

プッシュ型ニュースビューアにおける自動パーソナライズ方式

4 L-1

坂上 秀和 神場 知成 古関 義幸

NEC C&Cメディア研究所

1 はじめに

WWWの急速な普及により様々な情報に対するアクセスが容易になった一方で、膨大な情報の中からユーザにとって有益な情報を見つけ出すのが次第に困難となっている。効率的に有益な情報をユーザに提示する技術の研究が広く行なわれている中、プッシュ型情報配信が注目されている。1996年2月にニュース配信サービスとして開始されたPointcast Network(PCN)[2]に代表されるプッシュ型情報配信は、ユーザの操作なしにシステムが自動的に情報を配信・提示する形式を持つため、最新の情報に常にアクセス可能とした点と、予め好みのキーワードやジャンルを登録しておくことで関心の高い情報を見れるようにした点で情報アクセスの効率を画期的に向上させるものであった。現在、PCNの成功を背景に多くのプッシュ型情報配信サービスが展開されているが、今後これらの普及により配信される情報量が増大するに伴い、結局は配信された情報の中から必要な情報を見つけ出すことが困難になる危険性がある。

本稿では、ユーザ毎の嗜好に合わせて情報提供を行なうパーソナライズ技術によって本問題を解決する可能性を示すとともに、開発したプッシュ型ニュースビューアにおける自動パーソナライズ方式について報告する。

2 パーソナライズ技術の有効性

パーソナライズ技術とは、個人の興味や嗜好に合わせて情報の配信・提示を行なう技術である。本技術をプッシュ型情報配信に組み合わせることにより、配信される情報の選択段階や取得した情報の提示段階において個人の関心傾向を反映することが可能となるため、筆者等はパーソナライズ技術こそプッシュ型情報配信の今後を支える重要な要素技術の一つであると確信している。

パーソナライズ技術には、ユーザ嗜好情報の抽出と嗜好情報の反映の2つの側面がある。

2.1 嗜好情報の抽出

既存の多くのパーソナライズサービスでは、ユーザが予めキーワードやジャンル情報、指定の銘柄(株価)などを登録しておくことにより、それらにマッチした情

Automatic Personalization on a Push-style News Article Viewer

Hidekazu Sakagami, Tomonari Kamba, Yoshiyuki Koseki
C&C Media Research Laboratories, NEC Corporation

4-1-1 Miyazaki, Miyamae, Kawasaki, Kanagawa 216, Japan

報を提示する。しかしこの方式では、登録の際のユーザの負担が大きい、嗜好の変化に応じて修正が必要、予め登録した情報以外には対応できないなどの問題がある。可能な限りユーザの負担を減らしつつ、嗜好の動的な変化に対応できる機構が求められる。

2.2 嗜好情報の反映

現在の技術水準では(今後も当面は)、計算機による嗜好推定には一定の限界があることを認めざるを得ない。従ってパーソナライズ情報提示の際に重要なことは、「(計算機が予測した)重要な情報を押しつける」ことではなく、「重要な情報へのヒントを与え、アクセスしやすい環境を提供する」ことである。そのためには、関心が高いと推測される情報を優先的にアクセスできる機構とともに多くの選択肢中から必要な情報を取り出しやすい構造の提供が求められる。

3 開発したプッシュ型ニュースビューア

上記の課題を踏まえて、筆者等はプッシュ型情報配信によるニュースビューア ANATAGONOMY/SSを開発した。筆者等はこれまでに、インターネット上のパーソナル電子新聞[1]の研究開発を行なっているが、本システムは「プル型」情報アクセスを行なっていた従来システムを基に「プッシュ型」のニュースビューアとして構築したものであり、プッシュ型による情報配信、パーソナライズ機能、独自の情報提示・表現方式、嗜好情報の自動抽出、等の特徴を持つ。

3.1 全体構成

図1に本システムの全体構成を示す。

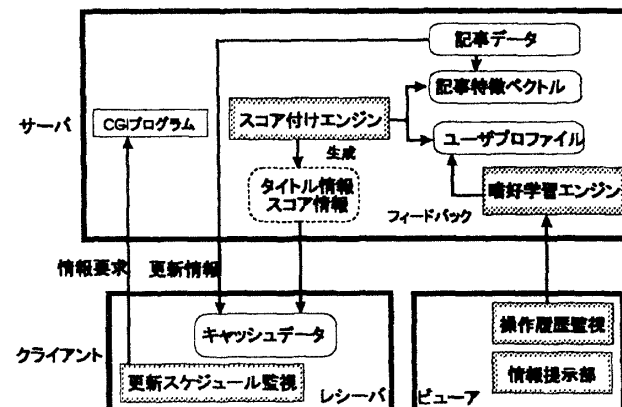


図1: ANATAGONOMY/SSの全体構成

サーバ側モジュールとしてスコア付けエンジン・嗜好学習エンジンを持ち、データとして記事データ・記事特徴ベクトル・ユーザプロフィールを保持している。スコア付けエンジンは各記事に対するユーザ毎の関心度をユーザプロフィールと記事特徴ベクトルとの比較から推定し、スコア情報として生成する。嗜好学習エンジンは、ユーザが記事に対して与えた関心度評価を基にユーザプロフィールを更新することで嗜好を学習する。

