

# 推薦の多様性を高めるためにユーザの関心度が 相反する概念に注目した手法の提案と評価

山村 優介<sup>1</sup> 中島 誠<sup>2</sup>

**概要:** 現在のウェブの世界では、ユーザの行動履歴や嗜好に基づく情報推薦システムが広く活用されている。しかしながら、多様な情報に触れる機会の制限や新たな関心の喚起が妨げられるといった問題がある。一般に、ニュース記事等の情報アイテムは、一つのトピックを扱う場合であっても、場所、時間、事柄や登場人物など様々な側面（ファセット）から捉えることができる。本研究では、情報アイテムが持つこれらファセットに属する概念（フォーカス）へのユーザの関心が一律ではないことに注目して、ユーザの新たな話題への関心を喚起する推薦手法を提案する。実験では、シミュレーションとユーザスタディの両面から、提案手法の有効性を検証した。

**キーワード:** 推薦システム, フィルターバブル, 人気バイアス, KENBUN

## A Method for Enhancing Recommendation Diversity by Focusing on Contrasting User Interests

YAMAMURA YUSUKE<sup>†1</sup> MAKOTO NAKASHIMA<sup>†2</sup>

**Abstract:** In today's web environment, information recommendation systems based on users' behavioral histories and preferences are widely utilized. However, such systems pose challenges, including restrictions on diverse information exposure and hindrances to spark new interests. Information items such as news articles can be perceived from multiple facets, including location, time, events, and people, even when covering a single topic. This study proposes a recommendation method to stimulate users' interest in new topics by focusing on the fact that users' interest in the concepts (focuses) belonging to these facets is not uniform. The effectiveness of the proposed method was clarified through both simulations and user studies.

**Keywords:** recommender system, filter bubbles, popularity bias, KENBUN

### 1. はじめに

近年の情報通信技術の発展により、膨大な情報へのアクセスが容易になったが、その一方で有用な情報の発見が困難になっている[1]。この問題に対処するため、Google ニュース[2]や X (旧 Twitter) [3]などの情報提供サイトでは、ユーザの行動履歴や嗜好に基づく推薦システムが広く利用されている。しかしながら、推薦システムにもフィルターバブル[4]や人気バイアス[5]といった課題がある。前者は、関心のある情報ばかりが提示され、多様な情報に触れる機会が制限される現象であり、後者は特定の人気アイテムが過度に推薦される傾向を指す。これらにより、ユーザは新たな事柄に関心を持つ機会を損なう可能性がある。この課題を解決するため、推薦の多様性が重要視されており[6]、過去の研究では精度を維持しつつ多様性を向上させる試みがなされてきた。しかしながら、これらの研究は新たな関心への喚起を、主目的としてはいなかった。

本研究では、ユーザの新たな事柄への関心を喚起する推

薦手法を明らかにすることを目的とする。一般に、ニュース記事などの情報アイテムは、一つのトピックを扱う場合であっても、その内容を複数の側面（ファセット）から捉えることができる。例えば、その出来事が発生した場所や日時、その出来事の背景や経緯などが挙げられる。提案手法では、一つのアイテムについて各ファセットに対するユーザの関心は一律でないことに注目する。具体的には、アイテムを捉える各ファセットに属する概念（フォーカス）に対するユーザの関心度の高低を予測し、それが高いものと低いものを併せ持つアイテムを推薦する。これにより、既存の関心をきっかけとして、普段では惹かれない話題に対してもユーザの関心を促す。

評価実験では、仮想ユーザを用いたシミュレーションテストと、実ユーザを対象としたユーザテストを通じて、提案手法を他の推薦アルゴリズムと比較検証した。その結果から、提案手法が多様でかつユーザの関心に合致したアイテムを推薦し、新たな関心を喚起する効果が確認された。

<sup>1</sup> 大分大学大学院 工学研究科  
Graduate School of Engineering, Oita University

<sup>2</sup> 大分大学 理工学部  
Faculty of Science and Technology, Oita University

## 2. 関連研究

本章では、推薦の多様性を向上させる研究の例として、リランキングアルゴリズム[8, 9], 推薦モデルへの正規化項の導入[10, 11], クラスタリングを活用した手法[12]について述べる。リランキングアルゴリズムは、既存の推薦アルゴリズムによりユーザの嗜好に合致した推薦リストを作成した後、そのリストを多様性の観点から再編成する手法である。具体的には、ユーザが頻繁に閲覧するジャンルや冗長なアイテムのスコアを下げ、普段閲覧されないアイテムのスコアを上げることで、リスト全体の多様性を向上させることができる。この手法は、既存の推薦モデルを再構築せず、後付けする形で実装が行える点で優れている。推薦モデルへ正規化項を導入する手法では、ユーザからのフィードバックを学習する式に人気アイテムに対する正規化項を追加することで、それらが過度に推薦されることを防ぐ。この手法では、モデルの内部で多様性を直接考慮するため、リランキングアルゴリズムと比較して計算効率が高いという利点がある。クラスタリングを活用した手法 ClusDiv では、アイテムを内容や人気度に基づいてクラスタリングし、人気クラスタ以外のアイテムも一定の割合で推薦することで、多様性を確保する。

これらの研究では、推薦精度を維持しながらも、多様性向上への有効性を示している。それ故、ユーザが新たなアイテムに関心を持つことを促す可能性もある。しかしながら、これらはユーザの新たな事柄への関心を喚起することを直接的な目的としているわけではない。そこで、本研究では、既存の手法が間接的にしか扱ってこなかったユーザの新たな関心の喚起を主目的とした推薦手法を提案する。

