

An Study on the Use of Student Models as Support to Virtual Adaptive Learning

Janalívia Carneiro
Programa de Pós-Graduação em
Informática
Universidade Federal da Paraíba
João Pessoa, Brasil
janaliviacarneiro@ppgi.ci.ufpb.
br

Danielle Silva, Tarsio Gomides
Centro de Informática
Universidade Federal da Paraíba
João Pessoa, Brasil
danielle@ci.ufpb.br,
tarsio.gomides@cc.ci.ufpb.br

Clairton Siebra
Laboratório de Inteligência Artificial
Aplicada
Universidade Federal da Paraíba
João Pessoa, Brasil
clairton@ci.ufpb.br

ABSTRACT

This paper emphasizes the importance of student models as the core of Intelligent Tutoring Systems and, as result of a systematic review, an overview of the current state-of-art on the use of computational approaches, combined with pedagogical and psychological theories, which are able to support customized education in Virtual Learning Environments (VLEs). Studies in this area has been searching for ways to offer personalized instructional material and adaptive support in education, according to the needs of each student profile, knowledge level, cognitive skills, abilities and competencies in the most diverse study areas. In this context, we investigate the current models used to identify the student characteristics, how the learning content can meet the student's profile in order to provide personalization of the learning tools and virtual courses in VLEs. We also analyse the current approaches on student modelling, so that they can be used as basis for a new student model that employs the best aspects of the current models in order to enhance the adaptability of e-learning systems.

Categories and Subject Descriptors

• Human-centered computing~User models • Applied computing~Distance learning • Applied computing~E-learning • Applied computing~Learning management systems

General Terms

Design, Human Factors, Theory.

Keywords

Adaptive learning, student profiles, learner modelling, virtual learning environments.

1. INTRODUCTION

According to the United Nations, the education quality is one of the main global problems, together with increasing inequality, food/water security, refugee crisis (65 million and rising), healthcare quality, food/water security, climate change, data privacy and others [7]. Nowadays, it is common in many colleges and universities to find classes with several hundreds of students and only one teacher. This lack of resources in education makes individual teaching very challenging. Thus, there is a barrier for students to reach their full potential due to the lack of, for example, personalized assigned learning tasks according to each student proficiency level [3].

One possible way to approach this problem is by means of the use of Intelligent Tutoring Systems (ITS). Such systems are able to offer adaptive support to students, providing individual guidance similar to a real tutor in an one-to-one teaching model. Presently, many approaches have been advised to improve the quality of e-learning. Between these approaches, are the adaptive and intelligent learning which has demonstrated to be crucial in facilitating student's learning in e-learning environments [2]. One of the resources coming with technological developments was the VLEs, which are an alternative to maximize the ability of virtual education. However, the vast majority of learning environments that are available to students are passive. These VLE's are mostly used only to transmit multimedia documents (videos, audios, texts, images etc.) and to provide ways to evaluate students through assessments. According to [21], such environments do not provide ways to customize the teaching according to students' profile. An example is the Modular Object-Oriented Dynamic Learning Environment (Moodle) [19]. Thus, displaying the educational content, according to students' needs, is still a challenge.

By means of the association of Psychology, Pedagogy and Computer Science concepts, it has been possible to develop several models, methods and tools to provide personalized education and adaptive support to different student's profiles [18]. The rapid evolution of Artificial Intelligence, as a multidisciplinary subject, involving researches in psychology and education, has become a trend for virtual learning systems [2]. In this context, the importance of a Student Model is very high for a successful ITS. The use of a student model and data stored in the database gives the crucial support to achieve the goal of offering tailored teaching. Through the student model, teachers can analyze the students' learning patterns and have detailed insights into the students' understanding and their misconceptions. After the topic is taught in the course, teachers can also identify the learning gaps of students and then individually intervene and assist with the concepts that a student is struggling with [20]. Several studies show that tailored teaching is the main aspect to achieve the learning potential of a student [14].

In this context, the objective of our study is to investigate the current student models that could fundament a new student model to maximize the Moodle's potentials in offering adaptive education. The remainder of this paper is structured as follow. Section 2 highlights the theoretical background for this research. Section 3 presents the related work (systematic review). Section 4 presents the discussions regarding our findings. Finally, Section 5

- analyses of the full text through the data of the papers, which could be retrieved from the database, to check if the study answers at least two of the research questions.

