Designing for Web Revisitation: Exploiting Structure from User Interaction and Navigation

Natasa Milic-Frayling
Rachel Jones
Kerry Rodden
Gavin Smyth
Anthony Frayling

30 September, 2004

Technical Report
MSR-TR-2004-97

Microsoft Research
Microsoft Corporation
One Microsoft Way
Redmond, WA 98052

Designing for Web Revisitation: Exploiting Structure from User Interaction and Navigation

Natasa Milic-Frayling
Microsoft Research Ltd.
7 J J Thomson Avenue
Cambridge, United Kingdom
natasamf@microsoft.com

Rachel Jones
Instrata Ltd.
62 Kingston Street
Cambridge, United Kingdom
Rachel.Jones@instrata.co.uk

Kerry Rodden[1]
Google Inc
2400 Bayshore Pkwy
Mountain View, California, USA
krodden@google.com

Gavin Smyth
Microsoft Research Ltd.
7 J J Thomson Avenue
Cambridge, United Kingdom
gavinsmy@microsoft.com

Anthony Frayling[1]
i2 Ltd.
Capital Park, Fulbourn
Cambridge, United Kingdom
anthonyf@i2.co.uk

ABSTRACT
High percentage of Web access are visits to pages that the user has already seen in the past. Currently available support for page revisitation, through standard browser features, and suggested improvements have been based on simple history models which do not fully incorporate information about the user's interaction with the Web and the resulting navigation structure. In our research we propose a syntax for parsing navigation history into structural elements that has proven useful for identifying patterns in the history and building features for supporting habitual Web usage. This approach is motivated by our exploratory user observation study which revealed the need for effective revisitation of key pages in the short navigation history and for supporting user in recurring Web activities observed over the long term history. We designed SmartNavigation features that address both scenarios and evaluate their usefulness through a contextual user study. Findings of the study provide further guidelines for designing revisitation support and validate the history model that we advocate.

Author Keywords
Web revisitation, navigation structure, navigation history, interaction, browser features, contextual user study.

ACM Classification Keywords
H.5.4. Hypertext/Hypermedia: Navigation, User issues.

INTRODUCTION
The Web has become an important media for publishing and accessing information. Thus both the designers of Web sites and developers of Web applications are concerned with the effectiveness of features for Web access. From the users' perspective, the importance of the Web is often tied to their success in finding relevant information, typically through search and browsing, and ability to consume that information in a timely manner: immediately, later in time, or often repeatedly, by revisiting the particular Web site.

Research on Web usage has reported astonishingly high rates of visits to pages that the user has already seen in the past [5][15]. Users seem to revisit pages for various reasons, accessing old and new web pages regularly over time [15]. Standard browsers create histories of user navigation and expose them through different interface features in order to support revisitation of past pages (i.e., URLs of these pages). For example, in the Microsoft Internet Explorer (IE) browsers, some of the most recently visited Web pages are accessible through the Back navigation button and its associated drop down menu. Furthermore, the IE History is devised to store and provide access to pages visited in the past through browsing, search, or filtering (by most recent URLs, by date, by site, etc.). In addition, the users can manually bookmark pages for future reuse. These are stored and accessible through Bookmarks in the Netscape browser or Favourites in IE.

[1] Work done while working on a project at Microsoft Research, Cambridge, UK.

All these features are designed to support the user in revisiting a desired page. However, studies of current features have shown that neither the current history models [5][15] nor the bookmarking facilities [1][5][10] are fully effective in supporting the user.

A number of studies suggest alternative history models [6][15] and bookmarking facilities [9][10][14]. Most of the design recommendations were supported by theoretical considerations of history models and their performance on data that was collected through empirical studies. The guidelines that resulted from this research have been used in very few systems (e.g., WebView [4]) and even then the resulting prototypes have not been evaluated in the users' natural setting.

Our own research of Web usage indicates that it is important to recognize transient revisitations of pages that are due to the underlying hyperlink structure of the Web [11]. As the percentage of such revisits can be significant, optimizing user revisitation should aim at reducing the number of transient navigation steps towards the key pages that the user needs to access. However, in order to identify such key pages one needs to enrich the model of navigation history with information about the user's interaction with the browser and the resulting navigation structure. We maintain that this is not only the case with short term history revisitation, as shown in [11], but that such an approach holds the key to effective use of long term history to support users' habitual Web activities.

In this paper we introduce a model of navigation history that incorporates structure of navigation induced by the user's interaction with the Web through the browser. We apply a simple syntax for parsing the user's navigation history into sub-structures, such as navigation sessions and navigation trails, and further exploit them to model the user's recurring Web activities. The immediate benefit of this approach is that we are able to capture recurring patterns beyond simple page revisitations. Indeed, we begin to identify navigation patterns which are a reflection of the user's recurring online tasks. This enables us to design generic features that support such habitual behaviour by complementing the existing browser features.

Our approach is motivated by an exploratory user observation study. Findings of the study informed the design of the SmartNavigation extension of the IE browser that consists of SmartBack [11], SessionOverview, and SmartFavourites features. The first two features address the revisitation of pages within the short term history while SmartFavourites support habitual revisits throughout the history. We conducted a contextual user study [8], focused on usefulness of the implemented features.

In the following section we give a comprehensive overview of Web revisitation research and set the stage for our own research program. We present the findings of the exploratory user study and motivate the structured history model. We follow this with a description of the prototyped features and their evaluation through a four week observational user study in the natural user setting. We conclude with a summary of the main findings.