## 3. 提案手法

本章では、ユーザに新たな事柄への関心を促す方法に関する仮説を心理学的な観点から述べた後、それに基づいた新たな推薦アルゴリズムについて説明する。

### 3.1 新たな関心の喚起に関する仮説

心理学的観点から見ると、新たな関心の喚起は「内発的動機づけ」や「好奇心」に関連すると考えられる。内発的動機づけとは、外的な報酬や圧力ではなく、自らの興味や満足感に基づいて行動を起こす動機のことを指す[13]。これについて、Berlyne[14]は、好奇心を「認知的不一致」や「新規性」と関連付けて捉え、これらの刺激が行動を内発的に動機づける要因であるとした。また、Loewensteinの研究[15]では、好奇心は「情報の空白」、すなわち既知の情報と未知の情報との間に生じるギャップを埋めたいという欲求から生じるとされる。つまり、人は既に知っていることを基準にして、まだ知らない情報を意識すると、後者についてより詳しく知りたいと感じる。また、このギャップが「適度」であれば、より効果的に好奇心が刺激されると指摘している。これは、ギャップが大きすぎると既存の関心

との関連が感じられず、逆に小さすぎると新鮮さを感じられないためである。以上から、既存の関心に合致した情報とそうでない情報を併せ持つアイテムを推薦することで、ユーザに新たな事柄への関心を促すことが期待できる。

### 3.2 方針

前節で述べた仮説を元に、ユーザの新たな事柄に対する関心を刺激するためのアイデアとして、一般に情報アイテムが、様々な側面（ファセット）から捉えられることに注目し、各ファセットに属する概念で、ユーザが好むものとそうでないものを併せ持つアイテムを推薦する。具体的には、アイテムのファセットに属する概念に応じて付加したタグごとにユーザの関心度を予測し、それが高いタグと低いタグを併せ持つアイテムを推薦する。ここで重要なのは、これらの併せ持つタグ同士が関心の度合いにおいて「適度なギャップ」を持つことである。関心を殆ど持たれていないタグには、関心がより高いタグを併せ持つアイテムを推薦する。一方で、少し関心の低いタグの場合は、少し関心が高いタグを併せ持つアイテムを推薦する。これにより、既存の関心と新たな情報とのバランスを保つよう調整する。

推薦には、ユーザが閲覧中のアイテム（以下、閲覧アイテムと呼ぶ）に基づいて、その関連アイテムを提示するものと、サイトのトップページ等で閲覧アイテムや検索キーワードには関係なく、おすすめのアイテムを提示するものがある。本研究では、前者におけるユーザの新たな関心の喚起を目指す。具体的には、閲覧アイテムと一つのタグを共有しながら、それと「適度なギャップ」のある別のタグを併せ持つアイテムを推薦することで、閲覧のコンテキストを保ちながら、新たな情報への関心を促進する。

### 3.3 タグ単位の嗜好予測

提案手法では、各アイテムに対してそのファセットに属する概念を表すタグを付加する。そして各タグに対して、ユーザが抱いていると予測される関心の強さを、評価値として持たせることで、ユーザの嗜好予測を実現する。任意のタグ  $\tau$  の評価値は、以下の式(1)で算出される。

$$\tau \text{ の評価値} = \tau \text{ の閲覧回数} \times \tau \text{ の記憶具合} \quad (1)$$

上式の  $\tau$  の閲覧回数は、過去に  $\tau$  を持つアイテムが閲覧された回数を表す。また、 $\tau$  の記憶具合は  $\tau$  に関する関心の時間経過による減衰を表す。これは、以下に示すエビングハウスの忘却曲線のモデル関数の式(2)により計算される。ここで、 $t$  は最後に  $\tau$  を持つ記事が閲覧されてから経過した時間（分）を表す。また、評価値の初期値、下限値は 0 とし、上限値は設定しない。

$$\tau \text{ の記憶具合} = \frac{1.84}{(\log_{10} t)^{1.25} + 1.84} \quad (2)$$

### 3.4 推薦リストの構築手順

本節では、一度の推薦でユーザに提示するアイテムのリスト（推薦リスト）を構築する手順について述べる。提案手法では先述の通り、閲覧アイテムと1つのタグを共有しつつ、そのタグと「適度なギャップ」のある別のタグを併せ持つアイテムを推薦する。

#### 3.4.1 テーマタグに基づいた推薦リストの構築

閲覧アイテムのタグの内、一部のタグ（例えば、評価値の高いタグ）を持つアイテムのみを推薦すると、多様性の低下を招く。提案手法では、閲覧アイテムが持つタグ毎に、それを共有する複数のアイテムからなるサブリストを作成し、それらを統合して推薦リストを構築する。また、各サブリストが含むアイテムの数は均等に作る。これにより、評価値の高低に関わらず、閲覧アイテムが持つより多くのタグに関連するアイテムが推薦され、ユーザが注目していなかったアイテムの側面に気づく機会を提供できる。これらのタグは、各サブリストにおける推薦のテーマとなるため「テーマタグ」と呼ぶ。また、各サブリストにおいて、そのテーマタグを持つアイテム（以下、候補アイテムと呼ぶ）の中から、推薦するアイテムを選定する方法については、次節で述べる。

