

Webブラウザにおける時間情報を考慮した 履歴機能の検討と実装

矢野 慎一郎

YANO Shinichiro

(指導教官 田中 二郎)

Abstract: Recently, the World Wide Web (WWW) has developed rapidly and the amount of daily information which we get from WWW has been increasing.

Current commercial web browsers such as Microsoft Internet Explorer and Netscape Navigator have functions to make it easier for users to revisit previously seen pages on the web. One of such functions is the bookmark system which enables users to mark favorite or valuable pages with bookmarks. Another is the history system which records locations of seen pages automatically and users can search history with items such as a date.

However, certain studies have proven that the history system is not used frequently in usual life. We think the reason for this matter is that the existing system has insufficient capacity as real history system.

In this thesis, we pay special attention to a better history system which can take advantage of time information each web page has, and developed a prototype system that the users can search and browse all information seen before on the WWW.

Keywords: World Wide Web(WWW), Browser, History, page revisitation, chronological navigation

1 まえがき

インターネットが成長するにつれて WWW (World Wide Web) は情報媒体の主流になりつつある。多くの人々が報道、エンターテインメント、教育、ビジネス、さらには親しい仲間同士でのチャットや掲示板などでの交流といった目的に利用するケースが増えている。

このように WWW が巨大化していくにつれてユーザーがそこから日々得る情報量も増えてきているのだが、それと同時に、ユーザが過去に得た情報を整理・活用する状況も増してきている。

過去に閲覧した情報に立ち戻る (page revisitation) 手段としては以下の 2 つの主な機能がある。

- ・Bookmark (栞) : 訪れたページがまた戻る価値があると考えた時に、ユーザーの命令によって場所を記憶しておくことができる。
- ・History (履歴) : ユーザーが訪れたことのあるページを記録・検索、各種の並び替え方法によって一覧表示することができる。

本研究では、これら二つの機能のうち後者である履歴に着目した。既存の Web ブラウザにある履歴機能は、過去の調査結果からその利用率が低いことがわかっている [1]。

そこで我々は、履歴の利用率が低い原因は履歴としての機能が不十分であるからだと考え、現在の履歴機能が持つ具体的な問題点を調査した。その結果、既存のままではユーザの履歴情報を最大限に活用することが困難なシステムであることが明らかになった。

この問題を解決すべく、我々は履歴としての情報をすべて活用し、それぞれの履歴が持つ時間情報を有効に利用することができる新しい履歴機能を提案する。

そして、それらの機能が実装されたシステム "MyHistory" を紹介する。

2 現在の履歴機能とその問題点

WWW の情報を閲覧するためのソフトウェアに Web ブラウザと呼ばれるものがある。これは様々な

ものが開発・公開されているが、現在の事実上の標準となっているものは Microsoft Internet Explorer(以下 MSIE) と Netscape Navigator(以下 Netscape) の二つであるといえる。

2.1 Page revisitation

Web ブラウジングをしていて、以前に訪れたページをまた見たいと思ったことはないだろうか? 例えば、日頃よく訪れるある新聞のサイトがあるとする。このサイトはその性質上、毎日ページ内の情報が更新されているのだが、過去に一度見たページを再度見たいと思ったとき、ユーザがとり得る行動はおおよそ次のようにあげられる:

1. そのページに対してブックマークをつける

これは Web ブラウジングに慣れているユーザにとっては簡単な行動であるが、そうではないものにとってはなかなか難しい。それに何気なくブラウジングしている最中に、そのページの後々の必要性に気付くことは稀であるだろう。

また、そのページ自体が更新されてしまっただけの方法はほとんど無力な行為になってしまう。ブックマークとは、そのページの内容に栞をつける作業ではなく、その場所 (URL) を記録する行為だからである。

