

利用者の対話的な操作を利用した情報共有方式の検討

4 Y-6

鷲崎 誠司 村本 達也

NTT 情報通信研究所

E-mail : {suzaki, muramoto}@isl.ntt.co.jp

1 はじめに

膨大な情報に対する効率的なアクセスを可能にする Social Filtering 手法が注目されている。TITAN[5]を始めとする WWW の検索サービスは、膨大なデータからの画一的な検索を行うため、個人毎に異なる探索要求に対して効果的な回答を出力することは難しい。現在では、Social Filtering の枠組みを利用して個人の興味に適合した出力を行うサービスも実現されている[1]。本稿では、利用者の認知的負荷の低い Social Filtering の実現のための、情報を視覚的に操作することで利用者の自然な操作を取り入れることを可能にした情報共有方式に関して述べる。

2 Social Filtering

情報の入手は検索サービスなどの網羅的なサービスから、一段進んだ情報の獲得方法を考慮すべき時期が到来した。情報検索を、情報の入手過程全般を意味すると捉えると、検索対象の増加による検索サービスの絶対的な有効性が下がりつつある現在、他者からの知識をいかに取り入れ利用者に提供することができるかが重要となる。このような手法の一つである Social Filtering は、その実現手法により大まかに以下の 2 種類に分類することができる。

● 推薦に基づくもの

利用者が参照情報を有効であると判断した際に注釈情報を加えておき、他者が参照できるようにしたもの[3]

● 相関に基づくもの

予め利用者の興味を蓄積しておき、相関に基づき興味と類似している他者を発見し、情報を推薦するもの[4]

既に firefly 社は、後者の技術を用いて Social Filtering 技術を用いたサービスを行っている[2]。上記の技術では、利用者は予め自分の音楽の興味を登録することで個人プロフィールをサーバ側に作成し、参照する情報に対する興味の度合いを利用者自身が主観的に判断し、その結果を利用して類似利用者を統計的に判定し推薦情報を決定する。決定の過程はいわゆるブラックボックスになっており、利用者にはどのように決定されたかが不明であった。

本稿では、利用者の興味の領域の作成と、他者の興味と自分の興味との関係を視覚的に表示することで、マウスのクリック、ドラッグなどの自然な操作を用いた各個人による適切な興味の領域の作成と、該当興味に対する他者からの有効な情報の提示を可能にする情報共有方式に関して次章で説明する。

3 情報共有方式の概要

3.1 個人の興味の構築

WWW ブラウザを通しての、利用者の自然な情報参照行動を利用した情報の共有の実現を考える。各個人の興味の領域を他者に公開するため、図 1 のような構成を持つ情報共有システムのプロトタイプを構築した。

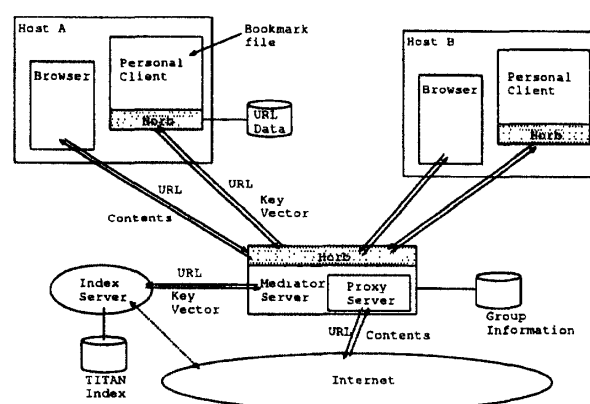


図 1: 構成図

上記システムでは、興味の関連の度合いを 2 次元平面上に表示するための個人クライアント画面（興味平面と呼ぶ）を以下のように作成する。まず、個人の興味を学習するための初期情報として、各個人が持つブラウザのお気に入り情報（いわゆるブックマークリスト）を取り入れる。このお気に入り情報は、興味のある情報を利用者が能動的に蓄積するものであり、各個人の興味と考えることができる。もしもお気に入り情報に階層が存在する場合、それらは興味の領域名を表すタグが付与され興味平面上に表示される。各々の領域は各領域に属する情報 (URL) が持つ単語情報に基づく特徴ベクトルにより類似度が計算され、その類似度に応じて画面上に配置される。

利用者がブラウザを通して新たな情報を参照すると、仲介サーバはまずプロキシサーバの役割を果たしインターネットから該当情報を取得する。そして、その情報が持つ特徴ベクトルの抽出後、興味平面上の自分の興味との間で類似度計算が行われ、類似した興味は画面上で近傍に表示される。

類似度算出処理時に必要な特徴ベクトルの抽出部は、TITAN が持つ WWW のインデックスを利用して特徴ベクトルを近似的に算出するインデックスサーバから成り、仲介サーバからの特徴ベクトルの抽出要求に応じて随時起動される。特徴ベクトルを用いた類似度の精度は近似計算である以上、あまり高くはないと考えられるが、我々は上記の分類結果の自動配置

Shared Information System with Interactive Communication

Seiji SUSAKI, Tatsuya MURAMOTO

NTT Information and Communication Systems Labs.

