Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 29 (2012) 721 – 725

Procedia Engineering

www.elsevier.com/locate/procedia

2012 International Workshop on Information and Electronics Engineering (IWIEE)

User Interest Modeling and Self-Adaptive Update Using Relevance Feedback Technology

Yi Zhu*, Li He, Xiaojun Wang

Jiangsu Radio and TV University, Nanjing, 210036, P.R. China

Abstract

User interest modeling is the core component of personalized service system. To improve the accuracy of the resulting model, this paper presents a method for modeling user's interest by collecting information of interactive behaviors and combining with implicit feedback techniques. This method also applies explicit feedback and Time-attenuation to self-adaptively update the interest model. It improves the accuracy of the model and will be pretty useful for personalized searching and filtering systems.

Keyword: User Interest; Personalized Service; Relevance Feedback; Time-attenuation

1. Introduction

User interest modeling is the core component of personalized service system. The accurateness of the interest model determines the quality of personalized service. This paper presents a method for modeling user interest by collecting information of interactive behaviors and combining with relevance feedback techniques. To increase the accuracy of the resulting model, time-attenuation factor is also introduced to update self-adaptive user interest.

2. Relevance Feedback

Relevance feedback is a feature of some information retrieval systems. The idea behind relevance feedback is to take the results that are initially returned from a given query and to use information about

* * Corresponding author. *E-mail address*: zhuyi@jstvu.edu.cn

doi:10.1016/j.proeng.2012.01.030

whether or not those results are relevant to perform a new query to get a better result. The procedure could repeat on and on. Relevance feedback technique overcomes the limitation of unidirectional relationship between user and the system and increases the interaction to improve the quality of search results. [1]

According to distinguishable feedback behaviors, there are three types: explicit feedback, implicit feedback, and blind or "pseudo" feedback. Explicit feedback is obtained from user' s direct identifying, evaluating and filtering of the results. Implicit feedback can be inferred from user behavior. Pseudo relevance feedback, also known as blind relevance feedback, assumes that the top "K" ranked results are relevant.

Among the three types, explicit feedback is more accurate because of user' s participation. Implicit feedback automatically collects interest information without requiring any special efforts by the user or interrupting their browsing. It's a more practical method to reduce the burden on users. Blind relevance feedback is less adapted because it's too simple and might cause severe Matthew Effect.

3. User interest modeling based on Relevance Feedback

Interests and demands are various. If a system tries to better serve the users, it should analyze characteristics of different user interests to provide a customized service that is called personalized information service. User interest modeling is the key component of the personalized service system. Generally, there are two major tasks of the model: to survey user operations to recognize its intention; to determine the next response or action of the computer system. [2]

The model is constructed in four steps: Data Collecting, Representation, Learning and Updating. Data collecting provides indispensable source for Learning and Updating. Quantity and quality of the data impact on the efficiency and effect of learning and updating procedure. Representation provides the user interest a formative description. The formatting of interest ends up with the learning process. Model maintenance and update keep track of latest interests to improve the effectively and practicability of recommender system. [3]

Data collecting is a procedure to obtain information on user's characteristics, preferences or behaviors. Similar to relevance feedback, there are two types of collecting methods, explicit and implicit. The explicit one is simpler. It requires user to input related information, such as comment on certain items, self-characteristic, and explicit intention. The other automatically collects and analyzes relevant behaviors without any additional work on the part of the user or interrupting user operations.

3.1. Retrieval of Relevance Feedback information

Nichols presented a list of categorical types of user behaviors that could be used as sources for implicit feedback (Nichols, 1997). Oard and Kim surveyed the state of the art in implicit feedback techniques with an eye toward their potential use for information filtering (Oard & Kim, 1998). Based on the sources of implicit feedback presented by Nichols, they identified three broad categories of potentially useful observations: Examination, Retention, and Reference. [4][5]

According to Kim's classification criteria, system-user interaction behaviors could be analyzed and applied to construct user interest models. Table 2 shows the behaviors: (1) Examination: Click, Vote (2) Retention: Bookmark, Share (3) Reference: Reply/Discussion. These behaviors reflect user's interest in item Pi to some extent.

Table 1. Observable behaviors for implicit feedback in computer system

Category	Observable Behavior
Examination	Click, Vote
Retention	Bookmark, Share
Reference	Reply/Discussion

3.1.1. Click behavior

Generally, click behavior is an effective implicit feedback to understand user's interest. Higher click rate on an item implies greater interest of the user. Therefore, many systems like to take click behavior as a very important parameter. However, experiments prove that the behavior could just be an assistant parameter. Clicking on an item might just mean that user is attracted by the title or the picture, but is not interested in the content. User's real interest must be determined by his further behaviors, such as reply, mark, recommend and bookmark.

3.1.2. Vote behavior

Many systems encourage user to vote on a certain item. Usually there are two types of vote behaviors: Simple Vote and Rating Vote.

Simple vote offers two options, like or dislike. The options might be more detailed sometimes, for example, there are six options in news.sohu.com: "cheering", "sad", "touching", "irritating", "funny" and "boring".

Rating Vote let user rate an item as he like. Commonly, the rate vote number is a range from 5 to 0. 5 means like it most, 1 do not like it at all, 0 no marking.

Vote behavior requires user to read the content of the item after click behavior, which makes vote a more effective parameter for user interest model.

Rating vote is more exact than simple vote.

3.1.3. Bookmark behavior

Bookmark behavior indicates user's very fond of an item. It is usually an immediate response when a user thinks the item is good and gives a "like" simple vote. But user bookmarks the item not only because he likes it but also because he wants to keep it and read it again. Compare with vote behavior, bookmark goes further in representing user's interest.