次に、推薦リストの長さについて説明する。マジカルナンバーという理論では、人が一度に把握できる情報の数の上限は、 $4 \pm 1$  あるいは  $7 \pm 2$  個であるとされている。そこで、サブリスト数の上限を 4、各サブリストで推薦するアイテム数の上限を 2 とする。つまり、推薦リスト長は 8 となる。このように、一度に表示されるアイテム数とテーマタグの数をマジカルナンバーの範囲に抑えることで、ユーザの認知負荷の軽減を図る。ただし、閲覧アイテムが持つタグの数が 3 以下また 5 以上である場合は、次の処理を行う。まず 3 以下の場合、不足分のサブリストは空として推薦を行う。5 以上の場合、意味的に最も近いタグ同士を統合し、いずれかのタグをランダムにテーマタグから除くという操作を、テーマタグ数が 4 になるまで繰り返す。これにより、推薦リスト内のテーマタグの多様性を向上する。図 3.1 に、閲覧アイテムのタグ数が 4 の場合の推薦リスト構築のイメージを示す。

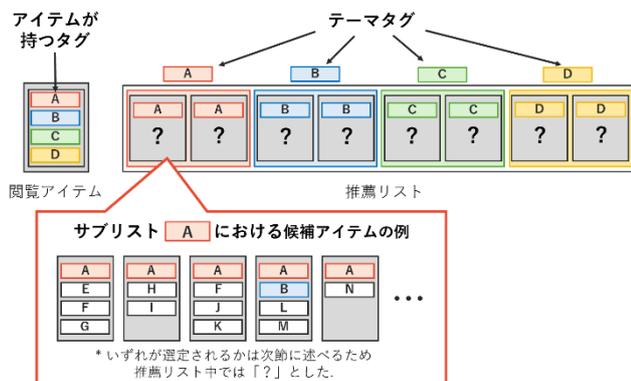


図 3.1 閲覧アイテムのタグ数が 4 の場合の推薦リスト

#### 3.4.2 各サブリストにおける推薦アイテムの選定

次に、各サブリストにおいて、そのテーマタグを持つアイテム（候補アイテム）の中から、実際に推薦されるアイテムを選定する手順について説明する。各サブリストでは、テーマタグが高評価である程、それ以外のタグの評価値がより低いアイテムを選び、逆にテーマタグが低評価である程、それ以外のタグの評価値が高いアイテムを選ぶ。これを実現するために、具体的には次の手順をとる。まず、候補アイテムそれぞれについて、テーマタグ以外のタグの評価値の平均を求める。そして、その平均評価値が、式(3)で計算される目標値により近い 2 個のアイテムを推薦する。

$$\text{目標値} = \text{評価値上限} - \text{テーマタグの評価値} \quad (3)$$

ここで評価値上限は、全タグの評価値から外れ値を除いた中での最大値とする。これにより例えば、評価値上限が 10 の状況で、テーマタグの評価値が 9（非常に高評価）であるサブリストでは、テーマタグ以外のタグの平均評価値が 1（非常に低評価）に近いアイテムが推薦される。また、テーマタグの評価値が 4（少し低評価）であれば、それ以外のタグの平均評価値が 6（少し高評価）に近いアイテムが推薦される。以上の処理を、全サブリストについて行う。なお、目標値に最も近いアイテムが複数ある場合は、その中からランダムに選定する。また、テーマタグしか持たない候補アイテムについては、平均評価値を 0 として扱う。

## 4. 評価実験 I シミュレーションテスト

様々な嗜好や記事の選び方をするユーザに対して、提案手法がどの程度多様な推薦を行えるかを評価するために、仮想ユーザを用いたシミュレーションテストを実施した。

### 4.1 データセット

実験を行うにあたり、推薦するアイテムやそれに対して付加するタグとして利用したデータセットについて述べる。

#### 4.1.1 新聞アーカイブシステム KENBUN

推薦するアイテムとして、新聞アーカイブシステム KENBUN[7]が所蔵する新聞記事を使用した。KENBUN は、主にブラウジングを用いて記事の探索を行うことができる。大分県立図書館から提供された 1876 年から 1967 年までに大分県内及び周辺地域で発刊された、計 25 社分の地方新聞の計 177,576 ページの新聞紙面画像を記録している。現在、大分県立図書館、中津市立小幡記念図書館、大分大学図書館、別府市立図書館にて稼働中である。先行研究[16]では、記事探索の円滑化に利用するため、記事内容を端的に表す見出しや広告、写真、図が抽出された。

#### 4.1.2 日本語 WordNet

タグ用語集データセットとして、日本語 WordNet[17]を採用した。WordNet は類語辞書データベースの一種であり、各語句が同義語グループ (synset) によって分類されている。また、synset 同士は概念的な関係に基づく階層構造を持つ。

日本語 WordNet には、57,238 の synset 及び 93,834 の語句が収録されている。

### 4.1.3 新聞記事へのタギング

実験にあたり、KENBUN が所蔵する新聞記事に対してタグを付加した方法について述べる。まず、先行研究で切り出された見出し画像に含まれるテキストを、Google Drive の OCR 機能を用いて抽出した。古い新聞のレイアウトの煩雑さや文字の掠れ等により、記事本文から自動的に語句を抽出してタギングに利用することは難しいことから、比較的テキスト認識しやすい見出しを用いた。次に、抽出した見出しテキストから名詞を取り出した。形態素解析及び品詞分類には MeCab[18]を使用した。併せて、旧字体を新字体に変換した他、SlothLib[19]を使用し、ストップワードを除去した。最後に、取り出した名詞が WordNet 内に存在する場合、その名詞が属する synset の ID を当該記事のタグとして設定した。このように synset をタグとして採用することで、見出しテキストに同義語が含まれる記事同士に同じタグを付加することができるため、表記ゆれの問題が軽減される。これらの処理の結果、1,212,815 の記事が 1 つ以上のタグを持ち（平均 4.87 個）、21,737 のタグが 1 つ以上の記事に付加された。

