

An empirical testing of user stereotypes of information retrieval systems

Xiangmin Zhang ^{a,*}, Hui Han ^b

^a *School of Communication, Information and Library Studies, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901-1071, USA*

^b *Department of Computer Science and Engineering, The Pennsylvania State University, University Park, PA 16802, USA*

Received 17 July 2003; accepted 28 January 2004

Available online 18 March 2004

Abstract

Stereotyping is a technique used in many information systems to represent user groups and/or to generate initial individual user models. However, there has been a lack of evidence on the accuracy of their use in representing users. We propose a formal evaluation method to test the accuracy or homogeneity of the stereotypes that are based on users' explicit characteristics. Using the method, the results of an empirical testing on 11 common user stereotypes of information retrieval (IR) systems are reported. The participants' memberships in the stereotypes were predicted using discriminant analysis, based on their IR knowledge. The actual membership and the predicted membership of each stereotype were compared. The data show that "librarians/IR professionals" is an accurate stereotype in representing its members, while some others, such as "undergraduate students" and "social sciences/humanities" users, are not accurate stereotypes. The data also demonstrate that based on the user's IR knowledge a stereotype can be made more accurate or homogeneous. The results show the promise that our method can help better detect the differences among stereotype members, and help with better stereotype design and user modeling. We assume that accurate stereotypes have better performance in user modeling and thus the system performance.

Limitations and future directions of the study are discussed.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Information retrieval; Stereotype; User modeling; Statistical analyses

1. Introduction

One way to improve the performance of information systems is to build user models into systems and customize the system to a user's specific need. User models can be either *group models* built for distinctive groups of users or *individual models* built for individual users. Stereotype is a widely used technique in user modeling for group modeling as well as for creating initial individual user models.

* Corresponding author. Tel.: +1-732-932-7500x8229.

E-mail address: xzhang@scils.rutgers.edu (X. Zhang).

A stereotype is a common user characteristic/trait that is shared by many users (Rich, 1979, 1989). Examples of stereotypes may be ‘expert users’ or ‘novice users’. Stereotypes are created on the assumption that the presence of particular characteristics in one member of the stereotype would imply that of others (Harvey, Smith, & Lund, 1998). Therefore, a stereotype normally contains the common knowledge about a group of users. A new user will be assigned into related stereotype(s) if some of his/her characteristics match the ones contained in the stereotype(s).

Though stereotyping usually involves intensive knowledge engineering on the part of the system administrator (Mostafa, Quiroga, & Palakal, 1998), the advantage of using the stereotype technique is that the knowledge about a particular user will be inferred from the related stereotype(s) as much as possible, without explicitly going through the knowledge elicitation process with each individual user. Another advantage is that the information about user groups/stereotypes can be maintained with low redundancy (Fink & Kobsa, 2000; Rich, 1989). Kuflik, Shapira, and Shoval (2003) have found that using a stereotype was better than using a personal-based user profile in information filtering systems.

Nevertheless, using stereotypes is not without problems. Most stereotypes are formed based merely on users’ external characteristics and on subjective human judgment, usually of a number of users/experts (Shapira, Shoval, & Hanani, 1997). The user’s knowledge about the system and/or task is not involved. It is common that such stereotypes do not represent their members accurately. The issue of inaccuracy of stereotypes has been pointed out by many researchers, e.g., Beaumont (1998), Bellika, Hartvigsen, and Widding (1998), Brajnik, Guida, and Tasso (1990) and Shapira et al. (1997).

The lack of accuracy is liable to lead to conflicts between the individual models and the assignment to various stereotypes (Shapira et al., 1997), which can affect accurate construction of individual user models. Consequently, system functions adapted to individual user models will fail to achieve their goals.

A common practice for the systems using stereotypes is to continuously check user responses, detect and resolve the conflicts between stereotypes and specific user knowledge values, and then update stereotypes and user models. However, we hardly see the evidence on the improvement of user models through such conflict-resolving and stereotype/model updating. Because of the inaccuracy problem, it is important that the user classes represented by the stereotypes be as homogeneous as possible, and this homogeneity should be based on the users’ knowledge of specific domain or task (Beaumont, 1998).

This paper reports the results of an empirical test on the accuracy of some common stereotypes of information retrieval (IR) systems. Since the user’s IR knowledge, the knowledge about IR system components, and the relationships among them are important to information searching performance (Allen, 1996), we argue that the stereotypes of IR systems be based on the user’s IR knowledge, as well as the domain knowledge that is related to the specific search task. We test the accuracy or homogeneity of 11 commonly used user stereotypes by exploring the differences among the members of these stereotypes in terms of their IR knowledge. A stereotype should include only those members with similar level of IR knowledge. We believe such a stereotype is homogeneous and is more accurate than the one based just on the user’s explicit characteristics. We assume such an accurate stereotype adapts better to the individual users of IR systems/services.

The fact that different people have different levels of knowledge is obvious to human experts. However, it is hard for a computer system to intelligently identify the differences. More important, we need to know not only that there exist differences, but also how the users differ and what the exact differences are. Without systematic investigations, such facts cannot be known even to human experts. By conducting this study, we propose a formal method for exploring the differences among the members of a stereotype. This kind of formal method is necessary for IR systems to build better, accurate stereotypes.

The remainder of this paper is organized as follows: Section 2 reviews literatures on stereotype-based user modeling; Section 3 describes the research design; Section 4 analyzes experiment results; Sections 5 and

Download English Version:

<https://daneshyari.com/en/article/10355302>

Download Persian Version:

<https://daneshyari.com/article/10355302>

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