AHP Supported Evaluation of LMS Quality

Bojan Srđević
University of Novi Sad
Trg D. Obradovica 8,
21000 Novi Sad,
Serbia
bojans@polj.uns.ac.rs

Matija Pipan
Jozef Stefan Institute
Jamova cesta 39,
1000 Ljubljana,
Slovenia
matic@e5.ijs.si

Zorica Srđević
University of Novi Sad
Trg D. Obradovica 8,
21000 Novi Sad,
Serbia
srdjevicz@polj.uns.ac.rs

Tanja Arh
Jozef Stefan Institute
Jamova cesta 39,
1000 Ljubljana,
Slovenia
tanja@e5.ijs.si

ABSTRACT

Learning Management System (LMS) provides a platform for an on-line learning environment by enabling the management, delivery, and tracking of the learning process and learners. Selection of the most suitable method is usually prolonged by the time and effort consuming evaluations of numerous features of LMS. To reduce the number of features and at the same obtain a reliable result from an evaluation, we propose a decomposition of this complex problem to more easily comprehended sub-problems that can be analyzed independently through a multi-criteria method called Analytic Hierarchy Process (AHP). To verify the approach, an expert is asked to use AHP on an originally developed reduced hierarchy of the problem of selecting the most appropriate LMS for the student target group. Results of the application are compared with the results obtained by the DEXi multi-criteria model.

Keywords: LMS, Evaluation, Analytic Hierarchy Process

INTRODUCTION

The Organization for Economic Cooperation and Development defined LMS technology as a technology used by instructors to build and maintain courses. It features personal communication via email, group communication via chatting and forums, posting content including syllabi, papers, presentations and lesson summaries, performance evaluation via question and answer repositories; self-assessment tests, assignments, quizzes and exams, instruction management via messaging, grade posting and surveys, and more.

There are many LMS systems on the market that can be obtained for free and are Open Source (i.e. Moodle, Sakai, Claroline, ATutor, etc.) or through payment (i.e. Blackboard, WebCT, Clix, and many others). All of them support many different features which can be used as evaluation criteria and analyzed from different aspects [6]:

1. Pedagogical aspect
2. Learner environment
3. Instructor tools
4. Course and curriculum design
5. Administrator tools and
6. Technical specification.

Pedagogical criteria can, for example, include [15]: Learner control, Learner activity, Cooperative/ Collaborative learning, Goal orientation, Applicability, Added value, Motivation, Valuation of previous knowledge, Flexibility and Feedback.

On the other hand, Kurilovas [12] groups technical criteria as follows:

1. Overall architecture and implementation: Scalability of the system; System modularity and extensibility; Possibility of multiple installations on a single platform; Reasonable performance optimizations; Look and feel is configurable; Security; Modular authentication; Robustness and stability; Installation, dependencies and portability;
2. Interoperability: Integration is straightforward; LMS standards support (IMS Content Packaging, SCORM);
3. Cost of ownership;
4. Strength of the development community (for open source products): Installed base and longevity; Documentation; End-user community; Developer community; Open development process; Commercial support community;
5. Licensing;
6. Internationalization and localization: Localizable user interface; Localization to relevant languages; Unicode text editing and storage; Time zones and date localization; Alternative language support;
7. Accessibility: Text-only navigation support; Scalable fonts and graphics; and
8. Document transformation.

It is obvious that selection of the most suitable LMS is a complex task that involves defining the evaluation criteria and selecting a method for criteria evaluation that will be systematic, comprehensive, easy to use, etc.

Once defined, the criteria can be evaluated using a self-evaluation questionnaire that employs a 7-point Likert scale 1 (strongly disagree) – 5 (strongly agree), 6 (not applicable), 7 (don’t know) [7, 13, 14, 15]. Other evaluation tools include MS-Excel spreadsheets application [1], fuzzy logic [6], an expert system shell for multi-attribute decision support DEXi [2], a hybrid Multi-criteria decision-making (MCDM) model based on factor analysis and DEMATEL [21] etc. The number of features for evaluation is usually very high in all these applications (e.g. 57 in Pipan [16]; 52

offered in Cavus [6]). To evaluate such a great number of features, a significant amount of time and effort is required of the evaluator.

