Quality of content delivery in computer specialists training system

Olexiy D. Azarov[a], Tetiana I.Troianovska*[a], Liudmyla A. Savytska[a], Tamara O. Savchuk[a], Larysa E. Nykyforova[b], Volodymyr A. Otryshko[a], Batyrbek Suleimenov[c], Konrad Gromaszek[d], Ainur Kozbekova[e], Azhan Sagymbekova[c]

[a]Vinnytsia National Technical University, 95 Khmelnytske shose, 21021 Vinnytsia, Ukraine; [b]Berezhany Agrotechnical Institute, st. Academic 20, 47501 Berezhany, Ukraine; [c]Kazakh National Research Technical University after K. I. Satpaev, 22 Satpayev street, Almaty, Kazakhstan; [d]Lublin University of Technology, ul. Nadnystrzycka 38D, 20-618 Lublin, Poland; [e]Institute of Information and Computational Technologies, 125, Pushkin St., 050010, Almaty, Kazakhstan

ABSTRACT

The given work is devoted to research of perception content features by computerized training system students on cluster analysis base. As a result of experiment was revealed that content giving in the offered computerized training systems is more effective in comparison with the traditional approach. The proof is the increase of good student's cluster.

Keywords: computerized training system, content delivering, indirect assessment, individual approach

1. INTRODUCTION

Exponential growth of educational content that needs to be mastered coupled with its constant aging generates the necessity to introduce new information technologies to the area of computerized training systems (CTS). Insufficient training level of specialists by CTS means is caused by low quality of content delivery in the system.

CTS - is a set of software and hardware tools, models, methods and information resources that provide content delivery to the entities of training by means of modern network and information technologies and also realize the supervision of received knowledge of the entities of training.

Quality of content delivery indirectly is determined by the results of the knowledge control of the entities of training. Low quality of content delivery reduces the quantitative indicators of success of the entities of training that is why it is necessary to consider in the immediate process of content delivery individual data of work characteristics of the entities of training.

These data affect the information processes of content delivery that allows to create adaptive training courses (ATC) the use of which gives higher percentage of absolute success and quality of training of entities of ATC in comparison with similar in content courses without individualized content delivery.

Nowadays when people value their time and strive to learn throughout life, CTS of professionals is not only the tool for improving qualifications, but also the acquisition of new specialized knowledge. Therefore, it is reasonable to improve constantly the quality of e-learning courses, especially the high competition and the demand for such systems is growing.

In general, most CTS are aimed at the supply of educational material (EM) to the entities on a predetermined principle without considering individual approach. Relevant is to analyze the perception of EM by the students of CTS that differ in systems of individual approach based on cluster analysis.

Formulation of the problem: set the task to analyze the peculiarities of EM CTS perception based on cluster analysis. To solve this problem, propose a model of cluster analysis and vague sets of numbers.

Photonics Applications in Astronomy, Communications, Industry, and High Energy Physics Experiments 2017, edited by Ryszard S. Romaniuk, Maciej Linczuk, Proc. of SPIE Vol. 10445, 104452S
 · doi: 10.1117/12.2281229

2. FORMULATION OF THE PROBLEM AND PLANNING OF THE EXPERIMENT

Generally, most CTS are aimed at the delivery of content to the training entities on a pre-defined principle without considering individual approach[1]. Relevant is to conduct cluster analysis of the perception of learning content by the CTS training entities that differ in principles of individual approach.

Formulation of the problem: set the task to investigate the peculiarities of perception of learning content of the training course using existing methods of cluster analysis.

Two training courses for discipline "Discrete Mathematics" were designed by the leaders of experiment. Both courses used the same raw data in other words the amount, nature and quality of the educational content are identical. The difference is its delivery. Content delivery in the first case can be attributed to the so-called traditional approach. The second variant of course is worked out in authorship CTS of experts using the principles of indirect assessment[2], that is, with elements of individual approach to perception of educational content by training entities[3]. Tests were made for the control. Tests to determine the level of residual knowledge were passed by students of 1st and 2nd years of study of the Faculty of Information Technology and Computer Engineering of Vinnytsia National Technical University who has studied the subject "Discrete Mathematics" and passed the exam.

