書籍を含む、何らかの情報推薦の方式には、一般的に、コンテンツに基づくフィルタリング (content-based filtering) と、協調フィルタリング (collaborative filtering) の 2 種類がある [2]。SNS (Social Network System) 形式の書評サイト『読書メーター』においては、「この本を読んだ人はこんな本も読んでいます」という協調フィルタリングを用いた書籍推薦を行っている。協調フィルタリングは、ユーザの過去の履歴を見て、対象ユーザと同じ書籍を多く選択しているユーザは嗜好が類似していると判断し、類似ユーザが読んでいて対象ユーザが未読である書籍を推薦することで、高い精度をもった推薦が行える有効な手法である。

1.2 問題点

一般的な協調フィルタリングの計算過程では、ユーザがアイテムを選択した、という事実のみを用いて推薦を行っている。しかし、ユーザがアイテムのどのような特徴を好んでいるかは考慮されていないため、ユーザの嗜好に合わないアイテムが推薦される可能性がある。例えば、アイテム i が持つ A という要素を好んで選択したユーザ集合 U_A と、アイテム i の持つ B という要素を好んで選択したユーザ集合 U_B が存在するとする。

対象ユーザがアイテム i を選択したとき、一般的な協調フィルタリングでは U_A と U_B は対象ユーザと同じアイテムを選択しているため、嗜好が類似していると判断される。しかし、対象ユーザがアイテム i の持つ A という要素を好んでそのアイテムを選択していた場合、 U_A は対象ユーザとの嗜好の類似性が高いと言えるが、 U_B は U_A と比べると、対象ユーザとの嗜好の類似性は低いと考えられる。つまり、このような場合であれば、 U_B に基づいて推薦されるアイテムは、対象ユーザの嗜好に合わないアイテムである可能性が高い。

1.3 アプローチ

書評 SNS におけるレビューを見たとき、同一の書籍に対するレビューであっても、各ユーザのレビューにおける評価の観点に違いが見られる。例えば、漫画書籍に対するレビューの場合、絵についての評価をするレビューもいれば、ストーリーについての評価をするレビューもいる。つまり、ユーザは異なる観点を持っている可能性がある。これは読者のパーソナリティ特性の違いが及ぼす評価の違いであると考えられる [4]。

本研究では、協調フィルタリングを利用した書籍情報の推薦の際に、同一の書籍を選択しているユーザの中でも、対象ユーザの評価の観点到類似したユーザを重視することで、従来よりも推薦の精度を向上できると考えた。対象ユーザの評価観点を推定するために、ユーザのレビューに対しての評価を、フィードバックとして利用する。レビューの評価観点の計算は、Latent Dirichlet Allocation (LDA) [5] を用いて計算する。具体的には、書評全体のトピックに対して、それぞれのレビューに含まれる単語とトピックの一致数の分布を求めることで得られた、レビューのトピック分布を評価観点とする。そ

して、対象ユーザの評価観点とレビュワーの評価観点を比較することで、対象ユーザとレビュワーの評価観点の類似度を推定する。評価観点の類似度に基づいてレビュワーの重要度を設定した協調フィルタリングによる計算を行うことで、推薦の最適化を行う。

本論文の構成は以下のとおりである。

まず、第2章で適合フィードバックによる推薦の最適化とトピックモデルを利用した観点抽出について、関連研究を挙げ、本研究の新規性について説明する。第3章では、提案手法についての概要を説明する。第4章から第7章にかけて、提案手法について各モジュールごとに具体的な説明を行う。第8章で実験の手法および結果と考察を行う。第9章で、まとめと今後の課題を述べる。

