Personalization of learning situations within a virtual environment for training based on fuzzy clustering

Mohamed Fahim*, Abdeslam Jakimi, Lahcen El bermi

Software Engineering and Information Systems, Engineering Team, Computer Sciences Department, Moulay Ismail University, FST Errachidia, Morocco
fahim.mohamed89@gmail.com, ajakimi@yahoo.fr, elbermi.lahcen@gmail.com

ARTICLE INFO

Article History:
Received 01 July 2017
Accepted 23 October 2017
Available 26 October 2017

Keywords:
Virtual environments for training
Personalization of learning
Fuzzy clustering

AMS Classification 2010: 68T05

ABSTRACT

Advances in networks, computers and multimedia technology have changed traditional methods for learning and skills training. Nowadays, Virtual Environments for Training (VET) have been popular, they can provide an environment where virtual reality can be used to create interactive interfaces and real-time software that can control every response and action made by the user. VET have proven to be advantageous to put learners into varied training situations to acquire knowledge and competencies, especially when these situations are taking place in uncontrolled circumstances, or when they are dangerous, unrealizable, or expensive to establish in reality. However individual learners find it difficult to select suitable learning situations for their particular situation because often, there is no personalized service to response to the user needs. Personalization of learning in a VET is a very important way of improving the effectiveness and the quality of the training, yet it also a complex process that requires consideration of several factors such as learners profiles. The goal is to associate suitable learning situations to each learner based on his profile. However, personalization of learning becomes an issue with the uncertainty and imprecision of data that may contains a learner profile. To address this issue, this study is an attempt to integrate a fuzzy clustering into the process of the personalization of learning.

1. Introduction

Advances in networks, computers and multimedia technology have changed traditional methods for learning and skills training and have opened up new possibilities for others methods to show up. Nowadays, VET have been popular, they can provide an environment where virtual reality can be used to create interactive interfaces and real-time tools that can monitor every response made by the trainee. A VET can be defined as Computer-Supported Environments for Human Learning (CSEHL), exploiting virtual reality technologies in order to immerse the learners in a virtual environment enabling them to learn by doing [2]. Training through VET provides significant benefits over other methods. The idea of a VET relies on simulating the actual behavior of a certain process which requires accurate modeling of each part in the process and integration of real-time 3D graphics. In this case, each instrument is electronically driven by real-time computers that generate output signals based on a software model describing the instrument behavior. Learning in open environments like a VET, demands even more personalization approaches to provide learners with individualized learning situations in order to assure the quality of learning. A learning situation contains several features of information necessary to achieve the objective of training,

*Corresponding Author
such as content description, lecture information, prerequisite information and so on. In a VET, the excellence of learners can be improved by recommending suitable learning situations (personalization), based on each learner’s profile. Personalization of learning situations (PLS) relies on the fact that the learning ability of each individual can depend on several factors such as age, gender, duration of training, personal preferences, content of the material. PLS become an issue with the uncertainty and imprecision of data that may contains a learner profile. In this paper an attempt to integrate the fuzzy logic theory and a fuzzy clustering algorithm into the process of the personalization of learning situations is presented. In the following, we first present the related works to the personalization of learning situation (adaptation). Then, we will talk about the fuzzy logic and Fuzzy C-Means (fuzzy clustering algorithm); Then, we present our methodology. Finally, we conclude and reflect on the future of the present work.

2. Related works

Each individual’s learning ability depend on several factors such as: age or gender. Adapting the learning situations is one of the main interests of current e-learning research in general. The goal is to associate suitable learning situations to each learner based on his/her profile. A profile contains information such as: learner’s knowledge level, desired difficulty level. To allow this personalization of learning, many solutions have been proposed. Current methods for personalization of learning can be divided into three groups: (i) oriented activities approaches [3]: where the learning process is represented by a graph in which the activities are identified and decomposed. (ii) Oriented resources approaches ( [4], [5], [6]): in which the learning process returns to select, assemble and present contents, (iii) oriented objectives approaches ( [7], [8]) : in which the learning process is seen as a process of satisfaction of pedagogical objectives already defined. These approaches use a set of algorithms and techniques from Artificial Intelligence and Web Semantics known as ant colony optimization ( [9], [10], [11], [6]), Bayesian networks [7], the algorithm of Support Vector Machines (SVM) [12], ontologies [13]. However, these methods are quite limited in term of handling uncertain and imprecise data. Under the paradigm of fuzzy logic, approaches have been proposed such as learner’s profile modeling ( [14], [15]), evaluation issue ( [16], [17], [18]), learning styles prediction ( [19], [20]), which cover different sides of e-learning systems.

