

福島 俊一 有吉 勇介  
NEC ヒューマンメディア研究所

## 1. はじめに

インターネットを活用した、いわゆるネットビジネスが急成長している。ネットビジネスに特徴的なのは、パーソナライゼーションやワントゥワンマーケティングと呼ばれる、各ユーザにカスタマイズしたサービスを提供する手法[1]である。特に、ショッピングサイトでは、各ユーザの嗜好に合わせてお薦め商品を提示するレコメンデーションサービスが注目されている[2]。

レコメンデーションサービスは、これまで主に、ソーシャル情報フィルタリング(協調フィルタリング)の応用分野として研究されてきた[3]。GroupLens[4]、Ringo[5]などソーシャル情報フィルタリングの代表的な研究プロジェクトの成果が、インターネット上のレコメンデーションサービスに適用され、実用化されている。

このような技術オリエンテッドな研究の流れに対して、本稿では、サービスオリエンテッドな視点からレコメンデーションの機能や手法、および、体系的な要件を改めて整理する。併せて、それらを具体化したシステムとして、筆者らが開発したレコメンデーションエンジン KnowledgeRadar/R を紹介する。

## 2. レコメンデーション機能

ソーシャル情報フィルタリングの手法によるレコメンデーションの直接的な形態は、ユーザ A と嗜好の類似した別ユーザ B のお気に入り、ユーザ A にも薦めるというものである。これは下記のパーソナル推薦機能に該当するが、パーソナル推薦機能は他の手法でも実現し得る。また、レコメンデーションの利用場面から考えると、パーソナル推薦以外にも様々なレコメンデーション機能が望まれる。このような視点からレコメンデーションの機能と手法を以下のように整理した。KnowledgeRadar/R では、これらすべての機能を提供する。実現手法が複数通りあるものは、それらの手法を組み合わせた、切り換えたりする形で機能を提供する。

### (1) パーソナル推薦機能

あるユーザに対して、そのユーザの嗜好に合った商品群を薦める。その手法には、ソーシャル情報フィ

ルタリング方式(SIF)、内容ベースフィルタリング方式(CBF)、ルールベース推薦方式(RBR)がある。SIF では、前述のように嗜好の類似する他ユーザのお気に入り、ユーザ A にも薦める。CBF では、各ユーザの嗜好を表現するキーワードベクトルと、商品内容を表すテキストとをマッチングする。RBR では、各ユーザの属性(性別・年齢など)と各商品あるいは商品属性(作者・メーカーなど)を対応付けるルールを作成する。さらに、各ユーザの各商品に対する評価(気に入ったか、購入したか、などを蓄積しておくことで、SIF における嗜好(評価傾向)の類似するユーザの検出、CBF におけるユーザ嗜好を表すキーワードベクトルの作成、RBR における対応付けルールの作成などを自動化できる。

### (2) 類似商品検索機能

ある商品に対して、それと類似する商品群を検索し、お薦め商品として提示する。(1)と同じ3通りの手法があり、SIF 的視点では、その商品を好む人たちが同じように好む商品群を薦める。CBF 的視点では、商品内容を表すテキストが類似する商品群を薦める。RBR 的視点では、商品属性に一定の共通性が見られる商品群を薦める。

### (3) 類似ユーザ検索機能

あるユーザと嗜好の類似するであろうユーザ群を検索する。これはお薦め商品提示ではなく、コミュニティ形成を促進するなどレコメンデーションサービスの補助的機能である。(1)や(2)と同じ3通りの手法があり、SIF 的視点では、商品評価の傾向が近いユーザ群を見つける。CBF 的視点では、ユーザ嗜好を表すキーワードベクトルが類似するユーザ群を見つける。RBR 的視点では、ユーザ属性に一定の共通性が見られるユーザ群を見つける。

### (4) ランキング機能

多くのユーザから人気のある商品群を薦める。各商品に対するユーザ評価を前述の通り蓄積しておくことで、各商品の人気の度合いを計算することができる。

### (5) コメント機能

各商品に対するコメント(推薦文など)をユーザが書き込むことができ、それをユーザ間で共有できる。また、あるユーザから別の特定ユーザに直接、お薦め商品やコメントを通知する。