2. 関連研究

2.1 適合フィードバックによる推薦の最適化

情報検索において、ユーザはクエリを正しく指定することができない場合、ユーザが欲しい情報を取得できないことは、重要な問題として認識されてきた。このことを解決する代表的な手法として、適合フィードバックが知られている。適合フィードバックは、情報検索において検索結果として出力された文章の内容に対するユーザの反応に基づいて、検索質問や検索戦略、検索式を修正することを指す [6]。例えば、検索エンジンの検索結果に対して、興味のあるページを指定することで、クエリが修正され、興味のあるページに近いページをユーザに再提示する、というものである。このとき、クエリを修正する基となるのは、ページ自体に対してのフィードバックである。クエリの修正は、Rocchio のアルゴリズムが用いられることが多く、次のように定義される。

$$q_m = \alpha q_0 + \frac{\beta}{|D_R|} \sum_{d_i \in D_R} d_i - \frac{\gamma}{|D_N|} \sum_{d_i \in D_N} d_i \quad (1)$$

ここで、 D_R と D_N は、それぞれユーザが閲覧したページのうち、興味のあるとしたページ、興味がないとしたページである。また、 α 、 β 、 γ はそれぞれ元のクエリベクトル、正の影響ベクトル、負の影響ベクトルにどれくらいの重みを与えるかを定めるパラメータである。推薦するコンテンツの種類がテキストである場合は、ベクトルはキーワードの出現頻度で表され、tf・idf などを用いて、キーワードに重みを付ける方法が適用される。

このように、推薦コンテンツ自体に対するフィードバックを利用する研究として、例えば、顔らは、スマートフォンでの効率的な商品選別を行うために、ユーザの振る舞いを基にした適合フィードバックを用いて推薦の最適化を行った [7]。このとき、基にした振る舞いは、閲覧中の商品に対する閲覧時間やスワイプ速度などである。

本研究では、書籍の評価における観点の違いを考慮したいため、ユーザの書籍に対する印象を知る必要がある。そこで、書籍のレビューに対するフィードバックを用いて、間接的に、ユー

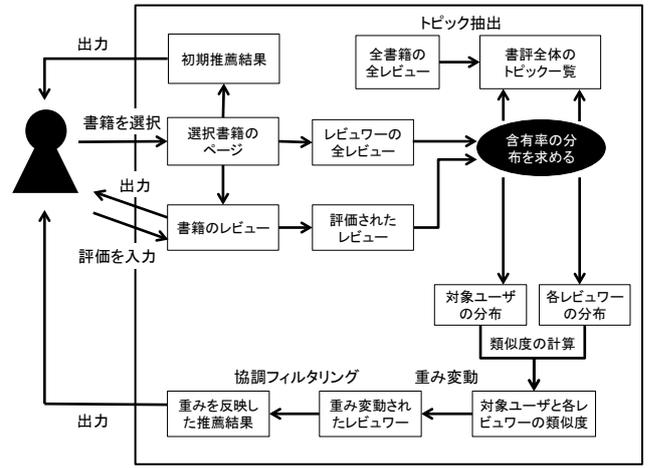


図1 システムの概要

ザの書籍評価における観点を推定することで、推薦の最適化を行う。一般的に、適合フィードバックでは、フィードバックを与えたアイテム自体が影響を与えるが、本研究ではレビュー自体ではなく、そのレビューを投稿したレビュワーの観点がユーザの観点到影響を与えるという点で、本研究の適合フィードバックの利用に関して独自性がある。

2.2 トピックモデルを利用した観点抽出手法

トピックモデルとは、文書が何について記述されているかという文書に潜在的に存在するトピックを考慮したソフトクラスタリングを行う手法である。トピックモデルを利用した研究例は近年増加し、様々なものがある。トピックモデルで観点の違いを求める研究として、大原らは、ユーザが閲覧している Web ニュースの記事の重要性をユーザに明示するために、閲覧記事と観点が同じであり主題が異なる対立記事を提示する手法を提案している [8]。この手法では、ニュースの観点には記事に記載されている明示的観点と記事に記載されていないがユーザが潜在的に知っている暗黙的観点があると考え、LDA を用いて記事からトピック抽出を行い、得られたトピックを記事の暗黙的観点とした。しかし、LDA は、大量の文章集合から確率的にトピックを求めるモデルであるため、短い文章に適用することができないという特性がある。そのため、ユーザによって文章量に大きなばらつきが生じるレビューに対して、それぞれのトピックを直接的に求めることは困難である場合が多い。そこで本研究では、書評全体のトピック群に対して、あるユーザの全レビューに含まれる単語がどれだけ一致するかどうかのトピック含有率の分布をそのユーザの観点と定義している。