3. Fuzzy logic theory

Human brain can deal with imprecise concepts. For instance, to answer a question about a hotel services, most time answers could be "Not Very Satisfied" or "Quite Satisfied", which are also fuzzy or ambiguous answers. Exactly to what degree is one satisfied or dissatisfied with some hotel services? These vague answers can only be created and implemented by human beings, but not machines. So, how can computers and machines handle those vague data? Based on this observation, Lotfi A. Zadeh ( [21], [22]) developed fuzzy set theory that generalizes classical set theory to allow the notion of partial membership. This invention was not well recognized until Dr. E. H. Mamdani, who is a professor at London University, applied the fuzzy logic in a practical application to control an automatic steam engine in 1974 [23]. The use of fuzzy logic allows working with quantitative and qualitative descriptions. In fuzzy set theory, an element can belong entirely to a set (degree of belonging is 1), or "almost" belong to it (with a degree of belonging equal to 0.9 for example). Fuzzy logic has been successfully employed in a variety of applications in recent years( [21], [25], [26]).

4. Fuzzy C-means

The Fuzzy C-Means (FCM) algorithm is a clustering algorithm proposed by Bezdek [27]. It is useful when the required number of clusters is predetermined; thus, the algorithm tries to put each of the data points to one of the clusters. In each iteration of the FCM algorithm, the following objective function J is minimized:

$$J = \sum_{i=1}^{N} \sum_{j=1}^{C} \mu_{ij} || x_i - c_j ||^2$$  \hspace{1cm} (1)

- \( N \) is the number of data points.
- \( C \) is the number of clusters required.
- \( c_j \) is the centre vector for cluster.
- \( \mu_{ij} \) is the degree of membership for the \( i \)th data point \( x_i \) in cluster \( j \)
- The norm, \( || x_i - c_j || \) measures the similarity (or closeness) of the data point \( x_i \) to the centre vector \( c_j \) of cluster \( j \).

For a given data point \( x_i \), the degree of its membership to cluster \( j \) is calculated as follows:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{C}(\frac{||x_i-c_k||}{||x_i-c_j||})^{2/m}} \hspace{1cm} (2)$$
where, \( m \) is the fuzziness coefficient and the centre vector \( c_j \) is calculated as follows:

\[
C_j = \frac{\sum_{i=1}^{N} \mu_{ij}^m \times x_i}{\sum_{i=1}^{N} \mu_{ij}^m} \tag{3}
\]

In equation (3):

- \( \mu_{ij} \) is the value of the degree of membership calculated in the previous iteration (equation 2).
- At the start of the algorithm, the degree of membership for data point \( i \) to cluster \( j \) is initialised with a random value, \( \theta_{ij} \), \( 0 \leq \theta_{ij} \leq 1 \), such that \( \sum_j \mu_{ij} = 1 \).
- In equations (2) and (3) the fuzziness coefficient \( m \), where \( 1 < m < \infty \), measures the tolerance of the required clustering. This value determines how much the clusters can overlap with one another: The higher the value of \( m \), the larger the overlap between clusters.

Termination condition:

- The required accuracy of the degree of membership determines the number of iterations completed by the FCM algorithm.

\[
\epsilon = \Delta^N \Delta^C_j | \mu_{ij}^{k+1} - \mu_{ij}^k | \tag{4}
\]

where, \( \mu_{ij}^k \) and \( \mu_{ij}^{k+1} \) are respectively the degree of membership at iteration \( k \) and \( k + 1 \), and the operator \( \Delta \), when supplied a vector of values, returns the largest value in that vector.

