

ユーザの嗜好を反映する図書検索システム

2Y-6

芝 温子 谷 幹也 市山 俊治

NEC 関西 C&C 研究所

1 はじめに

専門的知識から教養・趣味の情報まで広範囲をカバーする情報収集手段として、図書検索を行う電子図書館システムは今後の開発・実用化が期待されている。

現在一般に広まっている図書検索方式は、著者名やキーワード等の条件にあうものを検索する書誌情報検索方式であり、明確な目的意識・要求を持ったユーザを対象としている。

一方近年、対象情報に対するユーザの評価情報を蓄積し、検索者との嗜好類似度で重み付けした評価平均から推薦情報を決める方式：ソーシャルフィルタリング方式 [1][2][3] が研究・開発されてきた。この方式では、ユーザ評価を利用することで「質の良い情報」を得られる可能性も高くなる。またあらかじめ対象情報への評価を入力しておけば、陽に条件を指示しなくても好みに応じた情報が得られるため、要求が潜在しているユーザに対して有効と考えられる。

しかし一般図書館の図書検索システムでは、図書を十分に絞り込める明確な目的意識・要求をもったユーザや、要求が潜在的なユーザだけでなく、明確な要求と潜在的な要求を合わせもつようなそれらの中間に位置するユーザの利用を考える必要がある。そのようなユーザにとっては、書誌情報検索方式とソーシャルフィルタリング方式のどちらかを択一するインターフェースは有効ではない。

本稿では、ソーシャルフィルタリング方式に書誌情報検索方式を融合させて、より幅広いユーザが利用できる図書検索システムを提案する。

2 ソーシャルフィルタリング方式

2.1 概要

ソーシャルフィルタリング方式は、以下のような手順でユーザへの提示情報を作成する。

1. 対象情報に対するユーザ評価の収集

検索対象である各情報をユーザが採点し、評価情報として蓄積する。例えば音楽情報の推薦を行う MIT のシステム：'Ringo'[2] では、ユーザは 1～7 の離散値を各ミュージシャンへの評価として入力する。

2. 対象ユーザの嗜好を反映した予測評価の算出

上記の評価に基づき、ユーザ（検索者）と他のユーザ（検索者以外）との嗜好の類似度（ピアソン積率相

関係数等）を計算する。嗜好類似度が閾値以上のユーザ評価を、嗜好類似度で重み付けした平均を算出し予測評価とする。

3. 推薦情報の生成

予測評価が高いほどユーザの嗜好にそった情報と判断しユーザに推薦する。

- 1) ユーザがあらかじめ対象情報への評価を入力していて、
- 2) 十分に多くの他のユーザの評価情報が蓄積されていることが有効に動作する前提となる。

2.2 図書検索システムへの応用における問題とその解決の方針

ソーシャルフィルタリング方式を一般図書館の検索システムに取り入れる場合、以下のような問題がある。

1. 明示的要求をもつユーザは要求条件を反映する枠組がない。

例えば、新刊書、文学、随筆等のジャンルで柔らかい文体が好きなユーザが、「随筆の本を読みたい」と思った場合、ソーシャルフィルタリング方式だけでは、ユーザの意図から外れた他のジャンルの本等が多く含まれる。また、書誌情報検索方式では適切な図書の絞り込みができない。

2. システムが提示した推薦情報に不満があるユーザや、自分と異なるタイプのユーザの評価を知りたいユーザが情報を探索することができない。

過去に評価採点の蓄積が少なく精度の良い推薦情報が得られない場合、代替情報を探索する手段が要求される。また現在の嗜好以外の図書に興味を持ち、自分と異なるタイプのユーザに人気の図書を探してみたいユーザもいる。

上記問題を解決するために、1 に対しては、書誌情報検索方式と組み合わせて、要求条件を入力できるようにする。明示的要求をもつユーザは条件を入力して検索を実行すれば、その条件を満たし、かつソーシャルフィルタリング方式によるユーザの嗜好を反映した絞り込みによる結果情報が得られる。2 に対しては、対象情報への予測評価を算出する際に、比較対照ユーザ層を限定するためのパラメータと嗜好反映度を変更するパラメータを提供し、自分の嗜好を反映させず対照ユーザ層を絞った中での推薦情報を閲覧できるようにする。