RELATED RESEARCH

User revisitation of Web pages is one of the prominent characteristics of Web information access, as established through analyses of individual user logs in [5][11][15]. However, the studies also reveal suboptimal use of standard browser features such as Back and Forward, bookmarks, and history, although they were designed to enable return to pages seen by the user in the past. Analysis of existing designs and possible improvements of these features have been addressed in numerous studies: Back and Forward navigation in [6][7][11], bookmarks in [1][9][10][14], browser history in [2][4][15]. We shall here discuss in more detail three aspects of the past research: (1) approaches that exploit revisitation patterns in sequential history models to inform the design of short history revisitation, (2) the use of navigation structure as visual aids for browsing, and (3) studies of standard browser features.

Revisitation Patterns in the Sequential History Model

Revisitation of pages has been consistently observed in user navigation histories: on average 81% of Web visits are revisits to old pages, as reported in the work by Cockburn and McKenzie [7], 61% in the study by Catledge and Pitkow [4] and 58% in the study by Tauscher and Greenberg [15]. Thus, as much as accessing new pages on the Web is important for users, so are ways of returning to pages that they have seen in the past.

Tauscher and Greenberg [15] observed that the distribution of page revisits in the users' histories was highly skewed towards recent navigation. Indeed, there is 43% chance that the next page accessed by the user is among the 10 most recently visited URLs. Pages visited in the more distant past, on the other hand, have 15% of chance of being revisited in the very next step. Thus they consider a number of approaches for re-ordering URLs from the navigation history in order to present a short list of URLs which is most likely to contain the page the user wants to revisit. However, both the motivation and the theoretical evaluation of the proposed methods are based on a simple history model: the observed statistics characterize the history as a simple sequence of URL visits with the goal of predicting the likelihood of the very next revisit.

This model is natural for predicting the next step within a navigation sequence but not for capturing transitions to different Web activities. Furthermore, the effectiveness of the predictions methods is measured by treating all the revisits in the history as 'desirable' and worth predicting, which is not the case. Indeed, it was shown in [11] that the underlying hyperlink structure of Web content and the use of available means for navigation, such as Back button, yield a significant percentage of *transient revisitations* of pages which are not desired but necessary at the moment. For example, in a search scenario the user may need to

retrace a path of 2-3 pages back to the search result page in order to investigate the next recommended search result. Thus, retracing of the navigation path needs to be contrasted with the *targeted revisit* to the search result page in order to exploit its content, e.g., to access the next result on the list or to issue another search request.

Introducing Structure

More sophisticated models of History organization have been exploited in visualizing navigation paths. Research by Ayers & Stasko [2], Cockburn et al. [4], and Milic-Frayling et al. [12][13] considered two-dimensional visualizations of browser sessions, providing good insight into the structure of users' navigation paths. For example, work on the WebView [4] recognizes the importance of hubs and spokes and uses them as the basis for displaying the users' navigation history. Similarly, the visual displays in MosaicG [2] present a tree navigation structure, clearly indicating the branching points in the user's navigation. With each use of a Bookmark this system associates a new navigation tree, thus indicating the onset of a new browsing activity. Similarly, in WebScout [12][13] a new Web trail is started with the user explicit actions: typing in a URL, initiating navigation from another document, such as e-mail, executing a link from Favourites, etc. However, these types of structures have been used to a very limited extent to model the short and long navigation history. Tauscher and Greenberg [15] use context sensitive sub-webs, i.e., navigation sequences which begin by explicit specification of the URL by the user, to create a condensed hierarchical display of the recent navigation history. While this display optimizes the likelihood of including the URLs that will be next revisited, usability aspects of such a solution have not been addressed.

Standard Browser Features

Back Navigation

As the most prevalent method for revisitation is through Back, lots of effort has been put in understanding the user's mental models of Back navigation and benefits of various history lists associated with the Back button. One recurring theme has been the advantages and shortcomings of the stack model for Back button history which retains only pages from the main navigation path [6][7].

In [11] the Back button is complemented by the SmartBack feature, which allows the users to jump directly to key pages in the recent navigation sessions. Such are navigation hubs that serve as branching points in navigation and pages accessed through users' explicit specification of page address, i.e., by typing a URL, executing links from Favourites and history lists, and similar. Building on the SmartBack model of a navigation trail we further expand the syntax for parsing the history to generate useful structures for modeling the long term history and recurring navigation structures.

Bookmarks

Bookmarks have been viewed as an important component in creating users' personal information systems on the Web [1] and reducing information overload from exposure to the vast Web information resources. Although bookmarking is relatively easy and well adopted, it seems that assessing whether or not a page will be useful in the future cannot be done reliably by the user. As a consequence the usefulness of created lists has not been observed in practice.

A number of studies [15][5][14] show that users do not rely on bookmarks to return to pages. As Abrams et al [1] pointed out, maintaining the list or hierarchy of bookmarks creates an overhead not easily tolerated by the Web users. Furthermore, even when neatly organized, accessing a particular URL from bookmarks can be difficult as the list may be long and the hierarchy of bookmark folders may obscure the location of the link. Our study (see section on Exploratory User Study) shows that frequently accessed links are often missing from the Bookmark lists. Thus, there is a space for providing a feature to assist habitual behaviour that is missed by the current bookmark facility because of the inability to predict future use and the clutter that results.