### 4.2 実験方法

シミュレーションテストでは、様々な嗜好や記事の選び方を設定した仮想ユーザに対して、提案手法と比較手法による推薦を行い、その結果を比較した。まず、仮想ユーザの興味傾向を表すユーザプロファイルを 10 人分作成した。ユーザプロファイルは、ユーザの好きなタグと嫌いなタグのリストである。ユーザ毎にランダムに好きなタグと嫌いなタグを設定し、好きなタグには 1~10、嫌いなタグには -10~-1 の興味値をランダムに付与した。それ以外のタグの興味値は 0 とした。その上で、各仮想ユーザが推薦リスト内のいずれの記事を閲覧するかを決定する、以下 2 つの記事選択手法を用意した。

#### ランダム選択：

自身の嗜好（ユーザプロファイル）に関係なく、ランダムに記事を選択する。広いジャンルに対する好奇心が強く、様々な記事を閲覧するユーザを模している。

#### 興味優先選択

付加されたタグの興味値の合計が最も高い記事を選択する。該当記事が複数ある場合は、その中からランダムに選ぶ。比較的好奇心が弱いユーザを模している。

次に、提案手法による推薦を評価するため、比較対象となる記事推薦手法を用意した。

#### ランダム推薦

タグの評価値や閲覧中の記事（閲覧記事）に関係なく、ランダムに 8 つの記事を推薦する。

### 高評価優先推薦

提案手法と同様に、閲覧記事に付加されたタグをテーマタグとしたサブリストから推薦リストを構成する。各サブリストでは、そのテーマタグを持つ記事の内、合計評価値の高い記事を推薦する。該当記事が複数ある場合は、その中からランダムに推薦する。

シミュレーションの手順について述べる。まず、各仮想ユーザは、ユーザプロファイルで最も高い興味値を設定されたタグを持つ記事の中から、最初に閲覧する記事を選択する。その後、閲覧記事に基づいて生成された推薦リストから、次に閲覧する記事を選ぶ状況のシミュレーションを 500 回繰り返す。以上を、全てのユーザプロファイル、記事選択手法、記事推薦手法の組み合わせにおいて行った。また、1 つの記事を読むのにかかる時間は 1 分とした。さらに、コールドスタート軽減のため、好きなタグの興味値を、そのタグの初期評価値として与えた。

### 4.3 評価指標

推薦の多様性を評価するための指標として、以下の、網羅率と MCD を用いた。

#### 網羅率

全てのタグの内、推薦記事に一度でも付加されていたタグ数の割合。どの程度多くのタグを含む記事を推薦できたかを表す。

#### MCD (Majority Category Domination)

推薦記事に最も多く付加されていたタグが、推薦リストに占める割合。値が小さい程、一つのタグへの推薦の偏りが少ないことを表す。

上記の他、各推薦手法により、ユーザがどの程度既存の関心に合致する記事を閲覧できるかを検証するため、仮想ユーザが閲覧した記事のタグに設定されていた興味値の平均（平均興味値）を計算した。

### 4.4 実験結果

表 4.1、表 4.2 に、それぞれランダム選択、興味優先選択で各推薦アルゴリズムを用いた仮想ユーザに対する推薦結果を示す。各セル内の値は、ユーザプロファイル 10 人分の網羅率、MCD、平均興味値の平均値を、括弧内はそれらの標準偏差（±）を示している。また、提案手法と比較手法の結果に有意差があるかを Friedman 検定、Bonferroni 法（有意水準 0.05）で判定し、提案手法よりも有意に優れた結果は太字で、劣った結果は斜字で示した。

提案手法の網羅率は、いずれの記事選択手法でも高評価優先推薦よりも高く、ランダム選択ではランダム推薦よりも高かった。MCD は、いずれの選択手法でも、高評価優先推薦よりも優れていたが、ランダム推薦よりは劣っていた。平均興味値は、興味優先選択において、ランダム推薦よりも高く、高評価優先推薦よりも低かった。

表 4.1 ランダム選択の結果

	網羅率	MCD	平均興味値
提案手法	0.160 (0.00502)	0.00609 (0.00097)	0.0166 (0.0746)
ランダム 推薦	0.154* (0.00199)	<b>0.00375**</b> <b>(0.000318)</b>	0.0168 (0.0316)
高評価 優先推薦	0.121** (0.0110)	0.0562** (0.0110)	0.0568 (0.0491)

表 4.2 興味優先選択の結果

	網羅率	MCD	平均興味値
提案手法	0.159 (0.00346)	0.00675 (0.00102)	1.03 (0.581)
ランダム 推薦	0.156 (0.00391)	<b>0.00378**</b> <b>(0.000322)</b>	0.409* (0.147)
高評価 優先推薦	0.114** (0.0147)	0.0464** (0.0187)	<b>3.57*</b> <b>(1.98)</b>

