

International Conference on Communication Technology and System Design 2011

Design and Implementation of Web Usage Mining Intelligent System in the Field of e-commerce

B.Naveena Devi^a, Y.Rama Devi^b, B.Padmaja Rani^c, R.Rajeshwar Rao^d, a*

^aDepartment of CSE, Mahatma Gandhi Institute of Technology, Gandipet, Hyderabad, A.P.500075, India

^bProfessor, Department of CSE, Chaitanya Bharathi Institute of Technology, Gandipet, Hyderabad, A.P.500075, India

^cDepartment of CSE, Jawaharlal Nehru Technological University, Hyderabad, A.P., India

^dDepartment of CSE, Mahatma Gandhi Institute of Technology, Gandipet, Hyderabad, A.P.500075, India

Abstract

The rising popularity of electronic commerce makes data mining an indispensable technology for several applications, especially online business competitiveness. The World Wide Web provides abundant raw data in the form of web access logs. Now a days many business applications utilizing data mining techniques to extract useful business information on the web evolved from web searching to web mining. This paper introduces a web usage mining intelligent system to provide taxonomy on user information based on transactional data by applying data mining algorithm, and also offers a public service which enables direct access of website functionalities to the third party.

© 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of ICCTSD 2011

Keywords: Data Mining; Information Retrieval; Open Web Services; Web Usage Mining, Web Computing.

1. Introduction

The goal of Web Usage Mining is to find out extract the useful information from web data or web log files. The other goals are to enhance the usability of the web information and to apply the technology on the web applications, for instance, pre-fetching and catching, personalization etc. For decision management, the result of web usage mining can be used for target advertisement, improving web design, improving satisfaction of customer, guiding the strategy decision of the enterprise and market analysis [1].

Recently there are a large number of web services that we can use and many of them are open source based. Web services are APIs that facilitate the communication between applications for example RapidMiner, Digg.com, Amazon, eBay are opened access to their services and data through APIs, and we

* B. Naveena Devi. Tel.: +91-9441724900; fax: +91-40-24193067

E-mail address: veenamgit@yahoo.com.

can make use of their services for the development of web usage mining research applications. The concept of Web APIs enables direct access to the website functionalities in order to leverage third party efforts on value adding services [2]. However, the number of companies, services or web sites that gather information about users increasing continuously. These systems store private information about users and for that reason appears much controversy about the legitimacy. The main problem is that these companies don't share information with the rest of the world. In this paper, we present a public system to store information about their products and view details about user behavior.

Some of the problems about sharing information would be solved if there was a public service for user behavior information. If all people can access that information, all of them will have the same opportunities and will be at the same point in a commercial environment [2].

The rest of the paper is organized as various sections: section 2 will have implemented details about how Hierarchical Agglomerative Clustering applied on sample web log for mobile marketing. Section 3 elaborates how to provide public service (API) which enables third party to view their customer's behavior. Finally Section 4 demonstrates experimental result and Section 5 Conclusion with future work.

2. Hierarchical agglomerative clustering

In this paper we focus on, standard data mining techniques such as clustering a particular user may associate with other users exhibiting similar behavior pattern and preferences. Due to the heterogeneity of user's browsing features, the hierarchical agglomerative clustering algorithm is used to class user's browsing behaviors. Agglomerative hierarchical clustering starts with every single object in a single cluster. Then, in each successive iteration, it agglomerates the closest pair of clusters by satisfying some similarity criteria, until all of the data is in one cluster. However, it is necessary to define a suitable terminal condition when the agglomerative process should end [3].

In the hierarchical clustering, the general similarity measures are Euclidean distance function. In the initialization, every user is seen to be a cluster. The similar users' browsing feature will be found out and merged into a cluster until terminal condition is satisfied. Finally, the user clusters will be displayed based on browsing timings.

2.1. Pattern Representation

We have taken sample web log file for mobile marketing as shown in table 1 for illustration purpose. At this point in time, we assume that user sessions can be accurately determined. This log file contains details like user id, name of the company, name of the product, log in time, log out time, company session start time, company session end time, product session start time, product session end time, respective access time in seconds.

In the task of pattern representation, user sessions are created from web log files. User sessions can be reorganized as a $m \times k$ matrix as table 1, each row can be presented by $Session^u = (P_{u,1}, P_{u,2}, \dots, P_{u,k})$. The k is the number of clusters which is necessary to define a suitable terminal condition when the agglomerative should be end. We have taken parameter k value as 3.

One straightforward approach in creating an aggregate view of each cluster is to compute the centroid of each cluster. We have taken the dimension value for each session in the mean vector is computed by finding the ration of the sum of the session weights across transactions to the total number of transactions in the cluster.

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات