

Web 情報推薦システムにおける推薦精度の向上のための Web 閲覧行動に基づいた暗黙的フィードバック方式の検討

鷹野孝典[†] Kin Fun Li[‡]

神奈川工科大学情報学部情報工学科[†]

Department of Electrical and Computer Engineering, University of Victoria[‡]

1. 背景

World Wide Web (WWW) 上では、大量かつ多様な文書データや情報サービスが生産され続けており、これらの情報資源を対象とした情報獲得行動の重要性が増している。利用者の要求に沿った情報を効率的に獲得するために、多くの情報獲得方式が提案されている。特に、協調フィルタリング方式[1]は、e-commerce 等の応用分野で個々の利用者の情報嗜好に合致した情報を推薦するために広く用いられている。協調フィルタリング方式では、推薦情報に対する利用者の評価フィードバックなどから明示的に得られる嗜好情報や、Web 文書の閲覧履歴などから暗黙的に得られるものに基づいて、興味・関心の抽出や情報フィルタリングを行う。

個人の情報嗜好を満たす情報推薦を行うために、閲覧時間、マウスクリック、スクロール、テキスト選択、検索クエリおよび検索結果などの情報閲覧行動を対象とした個人プロフィールが活用されている[2, 3]。森田らの情報リマインダシステム[4]は、Web 閲覧セッションの間、印刷、コピー・ペースト等のブラウザ上の操作が記録され、それらから抽出される個人プロフィールを利用して、パーソナライズした情報提供を行う。また、平本らは、デジタル地図上の操作履歴から抽出される利用者の暗黙的意図に基づいて、適切な Web 文書を提示するシステムを提案している[5]。

しかしながら、Web 閲覧時の習慣には個人差があるため、利用者の関心を抽出する上で、影響のある Web 閲覧行動を決定することは困難である。例えば、マウスクリックは重要な Web 閲覧行動の要素であるが、利用者は特に深い意味もなく、ブラウザ上で文字列などを選択する場合がある。

表 1 典型的な Web 閲覧行動

ID	Web 閲覧行動	抽出される単語集合
I_1	Web ページを閲覧する	Web ページに出現する単語
I_2	Web ページ上でマウスクリックにより単語を選択する	選択された単語
I_3	Web ページ上で、単語をクリップボードにコピーする	クリップボードにコピーされる単語
I_4	Web ページ内で検索する	検索単語
I_5	Web ページを保存する	保存された Web ページに出現する単語
I_6	Web ページを印刷する	印刷された Web ページに出現する単語
I_7	Web をブックマークする	ブックマークされた Web ページに出現する単語
I_8	Web 検索エンジンで検索する	検索エンジンに入力された検索単語
I_9	Web 検索の結果として得られた Web ページを閲覧する	Web 検索の結果より、閲覧した Web ページに出現する単語

我々はこれまで、Web 閲覧行動、および利用者の評価フィードバックにより構築される情報嗜好シソーラスに基づいて、利用者の情報嗜好に適応する情報推薦システムを提案してきた[6]。本システムの特徴は、個々の利用者へパーソナライズした情報推薦を行うことを目的として、推薦された情報に対する利用者の継続的な評価フィードバックを利用して、利用者の情報嗜好と関連の高い Web 閲覧行動を積極的に利用することにより、利用者の情報嗜好を適切に反映した個人プロフィールを構築する点にある。

本稿では、この方式を拡張し、利用者の閲覧文書に対するブラウザ操作提示機能、および、提示されたブラウザ操作の実行による暗黙的フィードバック機能を備えた Web 情報推薦システムを示す。

2. 暗黙的フィードバック方式

図 1 は、提案する Web 情報推薦システムの基本構成を示している。提案システムは、3 つの反復的なフェーズより構成される。最初のフェーズの間、利用者の Web 閲覧行動 (表 1) が監視され、個々の閲覧行動より重要な単語群が抽出される。初期の情報嗜好シソーラスは、個々の閲覧行動により得られる単語群と単語スコアに

An Adaptive Recommender with an Implicit Feedback Method based on Web-browsing behavior learning

[†]Kosuke Takano, Department of Computer and Information Sciences, Kanagawa Institute of Technology

[‡]Kin Fun Li, Department of Electrical and Computer Engineering, University of Victoria

より構築される。第 2 フェーズにおいて、情報嗜好シソーラスの単語群と Web 文書間の類似性に基づいて、Web 文書がランキング形式で推薦される。最後の学習フェーズでは、推薦情報に対する利用者の明示的な評価フィードバックに基づいて、情報嗜好シソーラスが更新される。

本研究で示すブラウザ操作の提示機能と、提示されたブラウザ操作を対象とした暗黙的フィードバック機能は、次のステップで実現される。

Step-1: 明示的な評価フィードバックの履歴から、操作頻度の高いブラウザ操作を抽出する。

Step-2: 抽出されたブラウザ操作により得られた単語群と利用者が現在閲覧している Web 文書の単語群の相関値を算出した結果に基づいて、その Web 文書に対するブラウザ操作を推薦する。