また、図 4.1、図 4.2 にランダム選択、興味優先選択それぞれで各推薦手法を用いた際の網羅率の推移を示す。実践は各推薦アルゴリズムにおける網羅率の平均値を、薄く色づけた領域は最大値・最小値の範囲を表している。

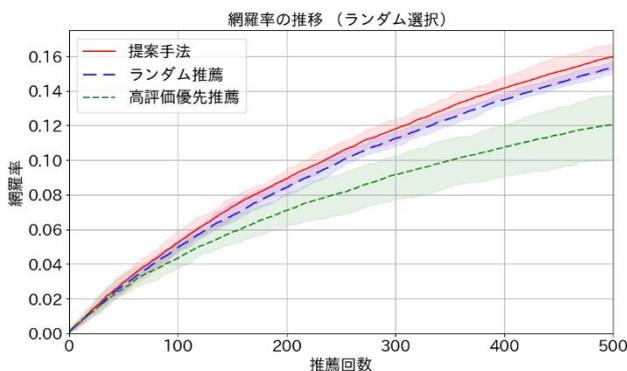


図 4.1 網羅率の推移 (ランダム選択)

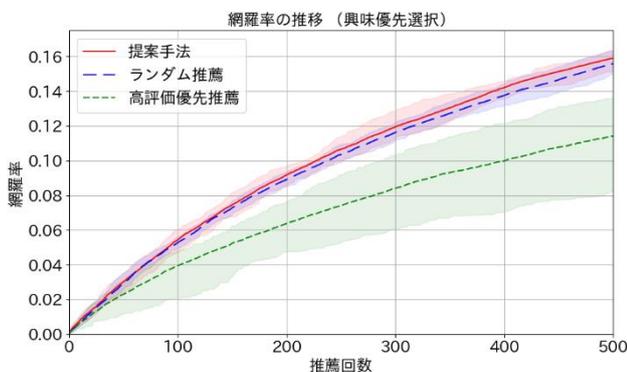


図 4.2 網羅率の推移 (興味優先選択)

## 4.5 考察

網羅率についての結果から、提案手法はいずれの記事の選び方をするユーザに対しても、多様な記事を推薦できたと言える。特に、高評価優先推薦では、網羅率の推移の状況から、推薦回数を増すごとに提案手法やランダム推薦との差が広がっていく傾向があり、フィルターバブルが発生していると考えられる。また、高評価優先推薦の結果を、記事選択手法で比較すると、興味優先選択の方が他の推薦手法との差がより早い時点で開きはじめ、さらに最小値が大きく低下していた。このことから、既存の関心と強く関連する記事を好むユーザに対して、フィルターバブルが顕著に発生していることが分かる。一方、提案手法では、常にランダム推薦と同等かそれ以上の網羅率で推移しており、記事選択手法による差もほとんど見られない。これらのことから、提案手法ではいずれの記事の選び方をするユーザに対してもフィルターバブルが発生しづらいと考えられる。

次に、MCDについては、いずれの記事選択手法を用いた場合でも、提案手法の MCD はランダム推薦よりも有意に数値が高く、高評価優先推薦よりも有意に低かった。このことから、提案手法ではランダム推薦には劣るものの、単に評価値の高い記事を推薦するよりは1つのタグへの推薦の偏りを軽減できるといえる。

最後に、閲覧記事の平均興味値については、ランダム選択の場合、提案手法の平均興味値はいずれの推薦手法とも有意差がなかった。また、興味優先選択の場合では、ランダム推薦よりも有意に高く、高評価優先推薦よりも有意に低かった。このことから、提案手法では、高評価優先推薦には劣るものの、単にランダムな記事を推薦するよりは既存の関心に合致する記事を推薦できたといえる。

以上から、シミュレーション結果においては、提案手法ではユーザに新たな事柄への関心を促すために必要な、推薦の多様性を確保できることが確認できた。しかしながら、人間の興味関心は常に変化するため、実際に各推薦手法で推薦記事にどの程度関心を持たせることができるかは次章で述べるユーザテストによって検証する。

## 5. 評価実験 II ユーザテスト

提案手法が新たな関心の喚起に有効化どうかを検証するために、実ユーザを対象としたユーザテストを実施した。

### 5.1 実験用システム

ユーザテストを行うにあたり、シミュレーションテストで使用した3つの推薦手法を実装した、実験用システムを開発した。なお、推薦手法によるデザインや操作方法の違いは無いようにした。システムのメイン画面を図 5.1 に示す。画面左に閲覧中の記事を含む新聞紙面を表示し、画面右の中央に、閲覧記事の見出しを、その周辺に実装した推薦手法に基づきユーザに推薦する見出し(推薦見出し)を表示する。推薦見出しの中から閲覧したいものを中央の領

域に移動させることで、表示する閲覧記事の見出しとそれを含む新聞紙面、推薦見出しを更新することができる。また、右上部に配置した検索ボタンを押すことで、図 5.2 に示す記事検索画面が別ウィンドウに表示される。記事検索画面では、入力された検索キーワードに関連する記事の見出しから、閲覧したいものを選択することができる。なお、最初の閲覧記事は、この検索機能を用いて決定する。