In the proposed experts' CTS apart from direct, indirect estimation is used. Indirect estimation is based on a scale that consists of real values as opposed to discrete scale. It will permit to estimate the training entities with higher accuracy.

This experiment is based on the assumption (hypothesis) that a cluster of advanced training entities (with excellent and good marks) of offered CTS of experts will increase compared to the results of testing of students who studied and passed tests in other similar systems.

3. DATA COLLECTION

On stage of cluster sampling 100 students were selected: for 50 people from different groups and the number of excellent, good and satisfactory students are approximately equal. So, you can talk about the equivalence of the two groups in relation to each other. All selected students according to the workload have learnt the discipline "Discrete Mathematics" and have passed the test to determine the level of residual knowledge. Half, i.e. 50 students have repeated the course material and passed the tests in authorship CTS of experts. The other half - in the system with the traditional approach to content delivery.

At the first stage is an array of real data are marks collected among the proposed CTS of experts using copyright specially developed software of client-server type - Deloss[4]. This software also implements a communication protocol between the software agent and the marks analysis system and contains the subsystem for making decisions based on indirect assessment[3].

The exterior of Deloss and its list of training entities is shown in Figure. 1.

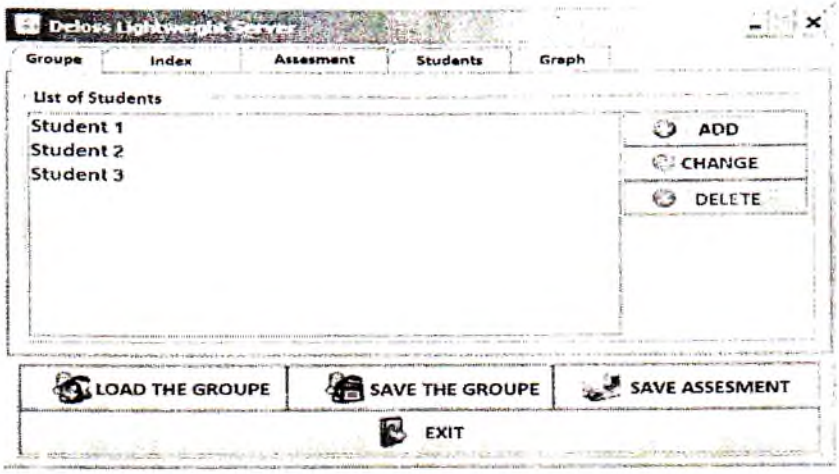

Figure 1. Formation of a list of training entities in Deloss environment.

Before the course training entity is offered a standard text after reading of which there are some nontrivial questions, answers to which would reflect his/her nature.

After reading the standard text the **"Text has been read"** button is provided which the training entity presses after the material has been read. The time spent by the training entity on perception of the content is fixed and also amount of text $V_{"et.\ text"}$. Results of compilation of $t_{et.\ text}$ (sec.), $t_{et.\ test}$ (sec.) and $V_{et.\ text}$ (symb.) are recorded in additionally developed client-server software **"Deloss"**. Software "Deloss" calculates the value of $t_{text\ n}$, $t_{test\ n}$ and $V_{et.\ text\ n}$ and evaluates the difficulty of perception by the training entity of educational content of ATC. The text perception difficulty - is an assessment which in the academic learning teacher makes intuitively.

The program «Deloss» performs the following actions:

1) assigning (saving and loading) of current list of training entities which are part of this group, is reflected in a tab **"Group"**. This is a list which is transferred to control applet for login ensuring of the training entity in the system (Figure 2, a);
2) simultaneous request for registration of several training entities - tab **"Indices"** (Figure 2, b);
3) tracks the progress of passing ATC by training entities - tab **"Indices"**;
4) fixation of tests solving by training entity - tab **"Testing"** (Figure 3, a);
5) it is possible to see individually on which computer training entity is working - tab **"Students"** (Figure 3, b);
6) tab **"Columns"** allows to see a visual diagram of his/her preparation course (Figure 3, c).

