

The Krakatoa Chronicle : WWW上のエージェント 機能を利用した、対話型パーソナル新聞

神場知成

NEC C&C研究所

概要 : WWW上で豊富な対話機能、個人向けの多彩なレイアウト機能を持つオンライン新聞 The Krakatoa Chronicleについて述べる。HotJavaブラウザを利用することで、サーバからクライアントへの対話エージェント送り込みを行う。これにより、ユーザが新聞のレイアウトのためのパラメータをダイナミックにコントロールすることや、サーバが各ユーザの記事購読状況を監視して自動的にユーザプロフィール生成、管理を行うことができる。

The Krakatoa Chronicle: A Highly Interactive, Personalized,
Newspaper on the WWW

Tomonari Kamba

C&C Research Laboratories, NEC Corporation

This paper describes The Krakatoa Chronicle, a highly interactive, personalized newspaper on the WWW. It is intended for Java-aware WWW browsers such as HotJava, and is architecturally quite different from conventional web-based newspapers. Its high interactivity and powerful personalization are the result of sending an interactive agent along with the text of the newspaper to function within the user's web-browser. Ours is the first web newspaper to attempt a realistic rendering of a newspaper, with a multi-column format and embedded, widgets for easy browsing. We provide dynamic layout control mechanisms that let the user specify how personal and community interests should be interpreted by the agent in designing the layout.

1. はじめに

World Wide Web (WWW)は、インターネット上でマルチメディア情報を受発信する方法として最も普及かつ有効な分散型ハイパーメディアシステムである。従来インターネット上では非商用の情報サービスが主体であったが、最近では商用サービスが急速に増えており、その中で最も重要な商用サービスの1つがオンライン新聞である。この分野で米国は日本に比べて大きく先行し、現在アメリカで提供されているWWW上のオンライン新聞は daily のもの、そうでないものを含めて100を超える[1]。オンライン新聞は課金方法に課題を残すものの、今後急速な発展をすることは確実である。マルチメディア(ビデオや音声)の利用、即時性など紙の新聞にはない多くのメリットがあるからである。

さて、オンライン新聞の重要課題の一つは personalization だろう。これには新聞の社会的側面も考慮するべきである。例えば、人々が新聞を読む時には自分が以前から興味を持っていることに関連する記事を読むだけでなく、「新聞の扱い方から見て重要そうな記事」や「ふと興味を引かれた記事」などを読むであろうことを考えれば、「ユーザがキーワードを登録しておき、それに基づいてシステムがパーソナル新聞を提供する」というような手段が新聞のメリットの一面を損なうことは明らかである。

本稿では、上記のような考察をもとに筆者等が開発した WWW 上のパーソナル新聞 The Krakatoa Chronicle について述べる（本システムは、筆者が1994年11月から1995年10月まで客員研究員として滞在したジョージア工科大学において、アシスタントの学生二人 - Krishna Bharat と Michael C. Albers - とともに開発した）[2]。技術的には WWW 上でエージェント機能（プログラム自体がネットワーク上を行き来して、ユーザに対してさまざまな機能を提供すること。）を実現する手段として普及し始めた Java 言語[3]をベースに、「紙の新聞を模したマルチコラムのレイアウトと、豊富なインタラクション」、「ユーザプロファイルの自動的獲得（ユーザがキーワード入力などを行わなくても、自動的にユーザプロファイルが生成される）」「パーソナル化された新聞と一般的新聞の連続的切り替え」など従来存在しなかった多くの機能を可能にした。以下では、システム構成、ユーザインタフェースなどについて述べる。

2. システム構成

図1に The Krakatoa Chronicle のシステム構成を示す。ユーザは、Java 言語を解する WWW ブラウザを利用してこのシステムの WWW ホームページの HTML (HyperText Markup Language) 文書にアクセスする（動作は HotJava ブラウザで確認した）。システムはサーバ側（WWW のサーバサイト）とクライアント側のプログラムから成る。ただしクライアント側のプログラムとは、ユーザがサーバ側にアクセスすると自動的にクライアントにダウンロードされ動作を開始するエージェント（Java 言語の枠組みの中では applet と呼ばれる。以下では applet で統一）を指している。なお Java 言語は C++ と類似した文法を持つオブジェクト指向言語であり、コンパイルしてマシンに依存しないバイトコードに翻訳され、そのプログラムはクライアント側のインタプリタ上で動作する。以下ではサーバ側、クライアント側の機能について述べる。