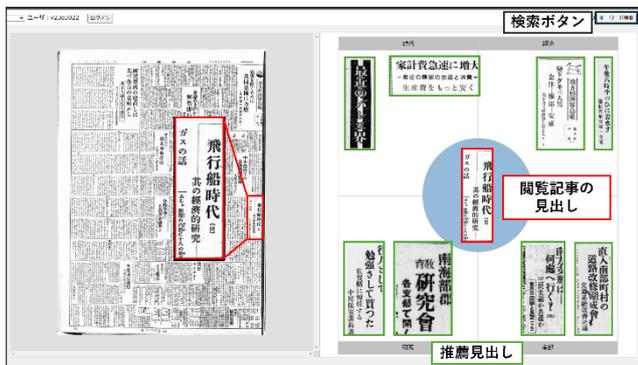


図 5.1 メイン画面

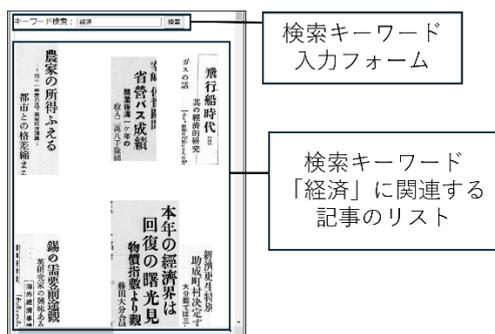


図 5.2 記事検索画面

## 5.2 実験方法

実験は、大学生 6 名（理工学部 4 名，大学院工学研究科 2 名）を被験者として実施した。各被験者に、実験内容、システムの操作方法についての説明を行った後、3 つのシステムを順に利用してもらった。それぞれのシステムの違いについては、説明していない。被験者毎にシステムの利用順を変えた。以下、ランダム推薦、高評価優先推薦、および提案手法を実装したシステムを、それぞれシステム 1~3 と書く。

実験では、まず初期設定として、各被験者に関心のある語句を 10 個、順位付けて指定してもらい、それらが属するタグ (synset) の評価値の初期値を  $10 - \text{順位} + 1$  とした。なお、これらの語句は、WordNet に収録された語句の内、新聞記事にタグとして付加された synset に属するものを使用した。また、ここで設定した初期評価値は、システム 1~3 で共通とした。

次に、被験者毎に設定した順番で、実験用システムを使用してもらった。最初に閲覧する記事は、検索機能を用い

て選択してもらった。このとき、初期設定で 1 位に指定した語句を検索キーワードとした。その後、被験者が推薦見出しから閲覧するものを選ぶ操作を 100 回行うまで各システムを使用してもらった。検索機能は自由に使用してよいこととした。

各システム使用後に、「多様性」、「関心との合致」、「新たな関心の喚起」の各項目についてのアンケートに答えてもらい、各システムを評価してもらった。各項目の評価を 5 段階リッカート尺度 (1. 非常にそう思う ~ 5. 全くそう思わない) で回答してもらった。肯定、否定意見が有意に多いか否かを、評価値 1, 2 を肯定意見, 4, 5 を否定意見とした二項検定で判定した (有意水準 0.05)。また、全タスク終了後に、それらの各項目における順位をつけてもらった。こちらは、各システムに順位スコア (1 位: 3 点, 2 位: 2 点, 3 位: 1 点) を付け、それらの合計値を比較した。

## 5.3 実験結果

### 評価項目①：多様性

まず、各システムで多様な記事が推薦されたかを聞いた。図 5.3 に各システムの評価をまとめたグラフを示す。結果として、いずれのシステムでも有意に肯定意見が多かった。また、図 5.4 に、推薦の多様性について各システムを順位付けてもらった結果を、表 5.1 にその順位スコアを示す。結果として、全被験者がシステム 1, 3, 2 の順に評価が高かった。

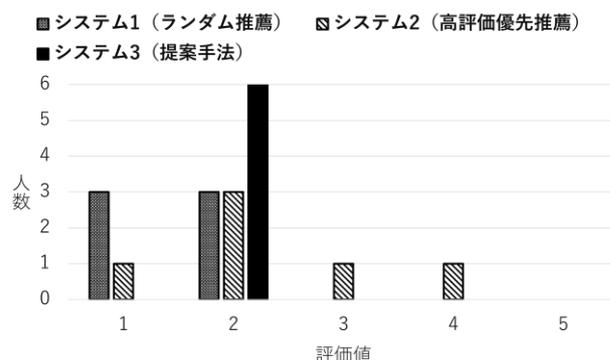


図 5.3 各システムで多様な記事が推薦されたか

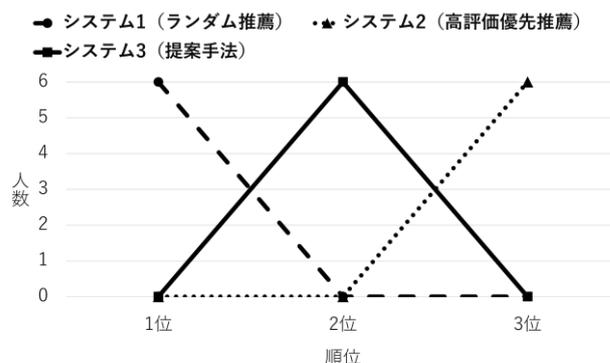


図 5.4 各システムを多様な記事を推薦した順に順位付け

表 5.1 多様性における各システムの順位スコア

システム 1	システム 2	システム 3
18	6	12

評価項目② 関心との合致

図 5.5 に、各システムで関心に合う記事が推薦されたかを聞いた結果を示す。結果として、システム 1 は否定意見が有意に多く、システム 3 は肯定意見が有意に多かった。また、図 5.6 に関心の合致について各システムを順位付けてもらった結果を、表 5.2 にその順位スコアを示す。結果として、システム 3, 2, 1 の順に評価が高かった。

