Chapter 7.28
Improving Mobile Web Navigation Using N–Grams Prediction Models

Yongjian Fu
Cleveland State University, USA

Hironmoy Paul
Cleveland State University, USA

Namita Shetty
Cleveland State University, USA

ABSTRACT

In this article, we propose to use N-gram models for improving Web navigation for mobile users. N-gram models are built from Web server logs to learn navigation patterns of mobile users. They are used as prediction models in an existing algorithm which improves mobile Web navigation by recommending shortcuts. Our experiments on two real data sets show that N-gram models are as effective as other more complex models in improving mobile Web navigation.

INTRODUCTION

Wireless users of the Web grow rapidly as more and more mobile devices such as PDAs, mobile phones and pagers are now equipped with browsing capabilities. Many current Web sites are optimized for desktop, broadband clients, and deliver content poorly for mobile devices due to display size and bandwidth. Moreover, the associated cost will prohibit maintaining two versions of a site, one for wired users and the other for mobile users. A viable solution is adaptive Web sites (Perkowitz & Etzioni, 1997). An adaptive Web site dynamically changes its contents or structure based on browsing activities.

Following the idea of adaptive Web sites, Anderson, Domingos, and Weld (2001) proposed shortcuts to improve mobile Web navigation. A shortcut is a dynamic link that provides a shorter path with fewer clicks for users to reach their desired pages. A shortcut to a destination page is dynamically created and inserted into the next page a user is going to browse. If that destination page is the one in which the user is interested, the user can access the page by following the shortcut. For example, assume a browsing session consists of A-B-C-D-E-F-G, where each letter represents a page. After browsing pages A and B, if a shortcut to G is created and inserted into page C, the user can follow the shortcut to reach G, without going through intermediate pages D, E, and F. The critical question is how to find shortcuts that are useful with only part of the session known. A shortcut C $\rightarrow$ H, for example, is useless in the previous example.

In order to provide useful shortcuts, Web usage mining techniques are employed. User browsing patterns are extracted from Web server logs. These patterns are built into prediction models that can be used to predict user browsing behaviors. Given a partial session, such prediction models will compute what other pages in which the user may be interested. These predictions are used to create and recommend shortcuts for mobile users.

A critical component in this approach is the prediction model. The model should be as accurate as possible with as little information about the session as possible. An accurate shortcut found earlier in a session is more worthwhile than one found close to the end of a session. Moreover, the prediction model should be easy to build and use. In their MINPATH algorithm, Anderson et al. (2001) used Markov models, which proved to be accurate (Anderson et al., 2001). However, those models require Web graphs. In this article, we propose to use a simpler prediction model, N-gram, for learning user browsing patterns.

N-grams are well known and are widely used in speech and text processing applications. Re-

searchers have found that accuracy increases with N, the order of N-grams. For example, 4-grams are more accurate than 3-grams, which is turn is more accurate than 2-grams. Though accuracy increases with higher values of N, it requires a larger number of training sessions to have a well trained N-gram model.

An N-gram based prediction model for Web browsing patterns is proposed by Su et al. (2000). The N-gram model has several advantages over other prediction models. It is simple, robust, and easy to use. Besides, N-gram does not use a Web graph. In our study, the same N-gram model with a slightly different lookup operation is used. Moreover, its effectiveness in improving mobile Web navigation is examined.

In our approach, first, Web server logs are preprocessed to identify sessions. A session is conceptually a single visit. The sessions are then used to train an N-gram model. A revised version of MINPATH algorithm, MINCOST, is proposed to find shortcuts. MINCOST uses a different function in calculating the saving and ranking of shortcuts. Our approach has been implemented and evaluated against two real data sets from NASA and EPA Web servers. Our experiments show that the N-gram prediction model is as effective as more sophisticated models in recommending useful shortcuts.

The article is organized as follows. The second section discusses related work in Web usage mining, adaptive Web sites, and MINPATH algorithm. Our approach is presented in Section 3. Experimental results with two data sets are reported in the fourth section The fifth section concludes the article and gives some future research direction.

RELATED WORK

We briefly discuss Web usage mining techniques and its applications in adaptive Web sites and mobile Web navigation.

Related Content

An Approach With Iterative and Incremental Development (IID) for Mobile Applications

Jona Ndrecaj, Shkurte Berishaand Erita Çunaku (2023). *Designing and Developing Innovative Mobile Applications (pp. 67-81).*

www.irma-international.org/chapter/an-approach-with-iterative-and-incremental-development-iid-for-mobile-applications/322065

Vision-Based Human Face Recognition Using Extended Principal Component Analysis

A. F. M. Saifuddin Saif, Anton Satria Prabuwono, Zainal Rasyid Mahayuddinand Teddy Mantoro (2013). *International Journal of Mobile Computing and Multimedia Communications (pp. 82-94).*

www.irma-international.org/article/vision-based-human-face-recognition-using-extended-principal-component-analysis/103970

Definitions, Key Characteristics, and Generations of Mobile Games

E. Jeongand D. Kim (2007). *Encyclopedia of Mobile Computing and Commerce (pp. 184-189).*

www.irma-international.org/chapter/definitions-key-characteristics-generations-mobile/17074

A Comparative Study of Shape and Texture Features for Finger Spelling Recognition in Big Data Applications

Yong Hu (2017). *International Journal of Mobile Computing and Multimedia Communications (pp. 66-78).*

www.irma-international.org/article/a-comparative-study-of-shape-and-texture-features-for-finger-spelling-recognition-in-big-data-applications/188624

Mobile Parenting and Global Mobility: The Case of Filipino Migrant Mothers

Ma. Rosel S. San Pascual (2014). *Interdisciplinary Mobile Media and Communications: Social, Political, and Economic Implications (pp. 194-212).*

www.irma-international.org/chapter/mobile-parenting-and-global-mobility/111723