

## WWWのお勧めホームページについての 情報共有システム

杉井俊彦 北英彦 林照峯

三重大学工学部

インターネットの普及に伴い、WWW (World Wide Web) のホームページの数が、急激に増加している。その結果、自分の興味ある情報が載っているホームページを、効率的に見つけることが困難になっている。

これに対して、我々は興味あるホームページにはアクセス回数が増えるであろうという考えに基づいたシステムを作成した。しかし、興味あるホームページを見つけた人が他の人にそのことを直接教えたいと思った場合には、WWW以外の手段を使う必要があった。そこで、お勧めのホームページについての情報をWWW上で直接入力し他の人に伝えることができるシステムを提案する。

## The information sharing system for recommended homepages on World Wide Web

Toshihiko Sugii Hidehiko Kita Terumine Hayashi

Faculty of Engineering, Mie University

The number of homepages on World Wide Web(WWW) increase rapidly with the spread of the internet. As a result, it is difficult for users to find their own favorite homepages efficiently.

We proposed the system based on the idea that the interested homepages had been accessed many times. However, if an user find an interesting homepage and want to tell about it to others, the user must use some other method than a WWW browser. In this paper, we propose the system that the user is able to tell others directly about a recommended homepage on a WWW browser.

## 1. はじめに

インターネットの急速な発展に伴い、情報公開、情報入手の主要な手段であるWWW (World Wide Web) の利用が急激に増えている。WWWの検索ホームページの1つであるAlta Vistaには、サーバの数は約20万サーバ、ホームページの数は約3,000万ページが登録されている[1]。すなわち、少なくとも全世界において、これだけのホームページの数があることが分かる。さらに、これらのホームページの内容は時々刻々と変動しており、その全貌を完全に把握することは不可能である。そのため、現在無数に存在するホームページの中から、自分の興味ある情報が載っているホームページを探し出すことが困難となってきている。

これに対して、我々は興味あるホームページにはアクセス回数が多くなるであろうという考えに基づいたシステムを作成した[2]。これは、proxy (中継サーバ) のアクセスログファイルの中のアクセス回数のみを取り出して、集計し、アクセス回数の多い順にURL (Uniform Resource Locator) をソートし、その一覧を提供するシステムである。特徴は、興味あるホームページを見つけた人に対して手間をかけることなく、他の誰かに興味あるホームページについての情報を教えることができるという点である。

しかし、興味のあるホームページを見つけた人が、他の誰かに直接そのことを教えたいと思った場合にはメモ、メールまたは口頭で伝えるなど、何らかの手段が必要である。WWWの使用中に他の手段を使わなければならないと、WWWによる検索作業の効率を低下させる。また、そのために教えてあげようという気が起こっても、実際にはそれを実行しなかったりする。

そこで、興味あるホームページについての情報を他の手段を使わずに教えられるように、お勧め情報の入力がホームページ上でできるように拡張を行った。

## 2. お勧め情報

### 2.1 お勧め情報とは

さまざまな情報があるとき、自分にとって価値ある情報がどれであるかを教えてくれる情報をここではお勧め情報とする。一般に、興味ある情報を教えてもらうより、他人からの情報を知ることによって探し出せれば、かなりの労力を削減できることは日頃経験することである。また、ここでは情報を教える人を提供者、その情報を利用する人をユーザと定義する。

### 2.2 情報の評価

興味あるホームページを教えるために評価が使われる。しかし、評価は提供者の主観によってつけられるため、ユーザと興味が必ずあうとは限らない。そのため、提供者とユーザの間にはギャップが存在する。その結果、提供者によってつけられた評価をもとに利用するユーザは、自分が期待していたものと結果が異なる場合がある。

### 2.3 直接的な評価と間接的な評価

評価といっても、2種類ある。それは、提供者が自分の主観をもとに評価を直接入力することによって、ユーザに教える直接的な評価と、提供者が意識することなしに、提供者の振る舞いをみてユーザが利用できる間接的な評価である。直接的な評価であれば、提供者にとって手間はかかるが自分の主観に基づいて、直接教えたい情報を伝えることができる。また、間接的な評価であれば、提供者の手間をかけることなく、ユーザは情報を知ることができるという利点がある。

### 2.4 対象とするグループ

グループ内で協同作業をしたり、情報共有するグループウェアでは、対象とするグループの大きさやタイプによってその支援の仕方などが異なってくる。グループには、不特定多数の大

