

Comparative Analysis of the Computerized Testing Systems in Education

Evgeny Tishchenko¹
celt@inbox.ru

Oleg Serpeninov¹
serpeninov53@mail.ru

Elena Zhilina¹
black-2@mail.ru

¹*Rostov State University of Economics, Rostov – on – Don, 344002, Russia*

Abstract

There are a number of controlling, educational innovative systems in the educational market at present. The selection of the computerized testing systems is a complex task, it is necessary to apply specific methods to complete it. The article performs the comparative assessment of the functional completeness of the computerized adaptive testing systems in education, whose functional possibilities are determined by the intersection of functions between the selected classes of test systems. The comparison of the testing systems according to the functional completeness allows to make a full list of the functions, to systemize the information on the composition and functional completeness, to quantify the computerized testing system compliance degree with the user's requirements, at the stage of the preliminary analysis to exclude the computerized testing systems where the necessary functions are not performed from the further consideration, to form a group of the computerized testing systems with the identical functional completeness, to expand the optimal choice for the consumer-user at the computerized testing system market providing the full list of the performed functions by each system, and to show the designer the position of the product among other existing computerized testing systems.

Keywords: computer system, adaptive testing, functional completeness, functions list, the “reference” system, absorption matrix, similarity matrix, similarity graph.

1 Introduction

At present a number of the innovative software applied in education has emerged. The selection of the computerized testing systems is a complex task, it is necessary to apply specific methods to complete it. The main challenge to face while performing performs the comparative assessment of the computerized testing systems is lack of certainty and probable complexity of modern educational products as well as lack of detailed researches in

Copyright 2019 for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

In: Jože Rugelj, Maria Lapina (eds.): Proceedings of SLET-2019 – International Scientific Conference Innovative Approaches to the Application of Digital Technologies in Education and Research, Stavropol – Dombay, Russia, 20-23 May 2019, published at <http://ceur-ws.org>

the sphere of the comparative quantitative and qualitative assessments of functional completeness the innovative testing systems in education. That is why, in our opinion, it is grounded to apply formal procedures to compare computerized systems according to the criterion of functional completeness.

2 Task

To assess the consumer quality of the analyzed the computerized testing systems (CST) is a rather complex task. The introduction of the innovative methods into the educational testing systems functional is taking place currently. CST are often differ in quality and price. In this connection it is expedient to accomplish the following tasks:

- to analyze the quality of the best known CST applied to automate the educational testing process and to make the list of the functions performed by these systems [Lip83];
- to give a comparative assessment of the systems under consideration according to the functional completeness [Kal99, Efi99];
- to single out the list of the compulsory and auxiliary functions for the analyzed systems and to design a “reference” model of a computerized testing system to meet the main users’ requirements .

[Wei73]

Computerized Adaptive Testing – CAT, where tasks (questions) are selected algorithmically depending on the previous answers adapting to the knowledge level of the tested one, can be referred to the innovative educational methods [Kim08, Tis18, Tis16, Sca07].

By computer adaptive testing we understand “computerized system of the scientifically grounded testing and assessment of the learning results highly efficient due to optimized procedures for generation, presenting and test evaluating test results” [Ser07, Ava02, Nik16]. CAT improves the motivation, reduces testing time, demands fewer tasks for every examinee without measurement accuracy reduction [Zhil10, Bus07, Nik14, Khu95].

To research and compare CST between themselves applying the formalized methods [Bue99] it is necessary to decompose targets which traditionally are split into subsystems and tasks (modules) [Tel02, Tit98, Smi02]. The design of the CST functional model started with modeled testing system in a whole in the form of the context diagram (fig. 1).

