

書籍販売サイト bk1 における レコメンドエンジン AwarenessNet の導入効果

市川 裕介[†] 田中 明通[†] 川村 亨[†] 中村 俊郎[†]

[†] 日本電信電話株式会社 情報流通プラットフォーム研究所
〒180-8585 東京都武蔵野市緑町 3-9-11

E-mail: [†] {ichikawa.yusuke, tanaka.akimichi, toru.kawamura, nakamura.toshirou}@lab.ntt.co.jp

あらまし 我々は独自開発した自動レコメンド手法によるレコメンドエンジン”AwarenessNet”をオンラインショッピングサイト bk1 に導入し、レコメンドサービスについての実証実験を行った。従来、レコメンドエンジンのアルゴリズムの有効性はシミュレーションによる予測精度の評価やサンプルユーザへのアンケート調査によって評価される。しかし、その評価結果から実際のオンラインショッピングサイトへ導入した際の導入効果を見極めることは困難である。本報告では、実運用サイトでのレコメンドシステムの実証実験を行うことにより、従来の評価手法では評価することができなかった、レコメンドエンジンのビジネス面での導入効果についてその事例を示す。

キーワード レコメンドシステム、協調フィルタリング、履歴分析

The Effect of introduction of Recommender Engine AwarenessNet on online book store bk1

Yusuke ICHIKAWA,[†] Akimichi TANAKA,[†] Toru KAWAMURA[†] and Toshirou NAKAMURA[†]

[†] NTT Information Sharing Platform Laboratories
3-9-11, Midori-cho, Musashino-shi, Tokyo, 180-8585 Japan

E-mail: [†] {ichikawa.yusuke, tanaka.akimichi, toru.kawamura, nakamura.toshirou}@lab.ntt.co.jp

Abstract We introduced the recommender system "AwarenessNet" by the automatic recommendation technique for developing originally into online shopping site bk1, and did the demonstration of the recommendation service. The effectiveness of the recommender system is evaluated by the evaluation of the predictive accuracy by the simulation and the questionnaire survey to the sample user so far. However, it is difficult to ascertain the introduction effect when introducing it from the evaluation result to an actual online shopping site. This research shows the case about the introduction effect on the business side of the recommender system where the recognition cannot have been done by doing the demonstration of the recommender system in the real operation site in a past evaluation technique.

Keyword Recommender System, Collaborative Filtering, History Analysis

1. はじめに

EC サイト等において、顧客個々の好みを分析し、顧客毎に興味のありそうな商品や情報を抽出しておすすめ(表示)することを One to One レコメンドと呼ぶ。One to One レコメンドは、ユーザにとっては好みの商品へ即座にアクセスできるように利便性が向上し、サイトの運営者にとっては商品購買率や生涯価値が向上することが期待でき、ユーザ・運営者双方にとってメリットのあるサービスとして注目されてきた[1]。

One to One レコメンドは、1)まず会員登録時のプロフィール登録、アンケート回答、購買・サービス利用履歴、Web アクセス履歴、等の顧客情報を収集し、2)

収集した顧客情報を分析、そして3)分析結果に基づき個人別推薦アイテムを提示、の3段階で実現される。この3段階の中核となるのが、2)で分析を行う「レコメンドエンジン」である。

レコメンドエンジンは、ネットニュースを対象としたレコメンドシステム GroupLens を代表に、1990 年始め頃から多くの方式が発表されている[4]。しかし、その評価としては、シミュレーションによる予測精度の比較が殆どであり、実際に EC サイト等へ適用した際にどれだけクリック率が得られ、どれだけ売り上げが上がるのかといった、サイトの運営者がレコメンドエンジンの採用を検討するために必要な評価が行われ

表 2-1 レコメンドシステムの分類

方式	利用情報	方法	メリット	デメリット
チェックボックス方式	ユーザからの事前登録情報	予めユーザが登録した興味カテゴリーに基づき、一致するカテゴリーのアイテムをレコメンド。	設定が単純で容易に使用できる。	ユーザにはアンケート登録の手間がかかる。
ルールベース方式	サイトの運営者が設定したルール	サイトの運営者がユーザのプロファイル毎に登録したルールに基づきアイテムをレコメンド。	サイト運営者の意思を反映し易い。	アイテム数が増加したり、プロフィールの細分化が進むと、設定が複雑になり困難。
自動レコメンド方式	ユーザの行動履歴	システムがユーザの行動履歴からユーザの嗜好を自動的に計算しレコメンドアイテムを抽出。	自動的に嗜好の推測を行うので、ユーザ・サイト運営者双方に負担がかからない。	レコメンドを得るにはある程度以上の行動履歴が必要。

