System Implementation and Adaptation Evaluation in Adaptive Web-Based Systems

Bujar Raufi

Abstract: This paper addresses the issue of implementation and evaluation of adaptation algorithm for Adaptive Web-Based Systems. The proposed architectural model of the system is consisted of two logical modes: Online and Offline mode, whilst the evaluation of the adaptation algorithm is done by using two criteria. The first one is measuring the performance of the Adaptation Algorithm through recall, precision and f-score measurements and the second is the measurement of the degree of the overall degree of navigation help offered to the user through web usability questionnaires.

Experimental results indicate a good performance of the algorithm which renders it suitable for performing online adaptation for Adaptive Web-Based Systems.

Key words: Computer Systems and Technologies, Adaptive Web-Based Systems, Adaptation Algorithm.

INTRODUCTION

Proliferation of the web in the recent years together with the increase of new applications and web phenomena’s has made it a place where information is not simply searched, browsed and read, but also can be adapted in various ways for the single goal, meeting user’s needs. The second part of the above mentioned clause is being done through the use of Adaptive Web-Based Systems. Adaptive Web-Based Systems represent systems that tend to arrange its internal link structure, content or both based on user access patterns.

In [5] we have introduced a layered framework which addresses the issue of complexity in adaptive web-based systems. The proposed framework ensures:

- **Flexibility**: that offers quicker discovery of new knowledge and the reuse of the existing.
- **Expressiveness**: allows extensive usage of semantic web technologies in the sense of good manipulation and reasoning with the knowledge at hand.
- **Interoperability**: offers sharing and accessing data from other resources for performing adaptation (In case of open corpus adaptation).
- **Modularity**: that allows a certain degree of independence between layers.

The proposed layered framework is consisted of five layers described as [5]:

- **Data Layer** where all the data and the site’s link structure resides. Also this is the layer where the atomic information is located with the precedence and next links as well as information weights.
- **Concept Layer** represents the semantic layer of the system. Consisted of concepts and concept relationship.
- **User Layer** where user preferences like access patterns and user behaviour are collected and used to perform adaptation.
- **Adaptation Layer** which performs the adaptation process based on knowledge gained from concept layer and user layer.
- **Finally the Presentation Layer** is what the user sees at the end as a final product of adaptation. In this stage, a rearrangement of atomic pieces of information or whole pages is done in order to meet the adaptation goal posed in adaptation layer.
In [6] we have introduced the work done in each layer of the above proposed framework. In this paper we are focusing on system implementation and the evaluation of the adaptation algorithm proposed in [7], [8]. In the section that follows, an approach on Adaptive Web-Based system implementation together with adaptation evaluation is introduced. The last section concludes this paper by proposing new directions concerning system implementation and adaptation evaluation.

SYSTEM IMPLEMENTATION AND ADAPTATION EVALUATION

The model architecture illustrates the main parts of the system, their interconnection and functioning, while adaptation evaluation measures the performance of adaptation algorithm proposed in [7]. The overall architecture of the system complies with the five layer framework presented in [5] in the sense that for each layer presented there, there is a specific module in the system's architecture that deals with that particular layer. The architecture of the system has been divided into two logical modes: the offline and online mode as illustrated in figure 1.

The offline mode performs computer intensive tasks such as data mining within web site’s repository (clustering and association rule mining) as well as navigation pattern mining and session reconstruction within web server logs. The reason for performing such tasks offline is the intensive calculation nature of the above mentioned functions considering that we are dealing with online web application with many requests and clicks during user visits. Such intensive calculations for every user visit would render the system inefficient. Therefore the offline mode utilizes the above mentioned techniques on the background, issued by the webmaster which performs the overhead tasks, analyzes them and the extracted knowledge is incorporated in the overall adaptation process. The incorporation involves putting additional information on the adaptive presentation panel besides the information retrieved by adaptation algorithm. Based on the above mentioned claim, the offline mode does not obstruct the overall functioning and performance of the system because it is being conducted on the background as a separated task from that of the adaptive web application.

The online mode on the other hand considers the adaptation "on-the-fly" presented through the adaptation algorithm elaborated extensively in [7]. The online mode traces the user visits through the links it clicks and presents adaptive content based on the visited information units (documents or atomic units). The characteristic of the online mode is that it offers two types of adaptive content. The first one is adaptive content delivery through the adaptation algorithm which was based on web site’s link structure and document similarity measures and the second one is adaptive content delivery by querying the
knowledge repository as presented in figure 1 which in fact is the concept layer of our framework. Querying is conducted on the generated RDF triples from Relational database, represented through RDF Repository. For this purpose SPARQL [2] query language was utilized for querying the knowledge base. The results of SPARQL queries can be presented as result sets, RDF graphs or as RDF triples presented for human consumption through web pages [1], [2].

The evaluation of the proposed adaptation approach in Adaptive Web-Based System has been conducted in terms of two criteria. The first criterion is the performance of the adaptation algorithm in the sense of the degree of retrieved documents from the algorithm itself. For this purpose, the recall-precision and f-score [7] measurements have been utilized. The second criterion is the measurements of the degree of navigational help that the proposed adaptation algorithm offers to the users. For this objective, web usability testing with direct visitor surveys has been conducted.

ADAPTATION EVALUATION

In adaptive web-based system, some effectiveness measures regarding adaptation such as recall, precision, and harmonic mean of recall/precision called f-score [4] can be calculated. The goal of such evaluation is two-fold: the first one is to show that the absolute effectiveness of the adaptation algorithm is acceptable for practical use. On the other hand, the evaluation can show that the algorithm has a better or worse effectiveness than another algorithm or a prior state of the web application where no adaptation algorithm is present.

