Monitoring and Evaluation Problems in Higer Education
Comprehensive Assessment Framework Development

Olga Cherednichenko[1], Olha Yanholenko[1], Iryna Liutenko[1] and Olena Iakovleva[2]
[1]*Department of Computer-Assisted Management Systems, National Technical University "Kharkiv Polytechnic Institute", 21 Frunze str., 61002, Kharkiv, Ukraine*
[2]*Informatics Department, Kharkiv National University of Radioelectronics, 14 Lenina str., 61166, Kharkiv, Ukraine*

Keywords: Monitoring and Evaluation, Information System, Comprehensive Assessment, Education Quality, Partial Credit Model.

Abstract: The work is devoted to evaluation component introduction into higher education management systems. Three classes of problems of comprehensive assessment are considered. The appropriate assessment models are suggested. The case study is related to comprehensive assessment of education quality based on the level of students' satisfaction.

1 INTRODUCTION

Information technologies (IT) are the powerful tool of increasing the efficiency of decision-making processes. The formalization of management problems and usage of appropriate mathematical models provide IT with tools for solving application problems. This allows to increase business performance in different domains. Higher education is a unique social and economical area. The quality of its functioning influences many processes of development of society. Therefore the elaboration of IT in higher education domain remains the important problem for specialists of different sciences.

The existing information systems (IS) of different higher education establishments (HEEs) can be classified by functionality, relation to educational process, producer and implementation technology (Krukov and Shahgeldyan, 2007).

IS functionality corresponds to definite kind of HEE's activities. IS of HEEs may be related to educational process or may automate some financial and administrative functions which are similar for different organizations and enterprises. IS can be elaborated by HEE itself to satisfy its needs. The commercial software is an alternative, it is created by IT-companies and is distributed on the software market. IS for HEE management can be realized based on a single or several technologies. Analyzing existing software for HEE management we can make a conclusion that the process of decision-making is still not enough automated.

Independently of the domain, the process of decision-making has the following stages: goal formulation, forming the set of possible alternatives, evaluation, and selection of the best alternative (Meyer and Booker, 2001). Monitoring and evaluation (M&E) subsystem provides measurement tools for estimation of different activities, projects and outcomes.

Automation of M&E is an urgent problem that has found many industrial solutions in different areas of public life. For example, environment monitoring IS provide data about ecological situation of some region, country or the earth that reflects the state of air, water, lands, threatened species, etc. (Athanasiadis and Mitkas, 2004). Education monitoring IS collect and process information on the level of HEE or some management agencies (Carrizo, et.al., 2003). Healthcare also needs IS of M&E (Health Monitoring, 2012).

M&E includes many subproblems (for example, indicators construction, data collection, comprehensive assessment). In this work we consider different classes of problems of comprehensive assessment (CA). Our aim is to improve decision-making process by means of useful CA components elaboration. CA software must be developed taking into account the following requirements of evaluation models: evidentiary and

Cherednichenko O., Yanholenko O., Liutenko I. and Iakovleva O..
Monitoring and Evaluation Problems in Higer Education - Comprehensive Assessment Framework Development.
DOI: 10.5220/0004412504550460
In *Proceedings of the 5th International Conference on Computer Supported Education* (CSEDU-2013), pages 455-460
ISBN: 978-989-8565-53-2

unification character, quantitative estimates, transparency and reliability.

The rest of the paper is organized in the following way. Section 2 describes the directions of researches devoted to education quality assessment. Section 3 represents three classes of problems of CA. The case study of the students' satisfaction evaluation is given in section 4. The conclusions and future work are presented in section 5.

2 ISSUES OF EDUCATION QUALITY ASSESSMENT

The necessity of quality assessment in higher education is not in doubt. The solution of this problem depends on two basic aspects: the understanding of the concept of education quality and methods of its evaluation.

The concept of education quality is interpreted in different ways. The most common way is to consider education quality as collection of knowledge and skills obtained during the educational process (Koenig, 2011). In addition from functional point of view education quality can be considered as service characteristic, process attribute or HEE resources feature (EFQM, 2003). Spacial aspect enforces analysis of education quality on different management levels: university, region, country (Kachalov, 2001). Time aspect leads to considering of education quality as feature suggested by HEE or expected and perceived by consumers (Oliveira and Ferreira, 2009).

