Using Parallel Web Browsing Patterns on Adaptive Web

Martin TOMA*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 2, 842 16 Bratislava, Slovakia
xtomam@stuba.sk

The possibility to use browser tabs as a tool for parallel web browsing is definitely not new. In recent years, however, more and more people tend to use this feature every day. Despite that, little research has been done to analyse why, and more importantly how, people use this browsing mechanism.

Parallel browsing occurs when user is using browser mechanisms that allow him or her to open and browse multiple Web pages at the same time [3]. Browser windows and tabs are examples of such mechanisms. Users have many reasons why they prefer to use tabs and why they practice parallel browsing in general, some of those were mentioned in a study on users of Mozilla browser [1]: short-term bookmark; parallel searching (branching from search results); opening links in background.

Users have decided to use tabs as the best tool to perform parallel browsing. According to [2], 57% of all browsing sessions involve some type of parallel browsing and users are multitasking by splitting their browsing activity into different tabs rather than viewing more pages overall. Research also shows that studies focused on this type of navigation behaviour have been lacking in Web and hypertext communities, despite the fact that parallel browsing covers fundamental interactions with hyper-links.

Our first goal was to design and implement a method, which will be used to collect accurate browser usage data with the primary focus on browser tabs. Existing solutions like Brumo browser extension framework lack the accuracy in some cases, because they are not utilizing the modern browsers APIs to the full extent. Collecting accurate data will lead to more accurate parallel web browsing analysis. We have implemented a Web browser extension called TabRec[1]. While using the improved Chrome API we are able to precisely detect eight different tab events. TabRec is already live and is capturing parallel browsing activity produced by about 25 users, having collected almost half million logs. Utilization of the most common patterns via the browser actions recommendation was our second goal. There are many techniques

* Supervisor: Martin Labaj, Institute of Informatics and Software Engineering
[1] http://tabber.fiit.stuba.sk

Spring 2015 PeWe Workshop, April 11, 2015, pp. 29–30.

dealing with sequential data analysis and pattern extraction. For pre-processing, we group several specific sequences into more meaningful items representing concrete user actions. This not only allows us to use those techniques and algorithms, which normally ignore duplicate items in transactions, but also express the real user intention much clearly. After the pre-processing, we use Generalized Sequential Pattern algorithm to find the most frequent sequences. We also consult the results with our custom sequence extraction process to better identify the most common patterns.

The most commonly shared sequences can be considered trivial and consist of repeating subsequences each in form of one or two events. Tab close and activate or create and update are examples of such trivial sequences. One of the most common non-trivial sequence we found, consists of five events. User firstly creates a new tab, then focuses it and navigate to a certain web address. This is actually a sequence of three events (create, activate, update). Sequence continues by switching to another, already opened tab and navigating there to another web address. This might represent a typical parallel browsing activity. User is keeping both tabs and regularly updates their content. Another example starts with user focusing an existing tab, then closing it and focusing another two tabs. This sequence continues by closing current tab and focusing another one. Because the time gap between these actions is only five seconds, we can imply that the purpose of this particular sequence was to find those tabs, user don't need anymore.

As the next step, we will now utilize some of these most common patterns and sequences we discovered to enhance users browsing experience. This can be achieved by recommendation of the most appropriate tab action for specific situation or by detecting characteristic sequences, e.g. multiple activation events, which can signalize that user is probably searching for a specific tab and then provide helpful action, e.g. rearranging tabs by domains. We focus on identifying characteristics of such situations and detecting corresponding sequences. TabRec will perform real-time sequence detection, notify users and execute appropriate action. We will record implicit feedback from accepting these actions, which will be the main evaluation metric.

Amended version was published in Proc. of the 11th Student Research Conference in Informatics and Information Technologies (IIT.SRC 2015), STU Bratislava, 393-394.

Acknowledgement. This work was partially supported by the Scientific Grant Agency of Slovak Republic, grant No. VG 1/0646/15.

References

[1] Dubroy, P., Balakrishnan, R.: A Study of Tabbed Browsing Among Mozilla Firefox Users. In: *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*. CHI '10, ACM, 2010, pp. 673–682.

[2] Huang, J., White, R.W.: Parallel Browsing Behavior on the Web. In: *Proc. of the 21st ACM Conf. on Hypertext and Hypermedia*. HT '10, ACM, 2010, pp. 13–18.

[3] Srivastava, J., Cooley, R., Deshpande, M., Tan, P.N.: Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. *SIGKDD Explor. Newsl.*, 2000, vol. 1, no. 2, pp. 12–23.