These criteria were used to filter the studies retrieved in the following search engines ACM *Digital Library* and the *IEEE Xplore Digital Library* which are the main search engines in technology; and in the WebOfScience, using the next search string:

((adaptive OR customized OR personalized) AND (learning OR education) AND (course OR environment) AND (on-line OR virtual) AND (student OR learner) AND (intelligence OR cognitive) AND ("profile" OR "model" OR "technique"))

It is important to emphasize that this analysis is not based in the quality of an article. The goal here is to select articles considering their relevance for our analysis. All the articles in the systematic reviews have been published, thus the quality is theoretically guaranteed by peer-review. In other words, the evaluation of significance, about the purpose of the systematic review was intended to ensure that the articles included are really focused on adaptive learning and contained enough details to answer the predefined research questions.

The research questions were:

- 1) *What are the models proposed to model the student profile?*
- 2) *What are the computational approaches used to determine the profile of student learning in virtual learning environments?*
- 3) *What kinds of recommendations are given after determining the student's profile?*
- 4) *Which are the techniques used in adaptive systems to match the teaching content with the learner's educational profile in Virtual Learning Environments?*

The results of the systematic review are shown in table 1.

Table 1. Summary of papers selection process

Database	Number of Selected Papers			References (Phase 3)
	Phase 1	Phase 2	Phase 3	
ACM	47	9	7	[2], [3],[8],[16], [17], [20], [24]
IEEE Xplore	20	11	5	[5], [6],[15], [22],[23]
Web Of Science	29	9(12)*	0	

* The total of 12 papers were selected in this phase, however 3 had already been identified in the IEEE Xplore search. Thus, only 9 first-time papers were recorded.

Table 1 show that the 96 articles returned initially. From the papers which could be retrieved from the databases, only 12 were really important for our analysis. This high initial amount could be restricted through a specific search string. However, such an approach is unusual because it could leave out important work for our analysis.

This paper covers the answers of the systematic review research questions one and two; and the answer to research question three is also shown in the analyse of some studies of the review. The answer to question four is not covered

because it is not relevant for the purpose of this paper. In this context, the student models of the analysed papers are shown bellow.

The work presented in [8] uses a variation of the model Bayesian Knowledge Tracing (BKT) that differs from the standard BKT because it individualizes parameter estimates for learners instead of individualizing parameters for Knowledge Components (KC), which are each unit of the curriculum. On the variation of the BKT shown, called Individualized BKT, individual differences weights (IDW) per student, through their performance in previous activities in an online course, are incorporated in the model. This study analyses how well BKT parameters in a tutor lesson can be individualized based on learners' prior performance in reading instructional text, taking a pre-test, and completing an earlier tutor lesson.

The study in [2] proposes a model for electronic reviews called PIAT (Personal Intelligent Assessment using Test) with the aim of finding out the student's prior knowledge and skills. The PIAT is based on learning style model Felder-Silverman Learning Style Model (FSLSM) [11] and the RBT-Revised Bloom Theory (RBT) [9]. While the FSLSM is used as basis to describes the learning style of a student considering his preferences in three dimensions: **input** (visual/verbal), **perception** (sensing/intuitive), **understanding** (sequential/global); the RBT theory is used to determine the learner cognitive domain in a given subject.

The study [20] proposes a Model Driven Assessment (MDA). The cognitive part of the MDA consists of 3 classes of cognitive processes, to interpret and access as students use the system tools for construction/evaluation of solutions. They are: (1) information search/purchase, (2) and (3) evaluation of solutions. In this model, the variables for the user rating system are represented by two metrics, which are calculated for each class of cognitive processes for each student: (1) "use of skill", which is the numerical average of evaluation questions, quizzes and explanations made by the user per minute; and (2) "effectiveness", which means the percentage of shares that helped the student to achieve the goal of the task.

The study [22] proposes an instructional system called Techno Neuro Pedagogy System (TNPS) which considers the student learning profile and its thinking style. The determination of the student learning style (visual, auditory and kinaesthetic) is based on the Neuro Linguistics VARK programming theory [11] while the determination of the student thinking style is based on the Total Brain neuroscientific theory [14], which proposes the division of the brain into four quadrants (software, processes, relational and creative).

In the study [3], the student's behaviour is modelled using the technique of selection of neighbouring node, based on empirical estimates of students' learning curve. The skill level is an approximation of the knowledge of the student and is represented as a decimal value.

The study [25] also uses variations of the model BKT. In this study, Multi-grained BKT and Historical BKT are presented for KCs modelling of KCs in those models. Multi-grained BKT is used to model the hierarchy between KCs while Historical BKT captures the relations between multiple quiz submissions, considering multiple submissions to a quiz and each submission is considered as a slice of time.

