



Full length article

Comparative analysis of relevance feedback methods based on two user studies

Stephen Akuma^{a,*}, Rahat Iqbal^a, Chrisina Jayne^b, Faiyaz Doctor^a^a Department of Computing, Coventry University, Coventry, UK^b School of Computing Science and Digital Media, Robert Gordon University, Aberdeen, UK

ARTICLE INFO

Article history:

Received 15 June 2015

Received in revised form

12 February 2016

Accepted 15 February 2016

Available online xxx

Keywords:

Implicit feedback

User interest

Explicit feedback

Implicit indicators

Explicit rating

Recommender system

ABSTRACT

Rigorous analysis of user interest in web documents is essential for the development of recommender systems. This paper investigates the relationship between the implicit parameters and user explicit rating during their search and reading tasks. The objective of this paper is therefore three-fold: firstly, the paper identifies the implicit parameters which are statistically correlated with the user explicit rating through user study 1. These parameters are used to develop a predictive model which can be used to represent users' perceived relevance of documents. Secondly, it investigates the reliability and validity of the predictive model by comparing it with eye gaze during a reading task through user study 2. Our findings suggest that there is no significant difference between the predictive model based on implicit indicators and eye gaze within the context examined. Thirdly, we measured the consistency of user explicit rating in both studies and found significant consistency in user explicit rating of document relevance and interest level which further validates the predictive model. We envisage that the results presented in this paper can help to develop recommender and personalised systems for recommending documents to users based on their previous interaction with the system.

© 2016 Published by Elsevier Ltd.

1. Introduction

As the amount of information available through the internet and various intranets continues to grow, effective retrieval of relevant information has become a challenging task. The tremendous amount of available information leads to information overload (Alhindi, Kruschwitz, Fox, & Albakour, 2015). This phenomenon has led to research on personalised information retrieval and recommender systems. Most of the current information retrieval systems are generic and do not provide task-specific information to users (Grzywaczewski & Iqbal, 2012; Jawaheer, Weller, & Kostkova, 2014). Hence, research in the area of personalisation has been found to be useful for addressing the problem of information overload by providing the users with relevant web documents based on their interest and current activity. This can enhance user search experience and improve efficiency. This can be achieved through relevance feedback based approaches. The process of gathering useful information about users in order to give them

feedback or recommend documents based on their previous interaction with the system is called relevance feedback (Zemirli, 2012). Relevance feedback is affected by contextual factors like task type (Kellar, Watters, Duffy, & Shepherd, 2004; Li & Belkin, 2008). Tasks are activities that people attempt to accomplish to meet a particular goal. Identifying an accurate task stream is difficult because the demarcation of boundaries is not clear enough. However, common streams have classified different task along features like fact-finding vs information gathering (Kellar, Watters, & Shepherd, 2007). Task type help to infer document usefulness and also influences the total time spent while performing the task (Liu & Wu, 2008). This work employs user task to examine the relationship between implicit parameters and explicit ratings.

Users of an information system present their intention/interest through the formulation of input queries; however this does not adequately capture their interest (White & Kelly, 2006) and the users have to go through several iteration to get accurate search results. Therefore, there is a need to augment user query input with additional sources of information obtained explicitly or implicitly from their post-click interaction with the system in order to provide users with accurate search results based on current activity and context (White & Kelly, 2006). With respect to relevance feedback, although the explicit approach (where users of a system

* Corresponding author.

E-mail addresses: akumas@uni.coventry.ac.uk (S. Akuma), aa0535@coventry.ac.uk (R. Iqbal), c.p.jayne@rgu.ac.uk (C. Jayne), aa9536@coventry.ac.uk (F. Doctor).

state their opinion of the system) is commonly used for movie, music and product review, it is intrusive and alters user browsing pattern (Claypool, Le, Wased, & Brown, 2001). Explicit feedback can be replaced with non-intrusive approach that tries to infer users' interest implicitly. The inference usually takes a sequence of steps which include observing user browsing behaviour, selecting the appropriate set of implicit indicators and modelling them as source of evidence for implicit relevance feedback. These implicit indicators are obtained by observing user activity, for example through user mouse event, dwell time, keyboard event, eye tracking and psychological measures (e.g. facial expression). Indicators from dwell time, mouse and key events are commonly used at low cost whilst eye tracking indicators are expensive and can be intrusive. Eye gaze has been said to have a direct link with human cognition (Buscher, Dengel, Biedert, & Van Elst, 2012), making it the most promising implicit indicator for predicting the user's perceived relevance of web documents. It has been shown to be useful in inferring cognitive states for personalisation (Conati & Merten, 2007). The cost, configuration and portability of an eye tracker mean that it is not easily applicable to real life applications. There is therefore a need to model low cost implicit indicators as a substitution for the eye gaze indicator.

