

# 商品サイト上での多角的な情報収集支援のための ブラウザ拡張機能の試作

李 偉健<sup>†</sup> 大園 忠親<sup>†</sup> 新谷 虎松<sup>†</sup>

<sup>†</sup>名古屋工業大学大学院情報工学専攻

## 1 はじめに

インターネット上での商品購入において、複数の情報源からの多角的な情報収集が有益である。例えば、多角的な情報収集により、より正確な情報を得ることが可能になり、偽レビューへの対策にもなる。しかし、そのための作業が手間であり、支援システムが求められている。本研究では、商品サイト上での多角的な情報収集支援として、複数の情報源等(他の商品サイトや動画配信サイト)から収集した情報を用いた推薦機構を持つブラウザ拡張機能を開発している。ここでは、多角的な情報源からの推薦を考慮して、推薦理由の提示に着目した。

本研究では、AHPという手法を使って推薦結果に影響を与える因子の優先度を確定して、各評価基準の重みを計算した。重みに基づいて総合得点を計算することで、透明性を持つ評価結果を推薦理由としてユーザーに提供することができる。

## 2 情報収集支援システム

本システムは、情報準備機構、評価機構、推薦商品展示機構の三つに分かれている。システムの構成図を図1に示す。

情報準備機構は、ユーザーの入力によって、類似度を計算するための商品情報を検索し、動画サイト上で関連ビデオの有無を確認し、その動画に関する情報を保存する。評価機構は、推薦結果に影響を与える因子を集め、ユーザーが設定した偏好によって各因子の重みを調整して推薦した商品の総合得点を計算する。また、上記の情報を用いて本商品を推薦した理由の生成を行い、推薦商品に関する情報を推薦商品展示機構に渡す。推薦商品展示機構は、上記の二つ機構から渡された情報を集め、ブラウザ拡張の情報展示のために必要な情

**Developing a browser extension to support multi-faceted information collection on product sites**

Weijian LI<sup>†</sup>, Tadachika OZONO<sup>†</sup> and Toramatsu SHINTANI<sup>†</sup>

<sup>†</sup>Department of Computer Science and Engineering, Graduate School of Engineering, Nagoya Institute of Technology.

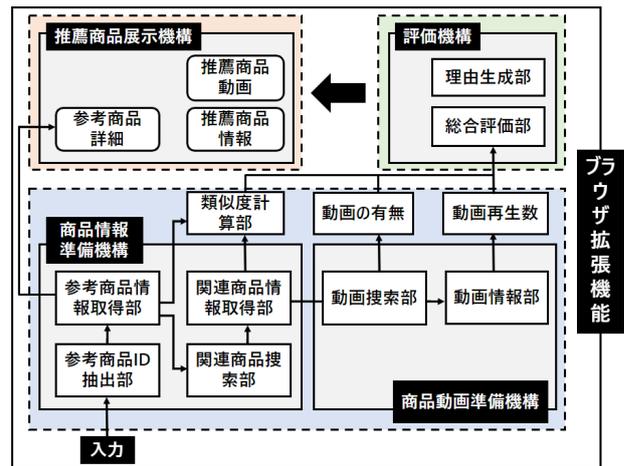


図1: システム構成図

報を使用するという機構である。

### 2.1 類似度計算部

類似度計算部は、入力としての参考商品を用いて、商品の属性を抽出して分析することで、関連商品と参考商品の互いの類似度を計算することを実現する。商品情報の取得については、ユーザーから入力した商品詳細ページのリンクを解析して商品IDを抽出し、APIを使用して商品情報を取得する。取得した情報を用いて検索APIを使用し、同じ種類の関連商品リストを取得する。システムの構成図を図1に示す。

商品サイトでは種類によって商品の属性が大きく違うので、システムの適用性と普遍性を確保するため、タイトル、価格、およびレーティングの三つの属性に基づいて商品の互いのコサイン類似度を計算する。そのために、任意の属性をベクトル化する必要がある。参考商品Aの属性ベクトル $V$ が $(1, 1, 1)$ であることを前提として、関連商品Bの属性ベクトル $V = (v_i, v_1, v_2)$ における価格とレーティングの類似度(それぞれ $v_1, v_2$ )の計算式を以下に示す。