In [15], the determination of the learner profile consists of the phases of “Discovery Behaviour” and “User Rating”. The user model was generated through the mining of data obtained from user interaction with the system. In the phase of “Discovery Behaviour”, the information obtained is pre-processed into vectors of characteristics, which summarize the user's actions in the interfaces. A machine learning algorithm groups the vectors according to their similarities, by identifying users who interact similarly with the interface. In the phase “User Rating”, the groups and the rules of association of classes, extracted from the phase 1, are used to build a classifier model of the user online. The learner level is measured using the score from before and after assessments collected. Students are classified into two groups of levels of learning: (1) Learning Low-Gain, Low Learning Gain (LLG) and (2) Learning High-Gain, High Learning Gain (HLG).

The next student models are used in adaptive solutions for the VLE Moodle, with is the focus of our research.

In the work of [17], an explicit student modelling is shown which is based on pre-established demarcations of the cognitive skills with components of the knowledge dimension within Revised Bloom’s Taxonomy [4]. In this adaptive approach the student model supports the construction of a framework to determine the order of the presentation of digital activities during the educational process in compliance with the learner’s model.

The study in [6] also uses the explicit student model to support the classification and retrieval of Learning Objects from different Learning Object Repositories according to the learner preferences. It is based on a naïve Bayesian classifier for the learning object classification under the supervised machine learning approach extracted from various Learning Object Repositories. This classifier assumes the following Learning Object Metadata elements : (1) Title/Topic; (2) Keyword; (3) Abstract/Description.

Another example of explicit student modelling is described in [5], where a Competency-Based Ontology model, derived from the Learner Information Packaging- LIP, is used to support a system in which the learner competencies are formed by a set of specific skills for each subject that the student is learning. This adaptive solution is able to identify what the student is struggling with and offer support in the form of suggesting a co-student that already has got the expected competency and is capable of helping with this matter.

In the study [15], a formative assessment model is presented, which is based on the formal model JISC. The JISC model is a model for learning assessments that integrates electronic assessments and effective learning through feedback to each student on his current skills and understanding at the end of each module. The variables that help in the detection of the profile of the student are the responses of the student’s assessments.

The study of [24] suggests the usage of a Business Intelligence (BI) based model called Learning analysts (LA). LA is a BI project in the educational domain which uses various data of educational organizations (students, the learning process, etc.) to build educational models. These models are divided in descriptive and predictive models. The descriptive model uses variables such as the amount of tasks defined, amount of participation in chats, messages received, amount of comments from the teacher to the

student, etc., and the predictive model uses variables such as ethnicity, gender, marital status, place of registration, etc.

An overview of the student modelling approaches and their learner aspects data is shown in table 2.

Table 2. Student modelling approaches and data

Student Modelling Approach	Approach’s Relevant data
Individualized BKT [8]	reading the instructional text, pre-test performance, Lesson 1 IDWs and features of student performance in completing Lesson 1
Multi-grained BKT/ Historical BKT [25]	hierarchical and temporal properties of the knowledge states
Competency based ontology [5]	Learner’s competency in a given domain
TNPS [23]	logical style, processes style, relational style, creative style
MDA [21]	skill use, effectiveness
PIAT (RBT and FSLSM) [2]	input (visual vs. verbal), understanding (sequential vs. global) and perception (sensing vs. intuitive)
Accumulative Digital Activities in RBT [17]	Cognitive skills
Naïve Bayesian classifier for the classification of LO [6]	Learner feedback
LA[24]	Personal data and educational data
Data mining [16]	Student interaction with the system

4. DISCUSSION

One characteristic of some student models is that the extraction of data that supports the model concentrates mainly on student’s feedback, as in [15]. Instead of that, it would be more interesting if the student’s learning profile could be detected dynamically through the observation of his interaction with the system, as in [23].

The adaptive approach in [17] focuses on de student cognitive skills combined with the learning objectives according the RBT. The accumulation of digital activities and the hierarchical model of learning objectives delimit the dimensions of cognitive processes, as well as the components of the knowledge dimension (conceptual, procedural, and metacognitive) of RBT. The Learning digital activities can be ordered in different cognitive levels of Bloom’s taxonomy according to their difficulty levels. Projecting a course according this taxonomy is a more complex task than one might think, because, to achieve that goal, it is necessary that the instructional objectives, the attitude objectives and cognitive skills objectives are known prior to the beginning of

the course [9]. The adaptive solution built in [17], with the goal of achieving dynamic order of digital activities during an e-course according with student's individual performances in this course. In this way, the incorporation of the educational objectives and a tailored course curriculum was successful.

