The Model of User Interest Update and User Classification in Personal Information Push Service

Cheng Chen\textsuperscript{a}, Gui-Ping Liao\textsuperscript{a}, Xiao-Hui Shi\textsuperscript{a}, Xing Zhao\textsuperscript{a},\textsuperscript{a*}

\textsuperscript{a} Agricultural Information Institute of Hunan Agricultural University, Changsha 410000, China

Abstract

Push technology as a proactive type of information acquisition method of communication, it has an effective way to achieve personalized information services. However, the effect of push technology is general in personalized service practice. Push technology still not resolved its two historical problems, which are user reasonable classification and update user interest. Facing these problems, this paper proposed a new method called information of three-dimensional concrete, and based on this new method, with reasonable mathematical methods, the authors proposed reasonable interest of users update and user classification mechanism, constructing the model which is suitable for personal classification and update user were effectively overcame.

Keywords: personalized information services; push technology; three-dimensional specific visualization of information; user interest update; user classification

1. Introduction

User-centric, in-depth understanding, learning the user classification, analysis of user requirement characteristics, user knowledge structure is required by personalized information service, which also provides user with information activity environment meeting individual demand. Therefore, how to dig the applicable, simple and refined interested information from contents which are multifarious and disorderly, and push it to user without delay are significant research subjects. As a search focus in active communication-service area today, a new direction for achieving personalized information service will be provided by information push technology. But currently, most of push products do not meet the requirements of individual user well for lack of direction, and its performance in the information service is

* Corresponding author. Tel.: +86-731-8463-8372; fax: +86-731-8463-8372
E-mail address: lgpxf@yahoo.com.cn.
also not ideal. Due to these issues before, a reasonable classification and user interest update are always the barriers of push service.

In order to find an effective way to solve the difficult issues, mathematical methods based on three-dimensional specific visualization of information and quantitative techniques are used to explore reasonable user interest update and reasonable user classification mechanism for facing the defects of push technology in information service.

2. Methods

2.1. Three-dimensional specific visualization of information

Under the background of information flooding, selecting and absorbing useful information are not easy. After meeting the user requirements of information service, more attention should be given for the connection between the key properties and users themselves. In that way, user can feel the objective things from an accustomed perspective, searching visual things from things from nonobjective status.

During the process of information push service, time attribute, relationship attribute and content attribute of information that paid close attention to is favorable for accurately summarizing information’s three key attributes, which can be interpreted as providing update time, the connection between information and users’ professions, and information’s content form.

In this paper, 8 pieces of representative information are chosen to express the specific and vivid process of information. Information will be reduced by the degree of care for the 8 pieces of information’s various attributes. Then, time attribute, content attribute, and relationship attribute of the reduced information will be projected to three-dimensional coordinate(x, y, z). Different time, content and relationship present different points, which also presents one piece of information of three-dimensional specific visualization of information (Fig. 1-a). In figure1, the yellow point presents the viewed information of time and content means only that the time and content attributes left after reduced information. The green point presents the viewed information of relation and content means only that the relation and content attributes left after reduced information. The brown point presents the viewed information of relation, time, and content means that the relation, time, and content attributes left after reduced information.

Figure 1-b, figure 1-c, and figure 1-d are the three-dimensional specific visualization of information from the perspective of three limits. The figures demonstrates that the three kinds of users show little care about the relation, time, and content attribute of information. In this situation, the information attribute cared by users can be considered rather than the information attribute that are not cared by users.
2.2. Quantization of three-dimensional specific visualization of information

Information of three-dimensional specific visualization which owns accurate generalization and vivid visualization, improves the information awareness of user, but still lacks in necessary mathematical characters. Hence, solving the issues during the process of information service with scientific mathematical method and improving the personalization degree of three-dimensional specific visualization of information carried out at a time. The specific quantization process as follows:

1) Digitalization of various axis coordinates in three-dimensional information space.
As the result of visualized information, time, content, and relation attributes are projected to x-axis, y-axis, and z-axis. Every point of x-axis, y-axis, and z-axis respectively corresponds to one attribute value of the three attributes, and data range of every coordinate axis is: x ∈ [0, ∞), y ∈ [0, ∞), z ∈ [0, ∞), respectively.

2) Spatial partition of three-dimensional information
Information space of three-dimensional is divided into divers relative independent cube (figure 2). Every point respectively corresponds to one space coordinate of three-dimensional. The tentative side length of every cube space is 100.

3) The classification of information
According to the results of classification, information will be progressively encoded, for instance, agricultural and rural information has been classified as: crop production, animal husbandry, forest industry, fishing industry and sideline production, which will be encoded with primary serial number presented as 01, 02, 03, 04 and 05. Crop production also contains rice production and wheat production and so on. Then rice production and wheat production will be also encoded with secondary serial number presented as 011 and 012. Following this way, swine production and sheep production belonged to animal husbandry can be encoded as 021 and 022.

4) Determination of information storage space
At first, all kinds of information will be independently distributed with storage space on the basis of primary serial number. The size of storage space will be decided by the secondary serial number of this kind of information. Secondary serial number under the same primary serial number will be intensively stored in its storage of primary. Assume that there are 01, 02 and 03, three primary serial number separately having 9 secondary serial numbers. The storage space of information can be arranged as presented in figure 3. The different surfaces of storage space belonged to primary serial number given with different colors are distinguished easily.

5) Quantification of information in storage space
Following the classification, information has been stored to the information space. The simple mapping process is left in the task of information quantification. In the information space, every piece of stored information distributed with only one three-dimensional coordinate (x, y, z), can be quantized a point in three-dimensional space.