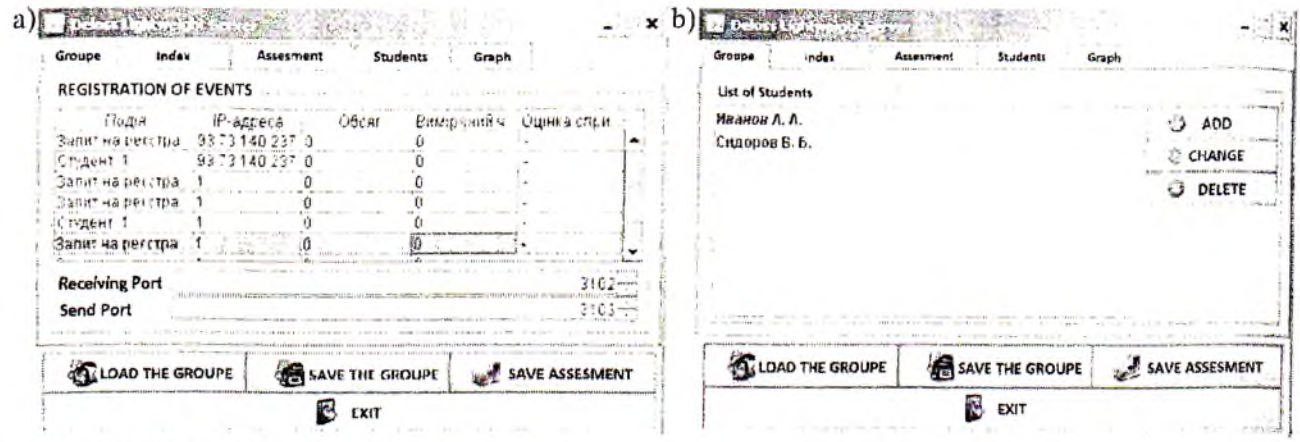

Figure 2. «Deloss» tabs.

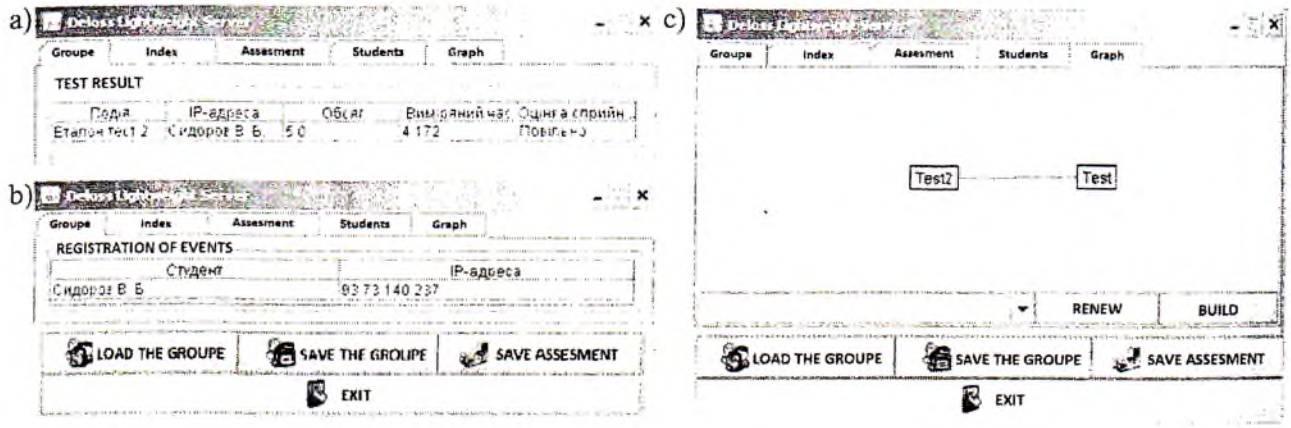

Figure 3. «Deloss» tabs.