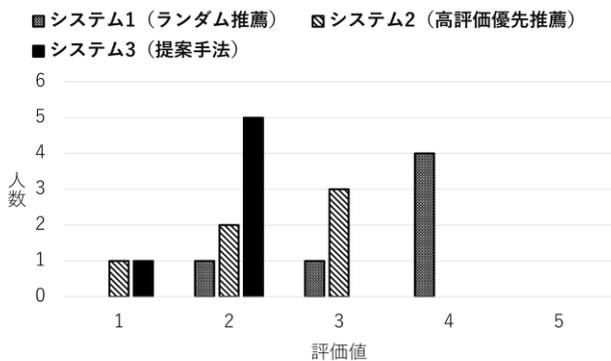


図 5.5 各システムで関心に合う記事が推薦されたか

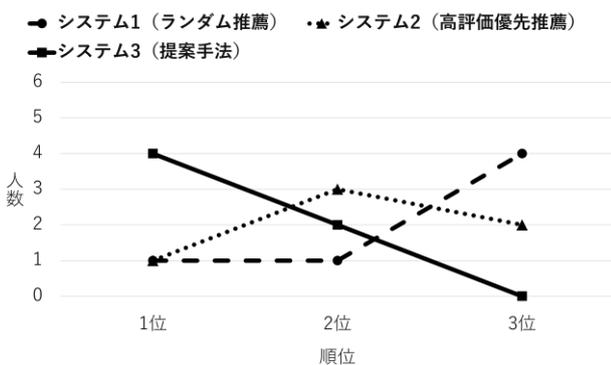


図 5.6 各システムを関心に合う記事を推薦した順に順位付け

表 5.2 関心の合致における各システムの順位スコア

システム 1	システム 2	システム 3
9	11	16

評価項目③：新たな関心の喚起

図 5.7 に、各システムで普段読まない記事に関心をもつことができたかを聞いた結果を示す。結果として、システム 1, 3 は肯定意見が有意に多かった。また、図 5.8 に新たな関心の喚起について各システムを順位付けてもらった結果を、表 5.3 にその順位スコアを示す。結果として、システム 1, 3, 2 の順に評価が高かった。

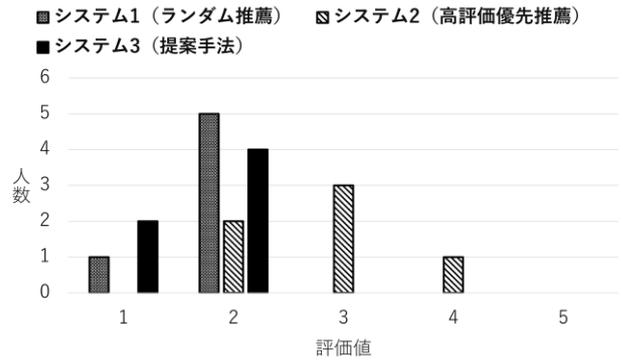


図 5.7 各システムで普段読まない記事に関心をもつことができたか

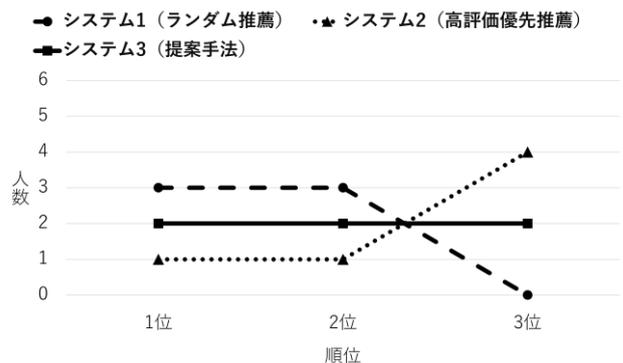


図 5.8 各システムを普段読まない記事への関心を促した順に順位付け

表 5.3 新たな関心の喚起における各システムの順位スコア

システム 1	システム 2	システム 3
15	9	12

5.4 考察

まず、推薦の多様性について評価する。図 5.3, 図 5.4, 表 5.1 の結果から、提案手法はランダム推薦には及ばないものの、多様な記事を推薦できることが示唆された。シミュレーションテストでは提案手法の方が、網羅率が高かったにも関わらず、ユーザテストではランダム推薦の方が高評価であった。提案手法では推薦リスト中の全ての記事が、閲覧記事とテーマタグを共有しているのに加え、各サブリスト内では同じテーマタグを持つ記事が推薦されることから、記事間の関連性を感じやすい。その一方、ランダム推薦では閲覧記事や他の推薦記事と関係なく記事が提示されるため、結果的に「より多様な記事が推薦された」とユーザに認識された可能性がある。

次に、関心の合致について評価する。図 5.5, 図 5.6, 表 5.2 の結果から、他の手法と比べて、提案手法によりユーザの関心に合致する記事を推薦できることが示唆された。ここで、高評価優先推薦の評価が提案手法よりも低かった点

については、アンケートの自由記述欄に「システム2では似たような記事が連続して表示されることがあった」、「同様の内容ばかりが推薦されて飽きてしまったため、興味の有無にかかわらず、様々な記事を閲覧できる方がよかった」といった意見が見られたことから、「内容が類似した記事が連続して推薦されることによる飽き」が理由であると考えられる。言い換えれば、この現象はLoewensteinの理論[15]における「既知情報と未知情報の適切なギャップ」が不足し、被験者の好奇心が十分に刺激されなかったことに起因するともいえる。