We believe that reliable results can be obtained with fewer criteria if the problem is decomposed in order to more easily comprehended sub-problems that can be analyzed independently, i.e. presented as a hierarchy. One of the most popular methods that deal with decision hierarchies is Analytic Hierarchy Process (AHP) [17] and we propose this method for the evaluation of selected LMS products because: (1) it supplies management in both education and industry with a less complex and more appropriate and flexible way to effectively analyze LMSs, (2) it supports their selections of an appropriate product, and (3) achievement of a higher level of e-learner satisfaction [18]. Other advantages of AHP that should be emphasized are that AHP provides a measure of consistency of the evaluator and that it can be used for participative evaluation of LMS product.

To verify AHP applicability, an expert is asked to use AHP on an originally developed hierarchy of the problem of selecting the most appropriate LMS for the student target group. Consistency of the expert is checked throughout the process. At the end, results of the evaluation are compared with results presented in [16].

AHP IN BRIEF

Main features

One of the key issues in decision making is eliciting judgments from the decision maker (DM) about the importance of a given set of decision elements. If a problem can be structured hierarchically, then a certain ratio scale can serve as an efficient tool to enable this hierarchy by performing pair-wise comparisons. The core of AHP [17] lies in presenting the problem as a hierarchy and comparing the hierarchical elements in a pair-wise manner using Saaty's 9-point scale, Table 1.

This way, the importance of one element over another is expressed in regards to the element in the higher level. The AHP is a multi criteria optimization method which creates so-called local comparison matrices at all levels of a hierarchy and performs logical syntheses of their (local) priority vectors. The major feature of AHP is that it involves a variety of tangible and intangible goals, attributes, and other decision elements. In addition, it reduces complex decisions to a series of pair-wise comparisons; implements a structured, repeatable, and justifiable decision-making approach; and builds consensus.

Judgment term	**Numerical term**
Absolute preference (element *i* over element *j*)	9
Very strong preference (*i* over *j*)	7
Strong preference (*i* over *j*)	5
Weak preference (*i* over *j*)	3
Indifference of *i* and *j*	1
Weak preference (*j* over *i*)	1/3
Strong preference (*j* over *i*)	1/5
Very strong preference (*j* over *i*)	1/7
Absolute preference (*j* over *i*)	1/9
An intermediate numerical values 2,4,6,8 and 1/2,1/4,1/6,1/8 can be used as well	

Table 1: The fundamental Saaty's scale for the comparative judgments

In standard AHP, an eigenvector (EV) method is used for deriving weights from local matrices; the EV is called the prioritization method, and the computational procedure is consequently called prioritization. After local weights are calculated at all levels of the hierarchy, a synthesis consists of multiplying the criterion-specific weight of the alternative with the corresponding criterion weight and summing up the results to obtain composite weights of the alternative with respect to the goal; this procedure is unique for all alternatives and all criteria.

AHP is aimed at supporting decision-making processes in both individual and group contexts. In later cases various aggregation schemes are applicable, e.g. AIJ and AIP [9], as well as various consensus reaching procedures are easy to implement. This issue is out of scope here; namely, the paper deals strictly with an individual application of AHP.

Measuring consistency

The DM makes judgments more or less consistently depending not only on his knowledge of the decision problem itself, but also on his ability to remain focused and to ensure that his understanding of the cardinal preferences between elements will always, or as much as possible, be formalized properly while using a verbal scale or related numerical ratios [20]. For example, if the Saaty's 9-point ratio scale is used, the question could be: will the DM put $a_{ij} = 3$, or $a_{ij} = 2$, if he considers element E_i slightly more important than E_j? Or, if there are seven elements to be compared, then matrix A is of size 7x7, and the question could be: is the DM really capable to preserve consistency while comparing head-to-head 21 times all pairs of elements? How is the DM to override the imposed difficulty with Saaty's scale when he compares elements E_i and E_k, after he has judged the elements E_i and E_j, and E_j and E_k? If he has already made the judgments $a_{ij} = 3$ and a_{jk}

= 4, he should logically put a_{ik} = 12 without any further judging because a simple transitivity rule applies: $a_{ik} = a_{ij}a_{jk}$ = 3x4 = 12. Because the maximum value in Saaty's scale is 9 for declaring the absolute dominance of one element over the other, there is a problem in attaining consistency while judging certain elements. The inconsistencies generally accumulate until the need for their measuring arises.