2. サイトが提供するバックナンバーを利用、もしくは検索サイトを利用して過去に見た情報と同じものを探す。

これはユーザが多大な労力と時間を費やす結果になりうる事である。さらには未知の情報の中から探し出すという心理的な負担を増やすことになる。

3. ブラウザに付属の履歴機能を利用して、過去訪れたページから見つけ出す。

これは本来、賢明な方法であると思われる。なぜなら、ユーザが欲しい情報とは一度見たことのある情報の記憶がきっかけとなっているので(逆を言えば、一度も出会ったことのない情報についてはユーザが思い出すことすらないのだから)、ユーザがたどった履歴を絞って見れば一番理想的な情報が見つかるはずである。また、

その方が新たに WWW 上の莫大な情報の中に迷い込み探し出すといった作業より数段と効率的なはずである。

しかし、Tauscher らの調査によると、ブラウザの履歴機能が利用される割合は Web ナビゲーション全体からみても 1% と低く、またブックマーク機能の利用率と比べても 5 分の 1 とかなり低いものとなっている [1]。また、この原因については履歴機能というものの自体が、長い間ユーザから見えないメニューの奥深くに隠されていたせいだと分析している [6]。

2.2 Netscape の履歴機能

Netscape に付属の履歴機能はツールという形で Web ブラウザ部とは別ウィンドウにして使用する (図 1)。過去に訪れた情報はリストとして一覧表示される。また、一覧はタイトル、URL や訪問日などによって並び替えることができる。検索機能ではタイトル・URL から、そして初回・最終訪問日からそれぞれ文字列、日付を指定することでページを絞り込むことが可能である。

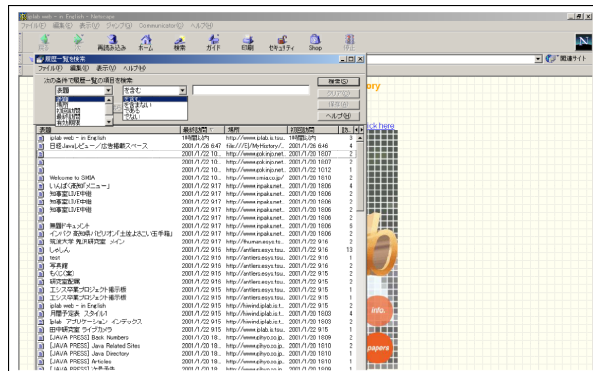


図 1: Netscape の履歴

2.3 MSIE の履歴機能

MSIE の履歴機能は、Netscape のそれに比べるとシンプルになっていて、Web ブラウザウィンドウに組み込まれる状態で使用される (図 2)。履歴の情報一覧は、今日表示したページ・サイトを訪れた回数

順によってリスト表示、または日付順・サイト順によってツリー表示される。検索部に文字列を入力することでタイトル、URL、またキャッシュされているページのデータよりページを絞り込むことができる。

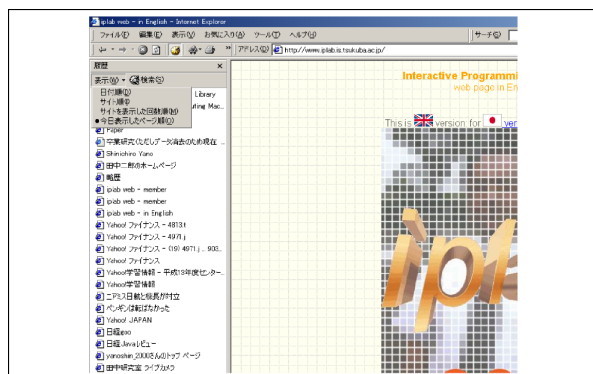


図 2: MSIE の履歴

2.4 現在の履歴機能が持つ問題点

以上より、Web ブラウザの履歴機能にある問題点についてまとめる。

1. 保存されるページの情報量に限界がある

最近に訪れたページは、一時ファイル（キャッシュ）という形でコンピュータ内に保管される。キャッシュは元々、履歴のためというわけではなく、再度そのページを訪れた時に短時間で表示できるように工夫されたものである。Netscape の履歴機能はキャッシュデータを積極的に利用してはいないが、MSIE の履歴機能では検索文字列をキャッシュされたページの本文中からも探し出すことができるようになっている。しかし、結局それは一時保存されたデータであり、ユーザによって設定されたある一定期間を過ぎると自動的に消去されてしまう。