In Figure 3 the result of the test completion by the training entity can be seen. Developers provide the ability to determine the IP addresses of hosts of training entities (Fig. 3, b). In the Figure 3, c, the trajectory of activity of training entity of ATC is visually evident. As shown in the Figure 3, c events in ATC, entities that activate these events and the IP addresses of hosts are recorded separately.

At the second stage of the experiment an array of real data are estimates collected among the CTS specialists with traditional content delivery. So, the necessary statistics for the experiment is collected. As data collected for the analysis of classical statistical methods is not enough, we apply cluster analysis[1].

4. JUSTIFICATION OF THE CHOICE OF MEANS OF DATA ANALYSIS

The main goal of cluster analysis is selection in raw multivariate data such homogeneous subsets so that objects within the groups were similar and objects from different groups - unlike[3, 5]. The advantage of this method is that it can work even when the data is not enough and the requirements of classical methods of statistical analysis are not fulfilled. Such studies use the basic steps [6, 7].

1) selection of samples for clustering. In our case the sample consists of 100 individuals that were tested. Moreover, 50 people were tested in offered CTS of experts, 50 people - in training courses organized like a traditional electronic textbook.
2) Determination of a plurality of signs, on which objects in the sample will be evaluated. For traditional approach a direct estimate obtained for testing and the time of testing is a plurality of signs; and for the offered CTS of

experts sets of signs are: direct assessment - DA; indirect assessment - IA; number of returns to a host in the course – N_{ret}; the way of training entity in ATC – W; reading time of a standard text $t_{em\ of\ the\ text}$; time of passing of a standard test $t_{em\ of\ the\ test}$; reading time of n-text $t_{of\ the\ text\ n}$; time of passing n-test $t_{of\ the\ test\ n}$[3].

3) calculation of values of one or another degree of similarity (closeness). When analyzing existing measures of similarity Euclidean distance was selected as it most clearly expresses the correlation dependence.
4) the use of cluster analysis method for creating groups of similar objects. For the purpose of more clear and visual presentation of analysis results of the data obtained one of the methods of hierarchical cluster analysis was selected
5) authentication of results of cluster solutions. Authentication of analysis results is the confirmation or refutation of the hypothesis.

Consider in more details the hierarchical methods, such as agglomerative cluster analysis method. An example of agglomerative clustering is hierarchical tree structure or dendrohrama (from Greek dendron - a tree), which reflects the agglomeration, merge of individual observations into a single final cluster. Dendrohrama shows the degree of proximity to certain objects of the cluster, and demonstrates graphically the sequence of their association. Dendrogram creation algorithm:

1) at the initial stage of the algorithm, each object is considered as a separate cluster;
2) matrix of distances between objects is defined (the most common method of determining distances - Euclidean distance);
3) two closer objects are united, creating a new cluster;
4) distance from this cluster to all other objects is determined and dimension of the matrix distances is reduced by 1;
5) the procedure is repeated until all the objects are not united in one cluster.

Cluster analysis allows to find different types of distances (Euclidean distance, Manhattan, Chebyshev, etc.).

Euclidean distance - is the most common type of distance, it is a geometric distance in the multidimensional space and is calculated as follows (1)[6]:

$$d(x, y) = \sqrt{\sum_i \left(x_i - y_i\right)^2}\ , \qquad (1)$$

where x, y – objects of testing in the offered CTS of experts.

So, with the help of mathematical package "Statistics" will analyze the data obtained and construct a dendrogram of distribution of training entities on clusters by test results by means of offered CTS of experts and traditional means of training courses.

Analysis of the data

Cluster analysis was carried out using mathematical package «STATISTICA» on experimental data that were collected as a result of test preparation of training entities. At the initial stage of analysis each student was considered as a separate cluster. For the traditional approach the content delivery features of the object is a direct assessment and testing time. Using offered CTS of experts there are six features of the object, such as: DA, IA, N_{rel}, W, $t_{el.\ text}$, $t_{el.\ test}$.