Variety of ways of education quality concept definition leads to elaboration of different methods of its assessment. There are many works in this direction, and the results may be divided into two subcategories: methods of experts' judgments and methods of psychometric theory. Experts methods are developing independently of application domain in decision-making theory (Brown 2005). The main disadvantage of these methods is experts' subjectivity. On the other hand, test theories apply statistical analysis for substantiation of knowledge testing results (Wright and Stone, 1999).

So education quality is a complex, multi-aspect, heterogeneous object. Its assessment must take into account the multidimensionality and heterogeneity of the object itself, dispersion of possible values and different measurement scales. Since the quality category covers different aspects, this work considers the peculiarities of the comprehensive quality assessment. Many problems in this domain remain unsolved. In most cases the comprehensive estimate is found as arithmetical mean not taking into account the heterogeneous structure of the complex object. All these issues make the CA problem interesting for our research.

3 CLASSIFICATION OF COMPREHENSIVE ASSESSMENT PROBLEMS

To assess quality of any object it is necessary to define the set of indicators, which reflect the state of an object, and the model which determine its quantitative measure.

We discovered that the problem of construction of set of indicators has some solutions. They include the approaches based on Qualimetry Theory (Azgaldov, 1982), which substantiate the rules of construction of indicators system, and Rasch theory (Wright and Stone, 1999), which considers the probabilistic models of estimation of latent variables for the substantiated set of observed indicators.

The model of quality assessment is determined by management goals. HEE management may be interested in solving the following problems: assessing the potential of existing facilities in HEE, assessing the actual quality of provided services and finally assessing performance of educational system. In most cases the solution of mentioned problems require comprehensive assessment of education quality.

There are different issues of CA which can be formalized in different ways depending on the object of assessment. From the point of view of goals and tasks of management we can distinguish three classes of CA problems.

The first class is represented by CA of stakeholders' requirements satisfaction. These requirements are described in normative conditions and specifications. The examples of tasks of this class are the assessment of HEE readiness to licensing or accreditation; the estimation of candidate while employment (e.g., on professor post).

The second class includes CA problems of quality as a characteristic that bears ability to satisfy potential needs. The problems of this class include: construction of HEEs rating, the estimation of learning results (examinations, testing), assessment of resources quality.

The third class of problems consists in CA of performance, which reflects the results of a considered object usage. The tasks of CA of performance may include the following: evaluation of outcomes of HEE activities, estimation of the profit of resources development, evaluation of HEE's management projects and programs realization.

In the first class of problems the CA value is strictly determined by requirements. The main goal of such assessment is to define whether the object satisfies all requirements from specification. The degree of how well the requirements are fulfilled is not considered. Evaluation process in this case can be modeled based on switch chains. We take the notion of switch chains from the Theory of Intelligence (Bondarenko and Shabanov-Kushnarenko, 2006). A switch chain consists of a set of basic Boolean functions (conjunction, disjunction and negation, etc.). The combinations of those functions allow modeling of different complex objects.

The main problem of CA of the second class is the way of aggregation of estimates by different criteria. In this class of problems the quality is expressed as a totality of object's features. Therefore in general each feature is evaluated separately and then the CA is done. From our point of view the most advanced approach to solve this problem is represented by Qualimetry Theory (Azgaldov, 1982). It provides theoretical basis of quality assessment. According to qualimetry the quality is represented as hierarchy of properties of assessment object. Based on the set of certain axioms the property tree of object's quality is constructed. The top point of the property tree is object's quality; it consists of a set of simple and composite properties and has a hierarchical structure. Qualimetry suggests estimation of all simple properties and calculation the CA value with the help of one of weighted mean methods.

In the case when we deal with heterogeneous object (for example, educational process resources, customer outcomes in HEE) the construction of property tree appears to be an unsolvable problem. As a rule, such objects can be represented as a set of separate elements which involve own quality features. Due to expert judgments used for the property tree construction it is impossible to represent a heterogeneous object by means of a set of simple properties. It leads to the idea of partitioning of evaluation process in two main stages. The first one is evaluation of separate elements, as a result partial estimates are defined. For this purpose qualimetry approach is applicable. The second stage is aggregation of obtained partial estimates into the CA.