2. 履歴としての時間情報が消去される

先に述べたキャッシュに関係することなのだが、更新の早いサイトを何度も訪れる場合、キャッシュデータは最近の情報に書き換わってしまい、それ以前に見た情報を検索することが不可能となる（図 3）。履歴という機能を充実するために

は、ユーザが今まで経てきた全ての情報を参照できることが望ましい。



図 3: キャッシュデータの消去

3. 検索機能が理想的ではない

ブラウザの履歴検索機能が持つ検索条件は、ユーザが利用する視点から見て無駄な部分が多いと考える。

Netscape を例に挙げてみると、ユーザが欲しい情報を絞り込むのに、「最終訪問日は〇月×日」で「このページのタイトルには～という文字を含まない」、さらには「訪問回数は△回のページ」などという条件で検索することはありえない。さらに MSIE では、キャッシュの設定の関係で検索対象である本文が消去されてしまい、ユーザのすべての履歴から結果を得ることができなくなってしまう。

4. 一覧表示された時にどのページかわかりづらい

訪れたページの一覧表示をすると、URL や訪問日などの文字列によってそれぞれのページの属性が表される。しかし、たとえば URL での一覧を考えた場合、ユーザにとって認識できるのはそのページのドメイン名（例：www.iplab.tsukuba.ac.jp/）までが大半であって、その先のファイル名まで記憶するのは困難である。また、タイトルでの場合を考えると、よくあるタイトル名（例：無題ドキュメント、FAQ）が並んでしまえば、どれがどのページであるのか区別がつかなくなってしまう。

3 時間情報を考慮した新しい履歴システムの提案

3.1 HDD の容量に関する考察

現在の履歴機能で不可能だった、Web ブラウジングで得た情報を全て貯めこむということは現実には可能なことであるかを考えてみたい。

まず、人はいったいどれほどのファイルサイズの情報を WWW から得ているのであろうか？

これについては著者の所属する筑波大学田中研究室において、調査をおこなってみた。調査の方法としては、研究室内の Proxy(代理) サーバ Squid のアクセスログを解析した。この Proxy サーバを利用していた 5 人が一日に WWW 上からダウンロードした情報量についてその合計と平均を割り出した(表 1)。

表 1: 一日にダウンロードされたファイルのサイズ

人物	一般的なページ	全ファイル
A	2.7	5.1
B	6.6	7.5
C	0.2	3.5
D	0.4	0.4
E	4.6	5.2
合計	14.5	21.7
平均	2.9	4.3

注. *単位 MB

一般的なページとは…HTML テキスト+画像ファイル (Jpeg,Gif)

全ファイルとは…一般的なページと動画, 音声,PDF ファイル等を含む

今回の調査では、ダウンロードされた情報を一般的なページと全ファイルの二種類に分けてそれぞれを計算した。ここで 一般的なページとしたのは、いまだに多くの Web ページはテキストで書かれた情報と JPEG・GIF 形式の画像ファイルによるものが多いという理由からである。

次にこれらの調査結果から、Web ブラウジングで得

た情報を貯め込むには、年間でどれぐらいの容量が必要か考えてみたい。

表 1 から一日に Web ブラウジングで得る情報のサイズは人によって差があることがわかるが、ここではその平均として一日 2.9MB であるとした。これより、年間を通してどれぐらいのサイズの情報を得るのか計算してみると

2.9MB x 365 日 約 1.1GB

これは、後者の全ファイルで計算したとしても年間で約 1.6GB である。この結果は、長い間 Web ブラウザの履歴機能が不十分であった原因のひとつとして考えることができるだろう。つまり、インターネットが一般に普及し始めた 1995 年当時のパソコンに搭載されたハードディスクの容量は最上位機種であってもせいぜい 1GB であったからである。現在は約 80GB のハードディスクが実売価格で 27,000 円 (2001 年 1 月 東京, 秋葉原) という時代である。数十年間分の Web ブラウジングで得た情報を貯めこむことが可能なのではないだろうか。