3. 提案システム

我々は、ユーザが書評レビューに対する評価を入力することで、推薦の内容を最適化してユーザに再提示するシステムを提案する。図1はこのシステムの概要を図解したものである。本システムでは、対象ユーザが選択した書籍のページにおいて、その書籍に対するレビュー、推薦書籍を表示する。初期の推薦書籍は従来の協調フィルタリングで計算され、各ユーザの重要度のパラメータは均一である。対象ユーザは各レビューを読む

でき、それぞれに対して共感できるかどうかについての評価を入力する。共感できると評価されたレビュー集合を解析することで、対象ユーザの重視している観点を求める。観点の解析は、あらかじめ用意した書評全体のトピック群に対して、レビュー集合に含まれる単語がどれだけ一致するかをトピックごとに調べ、その分布を求めることを行う。次に、各レビュー者に対しても同様に観点の抽出を行い、各レビュー者の観点を求める。そして、対象ユーザの重視している観点と各レビュー者の観点との類似度を求め、類似度をユーザの重みとして、推薦書籍の再計算および再提示を行う。

4. 書評トピックの作成

観点を求めるために用いる、書評全体のトピックを求める手法について述べる。

4.1 前処理

本研究で求めたいのは、書籍の評価におけるトピックであるため、書籍自体に関するトピックが抽出されるのは避けることが望ましい。そこで、書籍の登場人物名や作者名など、その書籍にのみ出現するような単語はあらかじめ除外する必要がある。そこで、以下の前処理を行う。

全書籍のレビューを Mecab [9] を用いて形態素解析し、名詞・形容詞・形容動詞を抽出する。そして、書籍ごとに、各単語を含む文書（レビュー）の出現頻度を求め、書籍のレビュー数での除算して正規化する。これは、対象とする書籍における単語を含むレビューの出現確率であり、書籍 b_i における単語 w を含むレビューの出現確率 $P(w_{b_i})$ を以下の式で表す。

$$P(w_{b_i}) = df(w_{b_i}) / \text{書籍 } b_i \text{ に関するレビュー数} \quad (2)$$

ここで、 $df(w_{b_i})$ は単語 w を含む書籍 b_i に関するレビューの数である。書籍集合 B における、単語 w の出現確率を合計したものが、その単語の書籍集合 B における出現確率 $P(w_B)$ であり、これを以下の式で表す。

$$P(w_B) = \sum_{b_i \in B} P(w_{b_i}) \quad (3)$$

なお、 $P(w_B)$ が閾値以下の単語は、書籍に依存する単語として書評トピックの解析の対象から除外する。

4.2 トピック解析

前処理が行われた全書籍の全レビューに対してトピック解析を行う。トピック解析には、LDA を用いる。LDA によって得られた書評トピックの例を表 1 に示す。このときの書籍数は 1055 冊であり、この書籍に対する全てのレビュー、計 88 万 7396 件を用いた。トピック数は 100 とし、各トピックあたりに含まれる単語数は 10 と設定した。

5. レビュー者の評価観点の抽出

レビュー者の評価観点を抽出するために、まず、レビュー者の全レビューを集積し、形態素解析を行い、名詞・形容詞・形容動詞を抽出する。4.2 で求めた各トピックの単語に対して、抽出した単語が一致する数を計算することで、そのレビュー者

表 1 書評トピックの例

topic1	topic2	topic3	topic4
登場人物	描写	姿	話
世界	毎回	成長	恋愛
それぞれ	作中	面白い	うまい
魅力	心情	熱い	いい
キャラクター	歳	読後	友情
業	ヒーロー	爽快	面白い
物語	考え	高校生	良い
面白い	地味	どんでん返し	ない
作品	価値	センス	感じ
個性	すごい	テンション	小説

