

個人適応型情報フィルタリングシステムの適用評価について

5 Q-5

橋高 博行 鈴木 英明 柴崎 雅史 曾根岡 昭直
NTTソフトウェア研究所 サービスソフトウェア方式研究グループ

1.はじめに

現在、インターネット利用者の増加とともに、発信される情報量も増加の一途をたどっている。発信される情報量が増加したことにより、情報取得者が興味にあった情報を取得することは困難になりつつある。こうした中、WWWで情報提供を行う際、画一的な情報の提供を行うだけでなく、WWWの特性であるリアルタイム性やインタラクティブ性をいかし、情報取得者の興味に合わせて情報の提供を行うことが望まれている。

筆者らは情報取得者の興味に合った情報を紹介するシステム(InfoBroket)を開発した[1]。本稿では、本システムをニュース記事を提供するWWWサイトに適用した評価実験の結果について述べる[2]。

2.実験の概要

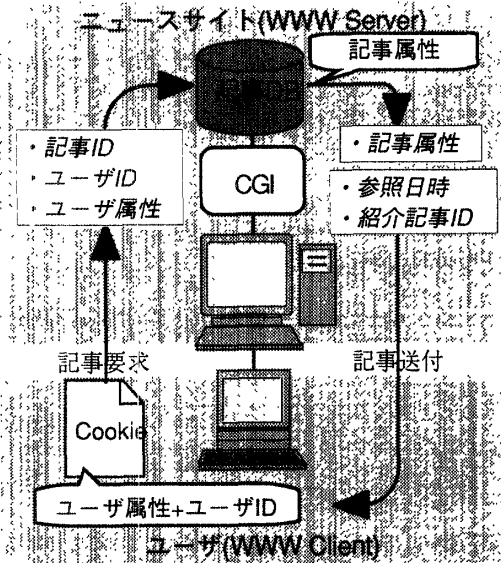


図1.評価情報の取得

今回の実験では、ユーザ興味モデルの妥当性、情報紹介方式の有効性を評価するための情報として、各ユーザごとにユーザのアクセス履歴、興味の変化履歴を収集する。具体的には、図1に示す流れで各評価情報をニュースサイト側で取得した。情報取得者(ユーザ)が情報を参照する際、参照する記事ID、ユーザ端末に保存されていたユーザIDとユーザの興味情報を数値化したユーザ属性値を取得した。次に参照する記事IDに対応する記事属性の記事DBより取得した。記事属性は記事の内容を数値化した情報

である。次に、ユーザの興味にあった記事を紹介する際、および参照記事と関連がある記事を紹介する際に、紹介する記事の記事IDと参照された日時を取得した。

ユーザが記事を参照するつど上記の情報を取得し、ユーザのアクセス履歴、興味の変化履歴をユーザごとに把握した。記事の総数は最終的に755で、総計で2348回参照された。

3.実験結果および考察

3.1記事属性値とユーザ属性値の比較

全部の記事に設定されている記事属性値の平均(平均記事属性値)と、ニュースサイトにアクセスしたユーザのユーザ属性値の平均(平均ユーザ属性値)の比較を図2に示す。本システムは参照した記事の記事属性値をユーザ属性値に加算してユーザの興味を把握する方式のため、ユーザがランダムに記事を参照したり、すべての記事を参照した場合は両者の値は一致することになる。実験結果では大きく異なっており、これは、ユーザは自分の興味にあった記事を選択的に参照すること、また、複数回数同じ記事を参照することがあることを示している。

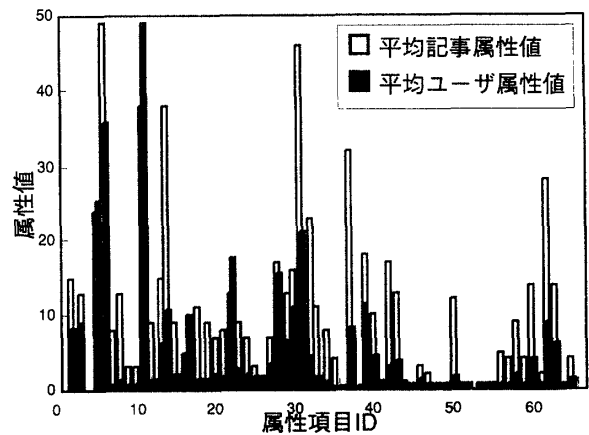


図2.平均記事属性値と平均ユーザ属性値の比較

3.2ユーザ属性値の比較

ユーザごとのユーザ属性値の比較を図3に示す。両者の値の分布は大きく異なっている。このことから、属性項目64項目に対して興味の強さに応じて値をつけた今回の単純なモデルでも、ユーザごとの興味の違いを十分表現できることがわかる。