History - User Interface Design and Evaluation

It has also been noted that the standard History feature which stores visited URLs and provides access to them through browsing, searching, and filtering (e.g., on dates, domain names) has not been frequently used: less than 1% of the user's activities observed in [5] and [15] are history access. This has been attributed to a number of user interface issues and led to design recommendations by Tauscher and Greenberg on devising history presentations [15]. They recognize that, although most likely revisits happen in the recent history, 26% of recurring pages are not covered by the most recent 10 URLs nor can they be reached by doubling and tripling the size of the history list that includes pages in the order of recency. "*History based on recency is not effective for all possible recalls because it lists only a few previous events. Alternative strategies must be supported"* [15, p132]. Thus it is necessary to address revisitation in a wider, longer term history context, moving away from recency of page visits as the main criteria for history representation. We agree with the view that "*these missed URLs[absent from the short term history] could help the user most since they occurred long ago and are, thus, more difficult to recall and/or locate"* [15].

We expect that, as search facilities improve, the users will access URLs in the navigation History through search queries. However, we also see an opportunity for expanding the existing modes of interaction with the History by proactively exposing archived URLs to support habitual user behaviour. This has a benefit of no overhead associated with searching or browsing the History and thus minimizing the distraction from the main user's task. This fits with the design principle outlined in [15]: "*It should be cheaper, in terms of physical and cognitive activity, for users to recall*

	Participants										
Type of Revisitation	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**K**	**Total**	
BACK	31	48	69	49	151	325	30	401	55	**1159**	**66.3%**
FORM SUBMIT	2	9	48	6	41	1	12	32	18	**169**	**9.7%**
LINK NAVIGATION	2	6	20	21	15	14	4	78	32	**192**	**11.0%**
FAVOURITE	1	0	2	4	0	3	2	7	0	**19**	**1.1%**
FORWARD	0	1	2	0	0	4	0	4	0	**11**	**0.6%**
HOME	0	11	0	24	0	2	0	9	2	**48**	**2.7%**
REFRESH	0	0	0	1	0	0	0	8	0	**9**	**0.5%**
TYPED	1	0	0	0	0	0	1	6	0	**8**	**0.5%**

Table 1. Breakdown of revisitation statistics for individual participants and different types of revisitations.

URLs from a history mechanism than to navigate to them via other methods" [15, p132].

RESEARCH AGENDA

In relation to the discussed work our research programme is scoped to:

- Enhance the browser to support page revisitation in short and long term history. Identify such opportunities through an observational study of Web usage.
- Incorporate information about the user interaction and resulting navigation structure into the analysis and the model of navigation history. Specify the syntax for partitioning the navigation history into meaningful and useful substructures.
- Revise and expand the objectives of revisitation support: instead of predicting page revisits that might be intended by the user in the very next step [15], explore predictive models to cover different types of recurring user behaviour. These types range from revisits to key pages in the current or recent navigation paths, to the *forward* looking predictions of pages that have been visited recurrently by traversing sequences of navigation steps.
- Create working prototypes for page revisitations, which exploit and expose the navigation structure of the history, to the extent that it aids the user's comprehension and effectiveness of URL access.
- Apply contextual studies to evaluate the prototypes with emphasis on the qualitative analysis of the user's experience and a feature's usefulness.

EXPLORATORY USER STUDY

We began our research program with an exploratory user study conducted to learn about navigation structures exhibited through the user's interaction with the browser.

Study Design

The user study spanned a period of 3 weeks, with 9 participants (A-K, see Table 1) from Cambridgeshire County Council. In the pre-screening interviews most of the participants were characterized as knowledge workers. They used both Internet and County Council Intranet. We observed them through a combination of a quantitative method, comprising automated client-side logging of web usage, and a qualitative approach, involving follow-up interviews with each participant twice a week to verify the quantitative data captured in the log files [8]. The logger recorded events from the Microsoft Internet Explorer browser and captured thumbnail images of the content viewed by the user, except for secure transaction pages and others that participants specified explicitly. The interviews were video taped and used to support data analysis and interpretation. This approach enabled us to capture details of every Web activity, while also gaining an understanding of the contexts in which an activity was carried out. We also supplemented our analysis with a half day observation of the participants' work environments, to understand factors such as their interactions with colleagues.

Analysis

Data from the user study provided us with valuable insights into a variety of Web usage issues. However, we here focus on the characteristics of the users' navigation patterns, as these are of most relevance to the objectives of this research. We analyze the user navigation logs using the following basic concepts:

Navigation Sessions

In contrast to the standard practices we do not tie navigation with the notion of application or windows management. *Navigation sessions are periods of the user's navigation activity that are separated from each other by noticeable periods of user inactivity*. This definition provides flexibility for looking at the navigation history at different granularity levels, by increasing or reducing the required idle time for session boundary.

Web Trails

As in [11][12][13] we use the concept of a *web trail* to designate a sequence of navigation steps consisting of link executions and Back and Forward navigations (Figure 1). The trail is initiated by explicit specification of a URL, e.g.,

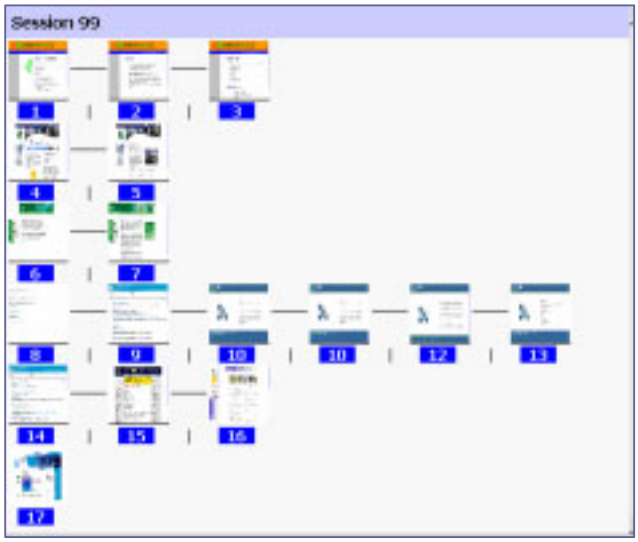


Figure 1. Example of multiple linear navigation trails within a single navigation session.

by typing a URL into the address bar, or by selecting a URL from bookmarks, navigation history or another document, such as an e-mail message containing a link.