図 1 に従来の方式と本システムの方式でカバーする範囲を比較し図示する。

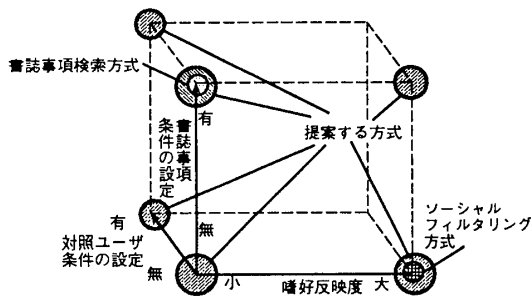


図 1: カバーする検索要求の範囲の比較

3 ソーシャルフィルタリング方式と書誌情報検索方式を融合した検索方式

ユーザが入力できる条件として、1) 書誌事項条件、2) 嗜好反映度、3) 対照ユーザ条件の3つを用いる。

ユーザのタイプと操作方式は以下ようになる。

- 潜在的な要求しかないユーザ
書誌事項条件や対照ユーザ条件を入力せずに、嗜好反映度を大にして検索要求を行う。システムはソーシャルフィルタリング方式により各ユーザの嗜好に応じた推薦情報を提示する。
- 明示的な要求をもつユーザ
書誌事項条件を入力し検索要求を行う。システムは書誌事項条件に一致する図書を検索し、それらの図書に対しソーシャルフィルタリング方式による予測評価を算出する。他のユーザとの嗜好類似度を算出する際は、ユーザが指定した条件から特定される範囲の図書に対する評価だけを利用する。
- システムからの推薦情報に不満なユーザや、自分と異なるタイプのユーザの嗜好が知りたいユーザ
嗜好反映度を下げ、対照ユーザ条件を指定し検索要求を行う。例えば若い世代に人気のある本が知りたいユーザは、嗜好反映度を下げ対照ユーザを若い世代に指定して検索要求を行う。システムは、
 - 嗜好反映度に応じた類似度の閾値の変更や類似度に比例した重み付けの値の変更を行う (嗜好反映度を下げると重み付けの差を小さくし、全ての評価を同等に扱うようにする)
 - 対照ユーザ条件にあうユーザを検索し、それらのユーザの評価のみを利用する

等をして評価予測を行い推薦情報を提示する。

本方式では対照ユーザ条件の設定を行うため、年齢や性別等のユーザ情報が蓄積されていることを前提とする。

4 システム構成

図 2 に前章の融合検索方式を実現するシステム構成を示す。対照ユーザ条件としては、一般ユーザでも比較的分かりやすい条件として、ユーザの年代、性別、職業等を考えている。書誌事項条件、対照ユーザ条件、嗜好反

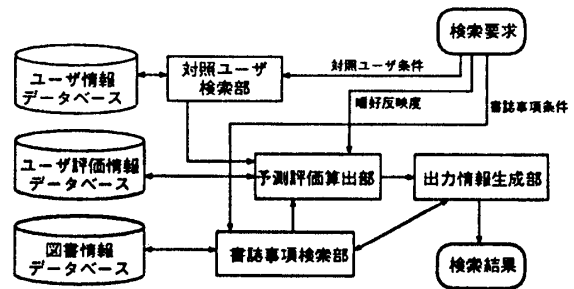


図 2: システム構成図

映度はそれぞれ、書誌事項検索部、対照ユーザ検索部、予測評価算出部に送られる。書誌事項検索部、参照ユーザ検索部ではそれぞれ、送られてきた条件に一致する図書 ID 集合、ユーザ ID 集合を検索する。予測評価算出部では、その図書 ID 集合内の図書に対する、ユーザ ID 集合内のユーザの評価情報を検索する。そして評価情報に基づき、検索者和其他のユーザの嗜好類似度を計算し、嗜好反映度を加味して閾値の設定と重み付けした評価の平均を求め予測評価とする。出力情報生成部では、図書 ID と各図書への予測評価を受け取り、書誌事項検索部を通じて図書のタイトルを取得し予測評価値の順に並べる等を行い出力情報を生成する。

5 おわりに

ソーシャルフィルタリング方式と書誌情報検索方式の融合により、潜在的な要求、明示的な要求の一方または両方をもつユーザに対応した図書検索システムを提案した。また、ソーシャルフィルタリング方式における予測評価の算出において、評価を参照するユーザ層と嗜好の反映度をユーザが調整でき、自分とは異なるタイプのユーザの嗜好が探れるようにした。

ソーシャルフィルタリング方式が抱える問題として、1) 各ユーザが評価を入力する手間、2) 精度の良い推薦情報生成には多くのユーザの評価情報の蓄積が必要、といった点がある。操作履歴からのユーザ評価の推量や、借りだし・返却等の図書館での動作の流れの中でユーザが自然に評価を入力できる仕組み作りが必要である。また、ユーザの評価情報が収集しきれない内は、一般的な人気という意味で貸出回数等に基づく推薦情報生成等の処置が必要と思われる。今後はそれらも考慮しながら、システムの構築を行っていく。

[参考文献]

- [1] Resnick, P., Iacovou, N., et al. "Group Lense: An Open Architecture for Collaborative filtering of News", CSCW 1994 Oct, 1994
- [2] Shardanand, U. and Maes, P. "Social Information Filtering: Algorithms for Automating "Word of Mouth"", CHI'95 Conference Proceedings, PP.210-217, May 1995
- [3] Hill, W., Stead, L., Rosenstein, M. et. al., "Recommending and Evaluating Choices in Virtual Community of Use", CHI'95 Conference Proceedings, PP.194-201, May 1995