ていない。

そこで我々は、ユーザからの評価入力等を必要とせず、アクセス履歴のみを用いて分析を行うことを特徴とした、独自開発のレコメンドエンジン”AwarenessNet”を、実際に EC サイトへ導入し、実証実験を行うことでクリック率の変化、売り上げに与える影響等を実測した。本報告では、AwarenessNet のアルゴリズムについて解説し、実証実験で得られた、その導入効果について報告する。

2. One to One レコメンドの実現方式

本節ではまずレコメンドサービスを実現する手法について解説を行い、続いて AwarenessNet のポジションを説明する。

2.1. One to One レコメンドを実現する 3 方式

現在、EC サイトにおいて利用されている方式は、表 2-1 に示す 3 方式に大別することができる[1][2]。本報告では、このうち「自動レコメンド方式」を対象とする。自動レコメンド方式は、他の 2 方式が手動で推薦アイテムを決定しているのに対して、自動的に推薦アイテムの決定を行う点で大きく異なる方式である。この為、狭義のレコメンドシステムと定義する文献もみられる[3]。

自動レコメンド方式の実現方法については次項で詳しく述べる。本方式は手動でレコメンド・アイテムを決める他の 2 方式に比べて以下の点で他の方式より優れており、現在も研究開発の中心的な話題となっている。

- マーケティング専門家によるルール設定等の準備作業が不要(人的コストの削減)
- ユーザ数やアイテム数の増加、ユーザの嗜好の変化等に動的に追従できる(高拡張性)
- 意外性のあるアイテムのレコメンドができる(潜在的な興味の掘り起こし)

2.2. 自動レコメンド方式を実現するアルゴリズム

自動レコメンド方式を実現するアルゴリズムを、利

用する情報に着目して分類すると、以下の 2 方式に大別できる[3]。

1. コンテンツベースフィルタリング
2. 協調フィルタリング

「コンテンツベースフィルタリング」は、アイテムのコンテンツ情報の類似度に基づき、ユーザの評価したアイテムと類似したアイテムを抽出する方式である。アイテム間の類似度の算出に履歴を使用していないため、履歴数が少なくてもアイテムの抽出が可能である反面、現状ではテキストマイニング技術が中心であるため、アイテム個々に対するテキスト情報が必須となる。また、評価したアイテムと類似したアイテムばかりをレコメンドする為、意外性のある情報のレコメンドは難しい。

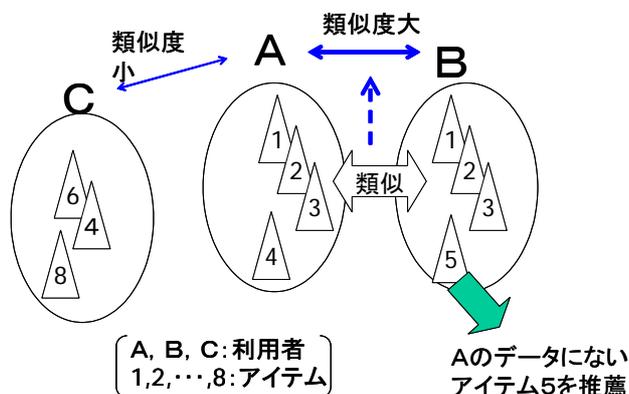


図 2-1 協調フィルタリングの基本動作

一方、「協調フィルタリング」は、ユーザとその他のユーザのアイテムに対する興味の類似度に基づき、レコメンド・アイテムを抽出する方式である。協調フィルタリングの簡単な概念図を図 2-1 に示す[5]。ユーザ A が興味を持つアイテム集合のなかからもっとも共通するアイテム集合を持つユーザ B を見つけ、ユーザ B が興味を持っておりかつユーザ A の興味未定なもの(ここではアイテム 5)をユーザ A にレコメンドする。

