Information Sciences 305 (2015) 302–319

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Recommending blog articles based on popular event trend analysis

Duen-Ren Liu [a,*], Hani Omar [a], Chuen-He Liou [b], Huai-Chun Chi [a], Cheng-Ho Hsu [a]

[a] Institute of Information Management, National Chiao Tung University, Hsinchu, Taiwan
[b] Center of General Education, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan

ARTICLE INFO

Article history:
Received 6 June 2013
Received in revised form 8 December 2014
Accepted 1 February 2015
Available online 7 February 2015

Keywords:
Blogosphere
Popular event-based recommendation
Trend analysis
Google Insights
Content-based filtering
Item-based collaborative filtering

ABSTRACT

Web 2.0 has become a popular social media on the Internet due to the fast evolution of Internet technologies, as well as increasing resources and users. Among the applications of Web 2.0, blogospheres are a new Internet social media for users to express their preferences and personal feelings. Most of the people tend to receive the newest information and articles related to popular issues. However, with the rapidly increasing number of active writers and viewers, it is hard for people to discover useful information that is beneficial or interesting to them. Accordingly, it is necessary to develop a recommendation approach that takes the emerging or popular events into consideration. In this work, we propose a novel event-based recommendation approach, which combines the event trend analysis and personal preference to recommend blog articles of popular events that suit user interests. We analyze blog articles to identify popular events, and then derive the popularity degrees of events based on blog-based popularity trend analysis and Google Insights-based popularity trend analysis. Our approach derives users' personalized preferences on target articles of popular events by considering user interests (article-push records) and the predicted popularity degree of the events. Our recommendation methods improve recommendation accuracy by enhancing content-based filtering (CBF) and item-based collaborative filtering (ICF) with the event-based preference analysis. Our experiment result demonstrates that the proposed approach can effectively recommend users' desired blog articles with respect to event popularity and personal interests.

1. Introduction

Web 2.0 has become a popular social media on the Internet due to the fast evolution of Internet technologies, as well as increased resources and users. Web 2.0 is a web application where people can collaborate and share information which has led to the creation of new business models of software [41]. Among the applications of Web 2.0, blogospheres are a new Internet social media for users to express their preferences and personal feelings. In fact, the number of blogospheres has been rising rapidly in recent years; people prefer the blogospheres because the free platform provided by blogospheres, allows people to share the latest information and exchange opinions easily without technical constraints, instead of just retrieving information passively. Currently, blogospheres have become an indispensable information exchange platform which enables users to publish articles about their daily life or emerging news. In addition, there are several emerging events

* Corresponding author.
E-mail address: dliu@mail.nctu.edu.tw (D.-R. Liu).

http://dx.doi.org/10.1016/j.ins.2015.02.003
0020-0255/

in blogospheres for people to access the most popular and up-to-date information such as new product launches, breaking news and hot events.

Most of the people using blogospheres tend to receive the newest information and articles related to popular issues. However, with the huge growth of bloggers and blog articles, the vast amount of blog information causes the problem of information overload and blog readers are difficult to find interested articles [28]. Blog readers are often interested in browsing an emerging or popular event's blog articles, since blog articles represent the opinions of the populace and constitute a reaction to current events on the Internet. Accordingly, it is important to provide recommendation service for people in the blog platform, especially for recommending blog articles of emerging or popular events that suit their interests. Very little research, however, has focused on this issue.

To solve the problem of information overload, Recommender systems have been, and continue to be applied in various applications to support item (e.g. movies, documents or music) recommendation [1,2,10,29,34]. Generally, the content-based filtering and collaborative filtering approaches are the most widely used for making recommendations. The content-based filtering (CBF) approach recommends an item to a user based on a description (content) of the article and a profile of the user's interests [43]. The collaborative filtering (CF) approach considers user neighbors or item neighbors based on the historical data, to make recommendations [12]. As mentioned before, several researchers aimed to solve the problem of information overload in the past, but most of those studies focused on analyzing the content of the recommending articles or user preferences, without considering the emerging events and the popularity degrees of the events. Since this problem is not handled appropriately currently, further research is needed.

Blog articles contain highly time-sensitive information of a wide range of topics, opinions and commentary. Thus, articles posted at different time usually reflect the information of different events. Current event identification approaches have been proposed by analyzing the words discovered in blog articles within a specific period. For example, Platakis et al. [46] consider timestamp as a very important characteristic in the blogosphere and tend to identify burst events by discovering the bursty terms and correlations between them. Bansal and Koudas [4] detect hot keywords by analyzing the unexpected popularity of a keyword within a temporal window. However, while blog users are more interested in articles which contain information about emerging or popular events, existing studies did not address the popularity degrees of the blog events as a notable issue for analyzing user interests. Additionally, the current popularity degrees of blog events cannot directly be regarded as the actual popularity score when recommended to users in the future. A new event may trigger emerging discussions such that the number of related blog articles is small at the beginning, but it may gradually become more popular when time goes on. Thus, it is important to analyze the trend of time-sensitive popularity of events to predict emerging hot blog events. In addition, blog readers may have different interests regarding the emerging popular blog events. Nevertheless, existing researches have not addressed such issues of how to predict the popularity trend of blog events and derive personalized popular events. In this work, we address the problem by predicting the popularity degree of the blog events and develop a recommendation approach in consideration of personalized popular events.

In addition to event detection in the blogospheres, there are several studies on blog recommendation, and some of them focus on the analysis of the contents of blog articles. Liu et al. [34] consider both popularity of articles and personal interests to recommend blog articles to the mobile users. Kening et al. [18] propose an effective blog clustering algorithm, modified from the K-means clustering approach, to cluster blogs based on link topology and content similarity; similar blogs are recommended to bloggers based on the clustering result. Ray and Singh [48] consider original themes of blog content as an input and recommend related themes of blog content to users. A blog article recommendation mechanism is proposed by Huang et al. [23], whereby relevant terms from blog articles associated with specific users are extracted and then used to recommend blog articles using Google's search engine. Some of the studies are based on user interests. Liu et al. [36] model the authors' interests as preference weights on predefined categories based on the probabilities of classifying their blog posts into the categories. Although several researches have proposed a content-based approach to recommend blog articles according to users' interests, they did not address the issue of identifying emerging/popular events and recommending blog articles of popular events.

In this work, we address the issues for recommending blog articles of emerging/popular events. Relevant questions include: how to effectively discover the events that blog users are discussing, and measure the popularity degrees of discovered events; and how to make personalized recommendation considering the variety of user interests. In general, an event is popular if the number of blog articles related to the event is reasonably large. Moreover, an event' popularity can also be derived from the searches of Google Insights. A popular event usually allures a large number of users to search information related to the event through Google search engine. We propose a novel approach of event-trend analysis to identify events and analyze the popularity trend of events. Our proposed event-trend analysis predicts the popularity of events based on the blog-based popularity trend analysis and Google-based popularity trend analysis. The blog-based/Google-based popularity trend analysis predicts an event's popularity trend by analyzing the trend effect of time-sensitive blog posts/Google searches of an event within a time window.

A novel personalized event-based recommendation approach focusing on popularity degrees of events is proposed to recommend blog articles of popular events that suit user interests. The event-based approach recommends blog articles by integrating popularity of events, content-based filtering (CBF)/item-based collaborative filtering (ICF) and event-based preferences of blog articles. A social bookmarking websites, funP, is used for evaluating our proposed recommendation approach. The website provides a free platform for Internet users to share information, such as by publishing a new article,