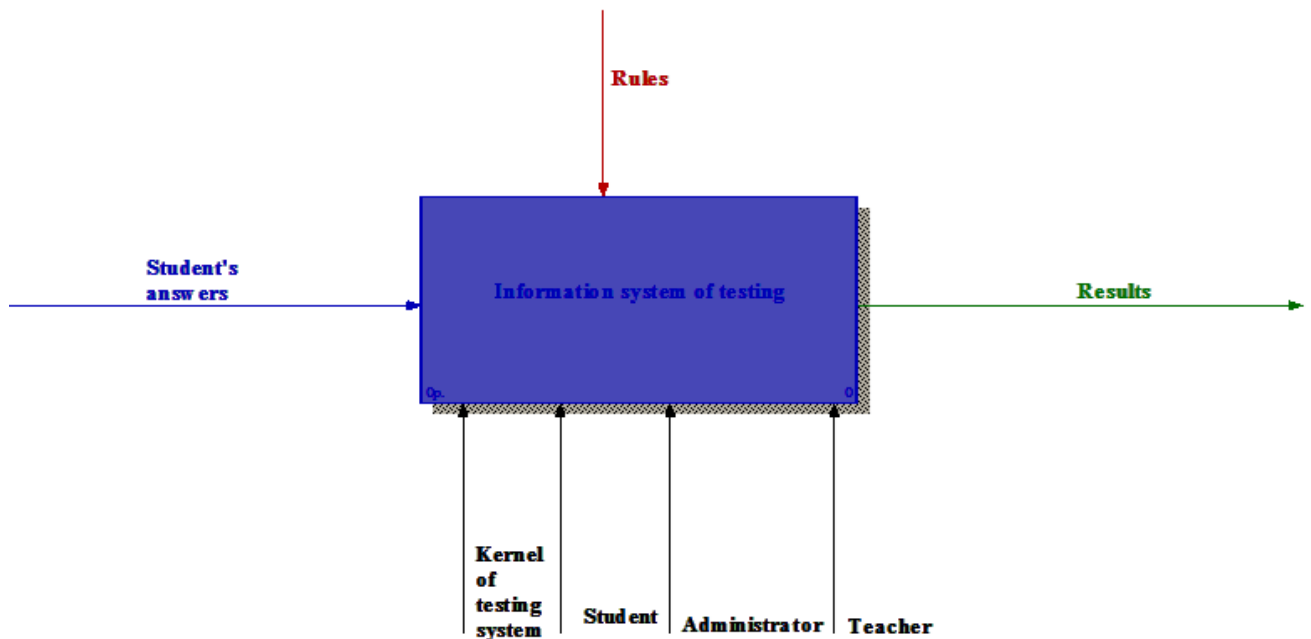


Figure 1: Context diagram of the visual model of computer testing system, model IDEF0

The system interacts with environment through entrance terms (fig. 1 «Student’s answers»), exit (the whole process results – «Results»), management («Rules») and mechanisms («Kernel of testing system», «Student», «Administrator», «Teacher»).

To conduct the functional completeness analysis CST model context diagram was decomposed into functions. By the function we understand the system's ability to satisfy a user's need. Another decomposition variant is to split CST into functional operations. The functional operation means the system user actions to perform a function. Function is a system's ability; it is performed when a user performs a corresponding functional operation.

In the context of the conducted research more than 90 educational CST, including "Expert", "AST-Test", "SInTeZ", "SunRav TestOfficePro", "WebTester", "WebTutor (module "Testing")", IBM Lotus LearningSpace, eLearning Server 3000 (module "Testing")" Moodle (module "Test")", etc. were analyzed. The preliminary list of the functions of the analyzed systems was made with application of demo and working CST versions, documents and literary sources. This list is rather full, but its further expansion is possible due to functions addition identified by expert surveys. CST analysis allows singling out more than 100 functions, whose top level of the hierarchy structure is presented Fig.2.

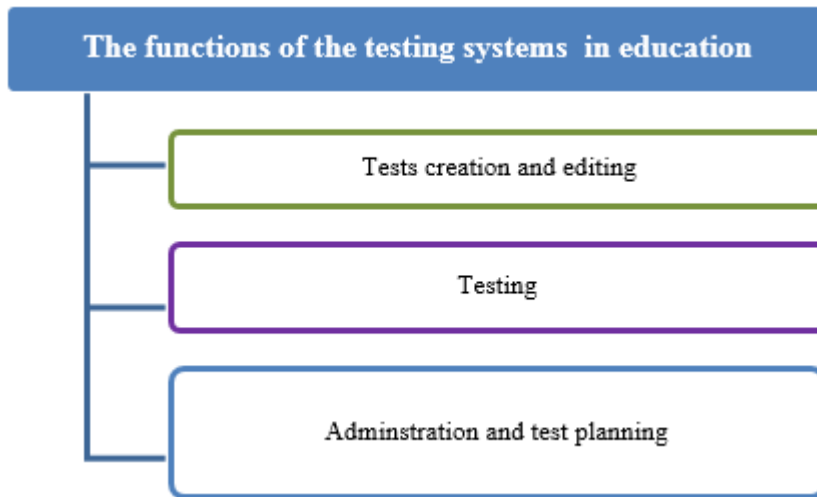


Figure 2: The top level of the hierarchy of the educational CST functions

The "Developing and editing texts" Group includes about 30 functions, whose fragment is presented at Fig.3. The "Administration and planning" Group comprises such functions as "Users authorization", "User groups", "Client-server mode", etc. The "Testing" Group has more than 20 functions whose fragment is presented at Fig.4.