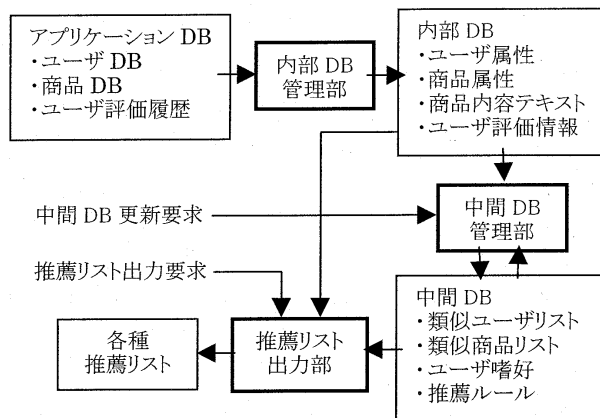


図1 レコメンデーションエンジンの構成

### 3. システム構成

レコメンデーションサービスを構築する上でのシステム的な要件として、以下のようなものが考えられる。これらの要件を満たすために、KnowledgeRadar/R は図1のような分散並列サーバの構成をとった。

#### (1)スケーラビリティ

ショッピングサイトへの適用を考えると、ユーザ数・商品数の各々が数十万～百万件という大規模データを扱える必要がある。そのため、差分計算のみで済ませられるように中間DBを設け、かつ、単純にはユーザ数×商品数の空間を必要とするメモリ量もスケーラビリティを考慮して削減した。また、図1の内部DB・中間DBなどは分散並列対応のDBMS上に構築した。さらに、商品ジャンルなどによってDBを分割することも可能とした。

#### (2)ノンストップ推薦

ショッピングサイトではユーザや商品データの追加・更新が頻繁に発生する。レコメンデーションサービスを停止せずに、これらのDB更新に対応できるのが望ましい。さらに、複数ユーザからの同時アクセス時のレスポンス低下も抑えたい。そこで、全体処理を内部DB管理部、中間DB管理部、推薦リスト出力部に分割し、これらを別プロセスとして動作させることで、並行処理を可能にした。

また、既構築のショッピングサイトに簡単にレコメンデーションサービスを追加できるインタフェースも望まれる。図1の内部DB管理部は、アプリケーション側で管理しているDB(ユーザDB・商品DBなど)との接続も受け持つ。

#### (3)リアルタイム推薦

ユーザによる新たな商品評価が得られたら、それを反映して再計算した推薦リストを、速やかにユーザに提示できるのが望ましい。これには中間DBの更新

が必要になるが、中間DB全体の更新は重たい処理である。そこで、推薦リストのリアルタイム更新が求められるケースでは、中間DBの部分的な更新による近似出力を可能とした。

### 4. 応用例

現状のKnowledgeRadar/Rは、SIFとCBFによるレコメンデーション機能が動作している。SIFおよびCBFは論文[6][7]のアルゴリズムをもとに実装した。RBRは未実装であるが、既に商品レコメンデーションサービスなどでの実用化が進んでいる。また、SIFとCBFの両手法を組み合わせていることから、商品レコメンデーションに限らず、ユーザの関心領域に合わせた文書配信にも適用可能で、KnowledgeRadar/Rを用いた論文クリッピングサービス[8]も構築し、社内運用も実施している。

### 5. おわりに

レコメンデーションサービスに求められる機能とシステム的な要件を整理し、KnowledgeRadar/Rに実装した。KnowledgeRadar/Rは、パーソナル推薦機能、類似商品検索機能、類似ユーザ検索機能、ランキング機能、コメント機能などの多彩なレコメンデーション機能を提供し、分散並列サーバ構成によってスケーラビリティやノンストップ推薦・リアルタイム推薦も実現とした。また、商品レコメンデーションサービス、論文クリッピングサービスの構築によって実用性を確認した。

### 参考文献

- [1]神場、パーソナライゼーション技術、情報学シンポジウム、1999年。
- [2]米国Webの新潮流:人気を集める「レコメンド・サービス」、日経マルチメディア、No.34、1998年。
- [3]P.Resnick, et al., Recommender systems, CACM, Vol.40, No.3, 1997.
- [4]J.A.Konstan, et al., GroupLens: applying collaborative filtering to Usenet News, CACM, Vol.40, No.3, 1997.
- [5]U.Shardanand, et al., Social information filtering: algorithm for automating "word of mouth", Proc. of CHI'95, 1995.
- [6]有吉、組合せ情報フィルタ方式の信頼度を用いた精度改良、情処研報、FI-53-4、1999年。
- [7]有吉、利用者毎に評価尺度変換を行うソーシャル情報フィルタ方式、情処58全大、5R-8、1999年。
- [8]有吉、組合せフィルタリング方式を用いた情報推薦システム、情処60全大、5U-5、2000年。