

個人の感性モデルを推定するインタラクティブな商品推薦システム

宮地 正大

Masahiro MIYAJI

1 はじめに

オンラインショッピングサイトやニュースサイトなどではユーザ個別に推薦内容を変化させるパーソナライズされた推薦システムが組み込まれている¹⁾。これらのシステムでは主に大規模なユーザの購買データをもとに、類似する嗜好を持つユーザや商品のパターンを識別する手法が用いられる。そのため、真に個人の嗜好や感性といった情報を用いた推薦がされているとは言えない。本研究では、ユーザの求める好みの情報を感性モデルとして定義し、対話型遺伝的アルゴリズムの最適化要素を取り入れることで個人に合わせた推薦システムを構築、提案する。

2 提案推薦システム

2.1 提案システム概要

本研究は対話型遺伝的アルゴリズム²⁾により、ユーザとシステムが対話することでユーザの好みのパターンや構成要素を推定し、感性に近い推薦を行うことを目標としている。提案手法の流れを以下に示す。

1. 推薦対象の中からランダムに商品を複数提示
2. ユーザの利用履歴から商品の設計変数同士を交叉
3. 推定したパラメータに類似する商品の提示
4. 手順 2, 3 を繰り返す

提示された商品の中からユーザが次に遷移・閲覧した商品を、遺伝的操作における評価とする。商品の設計変数定義には先行研究で多く用いられている手法と同様に、商品に付加されている説明文に出現する単語を TF・IDF 法を用いて解析したものを用いる³⁾。通常、異なる単語同士は、設計変数空間上で次元が異なるため、交叉することはできない。本研究では、設計変数間に重みとは異なる関連度を近傍として定義することで、異次元設計変数での遺伝的操作を可能とする手法を用いる。商品に含まれる単語間の関係性を数値で表すグラフ(感性語ネットワーク)を単語同士の類似度として用い、それらの組み合わせ及び重みを感性モデルとして定義し、最適化対象とする。

2.2 感性モデル

本研究では、ユーザに内在する物事に対する印象、感性的な好みの度合いなどを単語の組み合わせ及び重み

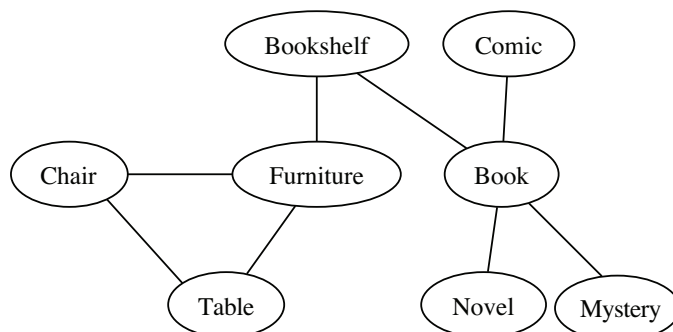


Fig. 1 感性語ネットワークの例

で表現する。それらの集合を感性モデルとして組み合わせ最適化による推定を行う。感性モデルは任意の数の単語及び重みで構成される。ユーザが商品へのアクセスを行った際に、ユーザの感性モデルと閲覧商品のもつ特徴パラメータから新たに感性モデルの候補となる複数の単語の組み合わせを生成する。

2.3 感性語ネットワーク

本稿で用いる単語間の関連を表したグラフを感性語ネットワークと呼ぶ。感性語ネットワークは単語同士の関連を人間ができる限り自然と感じる形で定義することを目的としている。Furniture(家具)とBook(書籍)を中心とした感性語ネットワークの例の一部を Fig. 1 に示す。例えば、'Chair(椅子)'と'Novel(小説)'は語の持つ意味や属性としては大きく離れているが、'Bookshelf(本棚)'という単語を介して関係性を表すことができる。このような語の関連度を表したシソーラスなどの語彙体系には、学術利用目的で公開されているものも存在するが、本研究では商品データの持つ属性情報から、出現単語の共起確率を用いて生成した⁴⁾。

3 実験

3.1 実験目的および実験環境

本システムによってユーザの履歴から嗜好を学習し、類似するキーワードを主題とする商品が推薦結果に現れることを明らかにする。実験は楽天市場における商品データ(以下、楽天公開データ)を用いた。楽天公開データの詳細を Table 1 に示す。楽天公開データの内、すべてのデータを同一に扱うと商品のドメインを横断した推薦がされるため、ある程度ジャンルによる絞込みを