The student's profile can be dynamically detected through the observation of his interaction with the system, as well as, by observing the personal characteristics of the student, as shown in the study [24]. Till a short time ago user's data retrieval by companies and universities faced plenty of limitations, such as authenticity of data, time requisite and scope. The main form of gathering data was by means of the application of questionnaires or interviews with a representative group of stakeholders. Nowadays, collecting data has become a more accessible task with the help of data mining projects. The information extracted with data mining is really accurate in terms of catching nonstop and real user behaviour. The data gathered through data mining can be compared to the retrieval of observational data instead of data retrieval through intrusive methods. It is not going to replace interviews and questionnaires, but will empower our perception of possible inconsistencies between the user understanding and user behaviour [26].

However, gathering this data is not a simple task, because, ethics and security reasons must be considered for data collection en data analyses. One of the biggest challenges for LA researches is the availability of public accessible datasets to evaluate their LA methods. This challenge is due the fact that the majority of the student data produced by educational institutions is protected. Besides, the application of LA in education originate new ethical and privacy matters which are extremely sensitive subjects related to data gathering and integration of educational databases. Privacy threats can create resistance to evolutions in LA and to avoid these problems, rules and policies need to be elaborated to avoid data abuse. For instance, data must be anonymous before usage. Another ethics question is the interpretation of the gathered data, it can generate confrontation between the values and interests of the stakeholders associated with the conclusions taken from the data analyse. For instance, if it appears from the result that foreign children have more difficulty in reading than non foreign children, it can lead to a variety of decisions than can go from creating extracurricular support to separated classes, till racism kinds [26].

Figure 3 shows an overview of the correlation between some of the student's models, the student's data considered to elaborate the model, and the resultant student profile.

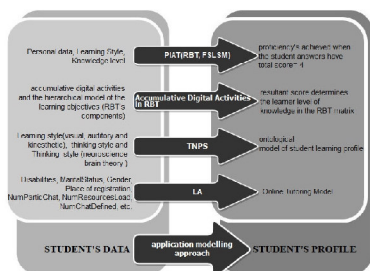


Figure 3. Relationship between student's data, student model and student profile

5. CONCLUSION

There is a variety of student modelling approaches and some of them can be found in more than one of the aforementioned studies, such as, RBT [17] and [2]; BKT [8] and [25]; and data mining [16] and [23].

Another distinction between the models is that some models concentrate the analyse of the student profile on the student learning skills, as in the study [17] while others concentrate more in the content of the learning material, as in the study [6]. The student's personal data are also considered for the student model in the papers [2] and [24].

In our systematic review, five papers [5], [6], [15], [17] [24] presented their adaptive solution for Moodle.

Besides the fact that Moodle is the LMS most used e-learning platform nowadays [7], a standard, consistent student model to support adaptively for this or other e-learning platform has not been found yet.

The Adaptive Education area seems still to be open and can offer many research opportunities. As a way to instigate the scientific community, we present two directions of research that can be exploited to advance the state-of-the-art discussed in this article.

A first direction is associated with a creation of student model that takes advantage of the Moodle possibilities to increase the level of adaptability in this platform. By combining the best aspects of the studied models, the new student model that we are going to propose is going to be a explicit student model that not only relies on the student collaboration but is also able to extract information by observing the student interaction, behaviour and pre-existing factors such as knowledge level and personal data and etc, with the necessary respect for security and ethics norms.

After examining the Moodle adaptive solutions in the works of the systematic review, the combination of two models were selected to form a robust basis for our model. The selected models were the ones proposed in [17] and [24].

The multidimensional database of the LA approach will be build based on the dimensions of the RBT matrix.

In this way, the chosen models are going to complete each other. The dynamism of the learner's data retrieval of the study [24] will support the deep coverage of the cognitive and knowledge aspects of the student of the work [17] in determining the student profile. Finally, our proposal is the development, verification and implementation of the combination of the student models described in [17] with the one described in [24].

6. REFERENCIAS

- [1] Alevin, V. (2010) Rule-Based Cognitive Modeling for Intelligent Tutoring Systems, Advances in Intelligent Tutoring Systems.
- [2] Al-Rajhi, L., Salama, R., Stamper, J., Gamalel-Din S. (2014) Personalized Intelligent Assessment Model for Measuring Initial Students Abilities. ACM, pp. 41, June 2014.
- [3] Andersen, P., Kråkevik, C., Goodwin, M. and Yazidi A. (2016) Adaptive Task Assignment in Online Learning Environments, ACM, Article No. 5, June, 2016.
- [4] Anderson, L., & Krathwohl, D. R. (Eds.). (2001). taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives: Complete edition, New York Longman (2001).