規模なグループや、会社の部署や研究室などの小グループなどいろいろある。

本システムでは、会社の部署や研究室などの小グループを対象した。つまり、小グループであればお互いのことをよく分かっているため、情報を教える提供者と、教えてもらって利用するユーザとのギャップを少なくすることができるからである。

## 2. 5 対象者の指定

一般的に提供者が情報を直接教えたい場合、教えることと同時に教えたい人が存在し、教える相手のイメージを持っているはずである。情報を教える相手が分かれば提供者としても、価値あるホームページを教えるために、それなりの責任をもつことになり、また評価をつけやすくなる。その上、教える対象者によってある程度内容を分類することができる。

一方、その情報を利用するユーザにとって各個人は、自分がどの対象者に属しているかは明らかに分かるため、お勧め対象者の表示があれば自分にとって興味のある情報があるのはどれかを選択しやすくなる。

## 3. 人気・お勧めホームページ案内システム「やじうまくんⅡ」の実装

### 3. 1 概要

2種類のお勧め情報のうち、アクセス回数という間接的な評価のみを利用したのが以前我々が提案した「やじうまくん」であった。

今回の「やじうまくんⅡ」では、提供者の直接的な評価も利用した。つまり、お勧めのホームページのURLと誰に情報を教えたいかというお勧めの対象者の入力ができるようにした。

### 3. 2 お勧め情報の入力

具体的には、もし提供者が興味あるホームページを見つけた場合、入力ウインドウ(図1)からそのホームページのURLとそのホームページ

を誰に教えたいかという対象者を指定して入力してもらう。そのことによって、提供者は誰に情報を教えたいのかということを伝えることができる。今回、研究室ということもあり、対象者を図1のような5つのグループとしている。

図1 入力ウインドウ

### 3. 3 システム構成

図2にシステムの構成図を示す。本システムは、proxyサーバの利用を前提としている。proxyサーバとは、セキュリティおよびネットワークのトラフィックの軽減のために外部のネットワークと隔離されたローカルネットワークとの中継の役割を果たすもののことである。proxyサーバを通ったアクセスに関しては、利用者が意識することなしに、いつ、どこのホームページを見にいったのか、などといった情報をすべてproxyサーバがアクセスログファイルに記録することができる。また、一度見にいったホームページに対しては、そのホームページを構成するファイルをproxyサーバがキャッシュとして残しておくことができる。

本システムでは、入力ウインドウから得られたお勧め対象者のデータとproxyサーバで取られたログファイルの解析、およびキャッシュからのタイトル抽出を行う。そして、その結果をホームページの1つとして作成し、ユーザに提供している。

### 3. 4 利用するデータ

一般的に1枚のホームページは、全体の構成を示すテキストファイルとその中に取り込まれる画像、音声などの複数のファイルから構成されている。WWWでは各々のファイルが別々にサーバからブラウザへ転送され、ブラウザ上で再構成される。そのため、proxyサーバ上では各々のファイルへのアクセスが別々のものとして記録される。本システムでは、画像、音声などのファイルを取り除いたログ、つまりホームページと対応するHTMLファイルのログのみを取り出して利用している。

また、データとして使われるログは、ホームページの利用の点から最近過去約3カ月間のログを使用している。さらに、ログというものは毎日次々と新しくできていくため、本システムでのデータの更新を毎日、計算機の利用の少ない深夜に自動的に行っている。

### 3. 5 インターフェース

実際の「やじうまくんⅡ」のホームページの画面を図3、図4に示す。

図3は、「やじうまくんⅡ」でのインデックスになるホームページである。本システムでは利用するデータを100位までのデータとし、それを5つのグループ各20位ごとに分けてある。

そして、アクセスログファイルの使用期間及び、データの更新日時をユーザに分かりやすいように上の方に表示している。

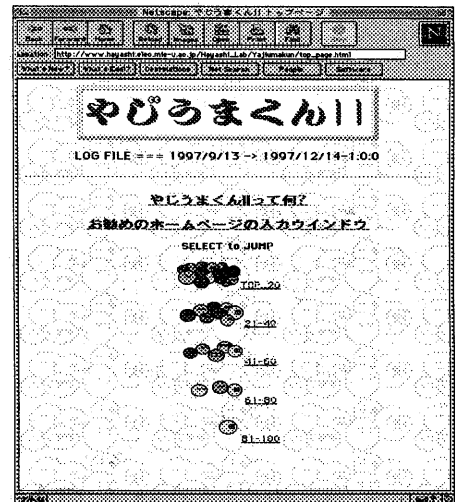


図3 「やじうまくんⅡ」のインデックスのホームページ

図4では、図3において”TOP\_20”をクリックした時に表示されるホームページである。このホームページでは、上位20位までのホームページを表示している。そして、ホームページに新たに加えられたお勧めの対象者の表示の他に各ホームページの順位、アクセス回数、一番最後