Navigational Hubs

Of particular interest are trails which involve revisitation of pages. They may contain a branching point, i.e., a page to which the user returned in order to branch out and explore alternative navigation paths. We refer to such a page as a *navigational hub.*

Navigational hubs are results of the user's navigation patterns and may or may not coincide with the *information hubs*, characterized by high connectivity with other Web pages (although, by definition, it should have at least 2 outgoing links).

The above three concepts comprise the basic syntax for parsing user navigation history that we proposed in order to study and design features to support for Web revisitation. We expect that further expansions and refinements of this syntax would be necessary to study a wider spectrum of user's interactions with the Web.

Findings

Similarly to the earlier logging studies [4][5][15], we find a high level of page revisitation. Much of that is use of the Back button. We observe that most of the navigation trails are linear (Figure 1), involving only forward link navigation. Out of total 4,038 trails, 339 are graphs, i.e., 8.3% contain hubs. Furthermore, out of 5,235 page visits 1,133 were visits to hubs and out of total 1,211 back navigations, 775 (64%) are visits to hubs. Thus, the remaining 436 (36%) of back navigation steps are simply traversing the branches of the graphs. Designating hubs in the short term history presentations should, therefore, help the user reduce the number of transient visits [11].

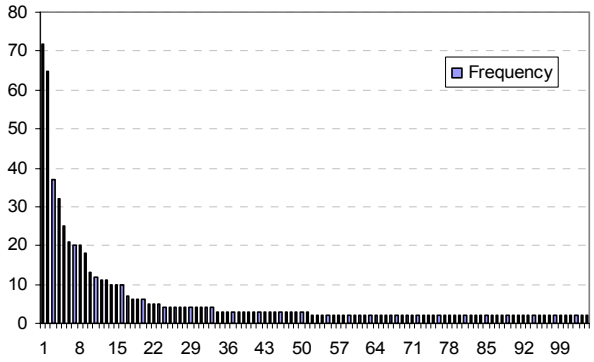


Figure 2. Distribution of URL visits for the participant F. It is evident that very small number of URLs is visited very frequently while very large number is visited quite infrequently.

Participant F

Visits	Sessions	URLs
72	31	http://camweb/
65	12	http://www.guardian.co.uk/
37	10	http://intranet2.camcnty.gov.uk/cambscc/pressc
32	8	http://www.guardian.co.uk/worldlatest/0,7722,3
25	9	http://camweb/staffbulletin/index.cfm
21	10	http://intranet2.camcnty.gov.uk/cambscc/pressc
20	9	http://intranet2.camcnty.gov.uk/db/NewsPool.n
20	6	http://www.guardian.co.uk/uklatest/0,7721,362
18	8	http://camweb/teldir/results1.cfm
13	8	http://camweb/teldir/index.cfm
12	10	http://www.blackstar.co.uk/
11	3	http://www.scambs.gov.uk/Scambs/council.nsf/
11	1	http://intranet2.camcnty.gov.uk/external/resour
10	8	http://camweb/businessinfo/index.cfm
10	6	http://www.google.co.uk/

Table 2. Most frequently accessed URLs by the participant F. The table shows three columns: the frequency of visit, number of sessions in which the URL is visited, and the URL string.

In terms of mechanisms used to return to other pages, the participants were comfortable with typing in URLs, especially with the assistance of IE auto-complete feature. They were also able to edit them appropriately in order to navigate, e.g., from an individual page to the site home page. Only three participants made regular use of bookmarks for returning to pages, and many felt guilty about not using them more. People tend to bookmark site home pages, but these are not usually the final destination; we observed a number of cases in which participants followed a predictable trail after initially clicking on a Favourite. Frequently accessed URLs tend not to be included in the Favourites.

	Number of observed repeated sequences per user								
Length	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**K**
2	8	9	15	19	23	32	6	74	12
3	2	6	7	8	4	6	4	24	7
4	0	1	5	1	4	3	2	3	4
5	0	0	4	1	1	0	2	3	2
6	1	0	0	1	0	0	0	1	0
Total	11	16	31	30	32	41	14	105	25

Table 3. Users traversed the total of 107 sequences of the length 3 or more, associated with 30 distinct URL sequences.

Seq. 1.: Length 3, frequency 8
http://camweb/ http://intranet2.camcnty.gov.uk/cambscc/presscut/presscut.nsf/index.htm http://intranet2.camcnty.gov.uk/cambscc/presscut/presscut.nsf/cuttings/today?Openview&count=50
Seq. 2: Length 3, frequency 2
http://camweb/ http://camweb/businessinfo/index.cfm http://camweb/staffbulletin/index.cfm
Seq. 3: Length 4, frequency 3
http://www.cambridge.gov.uk/services/planapps.htm http://www.cambridge.gov.uk/services/develop.htm http://www.cambridge.gov.uk/planning/devcont.htm http://www.cambridge.gov.uk/planning/dccontac.htm

Table 4. Example of repeated sequences of forward link navigations for the participant F.

