Chapter 10
Algorithms for Web Personalization

Abstract. Personalization is a process of gathering, storing and analyzing information about site visitors and delivering the right information to each visitor at the right time. A personalization technique can enable a website to target advertisement, promote products, personalize news feeds, recommend documents, make appropriate advice and target e-mail. In this chapter, an adaptive weight update algorithm based on the standard search engine, using the combination of query refinement, recommender system and user profiles is explored. The algorithm proposed captures and records the user interest with respect to the indicated keywords to the search engine. The user profile is built dynamically on the recorded interests of the user and used as a recommender system to refine the query on further hits by adding additional keywords depending upon the users interest. A dynamic weight update equation is used to adapt to the changing user interests. The algorithm is tested on varying keywords and the results obtained are better when compared with the Googles page rank system.

10.1 Introduction

Web Personalization is a set of actions that can tailor the Web experience to a particular user or a group of users. The experience can be like, users browsing patterns, sequence of pages visited by a user, usual query patterns, etc.. The actions can range from simply making the presentation more pleasing to anticipating the needs of a user and providing customized and relevant information to the user. To achieve effective personalization, organizations must rely on all available data, user behavior, the site content, domain knowledge, as well as user demographics and profiles [1].

One of the techniques to achieve effective personalization is by having visitors of a site fill out the forms with information fields that populate a database. The Web site then uses the database to match a user’s needs to the products or information provided at the site, with middleware facilitating the process by passing data between the database and the Web site. Consider an example of Amazon.com, which is used for online book purchase/selling. One of the facilities provided by Amazon.com for
its registered users is the suggestion of books/CDs depending on their previous pur-
chase history or his interests captured while browsing the Web. Customers tend to
buy more when they know exactly what’s available at the site and they do not have
to hunt around for it. Cookies may be the most recognizable personalization tools.
Cookies are bits of code that sit in a user’s Internet browser memory and inform Web
sites about a person. That is how a Web site is able to greet its users by name [2].

Web personalization is also achieved by collaborating filtering software that re-
sides on a web site and tracks the path of the users interest and viewing habits de-
pending upon various parameters like types of pages they visit, the amount of time
they spend on a page, etc,. Collaborative-filtering software compares the informa-
tion it gains about one user’s behavior against data about other customers with simi-
lar interests. In this way, users get recommendations like Amazon’s “Customers who
bought the book x also bought other books y and z”. Domain-specific search engines
are growing in popularity because of their increased accuracy and extra functional-
ity that is not possible with the general search engines. For example, scholar.com
of Google allows complex queries over Internet by type of file, prioritization of in-
tstitutions, size, and location. Yahoo is another example for domain-specific search
engines. Unfortunately, these search engines are difficult to use for naive users and
time-consuming to maintain.

The key challenge in personalization is to display to the visitor a sub graph that
contains both relevant content items, and also organizes them in a coherent and
meaningful manner. The problem of automatically constructing Personalized Site
Maps is discussed in [3]. The approach is based on the assumption that the best
way to indicate the relationship between a given pair of pages is to show the path
between them that has been most popular with past visitors. An ontology based
personalized retrieval of audio information is addressed in [4]. The crux of the work
is the development of an ontology-based model for the generation of metadata for
audio and the selection of audio information in a user customized manner.

The data mining methods are very much exploited to build the customer
profiles. The system constructs personal profiles based on customers transactional
histories and then uses data mining techniques to discover a set of rules describ-
ing customers behavior and supports human experts in validating the rules [5].
The study of personalized recommendation to build improved algorithm for learn-
ing Bayesian networks efficiently and effectively is proposed in [6]. The algorithm
reduces the number of independence tests and database passes, while effectively
restricting the search space and by means of the heuristic knowledge of mutual in-
formation. In [7], a web personalization system based on web usage mining and
automatic customization is proposed. It is an effective technique for capturing com-
mon user profiles based on association rule discovery and usage-based clustering.
The techniques for combining knowledge with the current status of an ongoing Web
activity to perform real-time personalization are also discussed.

The use of Agent technology for personalized recommendation, which is fuzzy
cognitive in nature, to give personalized suggestions based on the current users pref-
ferences, general users common preferences, and the experts knowledge is described
in [8]. Fuzzy cognitive agents are able to represent knowledge via extended fuzzy