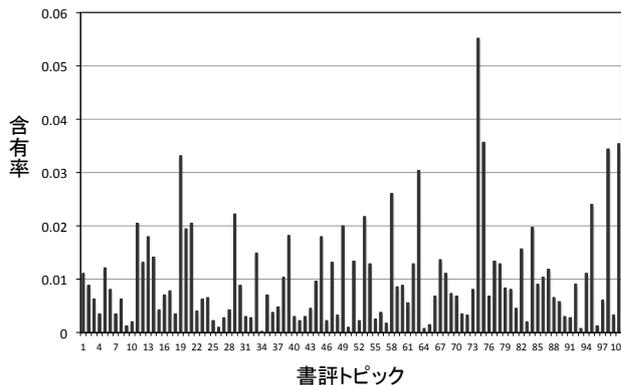


図 2 トピック含有率の分布の例

の各トピックに対する含有量を求め、一致した総数で割ることで正規化を行う。これを対象ユーザが選択した書籍の全レビュー者に対して行うことで、各レビュー者の書評トピック含有率の分布が求められる。この分布を「評価観点」として定義する。図 2 は、トピック含有率の分布の例である。

6. 推薦書籍の最適化

6.1 書籍の予測評価値計算

対象ユーザとレビュー者の観点の類似度を重みとして、協調フィルタリングによる書籍の予測評価値を計算する。書籍 a を選択したユーザ u に対する書籍 b の予測評価値は、以下の式で求める。

$$Pred(u, a, b) = \sum_{r \in \mathbf{R}(a)} score(b, r) \cdot sim(u, r) \quad (4)$$

ここで、 $\mathbf{R}(a)$ は書籍 a に対してレビューを投稿したレビュー者集合を表し、 $sim(u, r)$ はユーザ u とレビュー者 r の観点の類似度を表す。また、 $score(b, r)$ は、レビュー者 r が書籍 b を読んだかどうかによって値が決まる。

$$score(b, r) = \begin{cases} 1 & (\text{読んだ}) \\ 0 & (\text{読んでいない}) \end{cases} \quad (5)$$

書籍 a を読んだレビュー者集合のうち、書籍 b を読んだレビュー者の重みの合計値がユーザ u に対する書籍 b の予測評価値と



図3 インターフェース上での評価の入力例

なり、これにより、ユーザ u の観点を考慮した協調フィルタリングによる推薦が行える。

6.2 評価観点の類似度計算

対象ユーザの評価観点を表すトピック分布と各レビューの評価観点を表すトピック分布との間の類似度をコサイン相関値計算で求め、得られたコサイン相関値をそのレビューの重みとする。そこで、対象ユーザ u の観点を表すトピック分布を \mathbf{f}_u 、レビュー r の観点を表すトピック分布を \mathbf{t}_r とすると、対象ユーザ u とレビュー r との間の観点の類似度 $sim(u, r)$ は、以下の式で定義する。

$$sim(u, r) = \text{Cos}(\alpha \mathbf{t}_r + \mathbf{f}_u, \mathbf{t}_r) \quad (6)$$

$\alpha \mathbf{t}_r$ はユーザの擬似的な初期観点であり、2.1 で記述した Rocchio のアルゴリズムにおけるユーザの元のクエリベクトルに該当する。 \mathbf{f}_u はユーザのフィードバックに基づいた評価観点である。これについては次節で詳しく説明する。

7. 対象ユーザの評価観点の抽出

7.1 レビューへのフィードバック

対象ユーザは選択した書籍のレビューをシステムから提示され、それを読んだ上で、レビューに対してそれぞれ、評価のフィードバックを入力する。図3は、レビューに対する評価の入力インターフェースの例を表す。この図においては、レビュー文の右下に表示されている二つのアイコンが評価の入力ボタンであり、赤いアイコンはポジティブな評価、青いアイコンはネガティブな評価を意味するアイコンである。ユーザはレビューに対して、ポジティブな評価を下したい時には、赤いアイコンをクリックし、反対にネガティブな評価を下したい時には、青いアイコンをクリックする。また、レビューに対する評価がどちらとも言えない場合は、どちらのアイコンもクリックしないことで、中立の表現とする。システムはユーザから入力された各レビューへの評価を、ユーザプロファイル情報として登録する。

