

平成 30 年 6 月 5 日現在

機関番号：32613

研究種目：若手研究(B)

研究期間：2015～2017

課題番号：15K16091

研究課題名(和文) 推薦理由および予見情報を提示するユーザ操作可能な情報推薦システム

研究課題名(英文) A Operable Recommender System with Presenting Reason of Recommendation and Prediction Information

研究代表者

北山 大輔 (Kitayama, Daisuke)

工学院大学・情報学部(情報工学部)・准教授

研究者番号：40589975

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：以下の2つのテーマに分け、推薦理由および予見情報を提示するユーザ操作可能な情報推薦システムの開発を行った。

(A) ユーザ行動パターンに基づく推薦理由の明示化として主に、推薦理由を明示可能な推薦システムに関する研究、推薦理由の多様化に関する研究、推薦理由明示化の効果に関する研究に取り組んだ。

(B) ユーザ選択操作に基づく推薦方向の予見として主に、ユーザのレビュー選択に基づく観光スポット検索・推薦手法、リツイートパターンに基づくツイートの到達確率判定と編集支援手法、コーディネート投稿サイトのユーザタグを用いたコーディネート間の類似度に基づく検索ナビゲーションシステムを開発した。

研究成果の概要(英文)：By working on the following two themes, we developed a user-operable information recommendation system that presents reasons for recommendation and predictive information.

(A) The following study was conducted as a manifestation of reasons for recommendation based on user behavior patterns. (1) Research on the recommendation system that can clearly indicate the reason for recommendation. (2) Study on diversification of the reason for recommendation. (3) Study on the effect of clarification of recommendation reason.

(B) The following system was developed as a prediction of the recommended direction based on user selection operation. (1) Sightseeing spot search / recommendation method based on user's review selection. (2) Judgment of reaching probability and editing support method based on retweet pattern.

(3) Search navigation system based on similarity between coordination using user tag of coordination posting site.

研究分野：メディア情報学

キーワード：推薦システム ユーザ行動

## 1. 研究開始当初の背景

情報の個人化技術は様々な分野に活用されてきており、代表的なものでは Amazon.com や楽天市場などのオンラインショッピングサイトをはじめ Google などの Web 検索結果にも適応されている。個人化技術の基盤となるのは情報推薦アルゴリズムであり、利用者が好むであろう情報を、ユーザの利用履歴やアイテムの特徴を分析することで、推測し提示するものである[神鷹 2008]。2007 年より The conference series on Recommender Systems (RecSys) という国際会議が始まるなど、情報推薦分野は現在活発に議論されている。情報の個人化技術、情報推薦アルゴリズムは、意思決定支援の側面を持つため、当然ユーザの行動に影響を与えることとなる。この影響の 1 つにフィルターバブルの問題がある。

フィルターバブル(Filter Bubble)[Pariser 2011]は、Pariser が主張している、ユーザが接する情報の範囲に対する個人化技術の問題である。推薦アルゴリズムなどの個人化技術によって、知らず知らずのうちにユーザが感心を寄せる特定領域の情報にしかアクセスしなくなり、情報フィルタリングの「泡」に閉じ込められたような状態になる問題である。このような状態になると、感心領域外の新たな情報取得機会が失われるなどユーザの意思決定に対して少なくない影響を与えてしまうことを指摘している。この問題に対し、2011 年の RecSys 国際会議では、パネル討論が開催され[Resnick 2011]、フィルターバブル問題に対しての技術的な対応として、(1)推薦の多様性を確保する、(2)根拠を説明し推薦の透明性を高める、(3)ユーザが推薦の方向性を制御できるようにする、(4)短期的な嗜好と長期的な嗜好のバランスをとるといったことが議論された。

[神鷹 2008] 神鷹 敏弘：推薦システムのアルゴリズム(1)～(3)，人工知能学会誌，Vol.22，No. 6～Vol.23，No.2(2007-2008)

[Pariser 2011] Pariser, E.: The Filter Bubble: What The Internet Is Hiding From You, Viking(2011)

[Resnick 2011] Resnick, P., Konstan, J., and Jameson, A.: Panel on The Filter Bubble, The 5th ACM conference on Recommender systems (2011)

## 2. 研究の目的

本研究は、情報推薦システムにおけるフィルターバブル問題に対し、ユーザによる多様な情報の取捨選択を可能とする推薦アルゴリズムの開発を目的としている。この推薦アルゴリズムは大きく分けて、(1) ユーザ行動パターンに基づく推薦理由の明示化、(2) ユーザ選択操作に基づく推薦方向の予見からなる。すなわち、推薦理由の明示化により多様な推薦結果をユーザが理解可能な