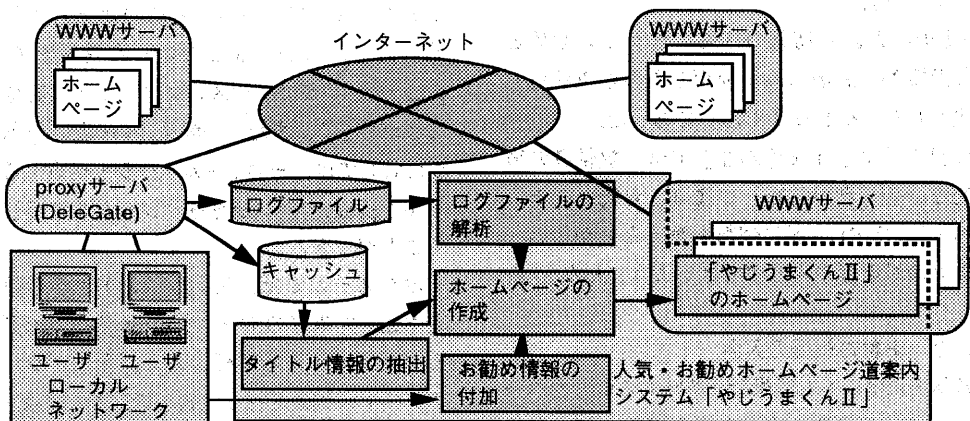


図2 システム構成図

にアクセスされた最終アクセス日、及びそのホームページのタイトル (URL) を表示している。順位は、各ホームページへのアクセス回数、アクセスされた日付、及びお勧め対象者の情報の有無に対してそれぞれ重み付けを行い、算出している。また、キャッシュからホームページのタイトルの部分を取り出している。タイトルは、キャッシュにファイルが残っていて、かつファイルにタイトルが書かれているホームページが対象となる。基本的にはホームページはタイトルでの表示を行っているが、タイトルのとれないホームページはURLでの表示を行っている。すべてのホームページのタイトル (URL) に対して、それぞれのホームページへリンクを張ってあるため、気に入ったホームページが見つければそこをクリックすることによって、すぐにそのホームページに飛ぶことができる。

また、本システムでは最近の動向を見るため、付加的に各ホームページの最終アクセス日を表示してある。そして、その表示を目安にして各ホームページの最近のアクセス状況が分かる。

#### 4. 試行実験

##### 4.1 実験方法

本システムを評価するため、1996年2月～1997年11月の22か月間の試行実験を行った。対象は本研究室の学生を中心に行った。

##### 4.2 実験結果

本システムの参照回数において、「やじうまくん」の試行実験期間である1996年2月～1997年7月と、「やじうまくんII」の試行実験期間である1997年8月～11月の結果を図5に示す。1997年4月から本システムへのアクセス回数が