Step-3: 利用者が Step-2 で提示されたブラウザ操作を行った場合、これを暗黙的フィードバックとみなす。この場合、提示されたブラウザ操作と現在閲覧している Web 文書の内容には、利用者の情報嗜好に関する有意な関連性が内在すると判定し、情報嗜好シソーラス上で、実行されたブラウザ操作に対応する単語群の重み値を増加させる。

Step-1~3 による利用者の暗黙的フィードバックにより、情報嗜好シソーラスが個人の情報嗜好に適應するように更新され、パーソナライズした情報推薦精度の向上が実現可能となる。

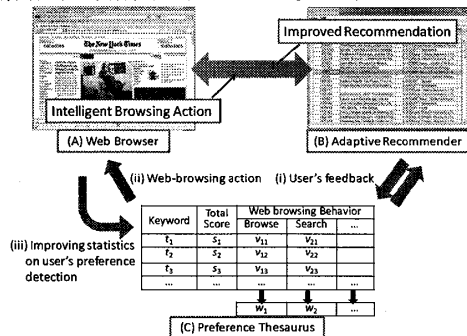


図 1 提案システムの基本構成

3. 予備実験とまとめ

プロトタイプシステムを用いて、約 18,000 件の RSS 文書を対象とした予備実験を行った。表 2, 3 は、それぞれ、明示的なものだけからなるフィードバックパターン(明示的 FB)、および、明示的なものと暗黙的なものの両方を混合したフィードバックパターン(混合 FB)を示している。また図 2 は、文書推薦結果の上位 15 件におけるジャンル分布をしめしており、フィードバック適用前の初期状態、初期状態に混合 FB を適用した場合、および、初期状態に明示的 FB を適用した場合の結果を比較している。図 2 より、混合

FB は、明示的 FB よりも、利用者の情報嗜好(テニス, ゴルフ)に適應した文書推薦を実現できていることが確認できる。また、表 2, 3 において、混合 FB における明示的フィードバック回数は、明示的 FB 中の実施回数よりも少ないため、混合 FB により、利用者の評価フィードバック実施に対するコストを軽減できることも確認できる。

今後、より詳細な評価実験を行うことにより、本方式の実現可能性、および、有効性を検証していく予定である。

表 2 明示的な評価フィードバックパターン

FB 回数	文書 ID	ジャンル
1	d_{18341}	テニス
2	d_{17771}	テニス
3	d_{18361}	ゴルフ

表 3 混合フィードバックパターン

FB 回数	文書 ID	ジャンル	FB タイプ	ブラウザ操作
1	d_{18341}	テニス	明示的	
2	d_{17771}	テニス	暗黙的	ページ閲覧
3			暗黙的	文書 d_{17771} に関する グーグル検索
4	d_{18361}	ゴルフ	明示的	

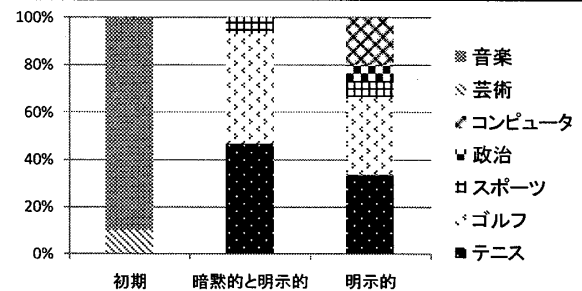


図 2 上位 15 件のジャンル分布

参考文献

- [1] J.S. Breese, D. Heckerman, and C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, pp. 43-52 (1998).
- [2] Dumais, S., Cutrell, E., Cadiz, J.J., Jancke, G., Sarin, R. and Robbins, D.C.: Stuff I've Seen: A System for Personal Information Retrieval and Re-Use, Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 72-79 (2003).
- [3] Marcialis, I. and Vita, E.D.: SEARCHY: An Agent to Personalize Search Results, Proceedings of the 3rd International Conference on Internet and Web Applications and Services (ICIW 2008), pp. 512-517 (2008).
- [4] Morita, T., Hidaka, T., Tanaka, A. and Kato, Y.: System for Reminding a User of Information Obtained Through a Web Browsing Experience, Proceedings of the 16th International World Wide Web Conference WWW2007, pp. 1327-1328 (2007).
- [5] Hiramoto, R., Sumiya, K.: Web information retrieval based on user operation on digital maps, Proceedings of the 14th annual ACM international symposium on Advances in geographic information systems, pp.99-106 (2006).
- [6] Takano, K., and Li, K. F.: An Adaptive Personalized Recommender Based on Web-Browsing Behaviour Learning, Proceedings of the 2009 IEEE International Symposium on Mining and Web (MAW09), pp.654-660 (2009).