

閲覧履歴を共有するウェブブラウザ

A web browser that shares browsing histories

高須賀 清隆†
Kiyotaka Takasuka

白井 雄一郎†
yuichiro Shirai

丸山 一貴‡
Kazutaka Maruyama

寺田 実§
Minoru Terada

1. 背景

インターネット上には今や膨大な量の情報が流れている。インターネットの利用者にとってはその中から有益な情報を選び出すことが非常に重要である。インターネット上で情報を探す手段として現在もっとも知られているのが、検索エンジンやディレクトリサービスである。これらのものはどちらも探したいものが明確に存在している場合には有効な手段である。しかし、新しいとか、多くの人が注目している、といったように対象を限定せずに情報を探し出すことも重要である。例えば、現実で言う路上の人だかりのような、特定の WEB ページへのアクセスの集中をリアルタイムに見つけ出す必要がある。以上のように、必要があるにも関わらず、このような観点から情報を探す有効な手段はまったくといっていいほどなかったと言える。

2. 目的

本研究では、どれだけ多くの人々が閲覧しているのかという情報を重視し、WEB ページの閲覧状況のある程度リアルタイムに知ることができるツールを開発し、評価する。

3. 本研究の着目点

本研究の着目点は以下の二つに分けられる (図1)。

- A) クライアント (ブラウザ) 側で閲覧情報を収集
- B) 収集した情報を P2P システムにより共有

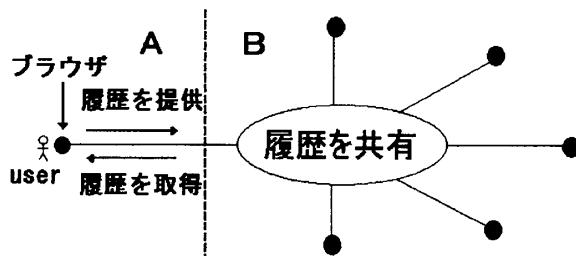


図1:履歴共有のモデル図

A は閲覧情報をどこで収集するかに関するもので、A の方法の他にもサーバでの収集と proxy での収集が考えられる。その中でもこの手段を選んだのは、利点として、あるユーザーが閲覧した全ての WEB ページを履歴共有の対象にできる点と、アンカーのクリックなどでユーザーが主体的に選択した URL と、それに付随して GET されてしまうアイコン画像などの区別が可能であるという点が挙げられるからである。

B は閲覧情報を共有するシステムについてである。履歴

を共有する上では匿名性が重要なので、その特徴として高い匿名性が挙げられる P2P ネットワークを用いた共有システムを選択することとした。また、P2P ネットワークにおいて自律的なクラスタ化が可能であるという点も考慮している。

4. システムの構成

本システムは履歴取得機能を備えたブラウザと履歴情報共有システムから構成されている。その動作としては、まず、ユーザーが閲覧した WEB ページの情報をブラウザ側で取得する。次に取得した情報を共有システムに送信する。そして、共有システムから共有している履歴情報を入手して表示する。図2にシステムの構成図を示す。

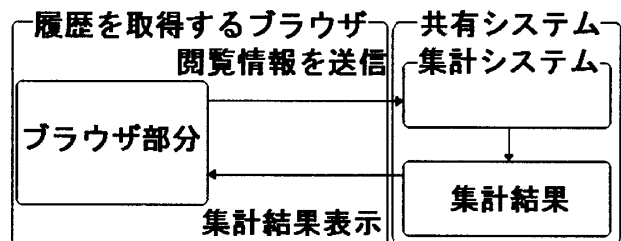


図2:システムの構成図

5. 実装

今回の実装ではブラウザ側の実装に重点を置き、共有システムとして P2P でのネットワークを用いるのではなく、その前段階として履歴を共有するために集中サーバを用いて実装を行うこととした。

5.1 ブラウザ

Visual C# [1] により実装を行い、動作には OS が Windows で .NET Framework をインストールしている必要がある。

ブラウザの一般的な機能を実装した。WEB ページを表示するレンダリングエンジンには IE コンポーネントを使用し、それを操作するためのボタンや URL 入力欄を準備した。

ブラウザは閲覧した WEB ページの情報 (タイトル、URL、閲覧した時刻) を共有システムに送信する。これらの情報は共有システムがランキングを作成する際に用いられる。送信する情報は先に挙げた三つの情報だけなのでユーザーを特定する情報は何もない。そのため匿名性を確保できる。また、データが重複することも考えられるので時刻は同一データを区別するのに用いる。まったく同じ時刻に違うユーザーが同じ WEB ページを閲覧する可能性もあるが、無視できる範囲内であると判断した。