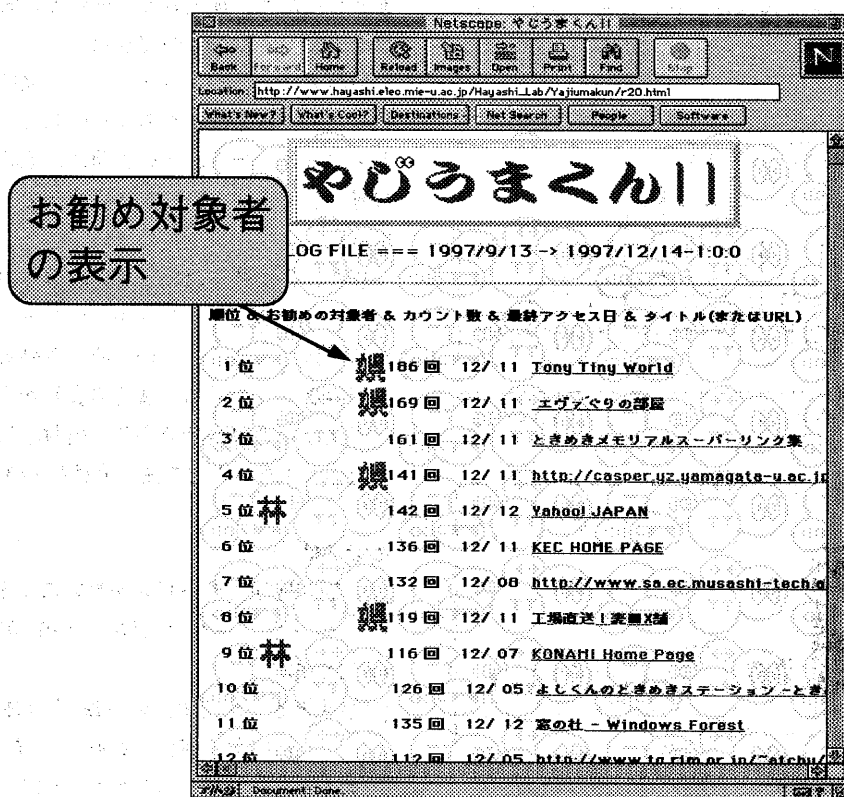


図4 「やじうまくんII」の上位20位のホームページ

増えているのは、大学では毎年研究室に4年生が入るため、メンバーが半数ほど入れ変わる。そして、今年はたまたまホームページを見る人が多かったためである。

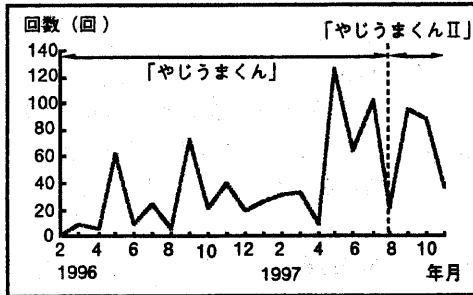


図5 本システムへの参照回数

「やじうまくんII」で使われる100位までのホームページのうち、お勧め対象者の表示が付いているホームページは全体の約20%であった。特に、上位20位ではお勧め対象者の表示が付いているホームページは約40%であった。

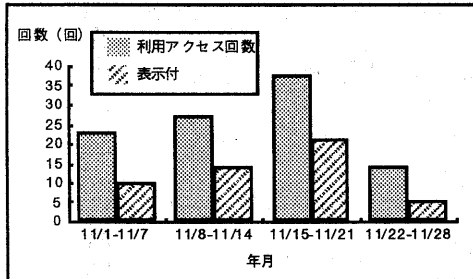


図6 本システムからのアクセス回数

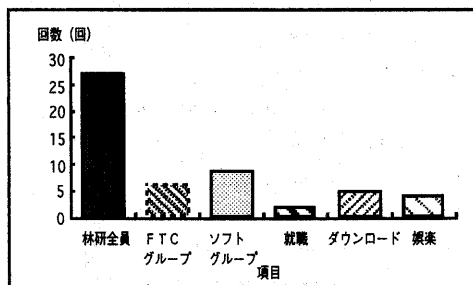


図7 お勧め対象者別お勧め入力状況

(1997年12月6日現在)

1997年11月1日～11月28日までに、おいて、「やじうまくんII」のホームページを利用してアクセスされた回数を図6に示す。その期間におけるアクセス回数の累計合計は102回であった。そのうち、お勧め対象者の表示が付いているホームページからアクセスされた回数の累計合計は50回であった。

さらに、1997年12月6日までに、各お勧め対象別に入力された回数を図7に示す。林研全員の表示が付いているホームページは、一般対象者向けのホームページが多く、FTCグループの表示が付いているホームページは研究関係のページが多かった。

## 5. おわりに

今回、新たに提供者の直接な情報を教えることができるようにした。特に、小グループを対象としたため、提供者とユーザとのギャップを少なくすることができた。また、他の手段を使わずに、ホームページ上で、提供者の直接教えた情報を伝えるようにした。さらに、利用するユーザは、そのお勧め対象者の表示を参考にしてホームページの選択ができるようになった。

現在、提供者におけるお勧め情報の入力に対して、URLをコピーアンドペーストという入力方法で行っている。そのため、お勧め情報の入力が少なく、お勧め対象者の表示が十分あるとは言えないのが現状である。提供者にとってより負担が少ない入力方法の検討が必要である。

## 参考文献

- [1] "WWW検索サービスを使いこなそう!"、インターネットマガジン、1997年9月号、pp.212-223
- [2] 杉井俊彦、北英彦、林照峯："お勧めのホームページに関する情報の共有"、電気関係学会東海支部連合大会(1997年)、p.279