A closer look at the URL vocabulary confirms observations from the previous studies. The users visit a relatively small set of URLs frequently and a large number of URLs infrequently (Figure 2). Furthermore, users continuously revisit old and access new URLs, i.e., both revisitation of old pages and growth of vocabulary due to newly visited sites are persistent throughout the navigation history. As an example, we show in Figure 2 and Table 2 the distribution of URL visits for the participant F.

Users do retrace their steps, in the form of repeated sequences, in order to access pages. A number of such sequences is relatively small compared to the total number of visits and navigation trails. As Table 3 shows, in the users' navigation logs we found 107 repeats of navigation sequences of length three and more, associated with 30 distinct page sequences. Several examples of such sequences for the participant F are shown in Table 3.

Opportunities

This analysis indicates that there is potential for novel ways of exposing URLs from the History in a context sensitive manner:

Figure 3. SmartNavigation Toolbar showing the SmartBack (JumpBack) button.

Figure 4. Multiple SmartBack links are shown in the SmartNavigation toolbar, allowing the user to jump back directly to key pages accessed in the recent past.

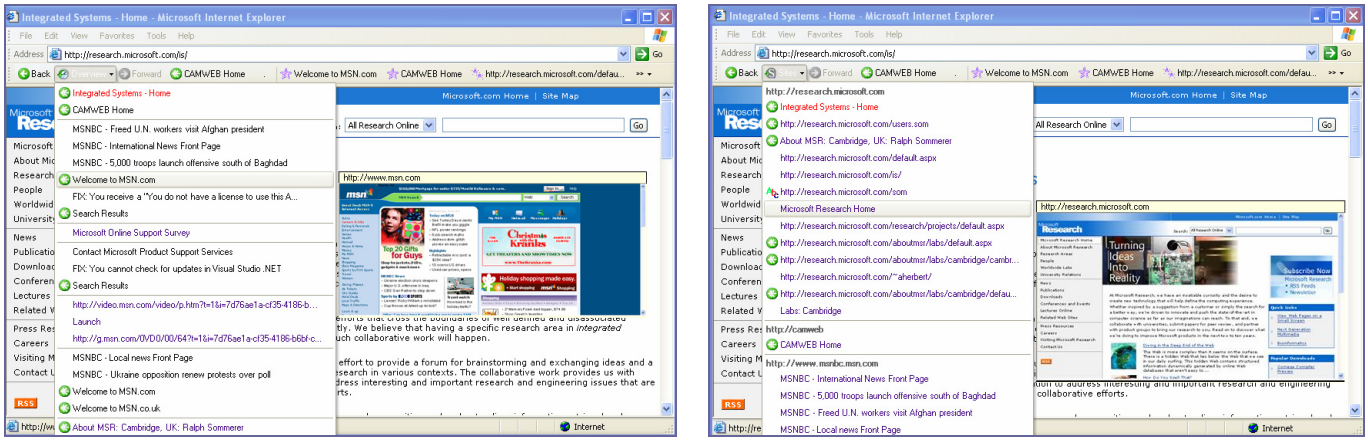
- Present to the user the URLs which have been frequently and recently visited. Select those that map onto the user's main activities, i.e., correspond to the beginning of navigation trails.
- In relation to the current navigation trail, present URLs that correspond to other on-line activities that the user tends to perform within the same navigation session.
- In relation to the currently viewed page, consider the past web trails that contain that page, Provide easy access to URLs of pages that often occur in those trails. This will typically capture pages within the same site that the user has seen in the past. It will also capture repeated sequences and thus reduce the number of steps to the pages that the user obviously perceives valuable.

SMART NAVIGATION FEATURES

The findings of the study motivated implementation of three features for Web revisitation. Each exploits information about the user interaction with the browser and navigation structure at a different granularity level: at the level of navigation sessions, navigation trails, or navigation branches, as seen in repeated navigation sequences.

SmartBack

The features allows the user to jump back to the key pages in the recent navigation sessions instead of using multiple Back button clicks to reach such pages [11]. The following types of pages are automatically designated as SmartBack targets: typed in or auto-completed URLs, URLs accessed from Favourites, and navigational hubs. In order to support easy revisitation of results pages in the search scenario, we



(a) Temporal order (recent first) **(b) Site based grouping of pages**

Figure 5. SessionOverview displays with the pages from the recent navigation history

also mark as key pages all those that are accessed as a result of a *form submission* (e.g., a search result page). In addition to automatic selection of SmartBack targets, the user could also designate a page to be accessible via SmartBack.

SmartBack targets can be accessed by a click on the SmartBack button in the SmartNavigation Toolbar (Figure 3) or presented as a link, showing explicitly the title and, on mouse hover, providing the user with the thumbnail image of the target page. In fact, multiple SmartBack links can be presented, allowing the user to quickly access key pages from the recent navigation history (see Figure 4). Note that SmartBack is persistent across browser windows, unlike the standard Back.

SessionOverview

This features essentially implements the short term history models discussed in [6][11][15] as a list of pages that have been seen recently during navigations, and presented in the order of page visit (most recent first) or grouped by the domain name (see Figure 5). The temporal list (Figure 5a) can be used with various levels of compression by eliminating duplicate URLs and keeping the first or latest access to the page. It includes pages from multiple windows and multiple browser instances, thus providing a single point of access to all recently seen pages. Thumbnail images of pages are displayed on mouse hover over the titles; SmartBack targets are designated in the list by the SmartBack icon.