We suggest to use a network model for CA of quality of heterogeneous object (Cherednichenko et. al., 2012). The CA is done using composite functions (for example, arithmetical or geometrical weighted means). Evaluation framework is represented as a graph with two types of nodes. Nodes-entries of this graph are associated with partial estimates. Nodes-aggregates express the estimate of group of elements based on particular composite function.

The third class of CA problems implies the estimation from the point of view of customer value. In this case the assessment object can be represented through latent variables that influence the observable attributes. Based on heuristic procedure the set of indicators is constructed. We think that values of these indicators have to be obtained with the help of statistical data collection. This causes application of statistical analysis for the CA value calculation.

The CA is done using probability-based reasoning. It is assumed that unknown value of latent variable is expressed through the function of probability of obtaining some definite value of each indicator. The probabilistic function is determined by statistical model. For example, to estimate learning results Rasch model can be used. It allows defining person's ability based on answers to questions of the test (Wright and Stone, 1999).

Therefore three classes of problems have different assessment focuses, ways of inputs definition and aggregation models (Table 1). The class of the problem defines assessment focus and aggregation model, but inputs may vary for each application case.

Table 1: Comprehensive assessment classes of problems.

Class	Assessment focus	Inputs	Aggregation model
I	Fulfillment of all stated requirements	Expert judgment	Switch chain based on Boolean functions
II	Possibilities of totality of quality features	Partial estimates of separate elements	Comprehensive assessment network
III	Performance of customers outcomes	Collected statistical data	Probability-based reasoning

4 CASE STUDY

Our case study represents the CA of students' satisfaction of education quality. Since we have already made researches in evaluation of students' satisfaction, we have chosen this case study to demonstrate applications of CA models considered above (Cherednichenko and Yangolenko, 2012).

We suggest to evaluate education quality as the quality of services based on SERVQUAL method (Parasurman et. al., 1985). According to it the service quality is considered in terms of five SERVQUAL dimensions: tangibility, reliability, responsibility, security and empathy. The SERVQUAL is targeted on revelation of expected and perceived service quality. We consider the adaptation of original SERVQUAL questionnaire for measuring education service quality (Oliveira and Ferreira, 2009). We suggest to use the single questionnaire with 19 questions that define the gap between the perceived and expected education quality as it is described in our previous work (Cherednichenko and Yangolenko, 2012). The questions are scored using 7-points scale. The scores range from 1, which means a strong negative difference between perceived and expected quality (so the expectations were not justified), through 4, which denotes the absence of any gap, to 7, which means a strong positive difference (the perceived reality turned out to be much better than expectations).

We have conducted a survey of 75 four-year students of our department. To process the students' answers we chose the following Item Response Theory models: Rasch model (RM) and Partial Credit model (PCM) (Reeve, 2011).

Since RM provides processing of dichotomous questionnaire data, students' answers have to be converted into dichotomous scale related to positive or negative gap. The probability $P(x_{ij})$ of i-th student to answer positively on j-th question is described by the following dependency:

$$P(x_{ij}=1\,|\,\theta_i,\beta_j)=\frac{\exp(\theta_i-\beta_j)}{1+\exp(\theta_i-\beta_j)}, \qquad (1)$$

where θ_i is a satisfaction level of i-th student; β_j is difficulty of j-th question.

According to PCM the probability of the event that i-th student gives x points for j-th question is expressed as:

$$P(u_i=x\,|\,\theta_i)=\frac{\sum_{k=0}^{x}(\theta_i-\delta_{jk})}{\sum_{h=0}^{m_j}\exp\sum_{k=0}^{h}(\theta_i-\delta_{jk})}, \qquad (2)$$
$$x=0,1,\ldots,m_j,$$

where θ_i is satisfaction level of i-th student; δ_{jk} is the difficulty of j-th question which defines the probability of selection of value x instead of x-1.

The overall estimation of perceived quality based on the answers on 19 questions according to both measurement models is given in Table 2. We find the descriptive statistics of obtained results (minimal and maximal values, mode, median, mean standard error – MSE and standard deviation – SD). The obtained values of students' satisfaction θ_i are measured in logits and are the input data for the CA.

Table 2: Overall students satisfaction estimate.

