

M-047

ネットワークによる人を介した情報推薦 Recommendation by Reliable People

山崎 宏幸[†] 太原 育夫^{††}
Hiroyuki Yamasaki Ikuo Tahara

東京理科大学大学院理工学研究科情報科学専攻

1 はじめに

情報過多と呼ばれる近年、インターネットユーザが Web に氾濫する玉石混交な情報の中から真に有用な情報を得るための手間は日々増大し続けている。その大きな要因は、Web で得られる匿名の発言の信頼性が、それ単体で見た場合、実世界において取得される情報の信頼性に比べ劣ることにある。この問題が生じる一因として、氏名や本籍、住所、職業といった社会的な情報や、年齢、性別、血液型、身長、体重といった身体的な情報によって暗黙的に身柄が担保されている実世界と異なり、発言の相手が見えず、身分や人柄が明示されない Web においては、その発言に責任が伴わず、更に、情報の不足からユーザは相手がどのような人間か判断することができない点が挙げられる。

そこで、本研究では協調フィルタリング [1] におけるユーザのアイテムへの評価を Web におけるユーザの自己表現、即ち Web 世界での認識票 (identification-tag) とみなし、ユーザと嗜好の類似する人間を情報フィルタとして利用する情報推薦手法を提案する。

2 提案手法

提案する情報推薦手法は主に以下の三段階からなる。

1. Web Identification Tag(ユーザの、任意のアイテムやコンテンツに対する評価履歴)の作成
2. 評価履歴を用い、あるカテゴリにおけるユーザの嗜好類似コミュニティを抽出
3. 抽出されたコミュニティの評価履歴から、概念において関連の見られる別カテゴリのアイテム等を推薦

3 Web Identification Tag

Web において一定の匿名性を保ちつつユーザの身柄を担保する最低限の要件として以下の二つに着目する。

- ハンドルネーム
- 商品、サービス、Web コンテンツ等への評価

ここで、ハンドルネームとはユーザの Web における匿名であり、実世界における芸能人の芸名のようなものである。また、商品等への評価とは具体的には以下の項目からなる。

- 対象 (商品、サービス、Web コンテンツ等の名称又は URI)
- 評価 (五段階評価等)
- Web 文書 (対象商品等に関する文書による説明、評価または自ブログへのリンク等)

提案手法では、このような項目によって明示される、ユーザが過去に経験、体験した商品等への評価を Web におけるユーザの自己表現とみなしユーザ間の身柄を担保する。

4 嗜好類似コミュニティの抽出

本提案手法では、後述する2つの協調フィルタリング手法によってユーザ間の類似度を求め、特に類似度の高いコミュニティをカテゴリにおけるユーザの嗜好フィルタとして提示する。またこの時、推薦するアイテム等を嗜好類似コミュニティの評価履歴の範囲に限定することで、本来ユーザが探索すべき膨大な情報をフィルタリングすることができる。

4.1 協調フィルタリング

今、ユーザ数を m とし、全ユーザを表す集合を $U=\{u_1, u_2, \dots, u_m\}$ とする。また、商品数を n とし、全商品を表す集合を $I=\{i_1, i_2, \dots, i_n\}$ 、ユーザ u_a が評価を行った商品の集合を $I_{u_a} \in I$ 、ユーザ u_a の商品 j に対する評価値を $r_{a,j}$ 、ユーザ u_a の評価値の平均値を \bar{r}_a とする。ユーザ u_b に対するユーザ u_a の類似度 $c_{b,a}$ を以下の式で定義する。

$$c_{b,a} = \frac{\sum_{j \in I_{u_b} \cap I_{u_a}} (r_{b,j} - \bar{r}_b)(r_{a,j} - \bar{r}_a)}{\sqrt{\sum (r_{b,j} - \bar{r}_b)^2 \sum_{j \in I_{u_b} \cap I_{u_a}} (r_{a,j} - \bar{r}_a)^2}} \quad (1)$$

4.2 条件付き協調フィルタリング

条件付き協調フィルタリングでは、ユーザーがシステムによって提示される嗜好類似コミュニティに対して、例えば以下に示すような条件を付ける。

- 『ショーシャンクの空に』を評価済み
- 『ソードフィッシュ』を高評価
- 『ハウルの動く城』を低評価

これは、ユーザー自身がデータに重み付けを行うことを意味し、特にこだわり (若しくは反感) のある対象に対し、ユーザが正または負の重み付けを行い適時に嗜好フィルタを切り替えることで、より能動的な推薦情報の取得が可能となる。

[†] 東京理科大学大学院, Graduate School of Tokyo University of Science

^{††} 東京理科大学, Tokyo University of Science

5 カテゴリを越えた推薦

人の嗜好はカテゴリによって異なると言われる [2]. そのため, 強調フィルタリングではカテゴリ毎に類似度の計算が行われ, 結果, カテゴリ毎に cold-start の問題が生じる [3]. しかし, 本提案手法では, あるカテゴリにおいて得られた嗜好フィルタ (嗜好類似コミュニティ) をカテゴリの概念において関連のある別カテゴリに流用することでこの問題を緩和し, 同時に推薦に根拠を与えることができる. 以下に例を示す.