In the temporal display we show elements of the navigation structure very subtly: we introduce demarcation lines that indicate the start of a new trail or a new branch in the navigation tree. We also use different font colour (appearing as gray or light blue in Figure 5a) for pages that are seen in the past or peripheral browser windows that may be open at the same time. Display is context sensitive and shows in red the page that is currently viewed in the browser window.

We also provide an alternative view of the recent navigation history by grouping pages by domain name (Figure 5b). We refer to this history view as *Sites*. The pages are shown as associated with the domain name, in the reverse temporal order in which the sites were encountered during navigation. SmartBack targets are still designated although the organization of pages is not based on the fine level navigation sequence.

SmartFavourites

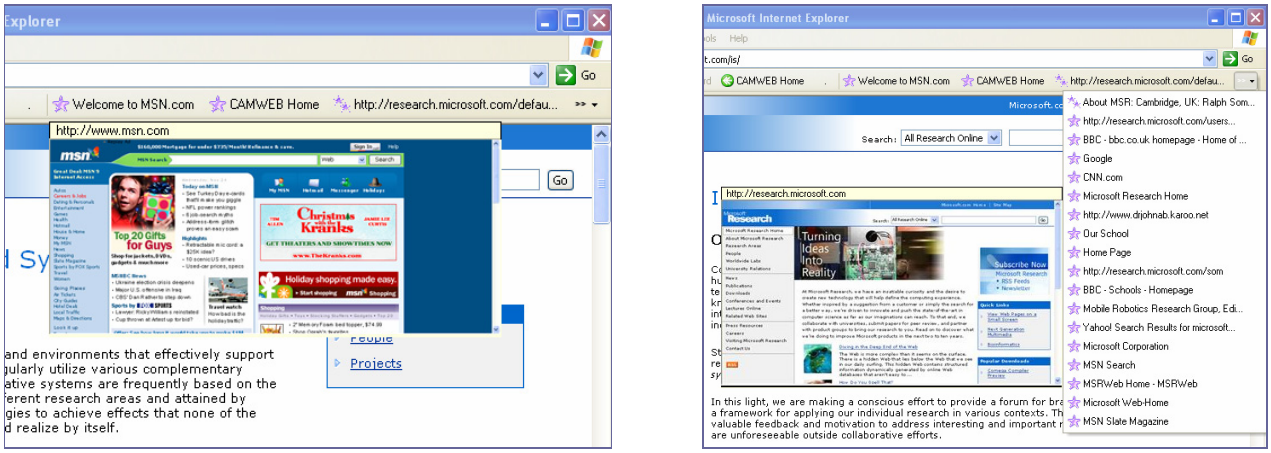
This feature is implemented as a sequence of links in the SmartNavigation Toolbar (Figure 5), using the *Overflow* drop down menu to the right, to show titles of additional pages. The main objective of SmartFavourites links is to capture and proactively expose links that the user has periodically visited in the past.

Several types of pages are presented:

(1) *TrueFavourites* - pages that have been visited frequently in the past, with time decay factor built into the algorithm to promote the pages that have been seen periodically in more recent past. In the calculation of TrueFavourites only the starts of trails are considered, assuming that navigation trails are very much a reflection of the user's habitual online activities.

(2) *Predictions* - pages that have been often visited in the same navigation session with the current page.

We implemented three algorithms that capture co-occurrence of pages at different granularity levels. In the first algorithm we take into account only trail start and special pages, such as search results. We present to the user those trail start pages that in the past navigation sessions co-occurred with the current trail start page. In the second and the third algorithm we exploit co-occurrence of pages within trails. More precisely, the second algorithm considers whether two pages appear somewhere in the same trail while the third algorithm looks at the 3-page sequences



(a) SmartFavourites links on the Link bar, with a page thumbnail **(b) Overflow menu with links and a page thumbnail**

Figure 6. SmartFavourites are presented in the Link bar, and change based on the currently viewed site. Thumbnails serve for disambiguation of links with the same title or those that are not easily remembered by the user.

that contain the currently viewed page and exposes the end-of-sequence page for easy access.

We present the user with a pre-defined but configurable mixed display of TrueFavourites and Predictions. TrueFavorites and those Predictions based on the first algorithm are designated by large purple star icons as they indicate start-of-trail pages that the user may want to visit after completing the current navigation trail. Predictions based on the other two algorithms suggest the pages that have been seen in the navigational proximity of the currently viewed and as such are designated by two smaller purple stars. These are the pages that the user may want to explore within the current navigation trail. For all the links on the Link bar and within Overflow menu, thumbnail images are displayed on mouse hover over the page titles.

EVALUATION THROUGH CONTEXTUAL USER STUDY

In this section we describe the contextual user study [8] that we conducted to assess the usefulness of the revisitation support provided by SmartNavigation features.

Study Design

The study involved participants in a workplace and a family (Table 6). The workplace chosen was a Chartered Accountancy Firm with 7 employees, including the owner. The family comprised 4 members: Mum, Dad, Boy of 16 and Girl of 11. This study lasted about 4 weeks and involved detailed logging and three interviews with the participants at the beginning, after approximately two weeks, and at the end of the study.