もちろん時代が進むにつれて、Web ページのファイルサイズ自体が大きくなると考えられるが、それでもハードディスクの記憶容量の拡大とは数段の違いがあることは明らかである。

3.2 時間情報を考慮した履歴システム

Time-Machine Computing ではデスクトップの作業状態を蓄積・再現することを可能にする [7, 8, 9]。ここでは、その考え方を履歴機能に応用してみたい。

いままでの履歴機能では、履歴の記録はそのページの場所 (URL) が基準であった。そのため、「訪れたことのあるページ」は記録されても「いつ訪れたページ」という情報が意味をあまり持っていなかった。

Web ページは生きているようにその内容を日々変化させるものが多い。履歴としての機能を実現する上で、同じページであっても、そのページには「時間的な厚み」があることを無視することはできない。

そこで、タイムマシンのようにそのページを時間移動することができ、過去に見たページそのままの情報を再現できるシステムがあるべきではないかと考える。

3.3 解決策の提案

2.4 節で既存の履歴機能の問題点をまとめた。本論文では、この中でも特に、

1. 保存されるページの情報量に限界がある
2. 履歴としての時間情報が消去される

の二点の解決に重点を置き新しい履歴システムを考える。

具体的には、3.2 節、3.3 節で述べた次の 2 点を備えたシステムを提案する

- ユーザーの訪れたページは全て貯め込むシステム
- 時間情報を考慮した履歴システム

前者により、現在の履歴機能では不可能だった、設定していたキャッシュ量以上の情報、またはキャッシュの書換えによって失われていたさらに過去の情報も全て有効に利用することが可能になる。

さらに後者によって、Time-Machine Computing[7, 8, 9] の考え方を履歴機能に応用することで、それぞれの Web ページに存在する時間情報を有効に活用することも可能となる。

4 システムの実装

現在、プロトタイプ”MyHistory”を開発中である。

4.1 開発言語について

本システムは Java を用いて開発を行なった。その理由としては、ネットワーク関連の機能が充実していること、データベースとの連携が容易であること、そして GUI の作成がしやすいことからである。

4.2 システムモデル

開発するシステムの概要は図 4 である。大きく、

- ・履歴サーバ部…ユーザーが WWW にアクセスをすると、その情報を記録する
- ・データベース部…履歴情報が蓄えられる

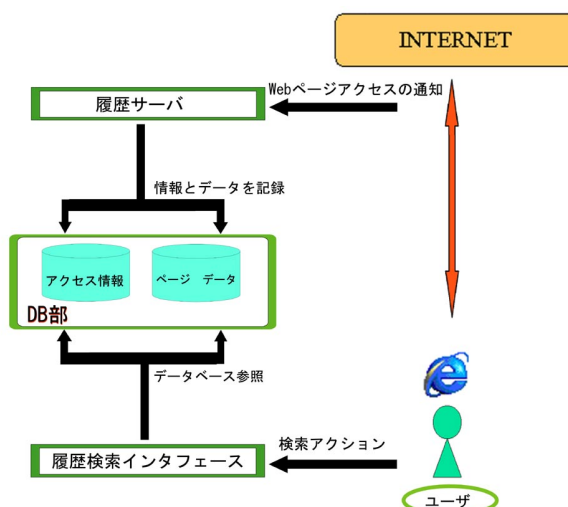


図 4: システムのモデル図

- ・履歴検索インタフェース部…ユーザーの操作により、蓄えられた履歴情報から情報を抽出する

の 3 つに分かれている。

4.3 履歴サーバ部

ユーザーはインターネット環境に接続するとき、履歴サーバ（図 4 左上部）を経由して WWW を利用する。

これは Proxy サーバの仕組みを応用したもので、Web ページへのアクセスが発生すると、WWW からダウンロードしたページの情報がデータベースに記録され、それと同時に見たページのファイルがそのまま HDD に保存される。