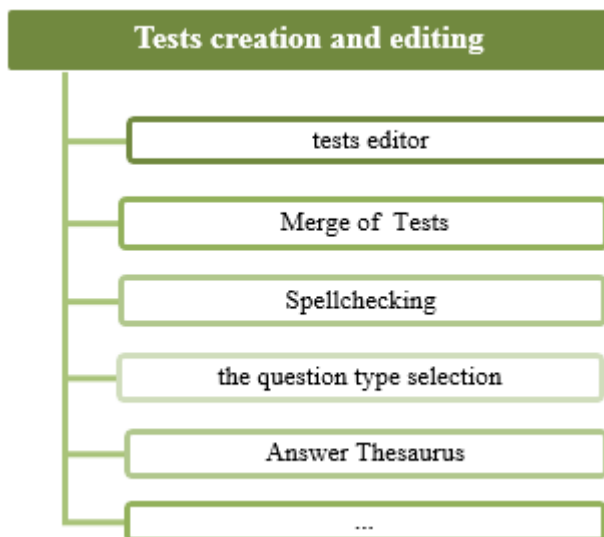


Figure 3: The "Developing and editing texts" Group functions

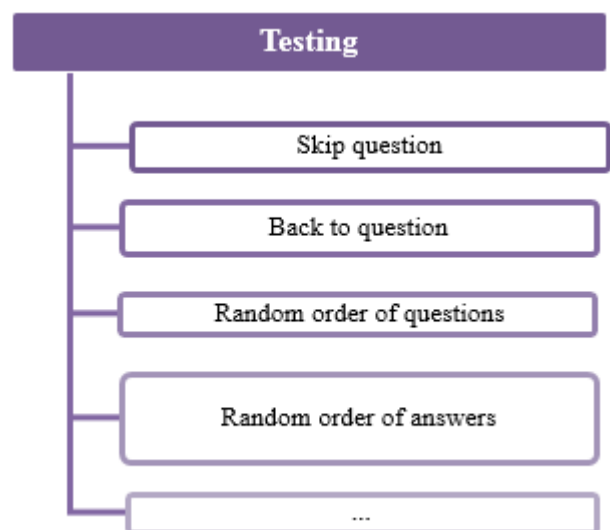


Figure 4: The "Testing" Group functions

3 Development Of Methodology

3.1 Making The List Of The Functions Performed By CTS

To analyze CST according to its functional completeness it is proposed to divide them into two groups:

- distributive testing systems (DTS) – the systems to be installed in each PC;
- web-systems (applications) for testing (WST) and testing modules in the educational systems.

The suggested groups allow identifying the common functions inherent to each group of the analyzed systems and the individual functions as well. Any system can be referred to : both distributive and web-system supporting adaptation functions. Functions Merge Scheme is presented at Fig. 5, it allows identifying two blocks of the analyzed functions of CTS of the adaptive testing:

1. block: , installed distributively;
2. block: web-systems of .

Further on the common functions present at CTS groups will not be considered.

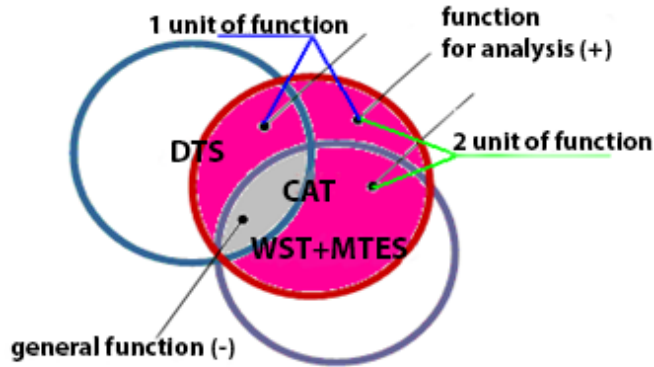


Figure 5: CST Functions Merge Scheme

The CAT functions list fragment (out of 53) to analyze the CTS functional completeness is presented in Table 1.