3.1.4. Reply/Discussion behavior

Reply/Discussion is a very positive behavior. User replies to an item and discusses it only when he feels very interested. Naturally, discussion is alternative. User might reply to agree on a point of view he likes or discuss something he opposes. Either case shows user's great interest in the item. Thus, user's interest indicated by reply behavior is the strongest.

3.2. User interest estimation

Based on the analysis, study and statistics of different behaviors, a user interest model could be constructed. Experiments show that unary linear regression could be adopted for the modeling to describe the relationships of user, item, user implicit behavior and user's interest towards the item.

Linear Regression Analysis is an approach to analyze variation tendency and establish the function model for investigation of relationships between variables.

Then, relationship between user behavior and interest may be written:

$$d(P) = \alpha \cdot L + \beta \cdot F + M + \gamma \cdot R \quad (1)$$

Where:

- P is the item
- L is the click rate with an initial value 0. User clicks once, then L will add 1
- F is a bookmark behavior initially 0. User bookmarks the item, then F will be 1
- M is a vote behavior. The value of M varies under different vote rules. In a simple vote, "Like" behavior gives M a value of 2, "Dislike" gives M -2. In a rating vote, M is among 0-5 given by user
- R is reply behavior initially 0. If the user attends the discussion of the item, R will be 1
- α、β、γ are constant parameters having nothing to do with d(P). After generalized least squares estimation and statistics on different behavior interests and practical application, the values of the parameters are proposed: α=0.1，β=5，γ=7. The former equation is adjusted:

$$d(P) = 0.1 \cdot L + 5 \cdot F + M + 7 \cdot R \tag{2}$$

The BI (Behavior Interesting) of each item could be calculated using this equation.
User interests in all items could be represented based on the equation:

$$B_i = \{d(P_1), d(P_2), d(P_1), \ldots, d(P_j), \ldots, d(P_m)\} \tag{3}$$

Hence, the characteristic matrix of all user interest models is proposed:

$$BI = \begin{bmatrix} (u_1, d(P_1)) & \cdots & (u_1, d(P_j)) & \cdots & (u_1, d(P_m)) \\ \vdots & & \vdots & & \vdots \\ (u_j, d(P_1)) & \cdots & (u_j, d(P_j)) & \cdots & (u_1, d(P_m)) \\ \vdots & & & & \vdots \\ (u_n, d(P_1)) & \cdots & (u_n, d(P_j)) & \cdots & (u_n, d(P_m)) \end{bmatrix} \tag{4}$$

3.3. Self-adaptive update of user interest

Usually, recommendation algorithm could be used in the above characteristic matrix. However, the interest model has two demerits:

Firstly, it is static. But user's interest varies over time. User ought to be more interested in an item just bookmarked yesterday than in a second one, which he discussed two years ago.

Therefore, a new parameter called Time-attenuation should be introduced to offer a self-adaptive update method for user interest. Because interest variation is a relatively long procedure, the calculation of time uses month as basic unit, three months as a period. The time-attenuation is computed:

$$T = 1 / (INT(t1 - t0) / 3) + 1) \tag{5}$$

Where t0 is the time when the user feels interested, t1 is current time, t0 and t1 uses month as unit. Function INT is integral round function.

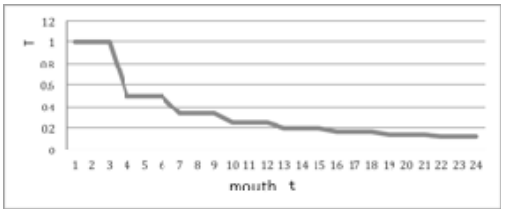

Fig. 1. Time-attenuation

Secondly, implicit feedback reduces the burden on users in comparison with explicit feedback, but it requires more operation behaviors. In other words, the interest model for new users may not be accurate. Thus, explicit feedback should be adapted to construct the initial model for new users. The algorithm is described as follows:

First of all, system classifies all the items into different categories. Then, it selects a certain amount of typical items in different categories and provides them to user when he is registering. Set the interest in items chosen by user as constant δ which is nothing to do with *d(P)*. The value of δ is proposed as 4 by statistics. Finally, $d_{base}(P)$, an initial interest model, is constructed. Of course, user could adjust the $d_{base}(P)$ to change the interest.

With the above solutions to the two demerits, the equation of user interest estimation is adjusted:

$$d(P) = (\alpha \cdot L + \beta \cdot F + M + \gamma \cdot R) + d_{base}(P) \quad (6)$$

4. Conclusion

Personalized information service is becoming more and more popular. User interest modeling is the basic and core of a system providing personalized services. This paper presents a method for user interest modeling using relevance feedback techniques. This method applies explicit feedback and Time-attenuation to self-adaptively update the interest model. It improves the accuracy of the model and will be pretty useful for personalized searching and filtering systems.

References

[1] Spink A, Losee R M. *Feedback in information retrieval*. Annual Review of Information Science and Technology, 1996(31), p. 33-78

[2] Sun T, Yang F. *An approach of building and updating user interest profile according to the implicit feedback*. Journal of Northeast Normal University, 2003,35(3).

[3] Wu L, Liu L, *User Profiling for Personalized Recommending Systems-A Review*. Journal of the China Society for Scientific and Technical Information，2006(2), p. 55-62.

[4] Kim J, Oard D, Romanik K. *Using Implicit Feedback for User Modeling in Internet and Intranet Searching*. University of Maryland CLIS Technical Report, 2000(00-01)

[5] Diane K, Jaime T. *Implicit Feedback for Inferring User Preference: A Bibliography*. ACM SIGIR Forum, 37 (2), p. 18-28