7.2 対象ユーザのトピック分布計算

7.1 で得られた、評価情報のうち、ポジティブフィードバ

表2 実験で使用したパラメータの組み合わせ

	α	β	γ
set1	1.0	1.0	1.0
set2	0.0	1.0	1.0
set3	0.2	0.8	1.0
set4	0.4	0.6	1.0
set5	0.6	0.4	1.0
set6	0.8	0.2	1.0
set7	1.0	1.0	0.0
set8	0.8	1.0	0.2
set9	0.6	1.0	0.4
set10	0.4	1.0	0.6
set11	0.2	1.0	0.8
set12	0.0	1.0	0.0

クを受けたレビューのトピック含有率を正の影響ベクトル、ネガティブフィードバックを受けたレビューのトピック分布を負の影響ベクトルとして、2.1 で記述した Rocchio のアルゴリズムを基に、対象ユーザ u の評価観点 \mathbf{f}_u を以下の式で定義する。

$$\mathbf{f}_u = \frac{\beta}{|R_{pos}|} \sum_{r \in R_{pos}} \mathbf{t}_r - \frac{\gamma}{|R_{neg}|} \sum_{r \in R_{neg}} \mathbf{t}_r \quad (7)$$

R_{pos} と R_{neg} はそれぞれ、対象ユーザ u がポジティブフィードバックとネガティブフィードバックを返したレビュー集合であり、 \mathbf{t}_r はレビュー r の評価観点、つまりトピック含有率の分布を表す。

8. 実験

プロトタイプを用いた被験者実験を行い、提案手法の評価と考察を行う。

8.1 実験環境

読書メーターにおける、2012年11月～2015年11月の月間ランキング上位の書籍、全1055冊の書籍のレビュー全88万7396件を対象として、書評トピックを求めた。書評トピックの解析から除外する閾値は0.35とした。日本語解析には MeCab を用いて、辞書は mecab-ipadic-neologd [10] を使用した。ユーザに提示するレビューは、読書メーターに投稿されているレビューを用いた。また、パラメータ検証のため、表2に示すパラメータの組み合わせによる推薦の最適化を行った。

8.2 評価手法

対象とする被験者は21歳～25歳までの男女15人であった。被験者には、まず、最近読んだ書籍の中で自由に書籍を選んでもらい、その書籍のレビューを表示する。レビューに対して、被験者は「このレビューを書籍選びの参考にしたいか」という基準で「参考にしたい」、「わからない」、「参考にしたくない」の3段階で評価を入力してもらった。プロトタイプシステムを通して得られた推薦結果上位20件のうち、未読書籍について「あらすじ」、「レビュー」を読んでもらい、「その推薦書籍を読みたいと思うかどうか」に関して「思う」、「どちらかと言えば思う」、「どちらとも言えない」、「どちらかと言えば思わない」、「思わない」の5段階でアンケート調査を行った。ベースライ

