

個人に適応する情報フィルタの個人プロフィールの相互利用は、ユーザの興味を広げる

芦川拓実[†] 小林暉[†] 金道敏樹[†]

金沢工業大学工学部情報工学科[†]

1. はじめに

情報化技術の発展により、SNS などでも誰もが情報を大量に取得でき、発信できるようになった一方で、欲しい情報が何か分からないという問題[4]や、インターネットの検索履歴が「フィルター」となって同じような情報ばかりが表示されてしまい、まるで泡の中のように、自分が見たい情報しか見えなくなってしまう「フィルターバブル」のような問題も起きている。

これらの社会的課題の解決を目指して、Amazonなどの商品レコメンドシステムに代表される協調フィルタに対して、計算リソースが少なくエッジ型サービスへの実装に適した個人適応型情報フィルタ INSOP[1]を用いて、個人の情報プロフィールを柔軟に相互利用するために、以下のような提案を行った(図1)。

- INSOPの出力の相関係数によって個人プロフィールの類似性を評価・利用することで、似通った興味を持つ他者を重視するような個人プロフィールの相互利用が可能であるのではないか。
- より柔軟な個人プロフィールの相互利用を目指し、コンテンツに紐づいた共起キーワードの興味度の推定値から個人プロフィールのクラスタリングを行い、選択的にユーザの興味の一部だけを利用すると意外性のあるコンテンツの推薦ができるのではないか。

今回、意外性のある推薦が求められる場合(映画)と、あたりさわりのない推薦(論文)が求められる場合の2つの観点から、ユーザの興味を拡張可能かを調べる実験を本大学の学生及び職員に行った結果を報告する。



図1: 我々のフィルタリングシステムのコンセプト

2. 提案手法

論文検索では、自分の研究分野に関連する論文を知りたいというニーズや備忘録としてのニーズが高く、既知のアイテムに近いあたりさわりのないコンテンツの推薦が望まれる。映画のような嗜好性の強いジャンルの検索では、自分の興味の少し外にある、意外性のあるコンテ

ンツを知りたいというニーズがある。したがって、少し意外性のあるセレンディピティ[4]の高いコンテンツの推薦が好まれることが予測される。

2.1. 似たユーザを利用する相関係数

個人プロフィールを実効的に広げるアイデアの1つはユーザ α の個人情報フィルタが興味の似ている他者ユーザ β の個人プロフィールを利用することである。今、ユーザ α 、 β の個人情報フィルタの出力がコンテンツ a に対する各ユーザの興味の度合いの推定値 $C(\alpha, a)$ 、 $C(\beta, a)$ であるとすると、このとき、ユーザ α にユーザ β の興味を加えた興味の度合い推定値を

$$C^+(\alpha, a) = C(\alpha, a) + \sum_{\beta} S(\alpha, \beta)C(\beta, a)$$

と置き換えられる[2,3]。ここで、 $S(\alpha, \beta)$ はユーザ α 、 β の興味の類似性を表す相関係数である。相関係数を用いてプロフィールを定量化することで、似通った興味を持つ他者を重視するような個人プロフィールの相互利用が可能である。

2.2. 意外性を加える興味を分割したサブプロフィール

意外性を持った推薦を実現するために、キーワードの共起を利用したサブプロフィールの利用を考える。

ユーザの興味を分割した、サブプロフィールはその人の特有の興味を示しているはずで、ジャンルを表す一般的な単語とそのユーザ固有の単語からなると期待している。個人プロフィールは、興味がある/ないと回答したコンテンツを持つすべてのキーワードの組み合わせから、各キーワードの興味の度合いの推定値を求め、閾値 θ を満たした場合を1、それ以外のキーワードを0とすることで共起単語の正方行列を作成できる。その正方行列をクラスタリングすることで、単語のクラスターが取得できる(図2)。

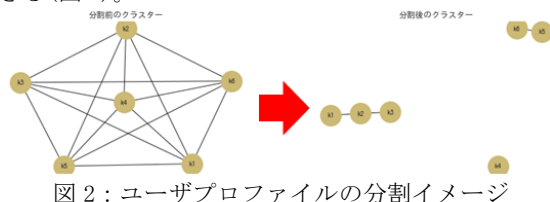


図2: ユーザプロフィールの分割イメージ

サブプロフィールはSFやホラーなどユーザの興味の一部を表していて、ユーザ同士のサブプロフィール間の類似度を推定し、類似度の高いサブプロフィールを利用することで意外性のある幅広いジャンルの推薦を試みた。

3. 推薦システム試作

3.1. 当たり障りのない推薦システム

試作した論文推薦システムは以下のものである。論文

Mutual use of personal profiles of personalized information filters expands user interests

[†]Takumi Ashikawa, [†]Hikaru Kobayashi, [†]Toshiki KINDO

[†]Department of Information and Computer Science, College of Engineering, Kanazawa Institute of Technology

のタイトルと概要を図3の様に表示し、ユーザの興味の有無を入力できるようにした。各論文には、概要から抽出した論文の特徴を表すキーワードが割り振られており、ユーザの興味の有無との関係をINSOPエンジンで獲得する。そして各論文に対するユーザの興味の度合いを推定し、他のユーザとの興味の類似度(相関係数)を推定する。