SmartNavigation Toolbar, including SmartBack, SessionOverview and SmartFavourites, was installed on the participants' computers and its use was logged for each participant separately. As the prototype is configurable and allows exploration of various history and prediction models we decided to change the configuration after the second week to receive feedback on different design options. The Back navigation model was held constant during the study, configured to eliminate duplicate visits and retain only the latest visit to a page. Similarly, the SmartBack was exposed as SmartBack button rather than a SmartBack link with the thumbnail preview. It enabled the users to jump back directly to the navigational hubs, search result pages, and starts of navigation trails. SessionOverview was configured for the same history model as Back and thus can be considered as an alternative user interface for the same short history model. In the second two weeks of the study we extended the SessionOverview to persist the history across windows and, as such, differ from the Back history which is constrained to individual windows. We did not evaluate the site based organization of the Overview at this time (Figure 5b).

SmartFavourites features initially included TrueFavourites and Predictions exposing start-of-trail pages only. In the second part of the study we included the pages that were predicted based on all four SmartFavourites algorithms.

Study Findings

We observed that participants used different combinations of features, depending on the type of Web activities and navigation habits (Table 5). Quantitative analysis of the navigation logs shows that the use of new features represents a significant portion of navigation activities, if we exclude the prevalent hyperlink execution and simple viewing of the Session Overview menu. From the statistics in Table5, one can see that the use of SmartFavourites through the Links Bar and Overflow menu (30.5%) is on the par with the use of the standard Back (33.5%) and much higher than the use of Favourites (7.3%). Table 5 shows relative usage of features for the sample of workplace and home participants separately. Here we refer to the statistics obtained by averaging over all the participants.

Access to the short term history through SmartBack (14.2%) and SessionOverview (5.3%) also amounts to a

	WORKPLACE						HOME			
WORKPLACE	**SIM**	**JAM**	**LYN**	**KAY**	**MAR**	**Total**	**ANN**	**DON**	**RIC**	**Total**
SmartBack	63	29	3	5	10	**110 (15.5%)**	5	57	25	**87 (12.8%)**
Overview Open	34	18	15	31	14	**112**	15	60	30	**105**
Overview Click	7	2	6	17	6	**38 (5.4%)**	5	26	4	**35 (5.2%)**
TrueFavourites Link bar Click	56	1	48	1	1	**107 (15.1%)**	14	48	137	**199 (29.4%)**
TrueFavourites Overflow Click	32	1	13	12		**58 (8.2%)**		3	33	**36 (5.3%)**
Prediction Link bar Click	2		2			**4 (0.6%)**		3	14	**17 (2.5%)**
Prediction Overflow Click									2	**2 (0.3%)**
Favourites	8	24	2			**34 (4.8%)**		2	65	**67 (9.9%)**
Back	55	25	68	59	36	**243 (34.2%)**	22	188	12	**222 (32.7%)**
Forward	12	1		9	1	**23 (3.2%)**		6	3	**9 (1.3%)**
Typed	45	5	31	12		**93 (13.1 %)**		2	2	**4 (0.6 %)**
Link Navigation	1007	208	1084	189	118	**2606**	138	859	2100	**3097**

Table 5. Statistics on usage of individual features by the participants observed at their workplace. Percentages in the Total columns are determined with respect to the sum of navigation steps excluding Link Navigation and Opening the Overview menu.

WORKPLACE					HOME		
SIM	**JAM**	**LYN**	**KIM**	**MAR**	**ANN**	**DON**	**RIC**
Employee	Bus Owner	Employee	Employee	Employee	Mother	Father	Boy
(Age 30) Accountant Technical	(Age 48) Accountant Non-Technical	(Age 29) Accountant Non-Technical	(Age 28) Trainee Accountant Non-Technical	(Age 26) Accountant Non-Technical	(Age 42) Secretarial Non-Technical	(Age 44) Engineer Technical	(Age 16) Technical

Table 6. Demographic details about the participants and simple classification of technical proficiency based on the initial interview and a questionnaire.

large proportion of page accesses when compared to the use of standard Back button (33.5%). We consider this a good adoption rate by the users and attribute it to the usefulness of the features, demonstrated in the user's natural setting.

From the interviews, we obtained further valuable feedback on the feature usefulness and design. We first summarize the observations related to each feature and then point out more general issues that should be taken into consideration in future redesigns.

SmartBack

- SmartBack is useful in hub and spoke navigation, specifically for navigating search and portal results (Sim, Jam, Don).
- Some users who are adept at using SmartBack do not understand how a hub is created (Sim, Don).
- Users for whom SmartBack does not demonstrate its usefulness within a couple of tries, give up using it.

Note, the alternative user interface that uses SmartBack link with the full title of the target page and associated thumbnail image is expected to address the uncertainty implied in the simple button interface.

SessionOverview

- This feature is seen by participants as a way to get to *sites* that they want to revisit (Ann, Mar, Kim, Lyn and Don). It is also seen as complementary to the Overflow menu of SmartFavourites as that is a list of most commonly visited sites, whereas Overview is a list of most recently visited sites.

- Two participants see it as a place to find pages they had been to before, but others would re-navigate from a site they immediately recognize in the Overview list.

As a result, it might be better to provide a list of sites, without intermediary pages. Our alternative Sites display (Figure 4b) may also offer a good alternative. Generally, the Overview seems to compete with History, Drop-down back, Back, SmartBack, Favourites and TrueFavourites, as it is sometimes used in place of these features.