この仕組みによって、**ユーザーが使用する Web ブラウザに依存しない、純粋な Web 履歴情報全てを貯め込むことが可能になる。**

4.4 データベース部

データベース部（図 4 左中部）は、Web ページにアクセスした情報を蓄える部分と、見たページのファイル自体を HDD に保存する部分から成り立っている。

サーバ部、検索インタフェース部との連携は JDBC-ODBC ドライバが用いられており、Microsoft Access



(左) 履歴検索・一覧ウィンドウ

(右) 履歴ブラウザウィンドウ

図 5: MyHistory の GUI

で作成されたデータベースにユーザの履歴情報を記録している。

4.5 履歴検索インタフェース部

ユーザが履歴情報を利用するためには履歴検索インタフェース部(図 4 左下部)を用いる。

GUIは具体的に図 5 のようになっており、それぞれのウィンドウで以下の機能を提供する。

○履歴検索・一覧ウィンドウ (図 5-(左))

ユーザによって指定された文字列で履歴情報を全文検索、または日付による絞り込をすることができる。日付による絞り込み方として、いつからいつまでといった日付の指定、もしくは過去何カ月以内、何日以内といった範囲指定を可能としている。

検索された結果はタイトルと訪問日によって一覧表示される。これは、2.4 節で述べた履歴の一覧表示の問題を踏まえた結果であり、ユーザにとって認識しにくい URL 表示などは省かれている。また、一覧の並び替えも、タイトル・訪

問日の昇順・降順のそれぞれが選択できるようになっている。

このウィンドウは、図 5-(右) のウィンドウと連動しており、一覧で選択された任意のページを図 5-(右) のブラウザ部に表示することができる。

○履歴ブラウザウィンドウ (図 5-(右))

このウィンドウでは、ユーザが過去に見たページの内容をブラウザ部により閲覧することができる。

またそれ以外に、以下のインタフェースを使用して、表示されているページを時間移動する機能がある。

・ナビゲーションボタン…2 種類の左右矢印ボタンにより、ブラウザに表示されているページの過去・未来方向への移動をする。中央の矢印ボタンでひとつだけ更新前・更新後のページに移動する。また、時計のアイコンつき矢印ボタンで過去・未来方向への自動移動(アニメーション)させることが可能である。

- ・**検索文字列**…入力された文字列で、表示されているページの過去・未来方向への全文検索をする。検索をする際に、左右矢印のどちらかを選ぶことで過去・未来のどの方向に対して検索をするか選択することができる。また、検索に該当したページはブラウザ部に表示される。
- ・**タイムスライダ**…現在から過去までを直線で表現しており、その直線上のスライダを動かすことで任意の時点への時間移動を可能としている。また、どの時点でページにアクセスしたのかが視覚的にわかるよう、スライダの直線に沿って印がつけられている。

5 関連研究

5.1 Web ページの再訪パターンに関する調査

Tauscher らによって、人々はどのように訪れたことのある Web ページに再訪しているのかに関する調査がなされている [1, 2]。

5.2 WebView

Web の再訪支援のためのグラフィカルなインタフェースの研究。この研究では、履歴情報の一覧にページのサムネイルを利用しているほか、その一覧の表示手法 (Hub-and-spoke, Site-map, Spatial, Temporal) によるそれぞれの利点などの調査が行われている [3, 4]。

5.3 PadPrint

ZUI (Zooming User Interface) を用いたグラフィカルな履歴のサイトマップを構築するシステム。ユーザが訪れたことのあるページをアイコン化し、Web サイトの地図 (サイトマップ) を構築する。その地図は、自由にズームインすることで詳細をみたり、またズームアウトすることで全体を見ることができるようになっている [5]。

5.4 ブックマークと履歴の統合

ブックマークと履歴の両機能を統合したシステムのプロトタイプを開発。彼らのシステムでは、ユーザが訪れたことのあるページはサムネイルとタイトルで一覧表示される。さらに、このシステムのユニークな点としてそれぞれのページは優位度という情報を持っている。そのページによく訪れるほどその優位度は高くなり、優位度は検索の際のスライダによるフィルタリングに対応して、検索結果が絞り込まれる。また、ユーザによって明示的に最優位度をつけることで、いわゆるブックマークされたページとしての区別が可能である [6]。

