Improving Web User Experience with Document Activity Sparklines

Sérgio Nunes[1], Cristina Ribeiro[1,2] and Gabriel David[1,2]

[1] Departamento de Engenharia Informática
Faculdade de Engenharia da Universidade do Porto
[2] INESC-Porto
Rua Dr. Roberto Frias, s/n 4200-465 Porto PORTUGAL
{sergio.nunes,mcr,gtd}@fe.up.pt

Abstract. The temporal dimension of the web has been mostly ignored when designing user interfaces for both searching and document browsing. The dynamic nature of the web is invisible to the typical user despite the fact that most web documents and sites change at a very rapid pace. In this paper we present and describe visual extensions applied to both individual web sites and search engines, that capture content activity over time. This idea is implemented with a prototype that processes publicly available web feeds to generate activity profiles and enhances selected web sites using sparklines. These proposals bring a temporal perspective to the user's browsing experience.

1 Introduction

Previous research on web dynamics has shown that web documents change at a very fast pace. Ntoulas et al. [4] have analyzed both content and link structure evolution of web pages and found that, after one year, 50% of the content on the web is new. Temporal information about web documents can be found either as document-based or as web-based evidence. A detailed survey about possible sources of temporal information on the web is presented in Nunes [5].

The use of temporal information in end-user interfaces has already been addressed. In the context of personal information retrieval Dumais et al. [3] found that time is an important information retrieval cue. In the field of web search, Alonso et al. [2] have used timelines to improve the functionality of search applications. They describe a prototype based on the DBLP bibliography collection. Our work addresses the same problem, proposing a different approach to bring temporal data to the web user's interface. Also, we present a prototype that works with live web data.

We have developed interface enhancements that use publicly available web feeds to present *content activity sparklines*. We leverage on the idea of sparklines, a type of information-rich graphics proposed by Tufte [6]. In the following sections we describe two types of UI improvements — website activity profiles and search engine results enhancements.

2 Activity Profiles

Web feeds are XML documents that provide information about content updates. The Atom Syndication Format is a popular IETF standard for publishing web feeds [1]. Typically, web feeds are structured in items containing a timestamp. Using the timestamp information we are able to aggregate items by day to produce an *activity profile*. Figure 1 shows the activity profile of a blog as a sparkline on the browser's top left corner. Even though this profile is only based on 13 updates, it captures very clearly the blogger's recent activity. Also highlighted in the figure is the additional info available as a tooltip when the mouse hovers the image.

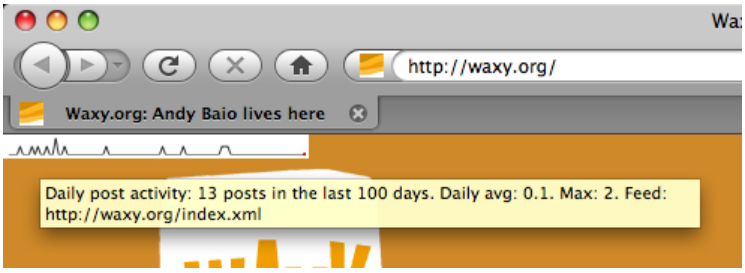

Fig. 1. A blog's activity profile with tooltip.

After conducting an informal survey over a large number of public websites, we found that activity profiles for blogs and news feeds are intuitive and represent an interesting extension to the standard UI. These client-side extensions bring a temporal perspective to the user's browsing experience.

3 Enhancing Search Results

Following upon the idea of using web feeds as a source of data for producing activity profiles, we apply the same idea to current search engine interfaces. We developed extensions for search engines focused in news and blogs, since these typically include public web feeds. Figure 2 shows a set of live results from the Icerocket Blog Search[3] with activity sparklines presented next to each result. We chose to use sparklines due to their high data density and small size, resulting in a low visual impact in the typical web search results layout.

This interface element informs the end-user about the recent activity of each result item. For instance, it is very easy to spot inactive blogs on a long list of search results. On the other hand, it is easy to identify activity peaks related to specific subjects.

