Culturally Enhanced Collaborative Filtering

Mahboobe Zardosht and Nasser Ghasem-Aghaee

Abstract—We propose an enhanced collaborative filtering method using Hofstede’s cultural dimensions, calculated for 111 countries. We employ 4 of these dimensions, which are correlated to the costumers’ buying behavior, in order to detect users’ preferences for items. In addition, several advantages of this method demonstrated for data sparseness and cold-start users, which are important challenges in collaborative filtering. We present experiments using a real dataset, Book Crossing Dataset. Experimental results shows that the proposed algorithm provide significant advantages in terms of improving recommendation quality.

Keywords—Collaborative filtering, Cross-cultural, E-commerce, Recommender systems

I. INTRODUCTION

The continued growth and rapid development of e-commerce and as a consequence the huge amount of information on the internet, has led to a proliferation of recommendation technologies of e-commerce. Personalized services, such as recommender systems, help visitors effectively locate their information on the internet, has led to a proliferation of personalization tools. Personalized services, such as Amazon, eBay, Taobao, etc., have been proved for data sparseness and cold-start users, which are important challenges in collaborative filtering. We present experiments using a real dataset, Book Crossing Dataset. Experimental results shows that the proposed algorithm provide significant advantages in terms of improving recommendation quality.

Many different kinds of recommendation technologies have been proposed that among all collaborative filtering (CF) is one of the most successful and widely used technologies in personalization and recommender systems. Numerous commercial web sites have adopted different kinds of recommender systems with various levels of personalization, such as Amazon, eBay, Taobao, etc. So it has practical importance to enhance the research on the personalized recommendation technologies of e-commerce.

The sparsity problem occurs when available data is insufficient for identifying similar users or items (neighbors) due to an immense amount of users and items [2]. Even in a large e-commerce site, the number of items rated by a user is usually less than one percent of total items [2]. Clearly the percentage of common items rated by two or more users is much less than that, which results in a very sparse user-item ratings matrix. With the large scale and sparse ratings matrix, the result of calculating similarities between users or items would be difficult and unreliable. As a result the recommendation quality may become poor.

II. RELATED WORK

A. Background on collaborative filtering

In a collaborative filtering (CF) scenario, generally we start with a list of m users \( U = \{u_1, u_2, \ldots, u_m\} \), a list of n items \( I = \{i_1, i_2, \ldots, i_n\} \), and a mapping between user-item pairs and a
set of weights. The latter mapping can be represented as a \( m \times n \) matrix \( M \). In the traditional CF domain the matrix \( M \) usually represents user ratings of items, thus the entry \( M_{ij} \) represents a user \( u_i \)'s rating on item \( j \). In this case, the users' judgments or preferences are explicitly given by matrix \( M \). This weight may be binary (representing the existence or non-existence of the item in the user session), or it may be based on the amount of time spent on the particular item during the session.

For a given active user (also called the target user) \( u_a \), the task of a CF system is to (1) predict \( M_{a,j} \), for a given target item \( j \) which has not already been visited or rated by \( u_a \); or (2) recommend a set of items that may be interesting to user \( u_a \).

In user-based CF algorithms, first a set of \( k \) nearest neighbors of the target user are computed. This is performed by computing correlations or similarities between user records (rows of the matrix \( M \)) and the target user. Then different methods can be used to combine the neighbors' item ratings (or weights) to produce a prediction value for the target user on unrated (or unvisited) items. A major problem with this approach is the lack of scalability: the complexity of the system increases linearly as a function of the number of users which, in large-scale e-commerce sites, could reach tens of millions.

In contrast, item-based CF algorithms attempt to find \( k \) similar items that are co-rated (or visited) by different users similarly. This amounts to performing similarity computations among the columns of matrix \( M \). Thus, item-based CF algorithms avoid the bottleneck in user-user computations by first considering the relationships among items. For a target item, predictions can be generated by taking a weighted average of the target user's item ratings (or weights) on these neighbor items.

**B. Background on culture and consumers behavior**

As people from all over the world can use the Internet as a purchasing channel the role of national cultures in the purchasing process assumes more and more relevance. Also efforts for developing ecommerce and globalization this industry, needs more attention to consumer behaviors and needs and this issue highlight the role of national culture more. Most of the research in this field aims at identifying attraction factors for buyers [4].