Model	Mean	Min	Max	Mode	Median	MSE	SD
Rasch	1,9	-2,44	4,44	4,44	1,51	0,94	1,86
PCM	0,1	-1,76	1,64	-0,25	-0,06	0,2	0,68

In the case when we evaluate education quality we need to make CA of students' satisfaction that can be defined through aggregation of θ_i parameters. The estimate of the quality criterion is calculated as mean of corresponding students' satisfaction estimates. Mean estimates are taken as intermediate aggregate estimates. CA value is calculated by aggregation of those intermediate estimates.

Quality criteria values and descriptive statistics of obtained estimates are given in Table 3 and Table 4. Each dimension is considered as education quality criterion (1 – tangibility, 2 – reliability, 3 – responsibility, 4 – security, 5 – empathy).

Table 3: Analysis of students satisfaction estimates according to RM.

Quality criterion	Mean	Min	Max	Mode	Median	MSE	SD
1	1,9	-4,4	3,77	3,77	2,04	1,7	2,12
2	1,23	-2,6	2,75	2,75	2,75	1,8	1,89
3	1,45	-2,52	2,53	2,53	2,53	1,52	1,33
4	1,65	-2,52	2,51	2,51	2,51	1,56	1,17
5	1,00	-2,54	2,56	2,56	1,15	1,48	1,61

We can see that PCM provides estimates of students' satisfaction in more differentiate manner. This is due to the bigger number of grades of answers to each question than in RM.

Table 4: Analysis of students satisfaction estimates according to PCM.

Quality criterion	Mean	Min	Max	Mode	Median	MSE	SD
1	0,16	-2,35	4,26	0,54	-0,04	0,52	1,18
2	-0,04	-5,57	3,5	-0,23	-0,23	0,68	0,68
3	0,28	-2,54	5,39	0,42	0,21	0,53	1,17
4	0,32	-1,48	2,74	0,09	0,09	0,52	0,99
5	0,03	-3,6	2,84	-0,26	-0,05	0,53	1,23

errors for RM are greater than for PCM which indicates a smaller dispersion of estimates for PCM. So estimates obtained with the help of PCM are more adequate and preferable.

Since PCM is more adequate we calculate weighted and unweighted arithmetical means (WAM and UWAM) and geometrical means (WGM and UWGM) only for this model. To find WAM and WGM the following weight coefficients were used: $\alpha 1 = 0{,}27$, $\alpha 2 = 0{,}2$, $\alpha 3 = 0{,}15$, $\alpha 4 = 0{,}25$, $\alpha 5 = 0{,}13$. The obtained results are the following: UWAM = 0,15, WAM = 0,16, UWGM = -0,11, WGM = -0,13.

So these results are close to overall satisfaction level equal to 0,1 which was found as the mean of θ_i (Table 2). Using such decomposition except aggregated estimate we can find intermediate estimates. Furthermore we can assign different weight coefficient to make the comprehensive result more suitable for the purposes of decision-making.

In the case when we evaluate whether the quality perceived by students corresponds to the given level, we deal with the CA problem of the first class. Such estimate can be found based on input data which are the estimates of students' satisfaction by five quality criteria. The switch chain consists of three layers of Boolean functions. The first layer is represented by the function that defines whether each student's satisfaction value is greater than defined level (it returns 1, if this requirement is fulfilled, 0 – otherwise). We take the median $\theta = -0{,}06$. The function of the second layer checks whether a single criterion is assessed positively, i.e. the most of satisfaction values of single criterion are greater than defined level. In our case the estimates given by more than 37 students have to be bigger than the defined level. The third layer function defines whether the requirement to satisfaction level over all criteria is fulfilled. In this case study at least 3 criteria must meet the requirement. Under the value $\theta = -0{,}6$ we found out that the perceived quality is satisfactory.

To accomplish our case study we made the CA of the third class. As CA we take the estimate of proposed quality. The main hypothesis is that proposed quality defines the estimates of perceived quality. We suppose that these estimates correspond to calculated θ_i. To find CA value we suggest to use Spearman Single Factor Model. We obtained value of proposed quality equal to 1,97 logits. This corresponds to enough level of educational services.

The obtained results showed the satisfactory education quality from three different points of view. Therefore we suggest to use described approach for implementation of M&E IS.

5 CONCLUSIONS AND FUTURE WORK

According to the functionality classification the following IS can be distinguished: systems of administrative, financial and economic management; systems of educational process management and support; systems of scientific and research work management; systems of information resources management.