The present challenge with implicit approach is that there are no standard and acceptable methods to determine how users' activities on the web relate to their interest. This study employs the use of an instrumented web browser to capture users' implicit and explicit data from client machines and store them in a central server for further processing. In this research we investigate the consistent and predictive indicators that are frequently used by a community of users with similar interests. The indicators can then be used to build a system which learns from users' behaviour and offers recommendations to them and future users with similar information needs while minimizing the time users spend finding relevant documents.

In this paper, we investigate the relationship between implicit and explicit feedback parameters in different tasks. We derive a predictive model from implicit indicators that can be used to estimate document relevance. Finally, we validate the predictive model with eye gaze tracker. The paper is focused on answering the following research questions:

1. Can specific task situations be used to derive a predictive function model from classical implicit indicators that will signify that a web document is relevant?
2. Can eye gaze predictive indicators be substituted by a predictive function based on classical implicit indicators?

The remaining part of this paper is structured as follows: Section 2 gives an overview of related work. Section 3 presents the methodology and two user studies. Section 4 presents the results and Section 5 provides the conclusion and outlines our future work.

2. Related work

Considerable research has been carried out to improve the quality of information retrieval systems by the use of relevance feedback. Particularly the focus of research has been on implicit relevance feedback or implicit feedback. The Implicit feedback approach uses implicit indicators to replace explicit rating for the development of recommender systems (Ding, Liu, & Tao, 2010; Iqbal, Grzywaczewski, James, Doctor, & Halloran, 2012). It is used unobtrusively to infer the user's information needs based on their interest. Although implicit feedback is widely available, it is considered a secondary option to explicit feedback (Jawaheer et al., 2014) and is noisy and less accurate compared to the explicit

method (Claypool et al., 2001). Current research investigates the best way of replacing explicit feedback measures with implicit feedback approaches (Alhabashneh, Iqbal, Doctor, & Amin, 2015). For instance, in a controlled setting, mouse and scroll movements have been found to exhibit some correlation with explicit rating, but it is somewhat difficult to interpret that in the real world (Buscher, Biedert, Heinesch, & Dengel, 2010). An advantage of the implicit approach is that a large amount of data can be collected ubiquitously without restricting a user to a particular place. The predictive strength of a number of implicit indicators has been investigated in the field of information retrieval. Among the implicit indicators previously investigated include: time spent on a document (also called reading time or dwell time), mouse movement, mouse distance, mouse clicks, amount of scroll movement, copy and paste, printing, highlighting, emailing and bookmarking (Akuma, 2014; Claypool et al., 2001; Iqbal, Grzywaczewski, Halloran, Doctor, & Iqbal, 2015; Konstan et al., 1997; Morita & Shinoda, 1994). Unlike explicit rating which is intrusive, expensive and alters users' browsing behaviour, implicit measures remove the cognitive cost of rating and these are not intrusive (Zemirli, 2012). The next phase of the review will focus on commonly used implicit feedback measures (classical implicit indicators) and gaze-based feedback measures.

2.1. Commonly used implicit indicators

Dwell time is one of the most researched implicit indicators. It is the duration that a document is in focus. In an attempt to effectively substitute explicit rating with implicit measures, Morita and Shinoda (Morita & Shinoda, 1994) introduced the use of dwell time as a behavioural characteristic for creating user profile and data filtration. They conducted an experiment with 8 users who were given a six week task to read articles in a news group they belonged to and explicitly rate them. The investigation was based on how the length of the document, its readability and the amount of unread article affects the reading time. They found that users spend longer time on articles they find interesting but the length of an article does not have any significant effect on the reading time. Using modified distributed software, Konstan et al (Konstan et al., 1997) repeated the study of Morita and Shinoda (Morita & Shinoda, 1994) in a natural setting. Explicit rating and reading time was logged from participants in a recommender system trial. Their findings show that a recommender system based on reading time is as accurate as an explicit recommender system. Both Morita & Shinoda, 1994 and Konstan et al (Konstan et al., 1997) research was based on a single implicit indicator (reading time) and users were restricted to certain news groups thereby limiting their 'true' web experience. In a related study with a focus on academic and professional journal articles, Kim, Oard and Romanik (Kim, Oard, & Romanik, 2000) developed a framework to investigate whether reading time is a good predictor of user interest and whether retention indicators like printing can be used to augment the predictive power of the reading time. Their findings indicate that users spent more time on documents they consider relevant. Research by Núñez-Valdéz et al (Núñez-Valdéz et al., 2012) found that display time and number of visits can be used to represent users' interest in recommending electronic books. Akuma et al (Akuma, Jayne, Iqbal, & Doctor, 2014) investigated the predictive strength of dwell time and mouse activity in a task-specific context. They correlated user generated implicit indicators with explicit relevance ratings on a set of documents and found a positive correlation between the dwell time and the explicit relevance rating. Yi et al (Yi, Hong, Zhong, Liu, & Rajan, 2014) examined the use of item-level dwell time to infer document relevance. They used both client and server-side logging to capture real-world user data from Yahoo across

Download English Version:

<https://daneshyari.com/en/article/6837223>

Download Persian Version:

<https://daneshyari.com/article/6837223>

[Daneshyari.com](https://daneshyari.com)