Table 1: CAT fragment - CTS Functions to Analyze Functional Completeness

Function name
.....
To rate of the tested (students)
To set the grading scale
To set the range of the number of questions in the test (total, min, max)
To diagnose the problem areas of the tested one
To form a knowledge bank
To set/ selects a starting point
To set test points
To assign the time (distribution)
To select the frequency of the unexpected answers by those to be tested
To set completion criteria
To form intermediate assessments
To set a number of the answers guessed and a number of the random errors
To calculate the statistical indicators of the variation of test tasks
To calculate the test task reliability coefficient

The features of the design, application and selection of the distributive systems and web development do not allow conducting the joint analysis of their functional completeness that is why this methodology we apply separately for each of the two groups Let us single out the systems supporting the analyzed functions, from the full list of the considered CAT (Tables 2, 3).

By the “reference” system we understand the system supporting all analyzed functions.

Table 2: CAT fragment - CTS Functions to Analyze Functional Completeness

Code	Name
S1	Test Constructor
S2	«The University chair tests» SSU
S3	ST-st
S4	«MasterTest»
S5	«SInTeZ»
S6	Universal test complex
S7	Test Edit
S8	Test Yourself
S9	UniTest System
S10	Revizor
S11	Mental Control
S12	Students' testing
S13	Vesta
S14	AUGUST
S15	CS Net
S16	MiniTestSL
S17	Quizzz
S18	VerlTest
S19	TestBOX Standalone Builder
S20	Softvea TestBOX
S21	SunRav TestOfficePro
S22	Attestation
S23	IT Test
S24	Your test
S25	FastTEST Professional Testing System
S26	Test4DL ,
S27	MSCat Demo
S28	Test 2002. Professoinal Edition
S29	-Test
S30	Reference system

Table 3: WST+ ES List

Code	mane
VS1	Test. Software environment «LEMON»
VS2	«KnowledgeCT»
VS3	«Teketesting»
VS4	«TESTOR.RU»
VS5	WebTester
VS6	WebTutor (module “Testing”)
VS7	Chamilo (module ”Test”)
VS	Claroline (module ”Tests”)
VS9	5+ (module “Testing”)
VS10	IS COA (module “Testing”)
VS11	eLearning Server 3000 (module “Testing”)
VS12	LAMS (module “Testing”)
VS13	Moodle (module ”Test”)
VS14	Sakai (module “Testing”)
VS15	distance education system Prometheus (module ”Test”)
VS16	distance education system “DOCENT” (module “Testing”)
VS17	Learn eXact (module “Testing”)
VS18	Microsoft E-Learning (module “Testing”)
VS19	IBM Lotus LearningSpace (module “Testing”)
VS20	Reference system

3.2 A Comparative Assessment Of The CTS According To The Functional Completeness

The fragment of the X matrix showing what functions performed by the researched DTS is presented in Table 4. If function j is supported by system I, then $X_{ij}=1$, otherwise $X_{ij}=0$. Applying the methodology proposed in [2, 16], it is possible to analyze the functionality of these systems: to assess the completeness of the automated functions with respect to the “reference” system as well as to assess the degree of their similarity and connection with each other.

Table 4: X Matrix (Source Data for Analysis). DTS Functions Performance (Fragment)

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	...	F52	F53
S1	1	1	1	0	0	1	0	1	0	1	0	0	1	0	...	0	1
S2	1	0	0	0	0	1	0	1	0	0	0	0	0	0	...	0	1
S3	1	0	0	0	1	1	0	1	1	0	0	0	0	0	...	0	1
S4	0	0	0	0	1	1	0	1	0	1	1	0	1	0	...	0	1
S5	1	0	0	0	1	1	0	1	1	1	1	0	1	0	...	0	1
S6	0	0	0	0	1	1	0	1	0	0	0	0	0	0	...	0	1
S7	0	0	0	1	0	0	1	1	0	0	0	0	0	0	...	0	0
S8	0	0	0	0	1	1	1	1	0	0	0	0	1	0	...	0	0
S9	1	1	1	0	1	1	1	1	0	1	0	0	1	0	...	0	0
S10	0	0	0	0	0	1	1	1	1	1	0	0	1	0	...	0	1
S11	0	0	0	0	1	0	1	1	0	1	1	0	1	0	...	0	0
S12	0	0	0	0	0	1	1	1	0	0	0	0	1	0	...	0	0
S13	0	0	0	0	1	1	1	1	0	0	0	0	1	0	...	0	1
S14	0	0	0	0	1	1	0	1	0	1	0	0	1	1	...	0	1
S15	1	0	1	0	1	1	1	1	0	1	0	0	1	0	...	0	1