協調フィルタリングは以下の点で他の方式より優

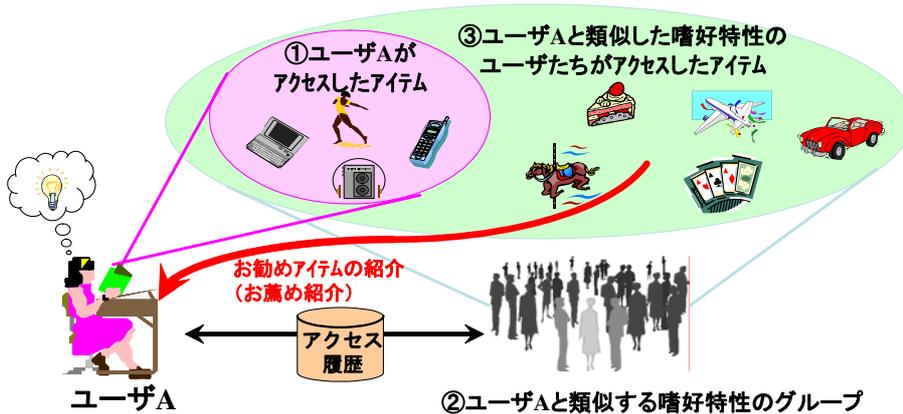


図 2-2 AwarenessNet の協調フィルタリングの動作概要

れている。

- コンテンツ情報やルール、アンケート入力情報を一切使用しない為、システム運営者、ユーザ双方に情報投入の負担が少ない
 - 属性やキーワードに縛られない為、予想だにしない意外性のあるアイテムを Recommend できる
- 協調フィルタリングを用いた初期のシステムでは Tapestry や GroupLens が代表的である[4]。ビジネスフェーズにおいては、GroupLens の商用版である NetPerceptions 等を代表に商用化が行われており、NetPerceptions が Amazon.com で採用されることで一躍有名になった。日本においては、1999 年に NetPerceptions が日本法人を立ち上げたほか、我々が開発した AwarenessNet が国産では初の自動 Recommend エンジンとして 2000 年に販売開始した[6]。

協調フィルタリングにおいて主要な技術は、ユーザ間のデータの類似度判定であり、さまざまな手法が提案されているが、類似度計算の元となる各ユーザの興味データにはユーザから入力された評価値を利用していることが多い。例えば、対象となるアイテムに対して、「非常に好き」から「非常に嫌い」まで 5~7 段階でユーザが入力するものである。しかしながら、評価入力には利用者の手間がかかり、データが得られにくいという問題(評価値疎ら問題)がある[4]。

2.3. 協調フィルタリングの問題点を解決した AwarenessNet

AwarenessNet では協調フィルタリングを用いる際に必要なユーザの興味データをアクセス履歴のみで行うことを前提とした。アクセス履歴はたとえば Web ページの参照履歴であるので「あり」と「なし」の 2 値でしかないが、そのかわりユーザからの入力の手間がないために、評価入力よりも多くのデータを得ることができ、評価値疎ら問題を解決することができる。

具体的な処理は図 2-2 のようになっている。

1. ユーザ A がアクセスしているアイテムを抽出

2. 抽出したアイテムにアクセスしている他のユーザ、すなわちユーザ A とよく似たアクセス傾向のユーザ達を抽出
3. 抽出したユーザ達がアクセスしていて、かつユーザ A がアクセスしていないアイテムをユーザ A に Recommend
アクセス傾向の類似度は、アクセス履歴から抽出した参照回数、参照時間、最新アクセス日時、重なり等の条件を組み合わせ

せて算出している[5]。

アクセス履歴のみで協調フィルタリングを行う際の問題に、人気の高いアイテムが推薦されやすいことや、誤ってたまたまアクセスされてしまったアイテムが推薦されてしまう可能性があることが挙げられる。特に、人気の高いアイテムは多くの人がアクセス(評価)しているため、評価値に基づく協調フィルタリングにおいても推薦の際に優先されやすい。しかしながら、人気の高いものを推薦することは特別な技術を用いずとも容易に行う事ができるので、わざわざ協調フィルタリングをして推薦する必要がない。

AwarenessNet では我々が独自に開発した「選考行動フィルタリング」というノイズカット技術を用いて推薦に相応しくないアイテムを取り除いている[6]。アクセス履歴等のユーザの意図が反映されている行動履歴を集計し、横軸にアクセス人数、縦軸をそのアクセス人数を持つアイテム数にとり両対数グラフにプロットすると、ある部分でべき乗分布に従うことがわかっている[6]。(図 2-3)