最後に、新たな関心の喚起について評価する。図5.7、図5.8、表5.3の結果から、提案手法はランダム推薦には及ばないものの、新たな関心を喚起できることが示唆された。ただし、ランダム推薦に関しては、先述したように「関心に合致する記事が推薦されづらい」という課題が示されている。このことから、ランダム推薦は、今回の実験のように記事の閲覧が強制される場面であれば、普段読まない記事に関心を持つきっかけを提供する可能性がある。しかしながら、現実の利用場面においては、ユーザの関心に合わない記事が提示され、利用意欲を下げってしまう可能性がある。今後はより現実的な環境での実験を行い、いずれの推薦手法が優れているかを総合に検証していく必要がある。

## 6. おわりに

本稿では、情報アイテムの複数の捉え方（ファセット）に注目し、それに属する概念についてユーザの関心が強いものと弱いものとを併せ持つアイテムを推薦することで、ユーザの新たな事柄への関心を喚起する推薦手法を提案した。また、その有効性を検証するため、シミュレーションテストとユーザテストを実施した。その結果から、提案手法が多様でかつユーザの関心に合致したアイテムを推薦することで、ユーザに新たな関心を喚起する効果が確認された。しかしながら、新たな関心の喚起についての評価が、ランダム推薦と比較すると低かった。このことから、今後はより現実の利用場面に近い形での実験を行い、いずれの推薦手法が優れているかを検証する必要がある。

## 参考文献

- [1] P. Maes, “Agents that Reduce Work and Information Overload,” *Communications of ACM*, Vol.37, No.7, pp.30-40, 1994.
- [2] Google, “Google ニュース,” Google, <https://news.google.com/home?gl=JP&hl=ja&ceid=JP:ja>, 参照 Jan. 24, 2025.
- [3] X Corp., “X,” X Corp., <https://x.com>, 参照 Jan. 24, 2025.
- [4] Yingqiang Ge, Shuya Zhao, Honglu Zhou, Changhua Pei, Fei Sun, Wenwu Ou, and Yongfeng Zhang, “Understanding Echo Chambers in E-Commerce Recommender Systems,” *SIGIR '20: The 43rd International ACM SIGIR conference on research and development in Information Retrieval*, pp.2261-2270, Jul. 2020.
- [5] Alejandro Bellogín, Pablo Castells, and Iván Cantador, “Statistical biases in Information Retrieval metrics for recommender systems,” *Information Retrieval Journal*, Vol.20, pp.606-634, Jul. 2017.
- [6] N.Hurley, “Towards diverse recommendation,” *RecSys Workshop: Novelty and Diversity in Recommender System*, Keynote, p.1, 2011.
- [7] 中島誠, 新聞アーカイブシステム KENBUN,” 九州地区大学図書館協議会誌, 第 57 号, pp.5-8, 2014.
- [8] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher, “Managing popularity bias in recommender systems with personalized re-ranking,” *The thirty-second international flairs conference*, Aug. 2019.
- [9] Carbonell J, Goldstein J, “The use of MMR, diversity-based reranking for reordering documents and producing summaries,” *In: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp.335-336, Aug. 1998.
- [10] Himan Abdollahpouri, Robin Burke, and Bamshad Mobasher, “Controlling popularity bias in learning-to-rank recommendation,” *Proceedings of the eleventh ACM conference on recommender systems*, pp.42-46, Aug. 2017.
- [11] Ziwei Zhu, Yun He, Xing Zhao, Yin Zhang, Jianling Wang, and James Caverlee, “Popularity-opportunity bias in collaborative filtering,” *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, pp.85-93, Mar. 2021.
- [12] Tevfik Aytakin and Mahmut Özge Karakaya, “Clustering-based diversity improvement in top-N recommendation,” *Journal of Intelligent Information Systems*, Vol.42, Issue 1, pp.1-18, Feb. 2014.
- [13] 鹿毛 雅治, “内発的動機づけ研究の展望”, *Japanese Journal of Educational Psychology*, 42 巻, 3 号, pp.345-359, 1994.
- [14] 西川 一二, “個人特性としての好奇心の領域とタイプについて—知的的好奇心と对人的好奇心—”, 関西大学審査学位論文, 2016.
- [15] Loewenstein G., “The Psychology of Curiosity: A Review and Reinterpretation,” *Psychological Bulletin* Vol.116, Issue.1, pp.75-98, Jul. 1994.
- [16] 中尾優太, 中島誠, “新聞アーカイブシステム KENBUN でのブラウジングのためのランドマーク表示,” 電気・電子関係学会九州支部連合大会 (第 73 回連合大会), Sep. 2020.
- [17] Hitoshi Isahara, Francis Bond, Kiyotaka Uchimoto, Masao Utiyama, and Kyoko Kanzaki, “Development of the Japanese WordNet,” *Proceedings of the Sixth International Conference on Language Resources and Evaluation*, 2008.
- [18] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto, “Applying Conditional Random Fields to Japanese Morphological Analysis,” *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, pp.230-237, Jul. 2004.
- [19] 大島裕明, 中村聡史, 田中克己, “SlothLib: Web サーチ研究のためのプログラミングライブラリ”, *日本データベース学会 Letters*, 6 巻, 1 号, pp.113-116, 2007.