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	...	F52	F53
S16	0	0	0	0	1	0	1	1	0	0	0	0	0	1	...	0	0
S17	1	0	0	0	1	1	1	1	0	1	0	0	1	0	...	0	1
S18	1	0	0	0	0	1	0	0	0	1	0	0	1	0	...	0	0
S19	1	0	0	0	0	1	0	1	0	0	0	0	1	0	...	0	0
...
S23	0	0	0	1	0	0	0	1	0	1	0	0	0	1	...	0	1
S24	0	0	0	0	0	0	0	1	0	0	0	0	0	0	...	0	1
S25	1	1	1	1	0	1	1	1	1	1	1	0	1	1	...	0	1
S26	0	0	0	0	0	0	0	1	1	1	1	0	1	0	...	0	1
S27	1	0	0	0	0	0	1	1	1	1	0	1	0	0	...	0	0
S28	1	1	0	0	1	0	1	1	1	1	1	0	0	0	...	0	1
S29	1	1	0	0	1	0	0	1	1	1	0	0	0	0	...	0	1
S30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	...	1	1

Let us single out the systems IS_i and IS_k and introduce the following notation:

- $P_{ik}^{11} = IS_i \cap IS_k$ - intersection power of systems with respect to automated functions;
- $P_{ik}^{01} = IS_k / IS_i, P_{ki}^{10} = IS_i / IS_k$ - the power of difference of the corresponding sets. As a mismatch measure between systems IS_i and IS_k we choose value $S_i k R_{ik} = P_{ik}^{01} / P_{ik}^{11} + P_{ik}^{10}$;
- to assess the extent of absorption by the system IS_k of the system IS_i - value $HikA_{ik} = P_{ik}^{11} / (P_{ik}^{11} + P_{ik}^{10})$;
- to assess the degree of similarity of systems - a measure of Jaccard similarity: $Gik L_{ik} = P_{ik}^{11} / (P_{ik}^{11} + P_{ik}^{10} + P_{ik}^{01})$.

Through the logical absorption matrices the degree of the interconnection and similarity between the compared systems IS_i and IS_k is identified. The indicators are calculated:

- the share of the common functions performed simultaneously by IS_i and IS_k in the total volume of functions of IS_i (matrix H);
- number of the common functions possessed by IS_i and IS_k (matrix 0);
- the share of common functions in the total volume of functions IS_i and IS_k (matrix G).

Selecting different threshold values for the elements of the matrices $P^{10}, S^{ik}, G^{ik}, H^{ik}$ it is possible to design the absorption matrices $P^{ik0}, S^{ik0}, G^{ik0}, H^{ik0}$ respectively.

The elements of the new matrices replace the respective element of the basis matrix for “1” if the element value is greater than or equal to the threshold one and “0” otherwise.

Elements of the main diagonal (whose column and line number coincide) in any case equal to zero.

The absorption matrix H^{ik0} and similarity matrix G^{ik0} help to determine if the analyzed test systems belong to the “reference” one (Tables 5, 6).

Table 5: Similarity Matrix G^{i0k} DTS (Fragment)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	...	S30
S1	1	1	0	1	1	0	0	0	1	0	1	0	0	1	1	0	0	...	0
S2	1	1	1	1	1	1	1	0	0	0	0	0	0	0	1	1	0	...	0
S3	0	1	1	1	1	1	0	0	0	0	0	0	0	0	1	0	1	...	1
S4	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	0	0	...	0
S5	1	1	1	1	1	1	0	1	1	0	1	0	1	1	1	0	1	...	1
S6	0	1	1	1	1	1	1	1	0	0	0	0	1	0	1	1	1	...	0
S7	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	1	0	...	0
S8	0	0	0	0	1	1	0	1	1	1	1	1	1	1	1	0	1	...	0
S9	1	0	0	0	1	0	0	1	1	1	1	1	1	0	1	0	1	...	0
S10	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	...	0

Table 7: Absorption Matrix $H^{ik_0} DTS$ (Fragment)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	...	S30
S26	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	...	1
S27	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	...	1
S28	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	...	1
S29	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	...	1
S30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	...	1
Σ	18	9	26	23	28	13	2	8	13	9	12	5	12	13	25	4	20		30

Similarly the analysis according to functional completeness criterion for the systems WST+ ES was conducted.