5.5 Time-Machine Computing

コンピュータデスクトップ環境の状態そのものを時間順に管理する方式を提案している。デスクトップ上での作業履歴を蓄積し、またそれを時間移動することにより作業状態の再現を可能とするものである。また、時間情報を視覚化する技法についても数々の試みがなされている [7, 8, 9]。

6 結論と今後の課題

本論文では、既存の履歴機能にある問題点を挙げ、それを解決するシステムの提案を述べた。特に、履歴としての情報を全て利用、さらにそれぞれのページが持つ時間情報を考慮することに重点を置いた。また、このシステムを実装した "MyHistory" を紹介した。

今後の課題として、特に以下の点について進めていく予定である。

- ・ネットワーク環境への対応
- ・個人情報の管理と保護

これらは将来に常時接続環境が普及することを見据えた課題と考えている。

システムがネットワーク上で利用可能になることで、いつでもどこにいても自分の履歴を記録・参照できるようになり、また、利用手段もパーソナルコンピュータのみではなく、PDA といった携帯型端末等からも

可能になり、自分個人だけの情報としてさらに有効に利用することができるであろう。それに関連して、個人が見たページがそのまま保存されることで、他者によってその情報を盗み見できる恐れがあり、その問題について慎重に考えていかなければならないと考えている。

インターネットの発展は近年目覚ましいものがあり、今よりもっと先の状況を見据えて研究・開発を行っていきたいと思う。

7 謝辞

本研究にあたり、貴重なご指導、助言をいただいた田中二郎先生に心深く感謝致します。

同じく、筑波大学 電子・情報工学系 田中研究室のみなさんからはゼミ等を通して数多くのアドバイスをいただきました。特に三浦元喜さんをはじめ、Webグループのメンバーのみなさまには本システムの研究・開発にあたり多くの有益な議論を頂くことができました。この場を借りて御礼を申し上げます。

また、共に励まし合い、時に貴重な助言をくれた友人たちにも感謝します。

参考文献

- [1] Tauscher, L. and Greenberg, S.: How People Revisit Web Pages: Empirical Findings and Implications for the Design of History System. In International Journal of Human-Computer Studies, 47(1), pages 97-138, 1997.
- [2] Tauscher, L. and Greenberg, S.: Revisitation Patterns in World Wide Web Navigation. In Proceedings of the Proceedings of the CHI '97 Conference Atlanta, GA, pages 399-406, 1997.
- [3] Cockburn, A. Greenburg, S. McKenzie, B. Jasonsmith, M. and Kaasten, S.: Web-View: A Graphical Aid for Revisiting Web Pages. In Proceedings of the OZCHI'99 Australian Conference on Human Computer Interaction, November 1999.
- [4] Cockburn, A. and Greenberg, S.: Beyond the 'Back' Button: Issues of Page Representation and Organisation in Graphical Web Navigation Tools.
- [5] Hightower, R. Ring, L. Helfman, J. Bederson, B. and Hollan, J.: PadPrints: Graphical Multi-scale Web Histories. In Proceedings of the UIST '98. San Francisco, CA, pages 121-122, 1998.
- [6] Kaasten, S. and Greenberg, S.: Designing an Integrated Bookmark/History System for Web Browsing.
- [7] 暦本純一.: Time-Machine Computing: 時間指向ユーザーインタフェースの提案. 日本ソフトウェア科学会 WISS'99, pages 55-64, 1999.
- [8] Rekimoto, J.: Time-Machine Computing: A Time-centric Approach for the Information Environment. In Proceedings of the 12th annual ACM symposium on User interface software and technology, pages 45 - 54, 1999.
- [9] Rekimoto, J.: "TimeScope: A Time Machine for the Desktop Environment". CHI'99 late-breaking results, page 180-181, 1999.