1-1, Hikarinooka, Yokosuka-Shi, Kanagawa, 239, Japan

は、利用者による分類作業の補助にすぎないと割り切り、利用者による対話的な興味を作成、削除作業による個人興味を作成機能を追加した。

利用者による興味の領域への情報の追加は、特定の興味に分類されていない情報を主観的に適当と思われる興味にマウスをドラッグすることで自動的に行われる。これは既に分類されている情報を他の興味の領域へ変更する場合も同様な操作で行うことが可能である。もしも適当な興味の領域が存在しない場合は、新たに領域を作成した上で追加する。また、不必要な興味の領域が存在する場合は、まずそれを削除して、どの領域にも属さなくなった情報は上記のように適当な領域に追加する。情報自体の削除は、マウスのクリックで簡単に行うことができる。このような作業の試行錯誤的な繰り返しにより、個人として適当な興味の領域の作成が可能になる。

3.2 情報の共有

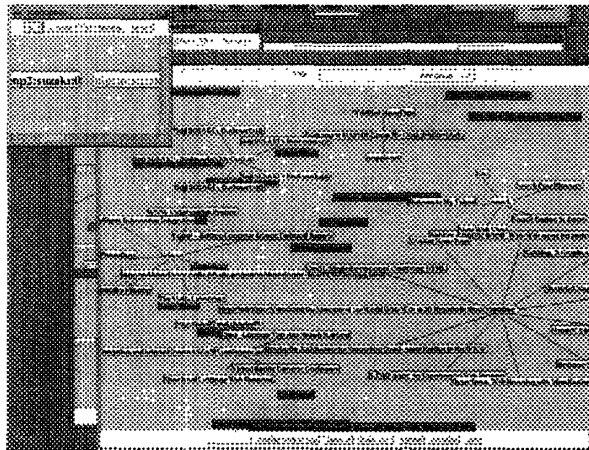


図2：情報共有時の画面例

図2は、Javaのアプリケーションとして実現した情報共有時の画面例で、左上はシステムにアクセスしている利用者をその類似度を元に表示した利用者管理ウィンドウ、その右には興味を表現する個人クライアントウィンドウが表示されている。

複数の利用者が自分の興味平面を構築すると、各利用者が持つ興味とそれらが持つ特徴ベクトルにより個人間の類似度が算出され、画面左上に反映される。利用者はその中で自分の近傍に表示された自分の興味と似た他者の興味を自分の興味平面上に流し込むことで、他者の知識の獲得が容易にできる。その際、同じ興味の情報は同ノードにまとめられ、他者との間で興味の共有状況を視覚的に把握することが可能である。そして、他者の興味の中で自分に不要な情報を削除すれば、自分の興味に合致した情報のみが自分の興味平面上に存在することになる。興味の領域に該当するノードは楕円で囲われ、各々の興味の領域に含まれる情報は該当興味からアークで結ばれた長方形のノードで表される。例えば新たに参照して未分類の情報は、ノードの背景が灰色で示され、利用者による分類操作に備えている。他者の情報は各ノードの背景色に変化し、利用者から容易に判断ができる。必要であれば、全体画面の拡大、縮小、そして各興味の領域毎の拡大表示も可能である。

本システムは、Java1.0.2、Horbl.3.1b1、perl5.002で記述されている。現在は、SGI Indigo2 Impact (IRIX5.3)、

Sun Ultra Sparc 2、Sun Sparc Station 20 (Solaris2.5.1, 2.4)上で個人クライアント、仲介サーバはSGI Indigo2 Impact、そしてインデックスサーバはSun Sparc Station 20 (SunOS4.1.4)上で稼働している。各個人の興味はローカルなオブジェクトとしてクライアント側に保持され、状況に応じてHorbが持つオブジェクト転送機能により仲介サーバを介して転送される。仲介サーバは、Horbの招待機構を用いてアクセス中のクライアントの管理を行い、利用者オブジェクトの生成、消滅を管理している。利用者が新たな情報を参照したり、新たな利用者が本システムにアクセスする際に、情報の特徴ベクトルを随時再計算させるため、ベクトル算出部分は描画部分とスレッドが独立しており、画面表示はベクトル値が算出されるのとは非同期で変更される。本枠組みによれば、個人の興味はあくまで各個人のローカルなオブジェクトとしてしか存在せず、各利用者により公開する興味を制限すれば、仲介サーバ上での不要なアクセスによる個人情報の漏洩を減少させることが可能であるため、セキュリティの面からも優位性がある。また、クライアントの増加によるサーバの性能低下が避けられるため、システムの拡張性も併せ持つ。

本システムの問題点としては、各個人が持つ特徴ベクトルの増加による類似度計算の遅延がある。例えば、他者の興味を自分の興味平面へ流し込むと、自分が保持する情報数は多くなり、その分特徴ベクトルの算出に時間がかかることになる。この問題は、情報の特徴ベクトルをJavaのHashtableクラスで管理していることに起因し、利用者が増加するに伴い必要記憶領域と時間が増加する。一例として、ある個人の興味領域が6つ、各興味に含まれる情報数の平均が10の場合、その人の興味を表す特徴ベクトルは約3,000個存在し、3利用者が同様の状態の場合、各個人の類似度計算には約180秒程度を必要とする。この問題点に対する改善方法としては、特徴ベクトルの保持部分を外部DBMSを使って管理し、JDK1.1以降が備えるJDBCのAPIを用いて再構築する、あるいは一度算出した各利用者の興味をオブジェクト永続化機構を用いて再利用可能な状態で保存することが考えられる。

4 まとめ

本稿では、利用者の対話的な操作を利用した情報共有方式に関して説明した。本枠組みでは、個人と他者の興味を視覚的に把握し、各々の興味の学習がスムーズにできる利点がある。

参考文献

- [1] <http://my.yahoo.com/>
- [2] <http://www.firefly.com/>
- [3] D. Goldberg, et. al.: *Using Collaborative Filtering to Weave an Information Tapestry*, Comm. of the ACM, Vol35, No.12, pp.61-70, 1992
- [4] U. Shadanand, et. al.: *Social Filtering: Algorithms for Automating 'Word of Mouth'*, CHI'95, pp.210-217, 1995
- [5] 鷲崎他: インターネット情報探索システム TITAN とその可視化手法, 情報処理学会第52回全国大会 4-185, 1996