In a Canadian experiment [5] using Amazon.com, have shown that providing recommendations to customers and providing consumer reviews increase the perceived usefulness of a website. In [6] broadband users in the USA have been studied and it has been found that consumers find emotional and practical benefits in participating in online discussions and that these discussions have profound commercial implications for sales of electronics products.

Also the way a website is designed can also influence consumer behavior and trust. In a study in the USA and Finland, [7] evaluated website elements which affect consumers’ perception of credibility. Elements that highlight the brick-and-mortar nature of organizations, such as listing physical addresses and contact phone numbers, enhance the website credibility. Also in [6] found that tailoring the website to the user experience leads to increased perceptions of website credibility. The general conclusion is that it is profitable for companies to tailor websites to local tastes by adapting content, language, and style [8], [9], [10]. These studies mainly focus on the appearance of the websites and online market places and its influence on consumers' behavior. Abilities of websites' other elements, including recommender systems, have not been studied yet [11].

In addition, most studies present in this literature focus on a single country, mostly the USA [12], [13], a trait that has been criticized by several authors. Other studies compare countries geographically and culturally very distant such as China and the USA or the USA and Finland [14], [9], [15]. In order to maximize chances of finding some differences, previous research interested in cultural issues in e-commerce highlights the importance of culture by investigating differences between countries that are clearly very different. With the severe competition between economic companies, it is obvious that only paying attention to the big differences between countries, in order to observe the role of national culture in consumer behavior, is not enough.

In many studies, [16], [17], [15], it has been observe that economic activities in internet is not boundary less and it is related to the cultures. As a result, studies about one country are not necessarily applicable for other countries, unless they have similar cultures. It is obvious that studies on special countries do not help that much in recognizing effects of cultures on buying behavior.

**III. CULTURE-BASED RECOMMENDER SYSTEMS**

In this section, we first introduce the national culture dimensions presented by Hofstede. Then the relation between culture and consumers behavior examine with real online purchasing data from different countries. We then present our approach to integrate the presented cultural dimensions into the item-based collaborative filtering framework.

**A. Hofstede’s national culture dimensions**

Geert Hofstede developed a model of six dimensions of national culture that helps to understand basic value differences. This model distinguishes cultures according to six dimensions: power distance, individualism/collectivism, masculinity/femininity, uncertainty avoidance, long-term orientation, and indulgence/restraint. The dimensions are measured on a scale from 0 to 100. The model is based on quantitative research and gives scores for 111 countries and regions [18].

1) **Power Distance (PDI)**

Power distance is the extent to which the less powerful members of organizations and institutions (like the family) accept and expect that power is distributed unequally. This represents inequality (more versus less), but defined from below, not from above. It suggests that a society's level of inequality is endorsed by the followers as much as by the leaders. Power and inequality, of course, are extremely fundamental facts of any society and anybody with some international experience will be aware that “all societies are unequal, but some are more unequal than others”.

International Scholarly and Scientific Research & Innovation 5(11) 2011 1648
2) Individualism/Collectivism (IDV)

Individualism on the one side versus its opposite, collectivism, is the degree to which individuals are integrated into groups. On the individualist side we find societies in which the ties between individuals are loose: everyone is expected to look after her/himself and her/his immediate family. On the collectivist side, we find societies in which people from birth onwards are integrated into strong, cohesive in-groups, often extended families (with uncles, aunts and grandparents) which continue protecting them in exchange for unquestioning loyalty. The word collectivism in this sense has no political meaning: it refers to the group, not to the state. Again, the issue addressed by this dimension is an extremely fundamental one, regarding all societies in the world.

3) Masculinity/Femininity (MAS)

Masculinity versus its opposite, femininity, refers to the distribution of emotional roles between the genders which is another fundamental issue for any society to which a range of solutions are found. The IBM studies revealed that (a) women's values differ less among societies than men's values; (b) men's values from one country to another contain a dimension from very assertive and competitive and maximally different from women's values on the one side, to modest and caring and similar to women's values on the other. The assertive pole has been called masculine and the modest, caring pole feminine. The women in feminine countries have the same modest, caring values as the men; in the masculine countries they are more assertive and more competitive, but not as much as the men, so that these countries show a gap between men's values and women's values.