[3] `http://www.icerocket.com/`

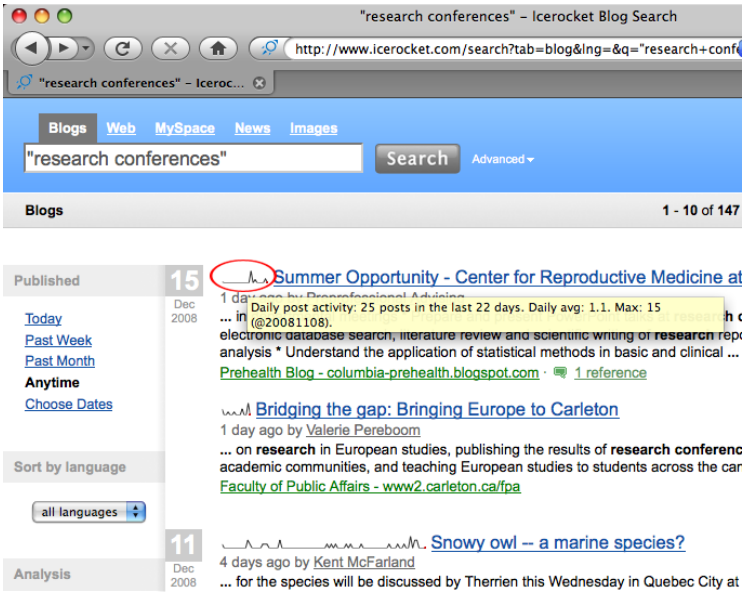

Fig. 2. Icerocket Blog Search with sparklines.

4 Prototype Development

The prototypes were developed using Greasemonkey[4], a Javascript client-side extension to the Firefox web browser. Our Greasemonkey scripts are activated each time a web page is loaded. If available, public web feeds are automatically discovered and parsed. This data is then used to produce a daily frequency count of the feed's activity. Finally, sparklines are drawn in real-time using the Google Chart API[5] and embedded in the document's layout.

As shown in the previous screenshots, a tooltip is added to each sparkline containing additional information about the feed's activity. For each feed, the tooltip provides details on the total number of posts, the total number of days covered, the daily average and maximum value, along with the day of highest activity. While the activity profile extension works on every site that publishes a web feed, the search results extension only works on selected search engines (e.g. Google Blog Search, IceRocket, Technorati).

All prototypes are available online at `http://irlab.fe.up.pt/p/sparkfeeds`. These are fully functional prototypes that were successfully used as a proof of concept in real-world scenarios. The main challenges faced during the development of these extensions were the parsing of dates (due to the plethora of possible "standards") and the parsing of web documents for embedding the sparklines.

[4] `http://www.greasespot.net/`
[5] `http://code.google.com/apis/chart/`

5 Conclusions

Currently available web feeds can be used as a source of temporal information to improve end-user web interfaces. We describe UI enhancements for both individual sites and search engine results. Also, we present fully functional prototypes based on client-side extensions. Given the strong temporal nature of the web, we feel that exposing temporal information to the final user presents several opportunities for UI developments. Google has recently launched a public experimental interface that presents results on a timeline[6].

Preliminary ad-hoc tests conducted with a small number of users resulted in very positive feedback. This was particularly evident in open problems, where the exploratory nature of the task is clearer. We plan to design and conduct users studies as future work. For instance, we plan to evaluate if these enhancements result in significant improvements when users try to complete specific tasks.

6 Acknowledgments

This work was supported in part by Fundação para a Ciência e a Tecnologia (FCT) and Fundo Social Europeu (FSE - III Quadro Comunitário de Apoio), under grant SFRH/BD/31043/2006.

References

1. The Atom syndication format. Technical report, Internet Engineering Task Force, December 2005.
2. O. Alonso, R. Baeza-Yates, and M. Gertz. Exploratory search using timelines. In *SIGCHI 2007 Workshop on Exploratory Search and HCI Workshop*, 2007.
3. S. Dumais, E. Cutrell, J. J. Cadiz, G. Jancke, R. Sarin, and D. C. Robbins. Stuff I've seen: a system for personal information retrieval and re-use. In *SIGIR '03: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 72–79. ACM Press, 2003.
4. A. Ntoulas, J. Cho, and C. Olston. What's new on the web?: the evolution of the web from a search engine perspective. In *WWW '04: Proceedings of the 13th International World Wide Web Conference*, pages 1–12, New York, NY, USA, 2004. ACM Press.
5. S. Nunes. Exploring temporal evidence in web information retrieval. In A. MacFarlane, L. Azzopardi, and I. Ounis, editors, *BCS IRSG Symposium Future Directions in Information Access (FDIA 2007)*, pages 44–50. BCS IRSG, BCS IRSG, August 2007.
6. E. R. Tufte. *Beautiful Evidence*. Graphics Press, 2006.

[6] `http://www.google.com/experimental/`