形式にし、推薦方向の予見により、システムにより更新される自身の嗜好情報を把握可能にすることで、推薦システムを利用するユーザが推薦情報をコントロールできる仕組みを提供する。開発した推薦アルゴリズムを実装したシステムを構築し、実証的に有効性を検証する。

## 3. 研究の方法

### (1) ユーザ行動パターンに基づく推薦理由の明示化

情報推薦においては、ユーザの行動履歴に基づいてそのユーザの嗜好を抽出する。このとき、行動履歴から解釈可能な嗜好は一意に定まるとは限らない。例えば、「映画」の情報をしばらく閲覧後、「猫」の情報を閲覧しはじめたユーザがいたとすると、そのユーザは「映画のトピック内の猫」を要求しているという解釈もあれば、「映画ではない猫」を要求しているという解釈もありえる。そこで、多様な解釈のどれを採用したのかを明示して、推薦結果を抽出するアルゴリズムを開発する。提案した方式に基づきプロトタイプシステムを構築し、有効性を検証する。また推薦理由の表現方法について検討する。

### (2) ユーザ選択操作に基づく推薦方向の予見

推薦システムでは、推薦結果の取捨選択によって、ユーザの嗜好は更新される。一般に、ユーザにはどのように更新されたのかは明示されず、また推薦アルゴリズムによっては理解可能なパラメータでは無い。そこで、更新された嗜好情報によって推薦される情報をユーザに対して仮想的に提示することで、その更新によって起こることを予見させる手法を提案する。例えば、ユーザが情報 A, B, C を閲覧した場合に、A, B, C によって更新された嗜好では情報 D が推薦され、A のみによって更新された嗜好では情報 E が推薦されるといったことが提示されれば、推薦アルゴリズム内の個々のパラメータは理解できなくても、推薦結果の方向性として適切な方向を選択することが可能となる。プロトタイプシステムを構築し、実際のユーザ行動を取得して検証する。

## 4. 研究成果

研究成果として、各テーマについての代表的な研究成果を報告する。

### (1) ユーザ行動パターンに基づく推薦理由の明示化

推薦理由を明示可能な推薦システムに関する研究

近年、多くの映画が公開されており、その数は日々増大している。そのため多くの映画の中からユーザの好みの映画を見つけるのは難しい。一方で、映画に関する情報を公開する場としてレビューサイトがあげられる。レビューにはその映画のタイトルやあらす

じなどの基本情報のほかに、視聴者による映画に対する感想や、映画を視聴して感じた印象、感情などが書かれて投稿されている。こうしたレビューにはその映画の特徴が多く含まれており、ユーザの好みの映画のレビューにはユーザの好む特徴が表現されている可能性がある。そこで、ユーザが選択した好みの映画のレビューを分析することで、ユーザの好みである特徴を含んだ映画を探すことができる考えた。具体的には、映画のレビューに含まれる視聴者が感じた印象などの情報から、その映画に含まれる要素を特徴ベクトルとして算出する。その特徴ベクトルを他の映画の特徴ベクトルと比較し、同じ要素を多く含んでいる類似映画を推薦する手法を開発した[9]。この手法では、ユーザは推薦に使うレビュークラスタを選択することができ、そのレビューの観点に基づいて推薦結果が表示されていることを把握可能である。また、映画全体を重視するか、選択したレビュークラスタを重視するかということの一つのパラメータで調整でき、どのような基準で推薦されるのかを把握可能である。また、映画以外にも書籍を用いて、同様の取り組みを行った。書籍では、レビュー中から印象を表す語を抽出し、レビュー間の類似度を得る。次に共通するレビュアーの投稿したレビュー間でのみ類似性を測ることで、同一の観点に基づく類似書籍を推薦する書籍推薦手法を開発した[4]。

#### 推薦理由の多様化に関する研究

推薦理由は推薦元コンテンツと推薦先コンテンツの関係や、ユーザと推薦先コンテンツの関係、ユーザと類似ユーザの関係に基づくものである。先の映画、書籍を用いた推薦システムは、推薦元コンテンツと推薦先コンテンツの関係をレビューを用いて表現したものである。ここでは、レビュー以外を用いたコンテンツとコンテンツの関係、ユーザとコンテンツの関係、ユーザと類似ユーザの関係を表現する試みについて報告する。