4 Results

Close to the “reference” one are the distributive systems: S3 (AST-Test), S4 («MasterTest»), S5 («SInTeZ»), S15 (Net), S17 (Quizzz), S25 (FastTEST Professional Testing System), S26 (Test4DL,), S27 (MSCat Demo), S28 (Test 2002. Professoinal Edition), S29 (-) and web-systems: VS6 (WebTutor (module “Testing”), VS11 (eLearning Server 3000 (module “Testing”), VS12 (LAMS (module “Testing”), VS14 (Sakai (module “Testing”), VS15 (distance education system Prometheus (module ”Test”), VS17 (Learn eXact (module “Testing”), VS18 (Microsoft E-Learning (module “Testing”), VS19 (IBM Lotus LearningSpace (module “Testing”). To illustrate it two groups of such CST whose graphs are presented at Fig. 6 and 7 were singled out.

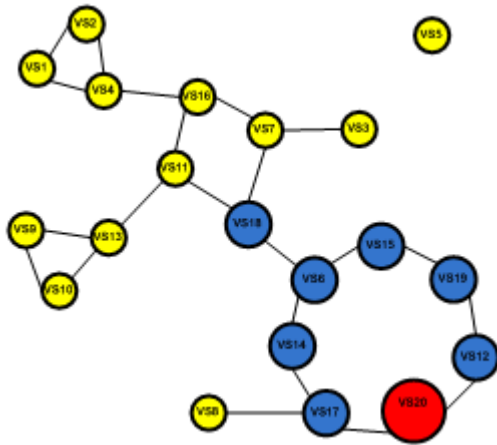


Figure 6: Similarity graph for DTS, grouped

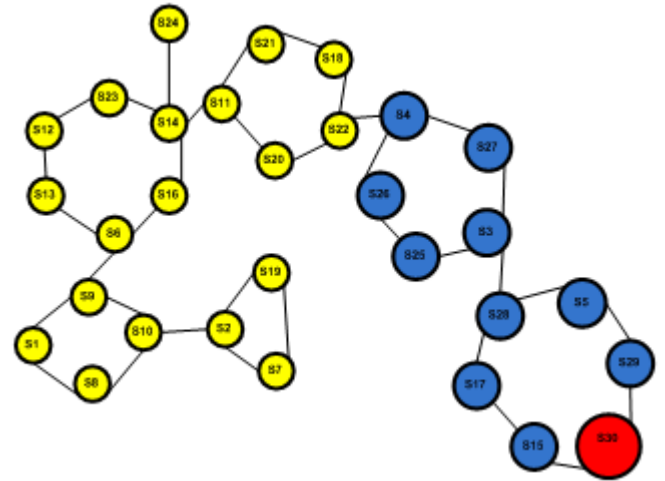


Figure 7: Similarity graph for WST + MTES, grouped

5 Discussion

CAT with CAT functions from one group can be compared with each other in the further, comparing such characteristics as price, producer, reliability, adaptation degree and others, giving the consumers the possibility to make a reasonable choice.

Threshold values for the elements of matrices and changed in the interval [0.25..0.8]. The optimal value was chosen option 0.4. Increasing the threshold values to 0.8, the number of the analyzed testing systems close to the “reference” one reduced dramatically. Only DTS FastTEST Professional Testing System turned out to be close to the “reference” one, and there was no one web-system to satisfy this criterion

6 Conclusion

There are a number of controlling, educational innovative systems in the educational market at present. It is rather difficult to select the right one among them.