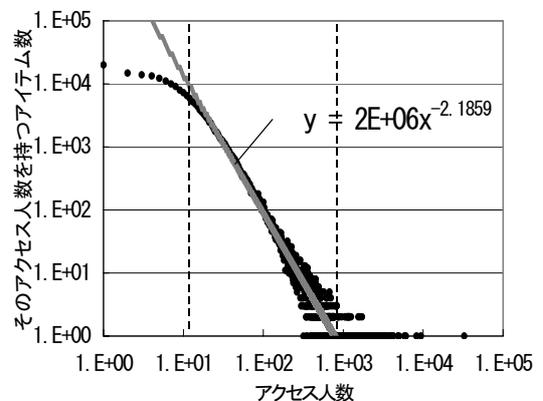


図 2-3 アクセス履歴におけるべき乗分布

ここで、べき乗分布を表す直線の部分から外れてい

る、アクセス人数が極端に多い右側の部分と、アクセス人数が少なくべき乗分布の直線から外れて湾曲している左側部分は、利用者の意図的な行為とは別の要因が働いていると考えられる。よって、これらを取り除き、それぞれのユーザが意図を持ってアクセスした領域だけを用いてレコメンドを行うことで、推薦に相応しくないアイテムの除去を実現しているのが選考行動フィルタリングである。このようにして、AwarenessNetでは評価入力に頼らずアクセス履歴のみでレコメンドを実現することで、評価値疎ら問題を解決している。

3. 既存研究における効果検証方法と課題

上で述べたように、レコメンドを実現する様々な方式が提案されているが、その方式の評価として最も多く用いられている手法がBreeseらのAllbut1[7]である。この手法は、ユーザのアイテムに対する評価の1つをマスクして、テストデータとする。それ以外の評価を学習データとして、マスクしたアイテムに対するユーザの評価を予測し、Mean Absolute Error(MAE)¹を利用して予測と実際の評価の誤差から予測精度を評価するものである。評価に使用するデータに、EachMovie collaborative filtering data set[8]を用いて、MAEを測定する方法は協調フィルタリングの研究において、ベンチマーク的に用いられている手法である。

本手法は異なるアルゴリズム同士を同じ条件のもとで比較を行う事が可能であるが、比較の尺度についてはユーザとシステムの評価の一致度(正確性)を測定しているに過ぎず、実運用面で評価されるべき以下の点での評価にはなっていない。

1. 潜在的な興味の掘り起こしの評価
2. サイトのニーズとの合致度

前者は、2-2 節において協調フィルタリングによるレコメンドの利点はレコメンドされるまでユーザが意識していなかったアイテムの「意外な発見」や「思わぬ買い物」であると説明したが、MAEによる評価はその点についてまったく評価できていないことを示している。

また、後者については確率的には人気の高いアイテム(TOP10 等)を紹介した方が正確性はあがる可能性が高いが、実運用ではレコメンドには人気の高いアイテムよりは、通常ルートでは人目につくことのない人気のないアイテムの販売促進を期待されている場合が多い(レコメンド機能を使わなくても人気の高いアイ

¹ ユーザによる評価投票値と予測値間の平均絶対値エラー。これは、統計的精度測定基準(Statistical accuracy metrics)の一種で、テストデータセットにおける数値的推薦スコアと、「アイテム-ユーザ」評価マトリクスにおける実際のユーザ評価値とを比較することによってシステムの正確さを評価する。

テムの販売促進をする方法は既にある場合が多い)という点で実情と剥離した評価ということである。

以上のように、MAEを用いた評価は、実サイトに適用した際の効果の指標としては大きな問題を含んでおり、実サイトにおける効果を実証している事例は見られないのが現状である。しかしながら、ECサイトへの普及を図る上で求められているのは、導入した際にどれだけ売り上げに対する効果を上げられるか、またそれが運用者が期待する効果であるのかということが計れるような検証データである。

そこで、我々は実サイトに対して AwarenessNet を適用・評価することで、既存の評価手法では測定できない実践的な観点での有効性を表す評価値の測定を行った。

4. 実証実験概要

株式会社ブックワン、NTT 東日本、NTT の三社共同実験として、ブックワンが運営するオンラインブックストア「bk1(<http://www.bk1.co.jp/>)」において AwarenessNet を利用したレコメンドサービスを提供した。実験期間は2003年5月～2004年3月の10ヶ月間。