Consistency analysis of the individual DM can be based on the consistency ratio (CR) defined by Saaty [17], and the total L^2 ED for each comparison matrix. Whichever method is used to derive the priority vector from the given local AHP matrix [19], if it already has all the entries elicited from the DM, measuring consistency is necessary in order to ensure the integrity of the outcomes.

Standard AHP uses EV, the prioritization method, and the consistency coefficient *CR* to indicate the inconsistency of the DM [17]. The other commonly used consistency measures are the total Euclidean distance, and minimum violations measure.

The *CR* is calculated as a part of the standard AHP procedure. First, the consistency index (*CI*) is calculated using the following equation:

$$CI = \frac{\lambda_{\max} - n}{n - 1} \qquad (1)$$

where $\lambda_{\max}$ is the principal eigenvalue of the given comparison matrix. Knowing the consistency index and random consistency index (*RI*) defined also by Saaty [17], the consistency ratio is obtained:

$$CR = \frac{CI}{RI}. \qquad (2)$$

Saaty [17] suggested considering the maximum level of the DM's inconsistency to be 0.10; that is, *CR* should be less or equal to 0.10.

EXAMPLE APPLICATION

Problem statement

The problem is stated so as to assess and rank by applicability the three e-Learning Management Systems based on three typical qualitative criteria and a number of qualitative sub criteria. An expert is asked to perform the decision making processes by applying the AHP model.

Hierarchy of the problem

An original hierarchy of the problem [16] consists of five levels: goal – criteria set – sub criteria set (4+4+3 per criterions in upper level) represented by specific groups of attributes – sub sub criterions (24 in total under sub criterions), represented by groups of more detailed attributes – and three alternatives (LSMs). In order to reduce the number of decision elements, the fourth level in the hierarchy (sub sub attributes) is avoided and thus the reduced hierarchy of the problem is created as shown in Figure 1.

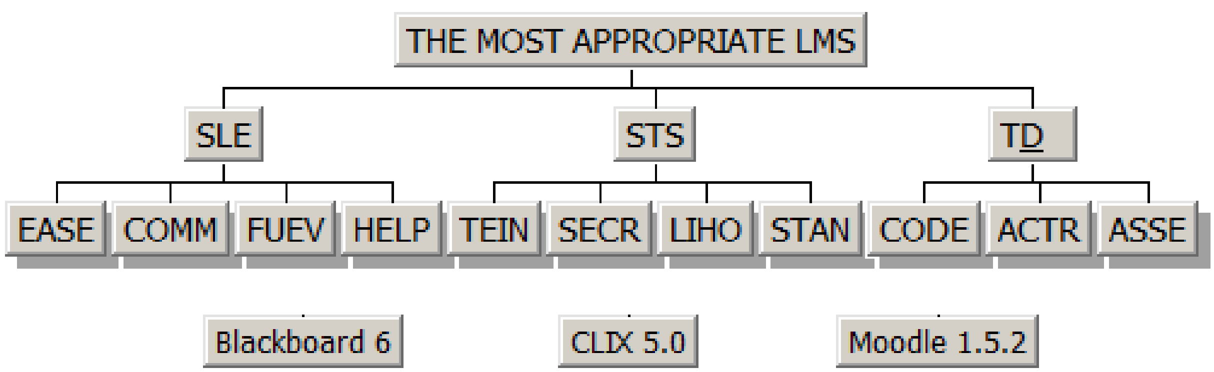

Figure 1: Reduced hierarchy of the decision problem

Identify LSM with the best applicability characteristics

Criteria set (with attributes as sub-criteria)

The set of criteria is the key component of the decision-making model. In creating the model [16], an attempt is made to meet the requirements set by Bohanec & Rajkovič [5] by taking into account the principle of criteria integrity (inclusion of all relevant criteria), appropriate structure, non-redundancy, comprehensiveness and measurability [4]. Comprehensiveness means that all the data about the subject are actually present in the database. Non-redundancy means that each individual piece of data exists only once in the database. Appropriate structure means that the data are stored in such a way as to minimize the cost of expected processing and storage [3].

