Web Page Recommender System using hybrid of Genetic Algorithm and Trust for Personalized Web Search

Suruchi Chawla, Shaheed Rajguru College Delhi University, Delhi, India

ABSTRACT

The main challenge to effective information retrieval is to optimize the page ranking in order to retrieve relevant documents for user queries. In this article, a method is proposed which uses hybrid of genetic algorithms (GA) and trust for generating the optimal ranking of trusted clicked URLs for web page recommendations. The trusted web pages are selected based on clustered query sessions for GA based optimal ranking in order to retrieve more relevant documents up in ranking and improves the precision of search results. Thus, the optimal ranking of trusted clicked URLs recommends relevant documents to web users for their search goal and satisfy the information need of the user effectively. The experiment was conducted on a data set captured in three domains, academics, entertainment and sports, to evaluate the performance of GA based optimal ranking (with/without trust) and search results confirms the improvement of precision of search results.

KEYWORDS

Clustering, Recommender System, Genetic Algorithm, Information Retrieval, Information Scent, Personalized Web Search, Search Engines, Trust

1. INTRODUCTION

Information on the Web is huge and information retrieval of relevant documents for a specific information need of web users is a big challenge for search engines. The search engines retrieve large collection of ranked search results for a specific information need out of which very few are relevant. It is found that relevant documents are present lower in ranking of search results due to imprecise search queries and therefore the precision of search results decreases. Research had been done for web page ranking in order to bring more and more relevant documents up in ranking for the improvement of precision of search results (Peng & Lin, 2006; Selvan et al., 2012; Page et al., 1999; Xing & Ghorbani, 2004; Ding et al., 2001; Jayanthi & Jayakumar, 2011). In (Chawla, 2016) GA was used for optimal web page ranking of High Scent Clicked URLs for personalized web search and experimental results confirmed the improvement of precision of search results. It is realized in this research that rank optimization of clicked URLs using GA can be more effective if trust is used to select the web pages for optimization. The optimal ranking of trusted clicked URLs using GA will retrieve more relevant documents up in ranking and therefore improves the precision of search results.

In this paper a novel approach is proposed using hybrid of GA and Trust for optimal ranking of clicked URLs based on clustered web query sessions. The benefit of using hybridization of Trust and GA together is because of the following reasons .GA is parallel in nature and well suited for solving problems where the solution space is huge and time taken to search exhaustively is very high.

DOI: 10.4018/JITR.2018040107

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
The trust value of web pages determines the relevance of web page based on web usage data and is calculated using clustered query sessions. Thus, the use of high trust threshold value will select those web pages which are relevant and satisfy the information need of the users most of the time when recommended in search results.

Thus, an algorithm is proposed for web page recommendation of optimal ranking of trusted URLs using hybrid of GA and Trust for effective personalization of user web search. The entire processing of the proposed algorithm is divided into two phase: Phase I (Offline) and Phase II (Online). In Phase I, the query sessions keyword vector is clustered to group the clicked URLs which satisfy similar information need. The clicked URLs are selected based on trust threshold value in a given cluster and GA is applied on the population of possible ranking of trusted clicked URLs in a given cluster for optimization. Thus, at the end of offline processing, each cluster is associated with optimal ranking of trusted clicked URLs.

During online processing, the input query issued for web search select the most similar cluster for the recommendations of optimal ranking of trusted clicked URLs. The recommendation of optimal ranking of trusted URLs continues till the search is personalized to the information need of the user.

The flowchart steps for the proposed approach of web page recommendations of optimal ranking of trusted URLs using GA is given below in Figure 1.

Experiment was conducted on the data set of user query sessions captured on the web in three selected domains Academics, Entertainment and Sports to evaluate the effectiveness of hybridization of GA and Trust for web page recommender system. The results were compared with PWS (with GA) (Chawla, 2016) /Classic IR and the improvement in the average precision of search results confirms that the rank optimization of trusted clicked URLs recommend relevant search results for effective Information retrieval.

2. RELATED WORK

In (Pera & Ng, 2013) group recommender system for movies based on content similarity and popularity was proposed. In (Choi, Jeong & Jeong, 2010) a hybrid recommendation algorithm using both collaborative and content based filtering was proposed. In (Beel et al., 2013) both online and offline evaluation of research paper recommender system were analyzed. Offline evaluation was found to be unsuitable for evaluation of research paper recommender systems. In (Herlocker et al., 2004) collaborative filtering recommender systems were reviewed based on user tasks, types of datasets, method of calculation of prediction quality and prediction attributes. In (Kadam & Gaikwad, 2015) cold start and sparsity problem found in content based filtering and item based filtering was overcome by using interpersonal interest, social profile in personalized recommendation system to recommend interested items to users.

It was found that recommender system can be more effective by incorporating trust than traditional collaborative filtering (Massa & Avesani, 2007; Levien, 2004; Lathia, Hailes & Capra, 2008; Hwang & Chen, 2007; Peng & Seng-cho, 2009). In (Massa & Bhattacharjee, 2004) trust based recommender system was proposed using both trust metric and similarity metric.

In (Xue & Fan, 2008) a new trust model based on social characteristic and reputation mechanism for the semantic web was proposed. In (Tian et al., 2008) Trust model based on reputation for peer-to-peer networks was proposed. In (Jamali & Ester, 2009) TrustWalker: A Random Walk Model was proposed combining both trust-based and item-based recommendation. In (Jianshu, Chunyan & Angela, 2006) Collaborative Filtering was improved with Trustbased Metrics where trust metric was defined for incorporating trust in similarity computation. In (Guha et al., 2004) method of propagation of trust and distrust was introduced. In (Baudry, Hanh & Traon, 2000) method was proposed for building trustable OO (Object Oriented) components.

In (Bedi & Sharma, 2012) Trust based Ant Recommender System was proposed. It used the concept of dynamic trust among users and the best neighborhood was selected based on genetic image
Legal Truth and Consequences for a Failed ERP Implementation
www.igi-global.com/article/legal-truth-consequences-failed-erp/53555?camid=4v1a

Bundling Processes Between Private and Public Organizations: A Qualitative Study
www.igi-global.com/article/bundling-processes-between-private-public/52822?camid=4v1a