Table 2 推定した感性モデルの例

遺伝子	推定した感性モデル：ソート済み (重み)
1	小説 (0.61), 小説 (0.56), 小説 (0.30), 小説 (0.24), 佐々木 (0.22), 部 (0.19), gt(0.15), 村上 (0.14), 地下 (0.11), 小説 (0.10)
2	図書館 (0.59), マキ (0.54), 小説 (0.34), 月 (0.26), 時間 (0.22), gt(0.19), 月 (0.14), 小説 (0.13), 情報 (0.11), 商品 (0.11)
3	春樹 (0.61), 小説 (0.56), 春樹 (0.31), gt(0.24), cmISBN(0.21), 月 (0.17), gt(0.14), 単行本 (0.14), nbsp(0.11), 月 (0.10)
4	エッセイ (0.64), 著者 (0.45), 時間 (0.30), 部 (0.27), 室 (0.25), 小説 (0.21), 文筆 (0.18), 小説 (0.14), 商品 (0.12), 小説 (0.12)

Table 1 楽天公開データ詳細

全登録商品数	60,123,534 件
書籍データ登録数	3,555,750 件
商品説明項目	商品コード, 商品価格, 商品説明文, 販売方法別説明文, 商品 URL, レビュー件数, レビュー平均, 商品画像 URL, 店舗コード, ジャンル ID, 登録年月日

Fig. 2 提案システムによる推薦書籍¹

行う必要があると考える。本稿の検証実験では、これらのデータの内、書籍データを対象に行った。

3.2 実験内容

タブレット PC 画面に複数の書籍のタイトル、カバー写真が表示してあり、提示される書籍の中から興味がある商品を選択するよう教示を与えた。被験者 22~25 歳男女 6 名、突然変異率 0, 世代数 5, 学習に用いる感性パラメータ数は 10, 一度に提示する推薦商品数は 8 とした。

3.3 実験結果

Fig. 2 に提案手法で得られた推薦商品一覧の例を示す。Table 2 に実験で被験者が選択した商品の履歴、推定した感性モデルの例を示す。この被験者は商品選択の過程で、村上春樹の‘スポーツニクの恋人’という書籍が推薦された。それを選択した結果、村上春樹の著作や

小説の中でもエッセイのジャンルを持つ商品が推薦される傾向が見られた。本手法による学習とは、現在閲覧中のレポートのキーワードと、過去に推定された感性モデルとなるキーワードを概念語ネットワーク上で最短経路となる単語の一つに、次世代の感性モデルを遷移することを指す。提案システムにおいて異なる単語間で交叉を行った場合、ネットワーク上での最短経路をたどり、それらからルーレット選択が行われる。そのため、親個体が本来保持していない単語である‘小説’や‘エッセイ’、‘月’といった異なる商品同士単語を概念上で結ぶような特徴語が生成されており、それらの情報をもとにした推薦が行われた。

4 まとめ

本稿では、人間の感性にもとづく推薦を行うために、対話型遺伝的アルゴリズムを用いた推薦システムを提案した。異なる単語同士の交叉に感性語ネットワークを近傍の定義として用いることで異次元間の交叉を可能にした。本手法を楽天の商品データに適用した結果、ユーザの過去選択履歴を考慮した推薦商品が現れる可能性を確認した。親個体の持つ特徴量から新たに類似する概念を生成した上で推薦を行うため、多様かつ感性的な繋がりを持つコンテンツの提示が期待される。

参考文献

- 1) G. Linden, B. Smith, and J. York. Amazon.com recommendations: item-to-item collaborative filtering. *Internet Computing, IEEE*, Vol. 7, No. 1, pp. 76–80, 2003.
- 2) 高木英行, 畷見達夫, 寺野隆雄. インタラクティブ進化計算, 遺伝的アルゴリズム 4. 産業図書, 2000.
- 3) M. Pazzani and D. Billsus. *Content-Based Recommendation Systems*, Vol. 4321 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2007.
- 4) G. Salton, A. Wong, and C. S. Yang. A vector space model for automatic indexing. *Commun. ACM*, Vol. 18, No. 11, pp. 613–620, November 1975.

¹商品画像は楽天市場 (<http://www.rakuten.com>) より引用