All of them should contain the CA unit. Due to goals and management tasks the different models can be used. We have realized three main classes of CA problems. The certain framework is associated with every problem's class.

We have discovered the most advanced procedures of CA. They are expert judgments, qualimetric practices or statistical analysis for initial estimates (inputs of CA). We suggest Switch Chains, Network Assessment and Probability-based Reasoning in order to construct comprehensive assessment model. Our researches are strictly devoted to implementation of CA procedures. On the other hand, we have tried to generalize our experience to provide some formal approach.

The investigation of the case-study shows potential possibilities of suggested frameworks usage. We should note that the estimation of students' satisfaction is not the clearest way to demonstrate the advantages of our approach. But we hope that the aim of illustration how different tasks influence comprehensive assessment is reached.

As a result of this work we can underline the following: 1) the process of CA is represented in two stages: estimation of separate elements and their aggregation into CA value; 2) three classes of problems and CA frameworks related to those classes are defined; 3) the set of experiments based on evaluation of students' satisfaction were done; 4) the principle role of probability-based reasoning methods for performance evaluation is proved.

Therefore, the suggested CA frameworks can be used for M&E Software elaboration. The future researches will be connected with the up-to-date CA tasks in HEE. Our researches are aimed at development of M&E models and IT that can be applicable in higher education as well as in other domains.

REFERENCES

Athanasiadis, I. and Mitkas, P. (2004). An agent-based intelligent environmental monitoring system. Management of Environmental Quality, 15(3), 238-249.

Azgaldov, G. G. (1982). Theory and practice of goods quality assessment (Basics of Qualimetry). Moscow: Economics.

Bondarenko, M. F. and Shabanov-Kushnarenko, U. P. (2006). Theory of intelligence: a Handbook. Kharkiv: SMIT Company.

Brown, R. (2005). *Rational choice and judgment: decision analysis for the decider*. Wiley-Interscience.

Carrizo, L., Sauvageot, C. and Bella, N. (2003). *Information tools for the preparation and monitoring of education plans*. UNESCO.

Cherednichenko, O., Yangolenko, O. and Liutenko, I. (2012). Issues of model-based distributed data processing: higher education resources evaluation case study. *The international proceedings volume of ICTERI 2012*, CEUR-WS Vol 848, 147-154. Retrieved July 15, 2012 from http://ceur-ws.org/Vol-848.

Cherednichenko, O. and Yangolenko, O. (2012). Towards higher education quality assessment: framework for students satisfaction evaluation. Proc. 4th International Conference on Computer Supported Education (CSEDU 2012), SciTePress, 2, 108-112.

EFQM Excellence Model Higher Education Version 2003. Retrieved August 14, 2012 from http://vpaa.epfl.ch/files/content/sites/vpaa/files/ACC-EFQM%20Excellence%20Model%202003%20ENG.pdf.

Kachalov, V.A. (2001). *ISO 9000 Standards and Problems of Quality Management in HEE (Notes of Quality Manager).* Moscow: IzdAT.

Koenig, J. (2011). *Assessing 21st Century Skills: Summary of a Workshop.* Washington, D.C.: The National Academies Press.

Krukov, V.V. and Shahgeldyan, K.I. (2007). *Corporative information environment of university: methodology, models, solutions.* Vladivostok: Dalnauka.

Meyer, M. and Booker, J. (2001). *Eliciting and analyzing expert judgment: a practical guide.* Philadelphia: SIAM.

Oliveira, O. J. and Ferreira, E. C. (2009). *Adaptation and Application of the SERVQUAL Scale in Higher Education.* POMS 20th Annual Conference, Orlando, Florida USA.

Parasurman, A., Zeithaml, V. and Berry, L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, 49, 41-50.

Reeve, B. (2009). *An Introduction to Modern Measurement Theory*. Retrieved May 1, 2012 from http://appliedresearch.cancer.gov/areas/cognitive/immt.pdf.

The Information System of the Federal Health Monitoring. Retrieved August 25, 2012 from http://www.gbe-bund.de/gbe10/pkg_isgbe5.prc_isgbe?p_uid=gastd&p_sprache=E.

Wright, B. and Stone, M. (1999). *Measurement Essentials (2nd ed.)*. Wilmington: Wide Range Inc.