The level of recall, precision and f-score has been tested on real life implemented web site consisting of 25 generated documents and atomic units out of which similarity measures for both Cosine and Jaccard have been produced. Similarity measures are tested for different vector spaces (term frequency – inverse document frequency-tf-idf and LSI – Latent Semantic Indexing). The reason for performing recall-precision and f-score measurements for each of the above mentioned documents is the requirement for calculation of the estimated "expected answer" [5] from the adaptation algorithm. The average recall, precision and f-score values are depicted in table 1.

The evaluation of the proposed adaptation algorithms has been conducted through the number of retrieved relevant, not retrieved relevant and retrieved irrelevant documents for both similarity measures (Jaccard and Cosine) and vector spaces (tf-idf and LSI). These values are illustrated in figure 2.

<table>
<thead>
<tr>
<th></th>
<th>recall</th>
<th>precision</th>
<th>f-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSI</td>
<td>73.77%</td>
<td>75.44%</td>
<td>47.65%</td>
</tr>
<tr>
<td>tf-idf</td>
<td>61.07%</td>
<td>86.16%</td>
<td>14.58%</td>
</tr>
<tr>
<td>Jaccard</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cosine</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>recall</th>
<th>precision</th>
<th>f-score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Cosine with LSI Vector Space](image1.png) ![Cosine with tf-idf Vector Space](image2.png)
Figure 2 Number of relevant retrieved, relevant not retrieved and irrelevant retrieved documents by the adaptation algorithm

From the mentioned values from figure 2, proper recall, precision and f-score measurements can be derived. Figure 3 depicts the recall, precision and f-score values for the named vector spaces and similarity measures, whilst table 2 illustrates the overall performance of the algorithm for both vector spaces and similarities.

Figure 3 Recall, Precision and f-score of the adaptation algorithm

| Table 2 Summary of the adaptation performance for various similarities and vector spaces |
|---------------------------------|----------------|----------------|----------------|
|                                | Cos.-LSI | Cos. –tf-idf | Jacc. - LSI    | Jacc. – tf-idf |
| #Relevant Documents Retrieved   | high     | low           | medium         | low           |
| #Relevant Documents NOT Retrieved| low      | low           | medium         | low           |
| Recall                         | high     | high          | high           | high          |
| Precision                       | high     | low           | medium         | low           |
From the table 2 it can be seen that Latent Semantic Indexing vector space in conjunction with Cosine similarity measure show outstanding results applied to the proposed Adaptation Algorithm. Its performance is seen in the sense of high amount of relevant documents retrieved, small number of relevant documents not retrieved as well as high recall and precision values.

MEASURING DEGREE OF NAVIGATION HELP

In order to measure the degree of help that the system offers to the visitors, we have adopted the post test questionnaires considering that we want to grasp user experience in the sense how they acquaint with the system to whom ultimately is dedicated for, as well as we try to minimize the biased question as much as possible. For this purpose an upgraded version of questionnaire used in [3] was utilized. The upgraded and added questions were regarding the adaptation help given in right navigation of the web site. All the questions were graded in scale starting from “Strongly disagree” as the lowest value in the scale up to “Strongly agree” as the highest value given from the user regarding the functionality of the adaptive web site.

The questionnaire was submitted for usability testing to 120 students to whom a task to find certain information in the site was given and later the usability questionnaire was presented. Various aspects of usability testing were analyzed. Figure 4 depicts the total percentage of user rating regarding the scale mentioned above for the total of twenty questions. The aggregated values below were calculated as a total number of user answers per scale mentioned above (i.e. from “strongly agree” to “strongly disagree”). In this sense, a particular care has been given to questions with opposite meaning in order to do a proper classification of answers. From the figure it can be seen that more than 67% of the surveyed users are in the range of partially agree (37%) and strongly agree (31%) in the contrary of 7% disagree and 6% strongly disagree ratings. Figure 5 on the other hand illustrates the total percentage of user answers regarding the required user clicks before it starts using the adapted right navigation bar. The results indicate that 25% of users needed more than five clicks before they start using the adapted content presented in the right navigation bar. It can be also seen that 20% of users started to use the adapted content immediately. The percentage of users that did not use the navigation bar at all is around 17% percents.

Interesting conclusion about the above mentioned question regarding user clicks is that those using the navigation bar after more than 5 clicks didn’t have the need for it because they were managing the task successfully which can be characterized as experienced users.

This claim is supported by other positively answered questions from this type of users regarding other usability aspects of the site. On the other hand, users with less experience regarding the site felt need for immediate help for finding relevant information from the adapted content presented in the right navigation bar.
CONCLUSIONS AND FUTURE WORK

In this paper we have introduced new approach and architectural model for implementing Adaptive Web-Based Systems. The model has been evaluated through the performance of the adaptation algorithm proposed in the fourth layer of the framework introduced in [7] and elaborated in introduction. The results indicate a good performance of the adaptation algorithm in the sense of providing relevant documents for the so called online adaptation.

Future work and directions for the development of Adaptive Web-Based Systems would involve:

- Exploring and developing other similarity measures besides the ones used in this dissertation. Considering that similarity measures used in this approach are ready made and did not yield satisfactory results except cosine similarity. The necessity for such research would be beneficial.
- Extending the knowledge representation through our proposed ontology with the possibility of semantic adaptation as well, besides simple triple querying with RDF and document description given here.
- Exploring emerging technologies, especially AJAX, for reviving the approaches presented in this dissertation with such technologies.

REFERENCES


ABOUT THE AUTHOR

As. Mr. Bujar Raufi, MSc, Faculty of Contemporary Sciences and Technologies, South East European University, Phone: +389 44 356 168, E-mail: b.raufi@seeu.edu.mk.