5. Proposal

The architecture of the system we are developing is shown in the figure 1. The first step in the operation of our system is to collect users data to build a learner profile. Information which forms a learning profile are the following:

- Learning session duration
- Gender
- Level of difficulty
- Desired language
- Current knowledge level
- Age

The prototype of our model was tested using twenty learners. The table 1 shows an extract of those learners and information that constitutes their profiles. We did not integrate the "language" and the "gender" information in the dataset, they will be taken in consideration later, during the execution on the platform. The FCM clustering algorithm was run on the dataset and iterated to converge and terminate at \( \epsilon = 0.00001 \). The predetermined number of cluster was \( C=3 \). In this work, the larger is the number of clusters, the higher is the personalization of learning. The output of the algorithm is shown in figure 2, which shows each learner and the cluster to which he belongs. Each learner has varied degree of
The membership value of belonging to each of the three dimensions (Clusters) and the highest value was inferred as the likely strong class the learner should belong. Table 2 shows degree of membership of each learner to each cluster.

<table>
<thead>
<tr>
<th>Learners</th>
<th>1st Cluster</th>
<th>2nd Cluster</th>
<th>3rd Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>lear1</td>
<td>0.046269342</td>
<td>0.105947344</td>
<td>0.84778331</td>
</tr>
<tr>
<td>lear2</td>
<td>0.254602401</td>
<td>0.470097456</td>
<td>0.27530014</td>
</tr>
<tr>
<td>lear3</td>
<td>0.258032639</td>
<td>0.180264197</td>
<td>0.56170316</td>
</tr>
<tr>
<td>lear4</td>
<td>0.762457186</td>
<td>0.115903344</td>
<td>0.12163947</td>
</tr>
<tr>
<td>lear5</td>
<td>0.005913503</td>
<td>0.978024006</td>
<td>0.01606249</td>
</tr>
<tr>
<td>lear6</td>
<td>0.904823898</td>
<td>0.033261791</td>
<td>0.06191431</td>
</tr>
<tr>
<td>lear7</td>
<td>0.396760067</td>
<td>0.797379013</td>
<td>0.13386092</td>
</tr>
<tr>
<td>lear8</td>
<td>0.186612866</td>
<td>0.395596766</td>
<td>0.41779037</td>
</tr>
<tr>
<td>lear9</td>
<td>0.012906196</td>
<td>0.952466943</td>
<td>0.03462686</td>
</tr>
<tr>
<td>lear10</td>
<td>0.080884083</td>
<td>0.091845524</td>
<td>0.82727039</td>
</tr>
</tbody>
</table>

The efficiency of this approach is obtained considering the output of which the algorithm is able to identify individual learners into various clusters each belongs to with the varied degree of membership values.

Figure 2. Clustering result with FCM.

6. Conclusion

This paper has outlined the development of a fuzzy based approach for the generation of learning situations within a virtual environment for training. The main advantage of this proposed methodology is that it is efficient in handling the uncertainty in the learners profile. The FCM is used to make clusters of similar learners in order to generate, suitable learning situations for each leaner or group of learners. However, a number of further data, in particular, learners feedback information is required to promote the functionality of the system. The ongoing work aims at including this proposed approach in our previous work [28], in order to generate adaptable and effective pedagogical scenarios for VET.

Acknowledgments

I must acknowledge the assistance and scholarly guidance of my supervisors who also happens to be co-authors of this paper.

References


Mohamed Fahim received the Master degree in Intelligent Systems and Networks in 2013 from Faculty of Science and Technology of Fes, Morocco. Currently, he is a PhD Student in Computer Science. He is also member of Software Engineering & Information Systems Engineering Team. His research interests include pedagogical scenarization and virtual environments for training (VET).

Abdeslam Jakimi is a teacher at the faculty of science and technology of Errachidia. His research interests include software engineering and e-learning.

Lahcen El Bermi is a teacher at the faculty of science and technology of Errachidia. His research interests include software engineering and e-learning.