- [5] Bremgartner, V., de Magalhães Netto, J.F. (2012) Improving Collaborative Learning by Personalization in Virtual Learning Environments Using Agents and Competency-Based Ontology, IEEE, pp. 1-6, 2012.
- [6] Chellatamilan, T., Suresh, R.M. (2012) Automatic Classification of Learning Objects through Dimensionality Reduction and Feature Subset Selections in an e-Learning, IEEE Conference Publications, pp. 1-6, 2012.
- [7] Dougiamas M. (2017). Moodle for the next ten years, presented at MoodleMoot Conference, Germany, 2017. <https://www.youtube.com/watch?v=rINpq6YmFps>. Accessed: 01/09/2017.
- [8] Eagle, M., Corbert, M., Stamper, J., McLaren B.M., Baker, R., Wagner A., MacLaten B., and Mitchell, A. (2016) Predicting Individual Differences for Learner Modelling in Intelligent Tutors from Previous Learner Activities, ACM, pp. 55-63, July 2016.
- [9] Ferraz, A.; Belhot, R. (2010) Taxonomia de Bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais, Gest. Prod., São Carlos, n., p.421-431, 2010.
- [10] Fleming, N. (2001) Teaching and learning styles: VARK strategies”, New Zealand: N.D. Fleming, 2001.
- [11] Graf, S., Viola, S., Kinshuk, and Leo, T. (2006) Representative Characteristics of Felder-Silverman Learning Styles: An Empirical Model. In Proceedings of the IADIS International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2006), pp. 235-242.
- [12] Greer, J. E. and McCalla, G. I., 1994. Student Modelling: The Key to Individualized Knowledge-Based Instruction. Berlin, Springer-Verlag.
- [13] Heift T., Schulze M. (2003) Student Modeling and ab initio Language Learning, Elsevier, Volume 31, Issue 4, December 2003, Pages 519-535.
- [14] Herrmann Ned, S.L. M. (1989) The creative brain. Búfalo: Brain books.
- [15] Hettiarachchi, E., Hertas, M.A. and Mor, E. (2014) E-Assessment High Level Cognitive courses: Improving student Engagement and Results, IEEE, pp. 43-48, 2014.
- [16] Kardan, S., Conati, C. (2015) Providing Adaptive Support in an Interactive Simulation For Learning: an Experimental Evaluation, ACM, pp. 3671-3680, April 2015.
- [17] Kostadinova, H., Totkov, G., Indzhov, H. (2012) Adaptive E-learning System Based on Accumulative Digital Activities in Revised Bloom’s Taxonomy, ACM, pp. 368-375, June, 2012
- [18] Lane, C. Grover S., Roschelle J. AI Applications in Education, <http://circlcenter.org/ai-applications-education/>. Accessed: 05 October 20107.
- [19] Modular Object-Oriented Dynamic Learning Environment, <http://moodle.org/>. Accessed: 01 April 2017.
- [20] Rajendran R. (2014) Enriching the Student Model in an Intelligent Tutoring System (Doctoral dissertation), Retrieved from: www.it.iitb.ac.in/~sri/students/ramkumar-thesis.pdf.
- [21] Segedy, J.R., Loretz, K.M., Biswas, G. (2013) Model-Driven Assessment of Learners in Open-Ended Learning Environments, ACM, pp. 200-204, April 2013.
- [22] Siebra, C. d. A. (2016) Agentes Inteligentes em Educação. Editora Universitária, 2016.
- [23] Silva-López, R.B., Reyes-Ortiz, J.A., Bravo, M., Fallad J., Rodriguez, J. (2015) Techno Neuro Pedagogy System: an instructional design methodology for customized online courses (Virtual Customized Environment Learning), IEEE Conference Publications, pp.281-286, December 2015.
- [24] Valdiviezo-Díaz, P., Cordero, J., Reátegui, R. and Aguilar J. (2015) A Business Intelligence Model for Online Tutoring Process, IEEE, pp. 1-9, 2015.
- [25] Wang, Z., Hu, Z., Zhu, J., Zhang M., Xiang, Li. (2016) Structured Knowledge Tracing Models for Student Assessment on Coursera, ACM, pp. 209-212, April 2016.
- [26] Wolfgang G., Hendrik D. (2012) Translating Learning into Numbers: Toward a Generic Framework for Learning analytics, Educational Technology and Society, vol. 15, pp. 42–57, 2012.