Stochastic Models for Tabbed Browsing

Flavio Chierichetti*
Dipartimento di Informatica
Sapienza Università di Roma
Roma 00198, Italy.
chierichetti@di.uniroma1.it

Ravi Kumar
Yahoo! Research
701 First Avenue
Sunnyvale, CA 94089.
ravikumar@yahoo-inc.com

Andrew Tomkins*
Google, Inc.
1600 Amphitheater Parkway
Mountain View, CA 94043.
atomkins@gmail.com

ABSTRACT

We present a model of tabbed browsing that represents a hybrid between a Markov process capturing the graph of hyperlinks, and a branching process capturing the birth and death of tabs. We present a mathematical criterion to characterize whether the process has a steady state independent of initial conditions, and we show how to characterize the limiting behavior in both cases. We perform a series of experiments to compare our tabbed browsing model with pagerank, and show that tabbed browsing is able to explain 15–25% of the deviation between actual measured browsing behavior and the behavior predicted by the simple pagerank model. We find this to be a surprising result, as the tabbed browsing model does not make use of any notion of site popularity, but simply captures deviations in user likelihood to open and close tabs from a particular node in the graph.

Categories and Subject Descriptors. H.3.m [**Information Storage and Retrieval**]: Miscellaneous

General Terms. Algorithms, Experimentation, Theory

Keywords. Tabbed browsing, random walks, branching process, stationary distribution, convergence

1. INTRODUCTION

The Opera web browser version 4, released in 2000, was the first to popularize browsing with multiple tabs available within a single window. This *tabbed browsing* paradigm became popular among the technical community, and appeared in Firefox by the following year, and in Safari by 2003. By 2006, all major browsers offered a tabbed browsing capability. Today, as Meiss et al. [22] and Viermetz et al. [27] show, it is increasingly unusual for an online user to access the web through a single tab.

User models like pagerank [23] provide simple, well-known, and mathematically well-understood approaches to thinking about browsing with a single tab in a single window. However, the situation becomes more complex when multiple tabs are introduced. The most salient distinction is that a user no longer traces a single path through a graph, and so may no longer be modeled accurately as a state in a Markov chain corresponding to the underlying graph of web pages or hosts. Rather, the user's browsing may be viewed as a series of "pebbles" that move from node to node in a graph. New pebbles may come into being, and existing ones may split or die out. The operation of moving from one configuration of pebbles to another is not naturally represented as a stochastic matrix, and it is no longer obvious how to think about the "steady state" of such a system.

In this paper, we present a model of tabbed browsing that represents a hybrid between a Markov process capturing the graph of web pages, and a branching process capturing the creation, splitting, and dying of tabs. We present a mathematical criterion to characterize whether the process has a steady state independent of initial conditions, and we show how to characterize the limiting behavior in both cases. We perform a series of experiments to compare our tabbed browsing model with pagerank, and show that tabbed browsing is able to explain 15–25% of the deviation between actual measured browsing behavior and the behavior predicted by the simple pagerank model. We find this to be a surprising result, as the tabbed browsing model does not make use of any notion of site popularity, but simply captures deviations in user likelihood to open and close tabs from a particular node in the graph.

Models of stateful browsing. Browser developers are actively engaged in providing users with new capabilities to improve online navigation. Bookmarks, back buttons (and the corresponding forward buttons), tabs, multiple windows, toolbars, URL bars, auto-completion, search, and many other mechanisms may be seen as offering users a way to move through the web graph using more contextual and stateful information than a naive browsing model would assume. Here we touch briefly on a few of these mechanisms.

Tabbed browsing has two primary manifestations. In the first, a user visiting a page opens a link on the page in a new tab. This is commonly accomplished by right-clicking the link and selecting "Open in new tab," or by holding down the control key while clicking the link. We will refer to this behavior as "control-clicking." In the second manifestation, a user explicitly requests a new tab without specifying the contents, and this tab then opens to the user's (possibly empty) homepage. From there, the user may enter a URL, perform a search through a toolbar or browser chrome search box, and so forth. These two behaviors must be handled differently, as the first depends on the current page and represents a form of transition, while the second is more akin to a restart.

In addition to tabbed browsing, users may also click the "back button" to revisit the previous page of the current

*Part of this work was done while the author was at Yahoo! Research.

WWW 2010, April 26–30, 2010, Raleigh, North Carolina, USA.
ACM 978-1-60558-799-8/10/04.