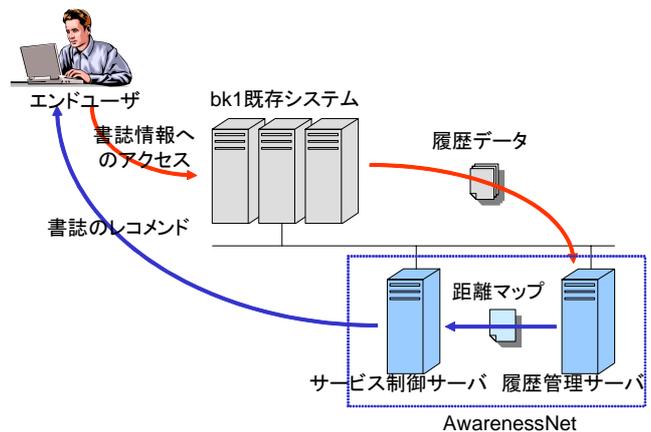


図 4-1 bk1 実験システムの構成概要

4.1. 実験システム

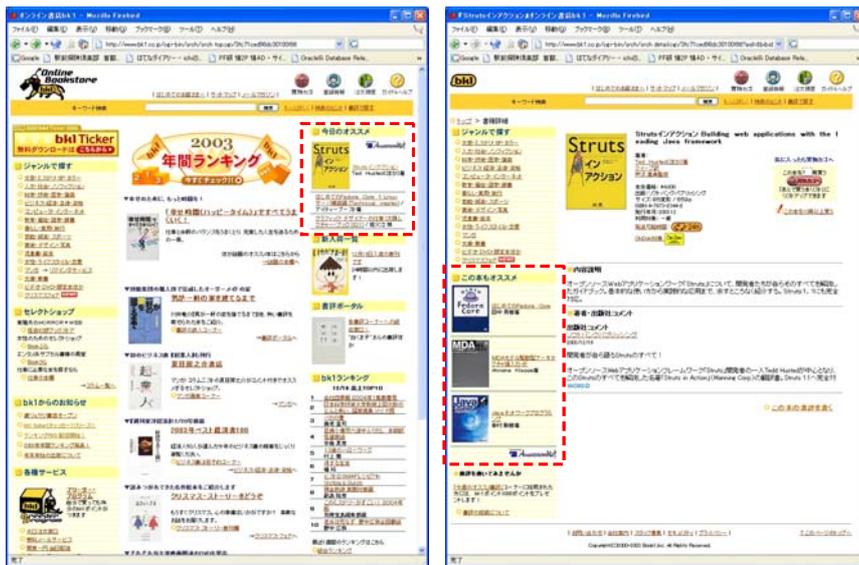
実験システムの構成概要を図 4-1 に示す。

AwarenessNet は履歴管理サーバとサービス制御サーバの2台構成となっており、既存の bk1 システムに追加する形で構築している。

1) 履歴管理サーバ

履歴管理サーバは1日1回 bk1 既存システムから履歴データを取り込み、履歴の事前加工を行い、ユーザとアイテムの関係を 2.3 節で解説している手法により、類似度を距離として数値化した「距離マップ」に変換する。変換された「距離マップ」は後述のサービス制御サーバに転送する。

取り込む履歴データは以下の3種類を対象として Web サーバのアクセス履歴から抽出、事前加工を行う日から過去20日間分を分析の対象として利用する。過



おすすめ紹介サービス

関連紹介サービス

図4-2 bk1で提供したレコメンドサービス

去 20 日以前の履歴は全て破棄している。

1. 書誌詳細ページへのアクセス
2. 書評の投稿
3. バスケットへの投入

2) サービス制御サーバ

サービス制御サーバは、ユーザからのアクセスをトリガにして、転送された距離マップを用いて図 2-3 の分析をリアルタイムに実行する。具体的には、Web アプリケーションサーバから送られてくるレコメンド要求に対して、レコメンド・アイテムを返し、その結果が Web ページに反映されてユーザに表示される。

4.2. ユーザへの提供サービス

bk1 では、従来からのサービスとして、キーワードによる書誌検索、ジャンルによる書誌分類、売り上げ TOP10 ランキングによる紹介、書評による紹介、新刊入荷書誌紹介、話題本の紹介等の様々なサービスによって書誌詳細ページへのアクセス、バスケットへの書誌投入を可能にしている。本実験において AwarenessNet を導入することで、以下の 2 つのレコメンドサービスがこれに加わることになる。(図 4-2)