The criteria set is stated under three main scopes: Student's learning environment, System, technology & standards, and Tutoring & didactics. These three scopes represent the global skeleton of the multi-attribute model with attributes (considered as sub-criterions) associated with each criterion.

(1) SLE (Student's learning environment): The first scope is adopted as the first criterion and declared as the Student's learning environment. It is composed of four basic attributes:

- (**EASE**) *Ease of use*
- (**COMM**) *Communication*
- **(FUEV)** *Functional environment* and
- *(***HELP***) Help*.

(2) STS (System, technology & standards category): The second group of attributes is grouped into the *System, technology & standards* category. These groups of criteria are assessed through four basic attributes:

- (**TEIN**) *Technological independence*. The attribute of technological independence is used for the evaluation of an LMS from the prospective of its technological accessibility, which is a pre-condition that has to be met if we wish to talk about system applicability and efficiency.
- (**SECR**) *Security and privacy. The Security and privacy* criterion focuses on two issues: *User security and privacy* and *security and privacy of an LMS*. User security and privacy should be at the forefront of attention; therefore an LMS must keep communication and personal data safe and avoid dangers and attacks on user computers. Application security and privacy assessment is made using authentication, authorization, logging, monitoring and validation of input.
- (**LIHO**) *Licensing & hosting*. Add description.
- (**STAN**) *Standards support.* It is also important to consider *e-learning standards* – standards for description of learners' profiles and standards for the description of learning resources [11]. In the context of e-learning technology, standards are generally developed to be used in system design and implementation for the purposes of ensuring interoperability, portability and reusability, especially for learning resources as they require for their preparation qualified professionals and are very time [10].

(3) T&D (Tutoring & didactics): Third group of criteria is merged into *Tutoring & didactics*. The tutor's quality of environment is assessed using the:

- (**CODE**) *Course development*,
- (**ACTR**) *Activity tracking* and
- (**ASSE**) *Assessment criteria*.

Activity tracking undoubtedly provides important support to the tutor in the learning process. Here we have focused on monitoring students in the process of learning and the possibility of displaying students' progress, analysis of presence data, sign-in data and time analysis.

Decision alternatives

The multi-attribute decision making model was completed with three learning management systems (LMS):

A1. Blackboard 6 (www.blackboard.com): Blackboard is among the most perfected and complex LMSs on the market. The system offers various communication options (both synchronous and asynchronous) within the learning environment. The Blackboard LMS is designed for institutions dedicated to teaching and learning. Blackboard technology and resources power the online, web-enhanced, and hybrid education programs at more than 2000 academic institutions (research university, community college, high school, virtual MBA programs etc.). Blackboard has 5,500 clients representing 200 million users (2.5 million from its largest, hosted client; 100,000 from its largest, self-hosted client) in 60 countries [8].

A2. CLIX 5.0 (www.im-c.de): CLIX is targeted most of all at big corporations because it provides efficient, manageable, connected and expandable internet-based learning solutions. This scalable, multilingual and customizable software aims at providing process excellence for educational institutions. For educational administrators, CLIX offers powerful features for course management and distribution. Additionally, it provides personalized learning paths for students, a tutoring centre for lectures and a whole bunch of innovative collaboration tools for both user groups, e.g. a virtual classroom. Altogether, CLIX makes planning, organizing, distributing, tracking and analyzing of learning and teaching a smooth and efficient process.

A3. Moodle 1.5.2 (www.moodle.org). Moodle is a free, open source PHP application for producing internet-based educational courses and web sites on any major platform (Linux, UNIX, Windows and Mac OS X). The fact that it is free of charge is especially attractive for schools and companies which always lack resources for the introduction of new learning technologies. Furthermore, the Moodle system is not only price-efficient – it can easily be compared to costly commercial solutions on all aspects. Courses are easily built up using modules such as forums, chats, journals, quizzes, surveys, assignments, workshops, resources, choices and more. Moodle supports localization, and has so far been translated into 34 languages. Moodle has been designed to support modern pedagogies based on social constructionism, and focuses on providing an environment to support collaboration, connected knowing and a meaningful exchange of ideas. It has nearly 54,000 registered sites (over 9,800 from the U.S.) representing over 200 countries, 44.3 million users, and 4.6 million courses. Moodle's wide spread international use, coupled with its continued growth over the past six years, has made it the leading open source LMS solution.