$$v_i = 1 - \frac{|I_{B_k} - I_A|}{\max I - \min I}, \quad (i = 1, 2)$$

ここで、 $B_k$  は関連商品  $B_1, B_2, \dots, B_n$  の一つである。 $v_i$  は、 $A$  と  $B$  の価格およびレーティングに関する類似度を表している。 $I_X$  は、商品  $X$  の価格もしくはレーティングである。 $v_i$  は、 $A$  のタイトルの語ベクトル  $W_A$  と  $B_k$  のタイトルの語ベクトル  $W_{B_k}$  の間のコサイン類似度である。タイトルの処理方法については、仮に二つのタイトルに含まれる単語が類似していれば、その内容も類似していると考え、単語の出現回数に基づいて二つのタイトルの類似度を計算する。具体的な流れは、(1) タイトルにおける単語の種類を並べてリストを作る、(2) 作った単語のリストによってそれぞれのタイトルをベクトル化する、(3) タイトルベクトルに基づいてコサイン類似度を計算する。(1)~(3)を行うことで、結果が関連商品のタイトルの類似度  $v_i$  として使用できる。関連商品の属性ベクトルを計算することで、参考商品の属性ベクトルを用いてコサイン類似度を計算できる。

### 3 評価機構

本機構は、理由生成部、総合評価部の二つ部分に分かれ、推薦結果に影響を与える因子によって、ユーザーが自分で設定した推薦偏好を用いて因子の重みを調整することで、商品の総合得点を計算して評価が一番高いもの推薦し、システムの透明性から考えて推薦理由を付けるという複合の評価機構である。

システムの透明性とは、ユーザーが推薦結果を生成する過程を理解することができ、もしくはあるアイテムが他のアイテムより人気がある理由を説明することができるような属性である。推薦理由が提示されることで、ユーザーがシステムの推薦結果に対する信頼度を高め、推薦結果を受け入れやすくなる。

本研究の推薦結果に影響を与えるのは、コサイン類似度、レビュー数、レーティング、動画の有無、動画の再生数という五つの因子であり、ユーザーが本システムを使用する前に、システムのオプションページで各因子の重要度を設定する必要がある。本研究では、AHPという手法を利用した [1]。重要度の尺度を表 1 に示す。

表 1: A と B の一対比較表

重要度の尺度	定義
1	A と B 同じぐらい重要
3	Aの方がBより若干重要・やや重要
5	Aの方がBより重要
7	Aの方がBよりかなり重要
9	Aの方がBより極めて重要

表 2: 例：評価基準の一対比較表

→比較	類似度	レビ…	レー…	有無	再生数
類似度	1	3	5	7	9
レビ…	1/3	1	3	5	7
レー…	1/5	1/3	1	3	5
有無	1/7	1/5	1/3	1	3
再生数	1/9	1/7	1/5	1/3	1

例えばユーザーにとって、重要度の順位はコサイン類似度>レビュー数>レーティング>動画の有無>動画の再生数とし、「(行) コサイン類似度」は「(列) レビュー数」より「若干重要」なのでこの一対比較値は「3」となる。「(行) レビュー数」と「(列) コサイン類似度」は、上記の比較値の逆数「1/3」を使う。以下、同じように全ての組み合わせの一対比較を行い、結果を表 2 に示す。

本機構では、重み計算法として調和平均 (Harmonic mean) を用いた。具体的な内容は、以下の通りである。

1. 一対比較行列  $A = [a_{ij}] \in R^{n \times n}$  の各行ごとに調和平均をとる。

$$h_i = \frac{n}{\sum_{j=1}^n \frac{1}{a_{ij}}}, \quad (i = 1, \dots, n)$$

2. 和が 1 になるように基準化。重みベクトルを  $(\omega_1, \dots, \omega_n)^T$  とすると、

$$(\omega_1, \dots, \omega_n)^T = \left( \frac{h_1}{\sum_{i=1}^n h_i}, \dots, \frac{h_n}{\sum_{i=1}^n h_i} \right)^T$$

### 4 おわりに

本稿では、商品サイト上での多角的な情報収集を支援するために、ある参考商品を提供する前提として、類似度等の因子に基づいて、より人気が高い商品を推薦することができるシステムを試作を示した。ユーザの推薦偏好を考慮して適切な商品を推薦することで、結果に対する推薦理由を付けることができる。ユーザは、本システムを利用することで、商品サイト上での目紛しい商品に困った際にも簡単な操作を行って適切な商品を発見できるようになり、より効率的な商品検索を行えるようになることが期待できる。

謝辞 本研究の一部は JSPS 科研費 19K12097, 19K12266 の助成を受けたものです。

### 参考文献

[1] Saaty T L. The analytic hierarchy process (AHP)[J]. The Journal of the Operational Research Society, 1980, 41(11): 1073-1076.