1) おすすめ紹介サービス

bk1 サイトのトップページ表示時に、図 2-2 に示す処理によって個々のユーザの嗜好に合わせた書誌を紹介する。紹介する件数は 3 件。

尚、アクセスしてきたユーザに過去の履歴が存在しない場合は、売り

上げ TOP11~20 位の書誌の中からランダムに選択したアイテム 3 件を返すようにした。

2) 関連紹介サービス

書誌の詳細ページを表示した際に、ユーザの履歴に基づき該当書誌と関連の高い書誌を紹介する。関連の高い書誌の抽出は、図 2-2 に示す処理において「①ユーザ A がアクセスしたアイテム」の代わりに該当書誌を用いることで行う。こちらも紹介する件数は 3 件。

これらのサービスにおいて、ユーザ毎に合わせたレコメンドを提供する為には、個々のユーザを識別する必要がある。本実験においては、bk1 サイトで以前から使用していた Cookie を利用してユーザの識別を行った。

5. 実験結果

5.1. 基本情報

実験期間の内、履歴からクローラー等のアクセスを除去する仕組みを導入した 2003/6/26 以降を測定開始日とし、データ収集の終了する 2004/2/29 までの期間の履歴について分析を行った。

5.2. レコメンドサービスのクリック率

図 5-1 に書誌詳細ページへのアクセス数の推移とレコメンドした結果クリックされた率の推移を示す。各推移は前 7 日分の平均値による 7 区間移動平均で示している。

尚、図 5-1 のグラフ中の灰色で四角に囲っている期間はシステムトラブルや外部要因等によりデータが取

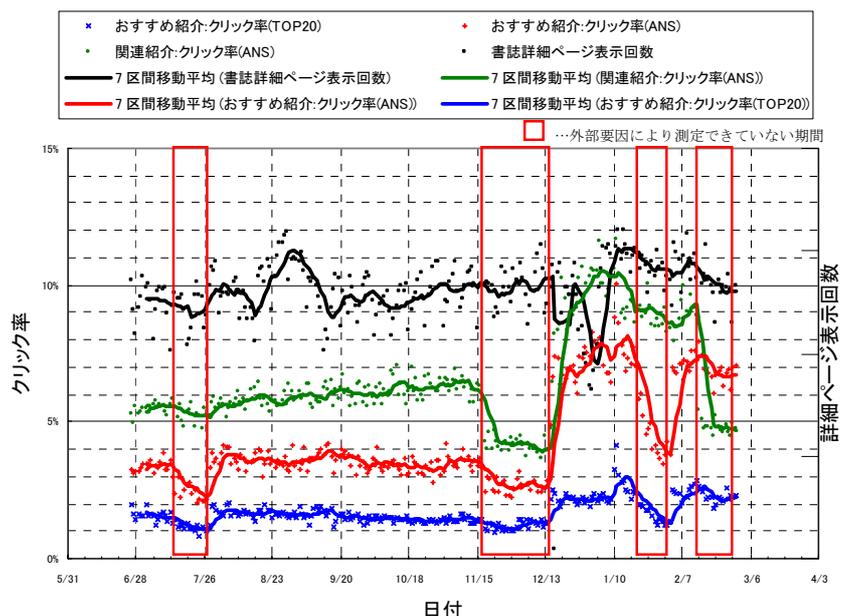


図5-1 レコメンドのクリック率の推移

集できていない期間を示している。また、2003/12/15 におすすめ紹介の表示箇所を右下の目立たない箇所から右上へ移動、同時に文字だけの表示から書影(表紙写真)を掲示するように画面レイアウトの変更を行っている。図 5-1 より以下のことがわかる。

- 表示方法の変更により、クリック率に大きな向上がみられる。
- 全体のアクセス数の変動には影響を受けず、関連紹介に関してはクリック率が増加傾向にある。
- 個人毎にパーソナライズされたおすすめ紹介と TOP20 から抽出したおすすめ紹介では、パーソナライズされたおすすめ紹介の方が大幅にクリック率が高い(ただし、TOP20 の紹介は初めて bk1 にアクセスしたユーザの為、もともとアクティブでないユーザに偏っている可能性が高い)。

画面レイアウト変更後の 2004 年 1 月の 1 ヶ月間のクリック率平均を表 5-1 に示す。

表 5-1 1 月のレコメンドのクリック率

項目	クリック率
おすすめ紹介クリック率 (ANS レコメンド)	6.5%
おすすめ紹介クリック率 (TOP20 からランダム抽出)	2.2%
関連紹介クリック率 (ANS による相関抽出)	9.6%