クライアント側モジュールは、レシーバプログラムとビューアプログラムから構成される。レシーバプログラムは記事更新スケジュールに基づいてサーバ側 (CGI プログラム) に情報要求を行なうスケジュール監視部と、サーバから取得した情報をローカルにキャッシュデータとして保持する機構を持つ。ビューアプログラムは、キャッシュされた記事情報を、ユーザの関心度や記事の新鮮度などを基に情報表示部によって表示する。

3.2 情報配信

記事更新時には、各記事のタイトル情報・スコア情報と、新規登録された記事の差分情報に分けて取得する。レシーバは常駐プログラムとして機能するため、クライアント端末上で他の作業を行なっている際にも更新スケジュールに基づいて記事取得を自動的に行なう。

3.3 パーソナライズ

記事への関心度と記事の新鮮度 (発信時刻からの経過時間と、既読か否かで評価) の2つの属性情報を表現することとした。これは、関心度が高くても古くなってしまった情報より、関心度が低くても新しい情報の方が有益な情報となる場合などを考慮したものである。

上記の関心度と新鮮度を提示情報に反映するために、本システムでは表示領域上に奥行き方向の広がりを持たせ、仮想的な3次元空間上に情報提示を行なうこととした (図2)。これにより、上下方向に関心度、奥行き方向に新鮮度を表現し、左右方向に記事タイトルが流れるという画面構成が可能となった。奥行きは、文字のサイズ・明るさ・移動速度の3要素を調整することで表現している。

また、記事情報の表示は、1. 通常時はタイトルが自動的に流れる、2. マウスカーソルをタイトルに合わせると本文の1行目を表示、3. タイトルをクリックすれば全文を表示、の3段階構成とした。これにより、通常時は大量の情報へのアクセスのきっかけを与え、1行

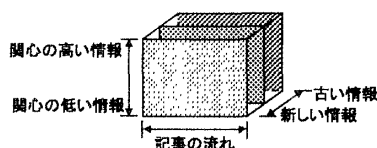


図2: 画面設計における仮想的3次元空間

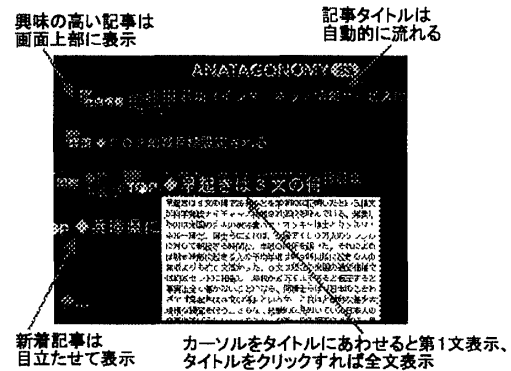


図3: ANATAGONOMY/SSにおける画面例

目表示により情報概要を把握させ、必要ならば全文表示によって詳細情報を与える構成となり、「多くの選択肢から必要な情報を取り出し易い構造」とすることが可能となった。図3に本システムの画面例の説明図を示す。

3.4 嗜好抽出

ユーザの負担を最小限に抑えることと、動的な嗜好変化に対応できることを目指し、ユーザの操作履歴から嗜好傾向を抽出する方式を本システムでは採っている。嗜好学習エンジンは、テキスト (記事) 情報に評価 (点数) を与えることでユーザプロフィールの更新を行なうため、嗜好抽出するには、読まれた記事に対して評価付けを行なう必要がある。そこで本システムでは、全文表示を行なった記事についてのみ、スコア付けエンジンが予想したスコアにボーナス点を加算しサーバ上の嗜好学習エンジンにフィードバックする機構とした¹⁾。本機構により、ユーザがキーワード登録などを行なわなくても、通常の購読操作を行なうだけで次第にユーザの嗜好がシステムに蓄積され、動的に次回以降のアクセス時の記事スコア付けに反映することが可能となっている。

4 おわりに

本稿ではプッシュ型情報ビューアにおけるパーソナライズ技術の有効性、課題等について述べ、実装したシステムについて報告した。本システムは嗜好の自動抽出、情報の関心度と新鮮度を反映した3次元情報表示に特徴を持つ。今後は、新聞情報以外のインターネット上の情報への対応と、テキスト以外のメディアをパーソナライズする方式等について検討して行く予定である。

参考文献

- [1] 坂上, 神場, 古閑: パーソナル電子新聞 ANATAGONOMY の開発と評価. インタラクティブシステムとソフトウェア IV (WISS'96), pp.21-30, 近代科学社, 1996.
- [2] Pointcast Network (<http://www.pointcast.com/>)

¹⁾ 1行目表示に関しては、ユーザの意図とは無関係にマウスカーソルがタイトルに入る機会が多いため、嗜好抽出には利用しない。