Evaluation of decision elements

After a brief explanation of basics and concepts of AHP, the expert compared in pairs first criteria versus goal, then sub criteria versus criteria, and finally alternatives with respect to each of the sub criteria. Comparison matrices and related calculated local weights of decision elements are presented in Figures 2-3.

Criteria vs. Goal: CR=0.021

	SLE	STS	T&D	Weights
SLE	1	5	4	0.683
STS		1	1/2	0.117
T&D			1	0.200

Figure 2: Criteria versus goal and their local weights

Sub criteria vs. Criterion SLE *(Student's learning environment):* CR=0.087

	EASE	COMM	FUEV	HELP	Weights
EASE	1	6	4	8	0.631
COMM		1	2	5	0.188
FUEV			1	4	0.135
HELP				1	0.046

Sub criteria vs. Criterion STS *(System, technology & standards category)* CR=0.021

	TEIN	SECR	LIHO	STAN	Weights
TEIN	1	4	7	2	0.501
SECR		1	3	1/3	0.129
LIHO			1	1/6	0.055
STAN				1	0.316

Sub criteria vs. Criterion T&D *(Tutoring & didactics):* CR=0.008

	CODE	ACTR	ASSE	Weights
CODE	1	1/2	2	0.297
ACTR		1	3	0.540
ASSE			1	0.163

Figure 3: Sub criteria versus criteria and their local weights

After the local weights (W) of all decision elements are calculated, a synthesis is performed to obtain composite weights of the alternatives with respect to goal (Table 2).

	Weights
Blackboard 6	0.257
CLIX 5.0	0.590
Moodle 1.5.2	0.152

HCR=0.059

Table 2: Final (composite) weights of alternatives

The alternative with the highest final weight is CLIX 5.0 (0.590) and can be considered as the most applicable LMS for the students. The second ranked alternative is Blackboard, while Moodle 1.5.2 is the least applicable LMS.

It is worthy to mention that the expert was very consistent during the whole evaluation process. Overall HCR is 0.059.

DISCUSSION AND CONCLUSIONS

One of the important problems in the field of e-learning is the selection of an appropriate LMS that will satisfy most of the users' preferences and requirements. The complexity of the problem is increased due to the growing number of LMS each year and also due to the number of features that should be taken into account while evaluating each LMS.

To reduce that complexity and facilitate selection of an appropriate LMS, we propose a decomposition of the problem to more easily comprehended sub-problems that the evaluator can analyze independently. The AHP methodology based on pair-wise comparison of decision elements on one hierarchy level was found to be appropriate for such analysis. Also, the final result of AHP application, which found CLIX 5.0 to be the most applicable LMS, proved that the proposed approach was justified: the reduced hierarchy and use of AHP led to the same result as the one provided by the DeXi evaluation of 57 criteria.

If AHP and DeXi are further compared, it should be also emphasized that:

a) AHP treats consistency of the DM (DMs), DEXi does not.
b) DEXi uses a simplified 3-point scale (linguistic semantic statements such as low, average and high); AHP most commonly uses Saaty's 9-points (fundamental) scale; other scales also in use are geometric (Lootsma's), balanced, Ma-Feng scale etc. In practical implementations the first seems easier, especially if many decision elements have to be considered (assessed). If one has to compare 7 or more elements at a time by using any AHP scale, it can be time consuming and inconsistent (e.g. due to 'short term memory' and/or 'brain channel capacity' limits).
c) AHP produces cardinal information represented by weights at all hierarchical levels of the decision problem; DEXi does it very approximately and with limited theoretical justification.
d) Both AHP and DEXi run easily on any standard PC platform.

Both AHP and DEXi can be used in individual and group d-m frameworks. In group contexts AHP enables the direct application of various aggregation schemes (e.g. AIJ, AIP; different weights allocated to DMs; different consensus reaching procedures) while in the use of DEXi, there are no implemented aggregation schemes.

