Personalized Web Advertising Method

Przemysław Kazienko¹ and Michał Adamski¹,²

¹ Wroclaw University of Technology, Department of Information Systems, Wybrzeże S. Wyspiańskiego 27, 50-370 Wroclaw, Poland
kazienko@pwr.wroc.pl, http://www.pwr.wroc.pl/~kazienko
² Universidade Nova de Lisboa, Faculdade de Ciências e Tecnologia, Monte de Caparica, 2829-516 Caparica, Portugal
michal.adamski@iol.pt

Abstract. Personalization of online advertising is a great challenge while the market is moving and adapting to the realities of the Internet. Many existing approaches to advertisement recommendation are based on demographic targeting or on information gained directly from the user. In this paper we introduce the AD ROSA system for automatic web banner personalization, which integrates web usage and content mining techniques to reduce user input and to respect the user’s privacy. Furthermore, the advertising campaign policy, an important factor for both the publisher and advertiser, is taken into consideration. To enable online personalized advertising the integration of all relevant information is performed in one vector space.

1 Introduction

In the age of disappearing borders and mixed societies the current demographic targeting of freely available Web content seems to be insufficient. The market consists of human beings, not demographics, so web personalization should depend on an individual’s behavior rather than on stereotypes created according to his or her geographical location or other demographic features (e.g. gender, age). Traditional advertising serving the same offers for everyone does not meet the current requirements of businesses. To increase the effectiveness, the right person should receive the right message at the right time and in the right context [1].

Web advertising is mainly done with banners – graphical elements on a web page, or with their ‘mutations’ – displayed in a new layer or new window of the browser. There are many other forms of online advertisements like sponsored links or articles, or mail-outs, but in this article we will only concentrate on banners and similar forms.

Users, showered with hundreds of advertisements, often pay less attention to banners appearing on a web page as bitmap images or animations, and this seems to be the main problem of web advertising. The solution is to increase the correspondence between user interests and the subject of the displayed advertisement [4].

Two significant research domains may be distinguished within Internet advertising: scheduling and personalization. The main goal of the former is to maximize the total click-through-rate for all advertisements by appropriately managing of exposition time and advertising space on the web page [3, 14].

The latter seems to be an important and difficult challenge for current advertisers. It aims to assign a suitable advertisement to the user, so it is necessary to have some
information about the user. Many web portals create user profiles using the information gained during the registration process or ask the user to answer some questions about their preferences. However, this requires a lot of time and effort, and that can discourage many users. Besides, users tend to give incorrect data when being concerned for their privacy [13]. Even reliable data becomes out-of-date with the evolution of the online customer’s interests. An alternative solution is to exploit information stored in the web server logs. This method is safe in regard to privacy fears and may also be useful for news portals or web sites where users do not need to log in to use the service [19]. Another approach to advertisement personalization is presented in [11]. Short-term and long-term interests of the user were identified. Short-term interests are derived from the keywords submitted by the user in searching services. However, such keywords may often have nothing in common with the user’s regular preferences. Long-term interests are proposed to be taken from user profiles, which are completed by users and stored in the database of the system. However, advertising personalization was performed using only short-term information.

A system based on web usage mining and clustering of navigation paths to create usage patterns was presented in [19]. Pages from both the publisher’s web site and the advertisements’ target sites are manually classified into thematic categories by experts. The assignment of appropriate advertisements to each active user is accomplished according to pages (categories) visited by the given user during the current session. This matching is based on fuzzy rules stored in the system. The fuzzy approach was also used in target advertising based on user profiles [21].

Three main advertising models can be distinguished [5]: broker, portal and advertiser models. In the broker model there exists an advertising broker that connects publishers (web sites in which advertisements would be displayed) and advertisers (companies providing banners to be emitted). The broker often provides some targeting options. This model is applied i.e. by DoubleClick [15]. The portal model (used by large web portals) is the special case of the broker model: the publisher owns the advertisement management software and cooperates with many advertisers. The advertiser model, in which the advertiser manages the advertisements, allows big online store to display its banners on pages of particular portals.

2 Advertisement Features

Nowadays most online advertising systems use the principle of the customer-based targeting. Each user is identified and classified according to his or her geographical location (IP address) and browser settings sent with the HTTP request, navigation habits and user profiles (preferences) completed by the user during the registration process. This data is used to personalize the displayed banner advertisement [2, 11].

Analyzing advertising offers of the greatest Polish portals (www.wp.pl, www.onet.pl), we observed many target criteria available for advertisers. Apart from the demographic data of a user (age, gender, location, etc.), advertisements can be targeted towards the user’s education, profession or interests. Furthermore, the publisher can choose the time of day of the emission, particular parts of the web page, and limit the number of emissions for a single user. An advertiser is usually charged on the basis of cost per month per one thousand emissions of advertisement (CPM). Another approach is the usage of click-through-rate (CTR) - the ratio of the number of clicks to the emission number [2]. It should be mentioned that the average CTR is