2. 1 サーバ

(1) 毎日の記事管理：記事は毎日、外部ニュースサイトからダウンロードする（ジョージア工科大学において The NandO Times [4] から実験用の利用許可を受けた）。次に SMART エンジン [5] を利用して各記事から特徴ベクトルを抽出する（記事はすべて英語）。特徴ベクトルは記事中に出現する単語とその重み（出現頻度をベースに計算される）をペアにしたものの集合であり、ファイルは各行が「currency : 120」などと記入されている。各記事本文は UNIX の文書整形機能 troff によってフォーマット変換し、クライアント側におけるレイアウトを容易にした。

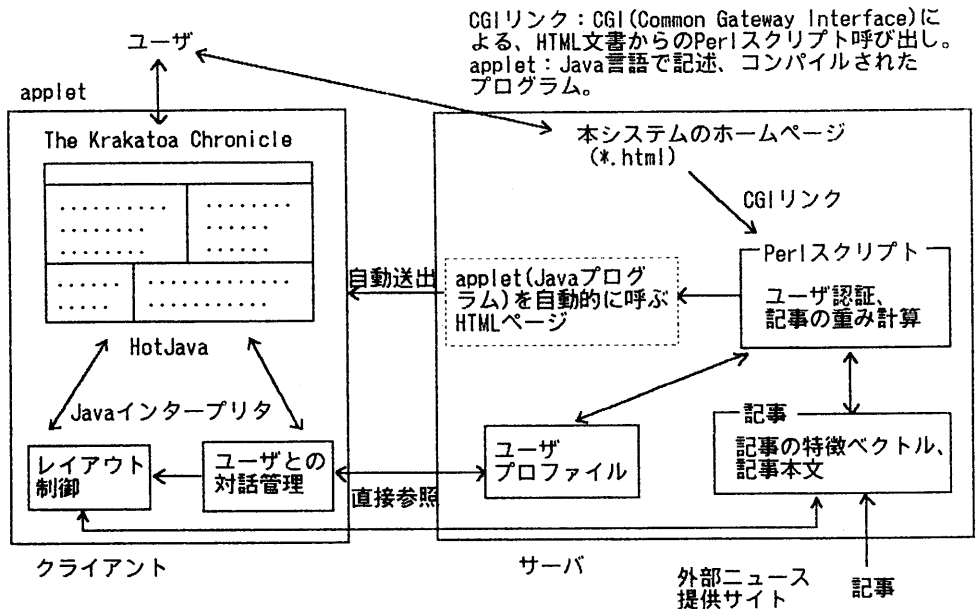


図1 The Krakatoa Chronicleのシステム構成

(2) ユーザプロフィールの管理: 各ユーザのプロフィールは記事の特徴ベクトルとほぼ同じフォーマットである。つまり、そのユーザが興味を持っているキーワードと重みのペアの集合である。ユーザが記事を読んでいる際には操作状況が自動的にクライアントからサーバに送付され、サーバはそれをもとにプロフィールを書き換える。

(3) 各記事の重みの計算: サーバは1人のユーザ、1つの記事に対して2種類の記事重み情報をクライアントに送る。1つは「そのユーザにとっての重み (personal score: 記事の特徴ベクトルと、ユーザプロフィールを示すベクトルとの類似度)」であり、もう1つは「システムを利用しているユーザ全体にとっての重み (community score: 全ユーザに対する前者の平均)」である。

2. 2 クライアント

Java対応のWWWブラウザは、アクセスしたHTML文書に特殊なタグが含まれていると自動的にappletをダウンロード、実行する。本システムの appletは以下の作業を行う。

(1) インタラクション管理: 用意されているインタラクションについては次章で説明を行う。ユーザが操作を行うと、行った操作内容が自動的にサーバに送付される。

(2) レイアウト制御: サーバから受け取った各記事のpersonal scoreおよびcommunity scoreをもとにさまざまなレイアウト制御を行う。詳細は次章で述べる。

(3) 記事の描画: troffによって変換されたフォーマットを元に記事描画を行う。

3. ユーザインタフェースと紙面レイアウト管理

ユーザは自動的にレイアウトされた新聞を表示する前に、自分のプロフィールを直接編集することも可能である。The Krakatoa Chronicleはユーザの記事購読状況を監視して自動的にプロフィール（キーワードと重みのペア）を作成、変更するが、ユーザが明示的に指定したキーワードについては、その重みを自動的に変更しない。

3. 1 各記事に対するユーザ操作とサーバへのフィードバック

図2に HotJavaブラウザで表示した画面例を示す。基本的なレイアウトに関しては、「従来の紙の新聞の外観を模すること」という方針に従った。採用した特徴は、マルチコラムであること、記事の重要性は左上から右下にかけて低くなる傾向にあること、重要記事ほどタイトルのフォントが大きくなること、などである。各記事に対して可能な操作は以下の通り。