The comparison of the testing systems according to the functional completeness allows implementing the following:

- to make a full list of the functions performed by the considered computerized educational testing systems;
- to systemize the information on the composition and functional completeness of the existing computerized testing systems;
- to quantify the computerized testing system compliance degree with the user's requirements for functional completeness;
- at the stage of the preliminary analysis to exclude the computerized testing systems where the necessary functions are not performed from the further consideration;
- to form a group of the computerized testing systems with the identical functional completeness;
- to expand the optimal choice for the consumer-user at the computerized testing system market providing the full list of the performed functions by each system, and to show the designer the position of the product among other existing computerized testing systems

At the same time the result does not speak on the complete analysis of the functional completeness of the computerized testing systems in education that is why in the future it is necessary to compare their prices and other features (for example, the applied technical means and the operating system, the user interface quality, the complexity of working with the system, reliability, etc).

References

- [Lip83] .V. Lipaev. Quality software. M.: Finance and Statistics, 1983. 263 p.
- [Khu98] .N. Khubaev. Comparison of complex software systems according to the criterion of functional completeness // Software products and systems. 2. pp. 6-9. 1998.
- [Kal99] .Kh. Kalugyan. Testing systems in high school as a tool for managing the educational process (for example, the discipline "Economic Informatics"): diss. - Rostov-on-Don: RSEU (RINH), 1999.
- [Efi09] .N. Efimov About approaches to the structure of an adaptive knowledge testing system // Problems of creating and using information systems and technologies: materials of the third scientific-practical conf. December' 11, 2009. - Rostov-on-Don: RSEU (RINH). 2009.
- [Wei73] .J. Weiss. The stratified adaptive computerized ability test. Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program. 1973.
- [Kim08] . Kim-Kang, D. J. Weiss. Adaptive measurement of individual change // Journal of Psychology. 2008. 216 (1). pp. 49-58.
- [Tis18] .N. Tishchenko, E.V. Zhilina, T.N. Sharypova, G.N. Palutina. Fuzzy models of results of the mastering the educational programs in the field of information security // Theory and Application of Fuzzy System and Soft Computing – ICAFS. 13th International Conference. 2018.
- [Tis16] .N. Tishchenko, T.N. Sharypova, E.V. Zhilina, S.E. Cherkezov. Economic and Mathematical modeling of complex cooperation of academic staff of educational cluster on the basis of fuzzy sets theory // Journal of Applied Economic Sciences (JAES). Issue 5 (43). pp. 905-907. 2016.
- [Sca07] . Scalise, M. Wilson. Bundle models for computerized adaptive testing in e-learning assessment // Proceedings of the 2007 GMAC Conference on Computerized Adaptive Testing. 2007.
- [Ser07] .V. Sergeev. Adaptive testing in distance learning systems // Science and Education. 4. 2007. Access mode: <http://technomag.edu.ru/doc/65577.htm>. 2019/03/10.
- [Ava02] .S. Avanesov. The composition of the test tasks. M.: From the Testing Center of the Ministry of Education of the Russian Federation. 2002.
- [Nik16] .N. Nikolaeva. Adaptive Testing System // Traditions and Innovations in Construction and Architecture. Natural sciences and technosphere safety. - Samara: SGASU. 2016. pp. 50-53.
- [Zhil10] .V. Zhilina, I.I. Miroshnichenko. Designing the semantic content core of an educational Internet resource with support for the functions of adaptive knowledge control // Bulletin of the Rostov State Economic University. - Rostov-on-Don, 2010. - 3 (32). - p. 219-226.

- [Bus07] .I. Bushmeleva Automated adaptive testing system //Modern problems of science and education. 2007. 2. pp. 48-50
- [Nik14] .Yu. Nikiforov. Use of adaptive computer testing systems // Humanitarian scientific research. 2014. 4. Access mode: URL: <http://human.snauka.ru/2014/04/6274>. 2019/03/26.
- [Khu95] .N. Khubaev. Methods of economic evaluation of consumer software quality // Software products and systems (SOFTWARE and SYSTEMS). 1. pp. 2-8. 1995.
- [Buc99] . Buch. Object-oriented analysis and design with examples of applications in C ++. M.: "Publishing Binom", St. Petersburg: "Nevsky Dialect". 1999.
- [Tel02] u.F. Telnov. Intellectual information systems in economics: manual for universities. - 3rd ed. - M.: SINTEG. 2002.
- [Tit98] .A. Titorenko Automated information technologies in economics. M. 1998.
- [Smi02] .N. Smirnova, A.A. Sorokin, Yu.F. Telnov. Designing economic information systems. M.: Finance and Statistics, 2002. 512 p.