コンテンツとコンテンツの関係として、ユーザの印象に残るといふ観点での特徴抽出を行った[15]。この研究テーマでは、ある商品の特徴語のうち、レビュー中で肯定的にも否定的にも述べられる特徴語を議論が起こり、印象に残る特徴と定義し、それを抽出した。また、学習用の参考書の特徴語を抽出し、それに関連する Web ページを検索することで、元の参考書と Web ページの関係を答えるような穴埋め問題を生成する応用システムの開発も行った[1]。これは、その特徴語を答えることで、学習参考書の内容の復習および関連 Web ページに書かれた補足知識を得ることを目的としている。

ユーザとコンテンツの関係として、ユーザの持つ金銭的価値観を推定することで、価値に基づく関係を表現した[6,8]。この研究では、価値工学に基づく価値観モデル参考に、

ユーザの購入履歴から価値観モデルのパラメータを推定する手法を開発し、ユーザの価値観に照らし合わせたときのコンテンツの価値を推定する手法を開発した。ユーザと類似ユーザの関係として、映画視聴時の心拍数を用いてユーザ間の類似度を測る尺度の開発を行った[5]。

#### 推薦理由明示化の効果に関する研究

推薦アルゴリズムは複雑になってきており、精度が高い推薦でもユーザに受け入れられるとは限らない。テーマでは、推薦対象に対しての推薦理由の明示化がユーザの満足度上昇に影響を与えるかの評価を行った[16]。ここでは動画共有サイト YouTube からのデータを取得して、他ユーザの再生リストを用いた推薦(協調フィルタリング)、同系統の動画の推薦(コンテンツベース推薦)、評価・人気に基づく推薦、およびそれらの統合による推薦の、計4種類の推薦システムの実装を行った。先の3つはシンプルでユーザに理解されやすいアルゴリズムとして、統合システムは複雑であり、ユーザには理解されにくいアルゴリズムとして用いた。また、これらは推薦理由を生成し、ユーザに提示可能とする。シンプルなアルゴリズムを持つ推薦手法と、複雑なアルゴリズムを持つ推薦手法、およびそれぞれの推薦結果に付加された推薦理由の明示の有無のパターンを組み合わせた計4種類のパターンを別々の被験者に利用してもらい、満足度評価の計測を行った。その結果、推薦精度によるコンテンツの違いよりも、推薦理由を示されているか否かの影響が大きく、推薦理由によってユーザの満足度に大きな違いが出るという知見を得た。

#### (2) ユーザ選択操作に基づくによる推薦方向の予見

##### ユーザのレビュー選択に基づく観光スポット検索・推薦手法[2]

レビュー文章はユーザの体験を表現しており、類似する体験を得たい場合、スポット名やカテゴリを元に検索・推薦を行うより、ユーザレビューを元に検索・推薦を行う方がユーザの要求をうまく表現でき、ユーザにとっても検索・推薦結果を予想しながら実行できると考えられる。そこで、観光スポットに関して、レビューのみを選択するという入力による検索システムを構築した。

まず、ユーザがランダムに提示されるレビューから好むレビューを選択する。次に、選択されたレビューと類似する特徴を持つ観光スポットを検索する。この時、レビューや観光スポットの特徴として、レビュー文やスポットの特徴、季節による特徴、同行者のタイプによる特徴のそれぞれの特徴ベクトルを用いる。特徴ベクトルの生成は、TFIDFを用いる方法と単語出現確率を用いる方法を定義した。

本研究では、スポット、季節、タイプの3種類の特徴を用いる。スポットに関しては、

そのスポットについて書かれたすべてのレビューを1つの文書として扱う。

本手法では入力として、ユーザがレビューを  $m$  件中  $n$  件選択する。それをまとめて1つの文書と見なし、他の選択されなかった個々のレビューも文書とみなすことで、TFIDF 値を算出し、選択レビューの特徴ベクトルとする。季節に関しては、各季節のレビューをまとめて季節ごとに1つの文書と見なす。その季節における単語の出現確率と全レビューにおける単語の出現確率の差を用いて特徴量として算出し、特徴ベクトルを生成する。タイプの特徴量も同様にして抽出する。

最終的に、ユーザの選択レビューの特徴ベクトルと各観光スポットの特徴ベクトルから類似する観光スポットを抽出する。

リツイートパターンに基づくツイートの到達確率判定と編集支援手法[11]

コンテンツそのものの推薦に限らず、コンテンツ生成時のコンテンツの部品を推薦することも考えられる。このとき、その推薦結果を採用するとコンテンツがどのようなものかを予見することが重要となる場合がある。