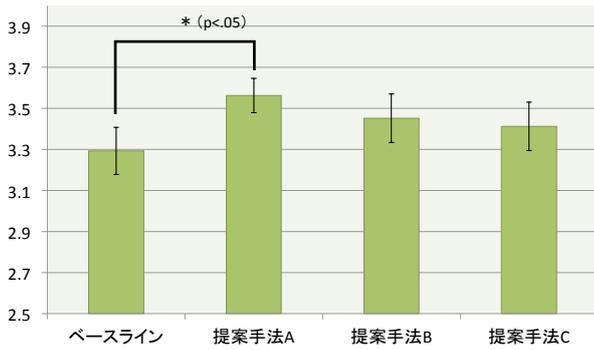


図4 スコア上位3件とベースラインの結果

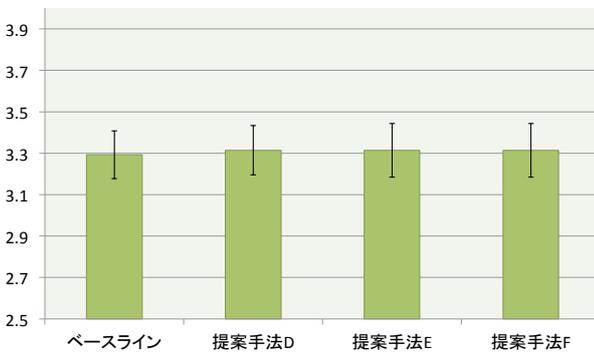


図5 スコア下位3件とベースラインの結果

は、重みを均一に設定した協調フィルタリングによる推薦とした。書籍の「1巻」「2巻」などは、まとめてシリーズとしての評価を行う。

8.3 実験結果

「その推薦書籍を読みたいと思うかどうか」に関して「思う」、「どちらかと言えば思う」、「どちらとも言えない」、「どちらかと言えば思わない」、「思わない」の評価を、それぞれ5点、4点、3点、2点、1点として、手法ごとに、推薦への評価の平均値を求めた。スコア上位3件の提案手法とベースラインのスコアを図4に示す。縦軸は推薦への評価の平均値であり、横軸は左からベースライン、提案手法A ($\alpha = 0.2, \beta = 0.8, \gamma = 1.0$), 提案手法B ($\alpha = 0.6, \beta = 0.4, \gamma = 1.0$), 提案手法C ($\alpha = 0.4, \beta = 0.6, \gamma = 1.0$), である。

ベースラインと提案手法において、推薦への評価に有意な差があるかどうかを、t検定で調べた結果、提案手法Aとベースラインとの間においてのみ、有意水準5%において有意な差が見られた。

また、各手法の評価結果が悪かったパラメータ (α, β, γ) のスコア下位3件とベースラインのスコアを図5に示す。縦軸は推薦への評価の平均値であり、横軸は左からベースライン、提

案手法D ($\alpha = 1.0, \beta = 1.0, \gamma = 0$), 提案手法E ($\alpha = 0.6, \beta = 1.0, \gamma = 0.4$), 提案手法F ($\alpha = 0.8, \beta = 1.0, \gamma = 0.2$), である。

8.4 考察

提案手法Aとベースラインとの間においてのみ、有意水準5%において有意な差が見られたことから、提案手法の有効性を示すためには、適切なパラメータを設定することが重要であると考えられる。スコアが高い時のパラメータとスコアが低い時のパラメータを比較すると、適切なパラメータに関する幾つかの考察が得られた。まず、対象ユーザの初期観点の影響の大きさを表す α の値に関しては、フィードバックによる影響の大きさを表す β と γ に対して、同等かそれ以上の値になると、レビューごとにもあまり類似度に差が見られず、初期推薦とほとんど変わらない推薦内容となった。そのため、提案手法においては、 α は、 β と γ に対して、比較的小さな値が適している。また、ポジティブなフィードバックの影響の大きさを表す β とネガティブなフィードバックの影響の大きさを表す γ との間においては、 $\beta < \gamma$ の方がスコアが良い傾向にあることが多く、 $\alpha \leq 0.2 < \gamma < \beta$ と $\alpha \leq 0.2 < \beta < \gamma$ の間でt検定を行った結果では、有意水準5%において、有意差が見られた。このことから、ユーザに合わない観点をもつレビューの、推薦における重要度を低くすることが、ユーザにとって精度の良い推薦の最適化につながると考えられる。

また、同一書籍に対して、複数の被験者が実験を行った際、それぞれが違う推薦結果になり、且つ、推薦に対する評価も高くなったことから、本研究の目的である、より対象ユーザに適した推薦の最適化が行えたと考えられる。