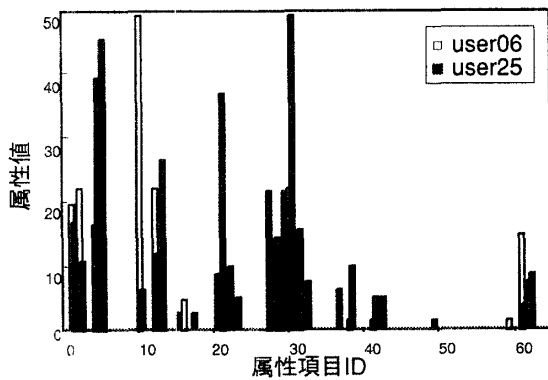


図3. ユーザ属性値の比較

3.2 ユーザ属性値の時間変化

ユーザ属性値の時間変化を図4に示す。今回の実験のモデルでは、参照した記事に設定されている属性値をユーザ属性値に加算することでユーザの興味を増加を表現している。また同時に、属性値を時間の経過に対して一定の割合で減少させることでユーザの興味の減衰を表現している。ユーザ属性値が時間の経過とともに変動しており、ユーザの興味は時間の経過とともに変化する事がわかる。一方で、例えば属性項目IDが10の属性値は10日目から20日目に高い値となり、20日目から30日目にかけて減衰するが30日目から再び高い値となっている。これはユーザの興味は細かい変動はあるものの、長期的に30日程度は持続することを示している。

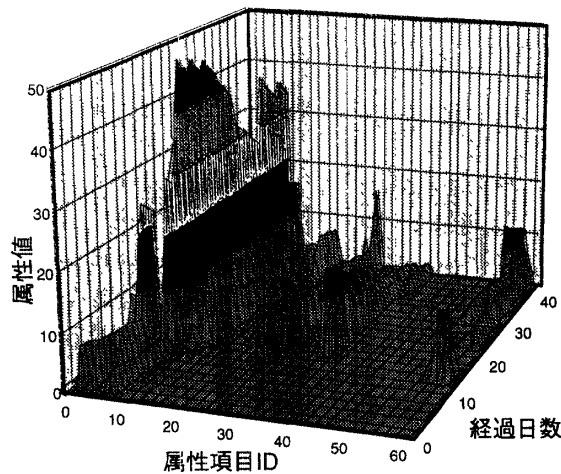


図4. ユーザ属性値時間変化

3.3 ヒット率

紹介した記事のヒット率を図5に示す。ここでのヒット率は記事を紹介した回数に対して、何回紹介した記事を参照したかを示す値であり、記事紹介の有効性を判断する指針となる値である。また、記事の紹介方法は二種類あり、一つが興味にあった記事を紹介する推薦記事紹介で、もう一つが参照した記事に関連する記事を紹介する関連記事紹介である。

図5より、推薦記事より関連記事の方がヒット率

が高いことがわかる。これは、推薦記事紹介で紹介される記事は最近の既読記事である確率が高いこと、推薦記事と新着記事を同じ画面で紹介したため鮮度のある新着記事の方が参照される確率が高いこと、今参照している記事に対して関連記事を紹介する方がよりその瞬間のユーザの興味にあった記事が紹介される確率が高いためだと考えられる。

また、全体のヒット率は18%ほどであったが、比較の対象となる実験を行っていないため、値の大小に対する考察ができなかった。今後対象実験として、紹介記事をランダムに選択した場合のヒット率を求める実験を行う必要がある。

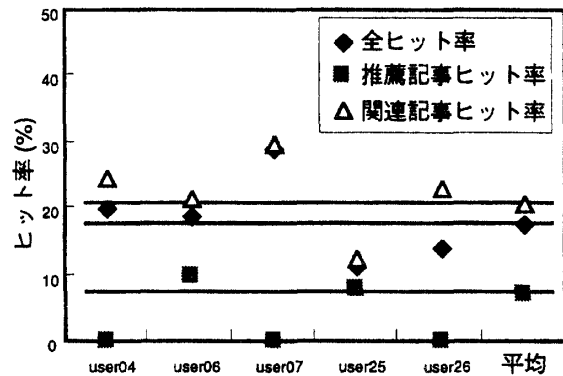


図5. 紹介記事ヒット率

4. まとめ

本稿では、InfoBroketシステムを用いたユーザの興味に応じた記事紹介サービスをニュースサイトに適用した評価実験の結果について述べた。結果として以下の知見を得た。

- ・ユーザは興味に応じ選択的に記事を参照している
- ・ユーザの興味は時間的に変化するが、概して長期的に持続する
- ・本システムで用いた単純なモデルでユーザの興味を表現することが可能
- ・参照した記事に関連する関連記事の紹介の方が興味に応じた推薦記事の紹介よりも有効である
- ・新着記事と同時に推薦記事を紹介する見せ方は効果が薄く、効果的な見せ方を検討する必要がある

今後はユーザの興味の把握、紹介記事の選択等の計算に用いる各種パラメータを変化させ実験を行い、最適となる値の検証をおこなう。また、今回のニュースサイトではニュース記事の鮮度が大きな価値をもつため、必ずしも興味にあった記事を参照するとは限らない問題がある。本システムをディレクトリサービスに適用した場合でも評価を行う。

参考文献

[1] 橘高他：インターネット上でのインタラクティブマーケティング、第54回情処全国大会,1997
 [2] <http://smartpage.watch.impress.co.jp>