(1)スクロール（図2のスクロールバー参照）：記事が1コラムから成る時は垂直にスクロールを行い、複数コラムから成る時は水平にスクロールを行う。

(2)Peek（のぞき見：図2の Peekボタン参照）：ボタン上でユーザがマウスボタンをプッシュしている間だけ、その記事を画面上に大きく表示する。

(3)スクラップブックへの保存（図2の Saveボタン参照）：スクラップブックは、ユーザが保存した記事へのリンクを持つページである。

(4)記事の personal scoreの変更（図2のスコアバー参照）：図2のスコアバーにおいて左側の白い領域の広さが、システムが予測したpersonal scoreを示しており、スコアバーの中に引かれた黒い線が community scoreを示している。personal scoreはユーザが直接変更可能である。その結果はユーザプロフィールに反映される。

ユーザが上記のような操作を行った場合は、サーバ側に操作内容のフィードバックが行われる。各操作内容に対して、サーバ側にはそれによって対応する記事が得る重みが記載されており、結果としてその記事に含まれるキーワードの重みが変わる。

3. 2 紙面のレイアウト制御

ユーザは大域的なレイアウト制御のために次の3つのスライダーを制御できる。

(1) density bar（図2の density bar参照）：画面上にレイアウトする記事の数を増減する。多数の記事を配置すれば1つの記事に割り当てられる面積は狭くなる。

(2) tendency bar（傾向を示すスライダー：図2の tendency bar参照）：記事を personal scoreによってレイアウトするか、community scoreによってレイアウトするかをコントロールする。

(3) sensitivity bar（記事の重みの、記事面積への影響度を示すスライダー：図2の sensitivity bar参照）：sensitivityが低いと記事のスコアは記事の位置だけに影響するが、sensitivityが高いと重要記事にはより大きな面積が与えられる。

4. 性能

1つの記事は troffによって変換後、イメージを除いて 100-200Kbyteのサイズであり、これらを appletが自動的にダウンロードするのに約6秒かかっている（サーバは Sunの SPARCstation 10、ネットワークは学内は EthernetLAN）。例えば8記事を含むページを表示するには約1分かかる。これはやや遅く感じるが、今後ハードウェア、ネットワーク環境、ソフトウェアの改良などによって向上すると考えている。

5. 関連研究および議論

従来、WWW上には2、3のパーソナル新聞システムがあったが[6]、既存のものはHTMLおよびHTTPという技術上の制約から、「クリックによって静的文書のリンクをたどること」以外のインタラクションを提供しておらず、本システムのように豊富なインタラクションおよびレイアウト制御は見られない。特に、新聞の社会的役割を考えると「パーソナル化」と「一般的に重要な情報の伝達」との調整は重要である。

ユーザプロファイルの獲得に関して、ユーザが直接キーワードを設定する方法は一定の成功を収めているが[7]、ユーザの無意識の興味や、短期間の興味変化が反映しにくいという限界がある。ユーザが各記事に対する興味の度合いを自分で選択する方法も[8]、ユーザには記事を読む以外の負荷がかかる。これに対し、本システムの手法は通常の記事購読以上の負荷をユーザに要求しない。今後プロファイル書き換えがユーザの満足の一部になっているかどうかの評価を行う予定である。

6. おわりに

WWW上のパーソナル新聞 The Krakatoa Chronicleについて述べた。サーバ側からクライアント側へ対話エージェントを送り込むことで、柔軟な自動レイアウト、豊富な対話機能、正確なユーザ利用状況のサーバへのフィードバックなどを可能にした。

参考文献

1. Dailies: U.S. Newspaper Services on the Internet, <URL:http://marketplace.com/e-papers.list.www/e-papers.us.dailies.html>
2. T. Kamba, K. Bharat, M.C. Albers. The Krakatoa Chronicle: An interactive, personalized, newspaper on the Web, The 4th Int. Conf. of WWW, 1995.
3. HotJava Home Page, <URL:http://java.sun.com>
4. The Nando Times, <URL:http://www.nando.net/newsroom/nt/nando.html>
5. SMART, <URL:ftp://ftp.cs.cornell.edu/pub/smart>
6. Fishwrap, <URL:http://fishwrap.mit.edu>
7. Tak W. Yan and Hector Garcia-Molina, SIFT - A Tool for Wide-Area Information Dissemination, USENIX Technical Conference, pp. 177-186, 1995.
8. Ken Lang, NewsWeeder: Learning to Filter Netnews, ML95, <URL:http://anther.learning.cs.cmu.edu/ml95.ps>