4) Uncertainty Avoidance (UAI)

Uncertainty avoidance deals with a society's tolerance for uncertainty and ambiguity. It indicates to what extent a culture programs its members to feel either uncomfortable or comfortable in unstructured situations. Unstructured situations are novel, unknown, surprising, different from usual. Uncertainty avoiding cultures try to minimize the possibility of such situations by strict laws and rules, safety and security measures, and on the philosophical and religious level by a belief in absolute Truth: “there can only be one Truth and we have it”. People in uncertainty avoiding countries are also more emotional, and motivated by inner nervous energy. The opposite type, uncertainty accepting cultures, are more tolerant of opinions different from what they are used to; they try to have as few rules as possible, and on the philosophical and religious level they are relativist and allow many currents to flow side by side. People within these cultures are more emotional and contemplative, and not expected by their environment to express emotions.

5) Long-term orientation (LTO)

Long-term oriented societies foster pragmatic virtues oriented towards future rewards, in particular saving, persistence, and adapting to changing circumstances. Short-term oriented societies foster virtues related to the past and present such as national pride, respect for tradition, preservation of "face", and fulfilling social obligations.

6) Indulgence/Restraint (IVR)

Indulgence stands for a society that allows relatively free gratification of basic and natural human drives related to enjoying life and having fun. Restraint stands for a society that suppresses gratification of needs and regulates it by means of strict social norms.

B. Culture and Ecommerce

It seems that with the existence of internet, the national and geographical boundaries should become irrelevant. Consequently, global expansion on the Internet could promise greater customer reach and profits. However, although the adoption rate of Internet shopping is relatively high in the West, it is still generally unpopular in the East. We show in this paper that despite the promises of greater global customer reach and potential profits, Internet shopping is still systematically affected by cultural differences.

To examine the effect of cultures on consumers' behavior in ecommerce we use “E-commerce by individuals and enterprises” dataset. This dataset released annually by Eurostat, the statistical office of the European Union, containing online purchasing information of 24 European countries. We calculate correlation between different cultural dimensions and the amount of online purchases in these countries in 2010. As the results shows in Table I, 4 of these 6 dimensions related to the ecommerce and we use them later in our algorithm to improve quality of recommendation. The Power Distance (PDI) and the Uncertainty Avoidance (UAI) have negative correlation and the Individuality (IDV) and the Indulgence/Restraint (IVR) have positive correlation with ecommerce.

| TABLE I CORRELATION BETWEEN ECOMMERCE AND CULTURAL DIMENSIONS |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| PDI             | IDV             | MAS             | UAI             | LTO             | IVR             |
| -0.62166        | 0.69715         | -0.37646        | -0.6975         | 0.01966         | 0.7197          |

C. Integrating culture with collaborative filtering

As noted earlier, the item-based CF framework provides a computational advantage over user-based approaches, since item similarities can be computed offline, prior to the online task of generating recommendations. But, this framework also provides another important advantage. Since the computation of item similarities is independent of the methods used for generating predictions or recommendations, other sources of evidence about items (in addition to item ratings or weights) can be used for performing the similarity computations.

Cultural dimension for each item i (CulDim(i, dim)) computed based on the cultural dimensions of users who have rated that item. For example if uᵢ, uⱼ and uᵣ rated item i, the cultural dimension for item i is computed based on cultural dimension of uᵢ, uⱼ and uᵣ that is presented by Hofstede regarding to these users’ country. Cultural dimension of each item is measured by the following equation:

\[ \text{CulDim}(i, \text{dim}) = \frac{\sum \text{CulDim}(u, \text{dim}) \times \text{item rating}_u(i)}{\sum \text{item rating}_u(i)} \]
where dim is one of the four cultural dimensions \( PDI, IDV, UAI, \) and \( IVR \). For each item these 4 dimensions should be calculated. \( U \) is set of users who have voted for item \( i \), \( M_{i,u} \) denote the vote of user \( u \) for item \( i \) and \( \text{cul}_{u}\text{dim} \) refers to the value of dim for user \( u \).  

The cultural dimensions show that users with which cultural attitude attracted to this item and also can be used as extra information when computing items similarities.  

The integration of cultural similarities for items with rating (or usage-based) similarities provides two primary advantages. First, the cultural attributes for items provide additional clues about the underlying reasons for which a user may or may not be interested in particular items (something that is hidden behind the rating values in the usual context). This, in turn, allows the system to make inferences based on this additional source of knowledge, possibly improving the accuracy of recommendations. Second, in cases where little rating (or usage) information is available (such as in the case of very sparse data sets), the system can still use the cultural similarities to provide reasonable recommendations for users.  