†電気通信大学大学院情報通信工学専攻修士課程

‡電気通信大学総合情報処理センター

§電気通信大学情報通信工学科

5.2 共有システム

共有システムに集められた履歴情報は集計されて、各種のランキングのHTMLが作成される(図3)。閲覧データに含まれる時刻から一ヶ月以内に閲覧されたWEBページのランキング、一週間以内に閲覧されたWEBページのランキング、最新1000のランキングの三つのランキングを作成する。このランキングはブラウザのボタンをクリックすることで全てのユーザが簡単に閲覧することができる。ボタンをクリックして閲覧できるランキングは設定で変更できるようにした。

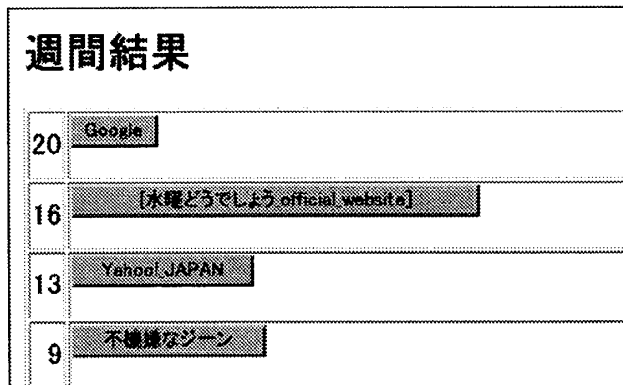


図3:ランキング画面の例

6. 議論

関連研究として、タブブラウザとして公開されているLunaspape[2]が挙げられる。Lunaspapeも本研究と同様にWEBページの閲覧数によるリアルタイムランキングを提案している。しかし、まだ実現が報告されていない。

サーチエンジンでのランキングとの違いを述べておく。第一に、サーチエンジンなどはリンク構成などの静的な情報を元にランキングを作成している。しかし、本システムはWEBページの閲覧履歴という動的な情報を元にほぼリアルタイムにランキングに反映される。第二に、本システムではユーザの閲覧頻度という尺度でランキングを作成しているという点もサーチエンジンとは異なる。第三に、時間変化を見ることができるとい点が挙げられる。閲覧情報に対応してランキングが常に変化するので、急激に閲覧数を増やしているページなどがわかる。それにより、たった今インターネットで注目を集めている記事を容易に判別することができる。

次にブログとの比較について述べる。ブログとは一つ一つの記事にWEBページのリンクをはり、それについてのコメントを書き込んだ記事が時系列に沿った構造となっているサイトを指すことが多い。このブログや、日記サイトなどを網羅的に検索するblogWatcher[3]というシステムが公開されている。WEB上の新しい情報を収集しランキングとして表示してくれるシステムという点で本システムと似通ってはいるが、ブログなどは人の手による更新を必要とし、その分だけランキングへの情報の反映に時間がかかるのに対し、本システムでは閲覧情報がランキングにすぐに反映されるため、リアルタイム性という点においてblogWatcherよりも優れている。

7. 評価実験

実際に電気通信大学の学生9人の方に協力してもらい4日間使用してもらった。その後アンケートをとり、加えてランキングページの使用の際にとっていたログを検証することで履歴共有の有用性や問題点を検証した。アンケートは5段階で評価してもらい、1に近い程肯定的で5に近い程否定的となる。表1にアンケート結果を示す。実験期間中のWEBページの閲覧総数は448回であった。そのうち70回がランキングページからの閲覧であった。

表1:アンケート結果

質問(1:そう思う 5:思わない)	平均
ブラウザとしての使い勝手はよかった	3.1
履歴共有は役に立った	2.0
他のユーザの履歴を見られるのはよかった	1.7
履歴共有に対して抵抗感は感じなかった	2.8

8. 結論

アンケートの結果は概ね履歴共有に対して肯定的な意見が多かった。実験前は、例え匿名であっても自分の閲覧履歴が公開されることについて抵抗感があると予想して第4の質問を尋ねたが、アンケートからそれは感じられなかった。これはユーザに意識せず使用してもらえたということだと思われる。また、実験は小規模であったため問題もあったが、それでも有用なシステムであるということが実証されたといってもいいだろう。

9. 今後の課題

より一層のブラウザ機能の強化やランキングの見易さの追及などの課題が挙げられる。また、ブラウザを作成してそれを普及させるのは非常に困難であるため、FirefoxのExtensionとして実装することも考えている。更に、個人情報の入力をさせられるページなどの場合、万が一URLに個人情報が入ってしまうことも考えられる。その場合の対策など、セキュリティの面での課題は多い。また、故意にアクセス数を伸ばすようなスパム行為への対策を講じる必要があるだろう。

今回は集中サーバを置いてのシステムであったが、P2Pでの実装を行うことで、さらに匿名性が高く、サーバを必要としない実用性の高いシステムにすることができるだろう。

10. 参考文献

- [1] Visual C#ホームページ:
<http://www.microsoft.com/japan/msdn/vcsharp/>.
- [2] Lunaspape:
<http://www.lunaspape2.com/>.
- [3] blogWatcher:
<http://blogwatcher.pi.fitech.ac.jp/>.