5.3. レコメンドサービスからの購買率

表 5-2 に全体の購買冊数に占める、レコメンドサービスによる購買冊数の割合を示す。全体購買冊数に占めるレコメンドサービス経由の購買冊数は 3.1% を占めている。尚、レコメンドサービス以外からの購買冊数の増加率が 0.03% であったのに対し、レコメンドからの購買冊数の増加率が 0.36% あり、12 倍の増加率があった。すなわち、レコメンドサービスからの購買冊数増加率は自然増以上の増加率であり、レコメンドサービスの導入が購買冊数増加に効果があったと言える。全体に占めるレコメンドサービスからの購買冊数の占める割合も、今後増加していく可能性があるといえる。

表 5-2 1 月の購買冊数の比率

項目	割合
レコメンドサービスから	3.1%
従来サービスから	96.9%

5.4. 潜在的なニーズの掘り起こし

自動レコメンドサービスにおいて、期待される効果の一つに潜在的なニーズの掘り起こしがある。図 5-3 は、横軸に各書誌の販売冊数を、縦軸に書誌の販売冊数毎の販売冊数の全体合計に占める割合を示し、全体の販売数とレコメンドからの販売数の割合を比較したものである。

図 5-2 から、販売冊数の少ない書誌における販売の割合がレコメンドによる販売の方が多くなる事がある。このことから、レコメンドを行うことで、よりニッチな書誌へユーザを導くことが出来ているといえる。つまり、AwarenessNet のレコメンドは潜在的なニーズの掘り起こしが出来ている可能性が高いといえる。

しかし、販売数の少ない書誌の販売数が必ずしも潜在的なニーズを表しているとはいえない。今後、より正確な潜在的ニーズの掘り起こし効果を測定する手法を検討していく必要がある。

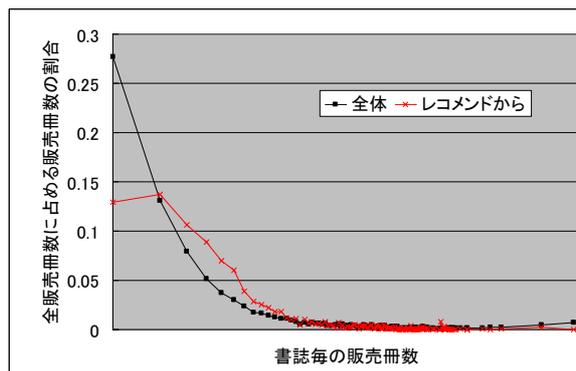


図 5-2 書誌販売冊数毎の全販売冊数に占める販売冊数の割合の比較(2004 年 1 月分)

6. まとめと今後の課題

本報告では、実証実験による実測で AwarenessNet によるレコメンドサービスの導入効果があること示した。

本報告で示した導入効果は、書籍販売という特定の分野の導入効果を示したに過ぎない。今後は、以下の点について更に検証を進める予定である。

- 意外性のあるレコメンドの評価手法の構築
- 他の分野での導入効果の測定

文 献

- [1] 新井、他、"ユーザーの好きと嫌いを聞き分けるパーソナライゼーション技術"、日経エレクトロニクス、pp.127-138、2003/2。
- [2] 寺野、"情報推薦システム"、e ビジネスの理論と応用、第 3 章、pp.59-87、電気大出版局、2003/1。
- [3] 宇田、他、"ユーザ投票と情報アイテム間類似度を併用した情報推薦システム"、情処研究報告、2004-DBS-132、pp.105-112、2004/1。
- [4] Resnick, P. and Varian, R., "Recommender System", in Communications of the ACM, Vol.40, No.3, pp.56-89, 1997。
- [5] 本橋、他、"アクセス履歴から利用者の嗜好を分析推薦するアウェアネスネット"、情処第 62 回全国大会、5C-2、2001 年前期。
- [6] 市川、他、"Web 検索サーバのログに基づいたアクセス傾向の分析"、信学技報、IN99-65、pp.13-18、1999/10。
- [7] Breese, J. S., Heckerman, D. and Kadie, C., "Empirical analysis of predictive algorithms for collaborative filtering", In Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence, pp.43-52, 1998/7。
- [8] <http://www.research.digital.com/SRC/eachmovie/>