Twitter を情報発信ツールとして活用している例は、個人や企業に関わらず多くなってきている。情報を拡散する手段としてリツイート機能(以下、RT)が存在するが、意図してRTをしてもらうことは難しく、思ったように情報が拡散できていない例は多く見られる。また炎上のように意図せず情報が拡散されてしまい、社会的な損失につながるようなケースも発生している。そこで本研究では、ユーザの RT パターンをナイーブベイズによって学習することで、任意のツイートの RT 確率を推測し、より RT されやすいツイートにするための編集支援手法を開発した。この手法により、ツイート内容を予測 RT 数に基づいて編集することが可能となり、ツイート結果についての予見が可能となる。

対象ユーザのフォロアーそれぞれについて、ナイーブベイズによってリツイート予測モデルを構築する。ツイートのテキストは形態素解析を行い、名詞・形容詞・形容動詞を抽出した。ツイートのメタデータ(投稿時間、RT、Fav、文字数、URLの有無、メディアの有無)については、それぞれユニークな語に置き換え、訓練データとした。例えば RT 数が0の場合は RT\_NONE RT 数が多い場合は RT\_MANY というような語に設定した。

時間帯については6~10時台を朝、11~14時台を昼、15~18時台を夕方、19~22時台を夜、23~翌5時台を深夜とした。なお、ユーザがRTしたツイートについてはTwitterAPIではRTした時間を取得することができないため、RT直前のツイートとRT直後のツイートの間の時間をRTしたツイートの時間とし、RTしたツイートが前後のツイートから4時間以上離れている場合は時間帯の特定ができないと

して訓練データには追加しない。

また、文字数については訓練データ全体のツイート本文(ユーザID、URLを除く)の文字長の中央値未満ならば「短いツイート」中央値以上ならば「長いツイート」とした。

上記のように作成した、RT 確率モデルを用いて、あるツイートに対し、各フォロアーの RT 確率がしきい値以上になるフォロアー数を予測 RT 数とする。このとき、フォロアーのフォロアーに RT されることも考えられるが、それらの確率モデルは、対象ユーザのフォロアーの確率モデルで近似できるものとし、予測 RT 数を加えた状態で、再帰的に予測 RT 数を求める。

このようにして、ツイートの編集結果が起こす影響を予見するとともに、RT 確率を高くする要因を推薦することで、コンテンツの部品の推薦も可能とした。

コーディネート投稿サイトのユーザタグを用いたコーディネート間の類似度に基づく検索ナビゲーションシステム[17]

ファッションコーディネートのように、ユーザが最終的な結果について明確な要求を持たず、探索的に決定するケースにおいては、探索の方向性とそれを段階的に探索できる仕組みによって、ユーザの推薦結果に対する予見を支援できると考えた。

近年、wear のようなファッションコーディネート投稿サイトが登場してきている。このようなサイトでコーディネート検索の際には、ブランドや特定のアイテム、ユーザ、タグなどから検索するのが一般的である。しかしユーザの中には、あるコーディネートに対し「もう少しカジュアルにしたらどうなるだろう」といった検索意図が存在するが、現状の検索方法ではそういった検索意図には対応していない。そこで、ユーザが任意のコーディネートを選択後、「シンプル」や「カジュアル」などのスタイルを選ぶことで、選択コーディネートより該当スタイルに近づく類似コーディネートを表示するシステムを提案する。

具体的には、コーディネートに付与されたタグを「コーディネートを表すタグ」と「それ以外」に分類する方法を開発し、また、タグを次元とした空間は疎な空間になるため、低次元空間に写像することで、低次元空間内でスムーズなナビゲーションができる方式を開発した。そのことにより、ユーザが徐々に自分のイメージに合うコーディネートを見つけられるようにナビゲートする。また、本システムの有用性を検証した。

なお、本研究テーマは第8回データ工学と情報マネジメントに関するフォーラムにおいて、発表学生が学生プレゼンテーション賞を受賞するなど、高く評価された。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者に

は下線)

〔学会発表〕(計 26 件)