提案手法による、推薦の最適化の性能を向上させるための、今後の課題を以下に示す。まず、書評トピックの抽出において、「登場人物」「魅力」「個性」など、書籍の評価における観点を表す単語として適している単語も抽出されているが、一方で、「面白い」「いい」「すごい」など、レビューにおいてよく使われる単語も抽出されていることが、観点の違いを求める上での精度を下げると思われる。これに対しては、前処理の段階において、ほとんどのレビューに出現するような単語を、観点に関連しない単語として除外するなどの対策が考えられる。また、正確なトピック解析のためには、LDAのパラメータの調節が重要な要素となっている。観点を求めることに適した書評トピックを抽出するためのパラメータについて検証を通した最適化を行う必要があるだろう。また、レビューへの評価は「このレビューを書籍選びの参考にしたいか」という基準で入力してもらったが、この基準だと、ユーザが注目していなかった箇所に対するレビューであっても、内容にユーザが納得すれば、「参考にしたい」と評価してしまったり、また、ユーザと同じ点に注目しているが、意見としては自分と反対であるレビューに対して「参考にしたくない」と評価してしまい、観点抽出の妨げになる可能性がある。そのため、何らかの一貫性のある基準でレビューへの評価を下すか、もしくはレビューから意見を表す表現を除外したものを提示するなど、何に対する、どのようなフィードバックを基に観点を抽出するかについて、様々な手法

の有効性の比較を行う必要がある。

9. ま と め

本論文では、書籍推薦における協調フィルタリングの手法として、よりユーザの嗜好に合った推薦を行うことを目的に、ユーザの観点との類似度に基づいてレビュワーの重要度を設定することで、推薦結果を最適化する手法を提案した。具体的には、まずユーザに書籍を選択してもらい、その書籍のレビューを提示する。そして、ユーザが個々のレビューに対する評価を行い、それらをフィードバックとして、ユーザの観点を推定し、ユーザと各レビュワーの観点との類似度を計算する。そして、計算した類似度を重みとした協調フィルタリングで、推薦の再計算を行うことで、推薦内容を最適化する。

プロトタイプを利用した被験者実験の結果、適切なパラメータを用いた提案手法は従来手法よりも推薦の満足度が高くなった。今後、観点を求めるために用いる書評トピックの正確な抽出や、観点抽出手法の有効性、観点の類似度計算手法の妥当性について、検証実験を通して改善を行っていく予定である。

文 献

- [1] “Google Books:Books of the world, stand up and be counted! All 129,864,880 of you.”, <http://booksearch.blogspot.jp/2010/08/books-of-world-stand-up-and-be-counted.html>, (2015/01/08 アクセス)
- [2] 土方 嘉徳, “利用者の好みをとらえ活かす-嗜好抽出技術の最前線- : 1. 嗜好抽出・情報推薦の基礎理論 1) 嗜好抽出と情報推薦技術”, 情報処理, Vol.48, No.9, pp.957-965, 2007.
- [3] <http://bookmeter.com/>
- [4] 三和義秀, “小説を対象とした読後の感情状態形成モデルの研究-読者のパーソナリティ特性と認知的評価に基づいて-”, 情報知識学会誌 Vol.23 No.1, pp.92- 110(2013).
- [5] David M Blei, Andrew Y Ng, and Michael I Jordan. “Latent dirichlet allocation”, the Journal of machine Learning research, Vol. 3, pp. 993-1022, 2003.
- [6] 土方 嘉徳, “情報推薦・情報フィルタリングのためのユーザプロファイリング技術,” 人工知能学会誌, vol.19, no.3, pp.365-374, 2004.
- [7] 顔 洪, 牛尼 剛聡, “スマートフォンでの効率的な商品選別を目的としたユーザの振舞いに基づく閲覧リスト最適化手法”, 情報処理学会論文誌データベース (TOD) ,Vol.8, No.4, pp.1-15, 2015.
- [8] 大原 正章, 真下 遼, 灘本 明代, “Web ニュースからの観点抽出手法の提案”, 情報処理学会研究報告, Vol.2015-DBS-162 No.27
- [9] <http://mecab.sourceforge.net/>
- [10] <https://github.com/neologd/mecab-ipadic-neologd/blob/master/README.ja.md>