In the following we describe our approach for integrating cultural similarities into the standard item-based collaborative filtering framework. Our approach involves first computing item similarities, both based on the items cultural dimension matrix, as well as based on the user-item ratings (or usage) matrix. Then, we use a combined similarity measure, as a linear combination of the two similarities to perform item-based collaborative filtering.  

1) Combined similarity measure  

The cultural similarity measure \( \text{CulSim}(i_p, i_q) \), for a pair of items \( i_p \) and \( i_q \), is measured by the following equation:

\[
\text{CulSim}(i_p, i_q) = \frac{1}{\sqrt{\sum_{\text{dim}} \text{CulDim}(i_p, \text{dim}) - \text{CulDim}(i_q, \text{dim})}}
\]

where \( D \) is set of the four cultural dimensions.  

Similarly, we compute item similarities based on the user-item matrix \( M \). We use the adjusted cosine similarity measure in order to take into account the variances in user ratings. We denote the rating similarity between two items \( i_p \) and \( i_q \) as \( \text{RateSim}(i_p, i_q) \),

\[
\text{RateSim}(i_p, i_q) = \frac{\sum_{k=1}^{n} (M_{k,p} - \bar{M}_p) \times (M_{k,q} - \bar{M}_q)}{\sqrt{\sum_{k=1}^{n} (M_{k,p} - \bar{M}_p)^2 \times \sum_{k=1}^{n} (M_{k,q} - \bar{M}_q)^2}}
\]

where \( M_{k,p} \) represents the rating of user \( k \) on item \( i_p \) and \( \bar{M}_p \) is the average rating value of user \( k \) on all items. Finally, for each pair of items \( i_p \) and \( i_q \), we combine these two similarity measures to get \( \text{Sim} \) as their linear combination:

\[
\text{Sim}(i_p, i_q) = \alpha \cdot \text{CulSim}(i_p, i_q) + (1 - \alpha) \cdot \text{RateSim}(i_p, i_q)
\]

where \( \alpha \) is a parameter between 0 and 1, specifying the weight of cultural similarity in the combined measure. If \( \alpha = 0 \), then \( \text{Sim}(i_p, i_q) = \text{RateSim}(i_p, i_q) \), in other words we have the standard item-based filtering. On the other hand, if \( \alpha = 1 \), then only the cultural similarity is used which, essentially, results in a form of content-based filtering. Finding the appropriate value for \( \alpha \) is not a trivial task, and is usually highly dependent on the characteristics of the data. We choose the proper value by performing sensitivity analysis for particular data sets in our experimental section below.

In order to compute predicted ratings, we use the weighted sum approach,

\[
M_{a,i} = \frac{\sum_{j=1}^{l} (M_{a,j} \times \text{Sim}(i, j))}{\sum_{j=1}^{l} \text{Sim}(i, j)}
\]

where, \( M_{a,i} \) denotes the prediction value of target user \( u_a \) on target item \( i \).

IV. EXPERIMENTAL RESULT

In this section, we empirically evaluate the culture based recommender algorithm and compare its performance against the performance of the benchmark algorithms.

A. The Dataset and Evaluation Metrics

The experimental data is Book-Crossing Dataset Contains 278,858 users (anonymized but with demographic information) providing 1,149,780 ratings (explicit / implicit) about 271,379 books. The dataset comprises 3 tables.

- **BX-Users**, contains the users. Demographic data is provided (‘Location’, ‘Age’) if available. Otherwise, these fields contain NULL-values. Between users, we select those who were from countries whose Hofstede’s cultural dimensions presented. And also have rated more than 20 books.
- **BX-Book-Ratings**, contains the book rating information. Ratings (‘Book-Rating’) are either explicit, expressed on a scale from 1-10, or implicit, expressed by 0. We choose explicit ratings, which were made by valid users.
- **BX-Books**, books are identified by their respective ISBN. ‘Book-Title’, ‘Book-Author’, ‘Year-Of-Publication’ and ‘Publisher’ are denoted. Books with more than 5 rates selected.

To measure the accuracy of the recommendations we computed the standard Mean Absolute Error (MAE) between ratings and predictions in the test data sets. The MAE is computed as:

\[
\text{MAE} = \frac{\sum_{i=1}^{n} | a_i - p_i |}{n}
\]
neighbors on the MAE, in Fig. 1, we plot the MAE with respect to the number of neighbors. In this case, the ratings are based on a discrete scale of 1 (lowest) to 10 (highest). Thus, the maximum possible value for MAE is 9 (indicating a maximum possible error on all predictions).