- [1] Web 検索による余談抽出に基づく穴埋め問題生成システム, 大貫 颯, 北山 大輔, 情報処理学会 第 80 回全国大会, 6L-03, 東京都, 3, 2018
- [2] ユーザのレビュー選択に基づく観光スポット検索手法, 潘 健太, 北山 大輔, 第 10 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2018), H1-2, 福井県, 3, 2018
- [3] Web 検索キーワードログと重回帰分析を用いたサッカーの試合検索システム, 久保田 岬, 北山 大輔, 第 10 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2018), P5-3, 福井県, 3, 2018
- [4] 共通レビュアーの観点の類似性に基づく書籍推薦手法, 宮本 達矢, 北山 大輔, 第 10 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2018), F2-5, 福井県, 3, 2018
- [5] 映画推薦のための心拍数を用いたユーザ類似度の設計, 江草 賢一, 北山 大輔, 第 10 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2018), P1-2, 福井県, 3, 2018
- [6] 中古市場のための価格差異を用いた適正価格の個人化手法, 佐藤 大樹, 北山 大輔, 第 10 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2018), P6-2, 福井県, 3, 2018
- [7] 非同期コミュニケーションのためのアニマルライフログの可視化手法, 尾野 恭優, 北山 大輔, 第 10 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2018), P7-5, 福井県, 3, 2018
- [8] 中古市場のための価格差異を用いた価値観に基づく書籍推薦システム, 佐藤 大樹, 北山 大輔, 電子情報通信学会技術研究報告, 117, 374, DE2017-24, 1-5, 東京都, 12, 2017
- [9] ユーザレビューを用いた全体的・部分的観点の類似に基づく映画推薦, 櫛見 圭司, 北山 大輔, 第 9 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2017), B5-4, 岐阜県, 3, 2017
- [10] Word2Vec と Web 検索を用いた検索クエリ置換手法, 鹿島 好央, 北山 大輔, 第 9 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2017), C6-1, 岐阜県, 3, 2017
- [11] ユーザタグとコメントの出現確率の差異に基づく意外な動画の推薦手法, 松山 卓矢, 北山 大輔, 第 9 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2017), I8-1, 岐阜県, 3, 2017
- [12] リツイートパターンに基づくツイートの到達確率判定と編集支援手法, 松林 侑輝, 北山 大輔, 第 9 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum

2017), P3-1, 岐阜県, 3, 2017

[13] レビューサイトの差異に基づくラベル伝播を用いたユーザレビュー分類手法, 浅野 祥汰, 北山 大輔, 第 9 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2017), D4-3, 岐阜県, 3, 2017

[14] 複数人による選択店舗の構造類似性に基づく飲食店推薦システム, 谷口 雄大, 北山 大輔, 第 9 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2017), P4-6, 岐阜県, 3, 2017

[15] A Feature Terms Extraction Method based on Polarity Analysis of Customer Reviews for Content-based Recommendation, Tomofumi Yoshida, Daisuke Kitayama, Proc. of The 18th International Conference on Information Integration and Web Based Applications & Services (iiWAS2016), 9-13, 11, 2016

[16] ユーザ満足度向上のための推薦理由透明性の影響評価, 関根 裕太郎, 北山 大輔, 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2016), B3-4, 福岡県, 3, 2016

[17] コーディネート投稿サイトのユーザタグを用いたコーディネート間の類似度に基づく検索ナビゲーションシステム, 吉越 優美, 北山 大輔, 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2016), B3-1, 福岡県, 3, 2016

[18] レシピ推薦のためのユーザの食材購入周期に基づく余剰食材・購入予定食材推定手法, 吉田 朋史, 北山 大輔, 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2016), P3-5, 福岡県, 3, 2016

[19] 画面選択操作を用いた雰囲気特徴量に基づくカフェ検索システム, 窪田 崇, 北山 大輔, 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2016), F2-1, 福岡県, 3, 2016

[20] レシピ間の対応度と相違性に基づく料理アレンジナビゲーション, 工藤 貴徳, 北山 大輔, 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2016), E2-5, 福岡県, 3, 2016

[21] ユーザの取捨選択行動に基づく嗜好推定による飲食店推薦システム, 児玉 礼, 北山 大輔, 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2016), F2-2, 福岡県, 3, 2016

[22] YARUKI: 行動継続支援のための協調的チェックシートシステム, 森 朱里, 北山 大輔, 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2016), P6-2, 福岡県, 3, 2016

[23] ツイートの感情推定に基づく対になる感情を誘発する行動の推薦手法, 村石 将嗣, 北山 大輔, 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2016), P2-2, 福岡県, 3, 2016

[24] 訪問数とユーザ評価に基づく穴場スポット抽出手法の評価, 北山 大輔, 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2016), P5-2, 福岡県, 3, 2016

[25] 映画レビューの印象タグの希少性に基づく協調フィルタリング, 北中 雄也, 北山 大輔, 第 8 回データ工学と情報マネジメントに関するフォーラム(DEIM Forum 2016), P2-1, 福岡県, 3, 2016

[26] Integration Method for Complex Queries Based on Hyponymy Relations, Daisuke Kitayama and Takuma Matsumoto, 26th International Conference on Database and Expert Systems Applications (DEXA 2015), 2, 359-368, Valencia, Spain, 9, 2015

## 6. 研究組織

### (1) 研究代表者

北山大輔 (KITAYAMA, Daisuke)

工学院大学・情報学部・准教授

研究者番号: 40589975