REFERENCES

1. 3Waynet Inc. *LMS evaluation tool user guide*. Under License to Commonwealth.
2. Arh, T., Blazic, J.B. Application of multi-attribute decision making approach to learning management systems evaluation. *J Comput 2*, 10 (2007), 28–37.
3. Awad E. M., Gotterer M. H.. Database management. Danvers, MA: Boyd & Fraser, 1992.
4. Baker, D., Bridges, D., Hunter, R., Johnson, G., Krupa, J., Murphy, J., Sorenson, K. *Guidebook to Decision-Making Methods*. USA, WSRC-IM-2002-00002, 2002, Department of Energy.
5. Bohanec, M., Rajkovič, V. Multi-Attribute Decision Modeling: Industrial Applications of DEX. *Informatica 23* (1999), 487−491.
6. Cavus, N. The evaluation of Learning Management Systems using an artificial intelligence fuzzy logic algorithm. *Advances in Engineering Software 41* (2010), 248–254.
7. Chen, Y. , Hwang, R., Wang, Ch. Development and evaluation of a Web 2.0 annotation system as a learning tool in an e-learning environment . *Computers & Education 58,* 4 (2012), 1094-1105.
8. Cobb, J., Steele, C. Association learning management systems 2011: *Special Blackboard edition*. Tagoras (2011), 30-31.
9. Forman, E., Peniwati, K. Aggregating individual judgments and priorities with the analytic hierarchy process. *Eur J Oper Res 108* (1998), 165–169.
10. IEEE Computer Society. Learning Technology Standards Committee LTSC, IEEE, Draft Standard for Learning Objects Metadata (LOM) (Tech. Rep. 1484.12/D4.0). Washington, DC, USA, IEEE Computer Society, 2002.
11. Jerman Blažič, B., Klobučar, T. Privacy provision in e-learning standardized systems: status and improvements. *Computer standards & interfaces 27* (2005), 561−578.
12. Kurilovas, E. Methods of Multiple Criteria Evaluation of the Quality of Learning Management Systems for Personalised Learners Needs. *Presented at the EC-TEL'2009 Workshop "Learning Management Systems meet Adaptive Learning Environments"*, September 29 (2009), Nice, France.
13. Lee, J., Lee, W. The relationship of e-Learner's self-regulatory efficacy and perception of e-Learning environmental quality. *Computers in Human Behavior 24,* 1 (2008), 32-47.
14. Naveh, G., Tubin, D., Pliskin, N. Student LMS use and satisfaction in academic institutions: The organizational perspective. *The Internet and Higher Education 13*, 3 (2010), 127-133.
15. Nokelainen, P. An empirical assessment of pedagogical usability criteria for digital learning material with elementary school students. *Educational Technology & Society 9,* 2 (2006), 178-197.
16. Pipan, M., Arh, T., Jerman Blažič, B. (2010). The Evaluation Cycle Management - Method Applied to the Evaluation of Learning Management Systems. *Chapter In Integrating Usability Engineering for Designing the Web Experience: Methodologies and Principles* (Eds. Tasos Spiliotopoulos; Panagiota Papadopoulou; Drakoulis Martakos; Georgios Kouroupetroglou), IGI Global, Hershey PA, USA, 2010.
17. Saaty, T.L. The analytic hierarchy process. McGraw-Hill Inc, 1980.
18. Shee, D. Y., Wang, Y. Multi-criteria evaluation of the web-based e-learning system: A methodology based on learner satisfaction and its applications. *Computers & Education 50,* 3 (2008), 894-905.
19. Srdjevic, B. Combining different prioritization methods in analytic hierarchy process synthesis. *Computers & Operations Research 32* (2005), 1897–1919.
20. Srdjevic, B., Srdjevic, Z. Bi-criteria evolution strategy in estimating weights from the AHP ratio-scale matrices. *Applied Mathematics and Computation 218* (2011), 1254-1266.
21. Tzeng, S., Chiang, C., Li, C. Evaluating intertwined effects in e-learning programs: A novel hybrid MCDM model based on factor analysis and DEMATEL. *Expert Systems with Applications 32*, 4 (2007), 1028-1044.