B. Experiment with real data

In this section we present detailed experimental results. Table II depicts the prediction accuracy of our culturally enhanced recommendations in contrast to those produced by standard item-based collaborative filtering. Here the MAE has been calculated with respect to the number of neighbors (similar items) in the k-nearest-neighbor algorithm. As it is obvious in the table, the culturally enhanced approach results in a significant improvement in accuracy.

For better understanding of the impact of number of neighbors on the MAE, in Fig. 1, we plot the MAE with respect to the number of similar items in KNN algorithm, for culturally enhanced approach. As it seems, when \( k = 60 \) the algorithm riches more accurate results.

A more telling picture emerges when we compare the range of values for the parameter \( \alpha \). Recall that \( \alpha \) is the parameter determining the degrees to which the cultural and rating similarities are used in the generation of neighbors. When \( \alpha = 1 \), then only cultural similarity among items is used, while \( \alpha = 0 \) represents the other side of the spectrum where only rating similarity is used (i.e., standard item-based recommendations). Fig. 2 shows the impact of \( \alpha \) on MAE for culturally enhanced approach.

With \( \alpha = 0 \), as it discussed, the algorithm is practically the item-based recommendation algorithm. Here the value of MAE when \( \alpha = 0 \), is very high comparing to other value of \( \alpha \). So we didn’t plot \( \alpha = 0 \) to carefully analyze the impact of \( \alpha \) on the accuracy. Better performance achieved when \( \alpha = 0.3 \).

<table>
<thead>
<tr>
<th>No. of neighbors</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
<th>120</th>
</tr>
</thead>
<tbody>
<tr>
<td>CULTURE-BASED CF</td>
<td>1.108</td>
<td>1.095</td>
<td>1.094</td>
<td>1.095</td>
<td>1.084</td>
<td>1.076</td>
<td>1.086</td>
<td>1.089</td>
<td>1.092</td>
<td>1.082</td>
<td>1.099</td>
</tr>
<tr>
<td>ITEM-BASED CF</td>
<td>2.355</td>
<td>2.347</td>
<td>2.349</td>
<td>2.364</td>
<td>2.349</td>
<td>2.34</td>
<td>2.365</td>
<td>2.337</td>
<td>2.369</td>
<td>2.344</td>
<td>2.358</td>
</tr>
</tbody>
</table>

where \( a_i \) is the actual rating and \( p_i \) is the predicted rating for item \( i \). Note that lower MAE values represent higher recommendation accuracy.

FIGURE 2

Fig. 2 Impact of the \( \alpha \) parameter on recommendation accuracy for culturally enhanced approach

As noted earlier, one of the problems associated with traditional collaborative filtering algorithms emanate from the sparsity of data sets to which they are applied. This sparsity has a negative impact on the accuracy and predictability of recommendations. This is one area in which, we believe, the integration of cultural knowledge with ratings data can provide significant advantage. To test this hypothesis, we created multiple training/test data sets in which the proportion of the training data to the complete ratings data set was changed from 90% to 10%. These proportions have a direct correspondence with the level of sparsity in the ratings data. In the case of each of the combination parameter values, we created ten random training and test data sets and computed average MAE’s over the ten folds. Fig. 3 shows the result with respect to the training/test data set proportion.

FIGURE 3

Fig. 3 Impact of training ratio on recommendation accuracy for culturally enhanced approach
As it is observed, the algorithm has better results with training ratio of 80% and test ratio of 20%.

V. CONCLUSIONS AND FUTURE WORK

In this paper we have extended the item-based collaborative filtering framework by integrating cultural information about items for similarity computations. We have used the Hofstede’s national culture dimensions which are presented for about 111 countries. We have calculated the correlation between these cultural dimension and real online shopping data set to see which dimensions are correlated with online shopping behavior of consumers. We have used these correlated dimensions in our algorithm. Our enhanced similarity measure combines cultural item similarities with item similarities based on the user-item mappings. Our experimental results show that the culturally enhanced approach significantly improves the prediction accuracies, while maintaining the computational advantages of item-based CF.

An interesting area of future work is to calculate these cultural dimensions for different state of a country with cooperation of sociologists and then apply this cultural recommendation system for national e-market places.

REFERENCES