インターフェース上ではユーザの興味に近い他者の候補を提示し、プロフィールを利用できるようになっている(図3)。

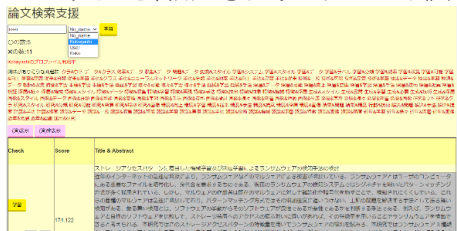


図3：論文推薦システムのインターフェース

3.2. 意外性のある推薦システム(映画)

個人情報フィルタ INSOP を CGI を用いた Web アプリケーションとしてバックエンドを Python で、フロントエンドを HTML/CSS/JavaScript で実装した(図4)。

300 作品の映画データをスクレイピングし、映画のジャンルは「sf」、「action」、「anime」、「comedy」、「suspense」、「horror」、「human_drama」、「fantasy」、「love_story」の9ジャンルにした。

1つの映画には約50~200単語ほど紐づいていて、それらの単語から共起ワードを生成し、単体ワードと共起ワードを合わせて、1つの映画に対して500~2000単語ほど紐づいたドキュメントプロフィールを作成している。



図4：映画推薦システムのインターフェース

4. 実験結果と考察

4.1. 論文推薦システム

学生と職員を対象とした実験を行い、入力結果から各ユーザ同士の興味の類似度(相関係数)を求めた。異なる人物の User1 と User8 が同一のカテゴリ(深層学習)に興味があるものとして推定した場合、興味の類似度は $r=0.403$ となった。また User1 が User8 のプロフィールを利用して、さらに学習をすると、各論文に対するスコアの相関のばらつきが小さくなった(図5)。これは User1 が User8 の視点を妥当なものとして取り入れたことを意味する。

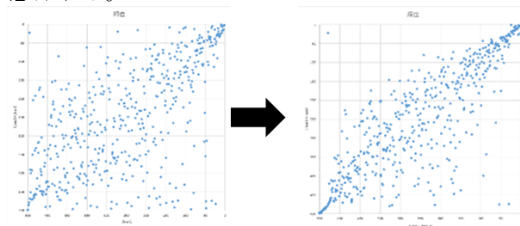


図5：ユーザの興味の变化を示す散布図

4.2. 映画推薦システム

6人の被験者に映画推薦システムを実際に使用してもらい、アンケートによる主観評価を行った。

映画評価方法は300作品の映画の中から30作品ほどに対して、ユーザの興味のある映画に○、興味のない映画に×の2値入力をしてもらいプロフィールを生成する。そのプロフィールを基に、2.1と2.2の技術の中で、BからFの5パターンの推薦結果を見て「意外性」、「興味度」の2つの項目において1~5段階で推薦システムを評価してもらった。

また、他者の興味を用いることで大きく推薦順位が上がった映画を赤色に、大きく推薦順位が下がった映画を紫色に分けて表示をすることでユーザにとって視覚的に順位の変動が分かりやすいように工夫をした。

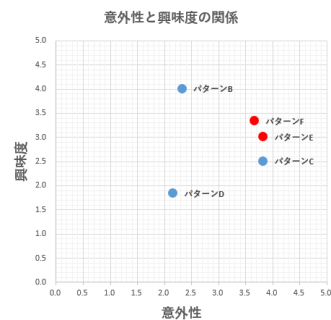


図6：映画推薦システムの利用ユーザの評価

その結果、正の相関があるユーザのプロフィールそのものを利用したパターンでは興味度が高い一方で意外性の値は低かった。サブプロフィールを活用したパターンに関しては、図6の赤点に見られるように興味度も意外性のスコアも高かった。

上記の実験結果から、ユーザプロフィール全体の相互利用は、ユーザの視野変化に繋がるあたりさわりのない推薦に有効であり、サブプロフィールの相互利用はそれにさらに意外性を付加できると結論できる。

5. まとめ

興味の似た他者の知見を利用することで、映画と論文というジャンルの異なるコンテンツを推薦する場合においても、ユーザにとって視野拡張のチャンスが生まれることが示された。

提案した推薦システムが利用者により受け入れられるように、インターフェースの改善や推薦精度の向上、プライバシー保護の観点からスマートフォンやタブレット端末などのエッジデバイスへの実装が今後の課題である。

参考文献

- [1] Toshiki KINDO *et. al.* Adaptive Personal Information Filtering System that organizes personal profiles automatically, Proceeding of IJCAI97, 1997.
- [2] 小林暉他, エッジ型情報収集支援における個人クエリ相互利用の利用者視野拡大効果, 情報処理学会第82回全国大会, 2020.
- [3] 芦川拓実他, 個人情報フィルタを使ったエッジ型情報収集支援における個人クエリ間の違いの定量化, 情報処理学会第82回全国大会, 2020.
- [4] 神島敏弘, 推薦システム-情報過多時代をのりきる(特集)情報のフィルタリング, 情報と科学の技術, 56巻(2006)