1. 料理カテゴリの嗜好類似コミュニティ

⇒ 旅行カテゴリのサービス推薦

2. 映画カテゴリの嗜好類似コミュニティ

⇒ 書籍カテゴリのアイテム推薦

1の例では, 嗜好類似コミュニティはユーザの味覚嗜好フィルタとして働き, 旅行の目的が食事であれば信頼性の高い推薦が可能である. 一方2の例では, コミュニティはユーザのストーリー嗜好フィルタとして働き, ストーリー性のある書籍に対して信頼性の高い推薦が可能である.

6 実験

提案した Web Identification Tag の要件を満たす Amazon[4] のカスタマーレビューを対象に実験を行った. 評価履歴として175作の映画と35作の書籍を持つ山崎の評価履歴に対し協調フィルタリングを行った結果, 得られた類似度の一番高いレビューの評価履歴が表1, アクション映画『ソードフィッシュ』を高評価したレビューに対し条件付き協調フィルタリングを行った結果, 得られた最も類似度の高い評価履歴が表2, 表1のレビューが高評価した書籍が表3である.

表1: 協調フィルタリング

高評価作品	中評価作品	低評価作品
ショーシャンクの空に	スピード	パニックルーム
レインマン		
サハラ	手紙	モブスターズ
メメント	パッション	薔薇の名前
ココシリ	アザーズ	戦火の勇氣

尚, Amazon レビューの評価は五段階評価であるため, ここでは便宜上1, 2点を低評価作品, 3点を中評価作品, 4, 5点を高評価作品としてある. また, 表1と表2の二重線で区切られた上段は山崎が評価済みの作品である.

表2: 映画『ソードフィッシュ』を高評価とした条件付き協調フィルタリング

高評価作品	中評価作品	低評価作品
ソードフィッシュ	エイリアン3	踊る大捜査線2
エイリアン		
エイリアン2		
M:I-2		
マイノリティリポート		
レオン		
紅の豚		
X-MEN 2		
ピースメーカー		
007 カジノロワイヤル		

表3: 推薦書籍

蒼き狼	天平の甍	大地の子
沈黙	砂の女	小沢一郎の政権奪取戦略
大地の咆哮	クロニク世界全史	昭和天皇日々の食

7 実験結果と考察

実験の結果, 単純な強調フィルタリングで得られたレビューはどちらかと言えば人生をテーマに扱った作品において山崎と評価の一致が見られ, その推薦書籍もやはり人間の人生をテーマに掲げたものが多いように感じられた. 一方, 条件付き協調フィルタリングで得られたレビューは, 主にアクション系の映画において山崎と評価の一致が見られ, 得られたレビュー二人の間には嗜好の偏向に顕著な違いが見られた. 書籍においても協調フィルタリングを行ったが, 評価アイテムが二以上一致するレビューが Amazon に存在せずフィルタリングを行うことが出来なかった. 以上の結果から, カテゴリにおいて得られた嗜好類似コミュニティを嗜好フィルタとして別カテゴリに流用する手法は, ユーザの嗜好を情報推薦に反映させる上で一定の効果があり, cold-start の問題に対し一つの解決策を与えると言える.

8 まとめ

今後は, 提案する Web Identification Tag の採用がユーザに与えるインセンティブや, ユーザの満足度に対する優位性といったものを調査していきたい.

参考文献

- [1] 沼尾 正行, 横山 甲: “階層化された知識の継承による情報フィルタリング,” 情報処理学会研究報告, vol. 99, No. 47, pp.43-48 (1999)
- [2] 石野 正彦, 八巻 直一, 市川 照久: “商品属性を用いたリコメンデーションシステムの提案,” 情報処理学会研究報告, vol. 2004, No. 116, pp.41-46 (2004)
- [3] 神島 敏弘: “推薦システムのアルゴリズム (2),” 人工知能学会誌, vol. 23, No. 1, pp.89-103 (2008)
- [4] Amazon <http://www.amazon.co.jp/>