

Title	Modeling and Predicting User Interests based on Taxonomy(Abstract_要旨)
Author(s)	Nakatsuji, Makoto
Citation	Kyoto University (京都大学)
Issue Date	2010-09-24
URL	http://hdl.handle.net/2433/126781
Right	
Type	Thesis or Dissertation
Textversion	none

(続紙 1)

京都大学	博士 (情報学)	氏名	中辻 真
論文題目	Modeling and Predicting User Interests based on Taxonomy (分類知識を用いたユーザ興味モデリングと予測)		
(論文内容の要旨)			
<p>本研究は、音楽に代表されるように、商品が細分化されたクラスによって管理されている場合に、被推薦者がまだ消費していないが、興味を持ちうるクラスの商品を推薦する方式を提案するものであり、以下の6章から構成されている。</p> <p>第1章は序論であり、本研究の背景、目的について概観している。</p> <p>第2章では本研究の背景技術を紹介し、現在広く用いられている推薦技術である協調フィルタリングについて述べている。また、本研究で、被推薦者と興味に近い利用者を抽出する際に用いるグラフ解析技術、Random Walk with Restarts (RWR) について述べている。さらに、既存の研究における推薦の意外性の定義について整理している。</p> <p>第3章では、利用者の興味を、商品の分類体系(タクソノミ)に沿って抽出する手法を提案している。本研究では、ある商品に興味を持つ利用者は、その商品が属するクラスにも興味を持つと仮定し、利用者の商品に対する評価値を、商品が属するクラスの評価値に反映させる。その上で、利用者間の興味の類似度を、共通して高く評価した商品のみでなく、共通して高く評価したクラスを考慮に入れて計算する手法を提案している。さらに、本研究では、利用者があるクラス配下の商品を消費したことがなければ、そのクラス配下の商品はその利用者にとって意外性のある商品であると定義している。そして、意外性のある商品を精度良く推薦することを狙いとしている。具体的には、まず、前述のタクソノミに基づく利用者間の興味の類似度計算により、クラス間の興味の相関関係を導く。そして興味は共通するクラスの商品で、かつ利用者が消費したことのないクラスに属する商品を推薦する。定量的な検証を通じ、提案手法が精度良く意外性のある商品を予測できることを示している。また実際に利用されているサイトで実験を行い、利用者の消費が意外性のある商品に集中することを確認し、提案手法の有効性、社会受容性を示している。</p> <p>第4章では推薦の幅を広げるため、商品のタクソノミの上位階層に位置する、粒度の粗いクラス(ドメインと呼ぶ)を超えた利用者の興味予測手法について述べている。例えば音楽では、洋楽と邦楽のように、タクソノミの上位で分岐するほど、商品間の関連性は遠くなる。しかし同じ音楽であるため、消費者はそうしたドメインを超えて商品に興味を持つ可能性がある。そのため、まず、ドメイン毎にタクソノミを分割し、それぞれに対し利用者の興味をタクソノミ上に表す。次に、ドメインを跨って商品を消費している利用者に着目し、興味の近い利用者を結んだ利用者グラフを構築する。そして、利用者グラフをRWRを用いて解析し、あるドメインの商品を利用する被推薦者と共通の興味を持つ利用者を抽出し、抽出された利用者が興味を持つ商品を被推薦者に提示する手法を提案している。邦楽と洋楽に対する利用者の興味を、ブログから抽出したデータセットを用いて評価をしたところ、従来の協調フィルタリングに比べ、提案手法が邦楽と洋楽に跨った興味に基づく商品の消費を精度良く予測できることを示している。</p> <p>第5章では、与えられたタクソノミが十分でないときに、利用者のコミュニケーションに基づいてタクソノミを補完する方法を提案している。例えばシステム開発では、開発対象の装置を体系化したタクソノミは構築されないことが多い。そのため、技術者が作成しうる程度の小規模なタクソノミを予め構築し、タクソノミ上のクラスと、技術者間で交換されるメール内で用いられている語句との関連を計算し、そうした語句を用いてタクソノミを補完している。さらに、補完されたタクソノミに専門知</p>			

を含むメールを対応させることで、専門知の蓄積を支援している。実際の開発プロジェクトで用いられるメール集合と技術者が構築したタクソノミを検証し、タクソノミを補完しない場合に比べ、提案手法が高精度に専門知を蓄積できることを示している。

第6章は結論で、本論文で得られた成果を要約している。

注) 論文内容の要旨と論文審査の結果の要旨は1頁を38字×36行で作成し、合わせて、3,000字を標準とすること。
論文内容の要旨を英語で記入する場合は、400～1,100 wordsで作成し
審査結果の要旨は日本語500～2,000字程度で作成すること。

(続紙 2)

(論文審査の結果の要旨)

本研究は、音楽に代表されるように、商品が細分化されたクラスを用いて管理されている場合に、被推薦者に、まだ消費していないが興味を持ちうるクラスの商品を推薦する方式を提案するものであり、得られた主な成果は次のとおりである。

1 利用者の興味を、商品の分類体系(タクソノミ)に沿って抽出する手法を提案した。その上で、利用者間の興味の類似度を、共通して高く評価した商品のみでなく、共通して高く評価したクラスを考慮し計算する手法を提案した。さらに、利用者が、あるクラス配下の商品を消費したことがなければ、そのクラス配下の商品はその利用者にとって意外性が高いと定義した。定量的な検証を通じ、提案手法が、クラスを跨る利用者興味の相関性を計算できるため、意外性の高い商品を精度良く推薦できることを示した。また、利用者が意外性の高い商品を数多く消費することを確認した。

2 邦楽や洋楽などのように、タクソノミの上位に属する、粒度の粗いクラス(ドメインと呼ぶ)を超えた利用者の興味の予測手法を提案した。まず、ドメインを跨って商品を消費している利用者に着目し、興味の近い利用者を結んだ利用者グラフを構築した。そして、あるドメインを利用する被推薦者から見て、他ドメインを利用する関連性の強い利用者を抽出し、彼らが興味を持つ商品を、被推薦者に提示する手法を提案した。実データを用いた検証を通じ、従来の協調フィルタリングに比べ、提案手法が精度良く、邦楽ドメインと洋楽ドメインに跨った商品の消費を予測できることを示した。

3 タクソノミが十分でないときに、利用者間のコミュニケーションに基づいてタクソノミを補完する方法を提案し、システム開発に関する専門知の蓄積に適用した。まず、タクソノミを用いて開発者のメールが含む専門知を同定し、その上で、開発者間の専門知の類似度を計算した。実開発プロジェクトで用いられるメール集合と開発者が構築したタクソノミを検証し、提案手法により高精度に開発者の専門知が蓄積できることを示した。

本研究の結果は、実際に利用されているサイトに実装されている。本研究は、効果の実証された方式を理論的に説明し、その一般化を図ったものであり、学術上、實際上寄与するところが少なくない。よって、本論文は博士(情報学)の学位論文として価値あるものと認める。また、平成22年8月11日に実施した論文内容とそれに関連した口頭試問の結果合格と認めた。

注)論文審査の結果の要旨の結句には、学位論文の審査についての認定を明記すること。更に、試問の結果の要旨(例えば「平成 年 月 日論文内容とそれに関連した口頭試問を行った結果合格と認めた。」)を付け加えること。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。
要旨公開可能日： 年 月 日以降