SmartFavourites

Users were asked about the usefulness of TrueFavourites and Predictions separately. From the participants' logs and interview feedback we conclude that:

- The benefits of TrueFavourites are that they are visible, created automatically and change over time. They are useful to people who go to a regular set of sites often, who act quickly (or have little patience), and who engage in project-based activities (Sim, Lyn, Don, and Ric).
- TrueFavourites are not useful for people who use the Favourites sidebar and have only about 15 Favourites. However, for many users TrueFavourites complement the existing Favourites.
- TrueFavourites could be improved by enabling the user to add and remove links, by making them a Favourite, and by making more of them visible. They compete with the auto-compete and Favourites.
- Predictions worked for people who go to a series of sites regularly, which in our studies were people who use forums or message boards (Lyn and Ric). There was a concern that this feature appears "Big Brother-ish".

Thumbnails are viewed as useful to distinguish between pages with the same title, but some participants would not wait for them to appear.

Design observations

- Accessibility and convenience

  From interviews with Lyn we learn how convenient the highly accessible links can be. When asked, if she uses True Favourites or types in the URL, she says, "If I can see them on the buttons, I'll use the buttons, because that's quicker, isn't it."

- Usefulness vs. privacy concerns and trustworthiness

  Lyn has also used the Prediction links on non-work browsing and says they are "quite useful" but also feel "Big-Brother-ish", as they know well what sites she has been visiting..

- Accessibility and behaviour reinforcement

  A spontaneous comment from Don raises the issue about behaviour reinforcement induced by the exposure of history through a highly accessible interface: "I never used this link [Weather forecast for Cambridge] but it showed up on the links and now I am using it everyday, just because it is there".

SUMMARY

Our work on Web revisitation builds on a comprehensive body of research that precedes it. We are extending it by introducing and exploiting a richer model of the user's navigation history. That model incorporates basic concepts of navigation sessions, web trails, and navigational hubs, defined by taking into account information about the user's interaction with the browser (typing in URLs, executing search, etc.) and the structure of the user's navigation that is induced by the underlying hypertext organization of the Web content.

We demonstrate how the model is used to build novel features for supporting revisitation of Web pages from the short and long term history. We evaluate the usefulness of these features through a contextual user study, by observing their use over an extended period of 4 weeks, on users' own tasks and within their natural setting. Our quantitative and qualitative findings justify further use of the proposed model. At the same time they point to interesting privacy, accessibility, trustworthiness, and behavior reinforcement aspects of the current designs. In our future work we shall aim to understand them further and address them adequately in subsequent iterations of features redesigns.

REFERENCES

1. Abrams, D., Baecker, R. and Chignell, M. Information Archiving with Bookmarks: Personal Web Space Construction and Organization. In *Proceedings of CHI'98* (1998), 41-48.
2. Ayers, E.Z., and Stasko, J.T. Using graphic history in browsing the World Wide Web. *Proceedings of the Fourth International World Wide Web Conference* (1995).
3. Catledge, L., and Pitkow, J. Characterizing browsing strategies in the World Wide Web. *Computer Networks and ISDN Systems* 27(6) (1995), 1065-1073.
4. Cockburn, A., Greenberg, S., McKenzie, B., Jasonsmith, M. and Kaasten, S. WebView: A graphical aid for revisiting Web pages. *Proceedings of the OZCHI'99 Australian Conference on Human Computer Interaction*, (1999).
5. Cockburn, A. and McKenzie, B. What Do Web Users Do? An Empirical Analysis of Web Use. International Journal of Human-Computer Studies 54 (2001), 903-922.
6. Cockburn, A., McKenzie, B. and Jasonsmith, M. Pushing Back: Evaluating a New Behaviour for the Back and Forward Buttons in Web Browsers. *International Journal of Human Computer Studies* 57(5) (2002), 397-414.
7. Greenberg, S. and Cockburn A. Getting Back to Back: Alternate Behaviors for a Web Browser's Back Button.

Proceedings of the 5th Annual Human Factors and the Web Conference (1999).

8. Jones, R., Milic-Frayling, N., Rodden, K., and Blackwell, A. Contextual Method for the Re-design of Existing Software Products. *Microsoft Technical Report* MSR-TR-2004-96 (2004).
9. Keller, R., Wolfe, S., Chen, J., Rabinowitz, J. And Mathe, N. A Bookmarking Service for Organizing and Sharing URLs. In *Proceedings of the Sixth International World Wide Web Conference* (1997).
10. Maarek, Y. and Ben Shaul, I. Automatically Organizing Bookmarks per Contents. In *Proceedings of the Fifth International World Wide Web Conference* (1996).
11. Milic-Frayling, N., Jones, R., Rodden, K., Smyth, G., Blackwell, A. and Sommerer, R. SmartBack: Supporting Users in Back Navigation. In *Proceedings of the Thirteenth World Wide Web Conference* (2004), 63-71.
12. Milic-Frayling, N. and Sommerer, R. MS WebScout: Web Navigation Aid and Personal History Explorer, Poster presentation in *On-line Proceedings of the Eleventh World Wide Web Conference* (2002).
13. Milic-Frayling, N., Sommerer, R. and Rodden, K. WebScout: Support for revisitation of Web pages within the navigation session. In *Proceedings of IEEE/WIC International Conference on Web Intelligence (WI'03)* (2003), 689-693.
14. Takano, H. and Winograd, T. Dynamic Bookmarks for the WWW. In *Proceedings of Hypertext'98* (1998), 297-289.
15. Tauscher, L. and Greenberg, S. How people revisit web pages: empirical findings and implications for the design of history systems. *International Journal of Human Computer Studies* 47(1) (1997), 97-137.