![Fig.2. The division of space](image1)

![Fig.3. The division of information storage space](image2)
2.3. Definitions of relative conception

Considering this new conception of three-dimensional specific visualization of information, some conceptions involved in the following study are given with some new definitions, which are favorable for expressing and solving issues.

Definition 1. Let \( I \) be a piece of information in reality, if under the condition \( F \), \( I \) can be reduced to three information owning different attributes, and this information can be projected to three-dimensional information space \( (V_1, V_2, V_3) \), this process called three-dimensional specific visualization of information.

Definition 2. Visualized information \( I_v \) owning \( V_1, V_2 \) and \( V_3 \), \( I_v \) can be projected to three-dimensional space \((x, y, z)\), \( x \in [0, \infty), y \in [0, \infty), z \in [0, \infty) \), every point of \( x \)-axis, \( y \)-axis and \( z \)-axis respectively corresponds to one attribute value of the three attributes. This process is called quantification of three-dimensional specific visualization of information.

Definition 3. Let \( I_q (x, y, z) \) be the quantified three-dimensional specific visualization of information, \( x \in X, y \in Y, z \in Z \).

Definition 4. Let \( c_1, c_2, c_3 \) be the random variables of three-dimensional information space in definition 2. If \( c_1, c_2, c_3 \) are separately displacements presented in \( x \)-axis, \( y \)-axis and \( z \)-axis from differences between current interested information of user and previous interested information of user. Then vector \( c = (c_1, c_2, c_3) \) is the transport vector of user’s interest. \( S_m = \sqrt{c_1^2 + c_2^2 + c_3^2} \) is called the transport strength of user’s interest.

Definition 5. Let the interested information \( I_s = I_q (x_0, y_0, z_0) \), let \( t \) be time variable, \( 0 \leq t < \infty \), the next moment interested information of user is defined as \( I_s = I_q (x_0 + c_1t, y_0 + c_2t, z_0 + c_3t) \).
3.2. Classification model of user

According to analysis before and the strength \((S_m)\) of interest transport, interest transport of user can be classified to three kinds.

1. \(C_1, c_4\) and \(c_7\) present transportable users in the same small industry. For example, the user who are growing rice may be always care about the related information of rice. There is a little transport, this kind of users are called “quiet” user.

2. \(C_2, c_3\) and \(c_8\) present transportable users in different and small industries, but under the same big industry. For example, users working in animal husbandry may currently care about the information of swine production, and at the next moment, more attention will paid on the information about dairy husbandry. However, the two industries also belong to animal husbandry. These kind of users are called “active” users.

3. \(C_3, c_6\) and \(c_9\) present transportable users in different and big industries. For example, the users are caring information about rice, but at next moment, more attention will be paid for swine production. This kind of users are called “naughty” users for very large transport.

According to strength of user interest transport, given match table of user (table 1) and matching table of information (table 2), users are classified to three main classes presented as “quiet”, “active” and “naughty” user. Accurate matching of information will be carried out by using the recognition attribute of information and user.

<table>
<thead>
<tr>
<th>Name</th>
<th>Field name</th>
<th>Data type</th>
<th>Field length</th>
<th>Permitting empty</th>
<th>remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>User name</td>
<td>Userid</td>
<td>Varchar</td>
<td>20</td>
<td>N</td>
<td>Recognition attribute of information (describing the field related to information, matching with users classes in users matching table)</td>
</tr>
<tr>
<td>User classes</td>
<td>Type</td>
<td>Char</td>
<td>20</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Users matching table

<table>
<thead>
<tr>
<th>Name</th>
<th>Field name</th>
<th>Data type</th>
<th>Field length</th>
<th>Permitting empty</th>
<th>remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information type</td>
<td>Type</td>
<td>Char</td>
<td>20</td>
<td>N</td>
<td>Recognition attribute of user (describing the field cared by users, matching with users classes in users matching table)</td>
</tr>
<tr>
<td>Relation</td>
<td>Relation</td>
<td>Varchar</td>
<td>255</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Updating time</td>
<td>Time</td>
<td>Varchar</td>
<td>20</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Content Type</td>
<td>Content</td>
<td>Varchar</td>
<td>20</td>
<td>N</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Information matching table

Not influencing the matching accuracy of user and information, strength of interest transport is used for the first classification rule. Differential management will be carried out to aim at the different characteristics of “quiet”, “active” and “naughty” user. In the process of information push service, the
difficult issues will be solved, such as classification rule not meeting the information push accuracy and updating user’s interest in time. The specific classification process of user is shown in the figure 5.
4. Conclusion

In this paper, the primary restraining factors has been discussed in the growth of personalized information push technology. The new approach of three-dimensional specific visualization of information has been presented to aim at these issues in the interest of updating user and user’s reasonable classification. The suitable model of user interest update and user classification has been constructed by the new approach combined with scientific mathematics methods. Pushing interested information to user in time, users automatically searching applicable, simple and refined information from the boundless information sea, the model can effectively solve the issue of user interest update and user’s reasonable classification, enhancing the level of individuation. In this respect of research results popularization, the established model owns strong universality because of explicit mathematics methods used for the basis of model construction. Research results are also suitable for personalized information fields to some extent, having profound scientific and practical significance for the personalized information service popularization of this model.

Acknowledgements

This research is funded by the Graduate Science and Technology Innovation Foundation of Hunan under Grant No.CX2010B280. We would also like to thank Xing Zhao, Leo Christopher C. Viray, and Wen-Pu Li from Hunan Agricultural University of China for their help. Finally, authors are especially grateful to the anonymous reviewers of this paper for their invaluable comments.

References