Since the signs of objects are measured in different measurement units, it is necessary to conduct rationing of the data by the formula (2):

$$n = \frac{x_{ij}}{x_{j\max}}, \qquad (2)$$

where x_{ij} – meaning of the feature itself; $x_{j\max}$ – the maximum value of the same feature.

Then n_{ij} – is the normalized meaning of the object feature, where i - number of features of the object, j - set of objects.

Determination of the closeness between objects are based on Euclidean metrics, as a result distances matrix was obtained (Fig. 4). With the help of distances matrix, we build dendrograms on each type of test: figures 5 and 6. As a result of the experiment it can be seen on dendrograms formed clusters of training entities according to the successful passing of the test[13-16].

	Euclidean distances (Spreadsheet3)																		
	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_10	C_11	C_12	C_13	C_14	C_15	C_16	C_17	C_18	C_19
C_1	0,00	0,03	1,21	1,23	0,57	0,57	0,57	0,59	1,03	1,03	1,27	1,27	0,91	0,93	0,97	1,00	0,75	0,69	0,71
C_2	0,03	0,00	1,19	1,21	0,57	0,57	0,57	0,59	1,02	1,03	1,27	1,26	0,91	0,93	0,97	0,99	0,74	0,68	0,69
C_3	1,21	1,19	0,00	0,07	1,15	1,13	1,11	1,12	0,81	0,77	1,29	1,27	0,99	0,93	0,89	0,89	1,05	1,08	1,03
C_4	1,23	1,21	0,07	0,00	1,16	1,15	1,12	1,13	0,80	0,76	1,30	1,28	1,01	0,95	0,88	0,87	1,05	1,08	1,04
C_5	0,57	0,57	1,15	1,16	0,00	0,03	0,05	0,09	0,64	0,65	0,72	0,72	0,51	0,52	0,62	0,65	0,54	0,46	0,49
C_6	0,57	0,57	1,13	1,15	0,03	0,00	0,05	0,06	0,63	0,63	0,72	0,72	0,51	0,51	0,59	0,63	0,51	0,44	0,48
C_7	0,57	0,57	1,11	1,12	0,05	0,05	0,00	0,08	0,62	0,63	0,72	0,71	0,50	0,51	0,60	0,63	0,50	0,43	0,45
C_8	0,59	0,59	1,12	1,13	0,09	0,06	0,08	0,00	0,60	0,61	0,71	0,71	0,52	0,51	0,54	0,58	0,48	0,40	0,45
C_9	1,03	1,02	0,81	0,80	0,64	0,63	0,62	0,60	0,00	0,06	0,57	0,58	0,39	0,27	0,32	0,34	0,76	0,75	0,77
C_10	1,03	1,03	0,77	0,76	0,65	0,63	0,63	0,61	0,06	0,00	0,58	0,58	0,39	0,28	0,34	0,36	0,75	0,75	0,76
C_11	1,27	1,27	1,29	1,30	0,72	0,72	0,72	0,71	0,57	0,58	0,00	0,09	0,52	0,50	0,70	0,72	0,97	0,94	0,96
C_12	1,27	1,26	1,27	1,28	0,72	0,72	0,71	0,71	0,58	0,58	0,09	0,00	0,51	0,50	0,73	0,76	0,97	0,94	0,95
C_13	0,91	0,91	0,99	1,01	0,51	0,51	0,50	0,52	0,39	0,39	0,52	0,51	0,00	0,12	0,62	0,66	0,88	0,84	0,85
C_14	0,93	0,93	0,93	0,95	0,52	0,51	0,51	0,51	0,27	0,28	0,50	0,50	0,12	0,00	0,51	0,54	0,83	0,79	0,81
C_15	0,97	0,97	0,89	0,88	0,62	0,59	0,60	0,54	0,32	0,34	0,70	0,73	0,62	0,51	0,00	0,04	0,55	0,56	0,60
C_16	1,00	0,99	0,89	0,87	0,65	0,63	0,63	0,58	0,34	0,36	0,72	0,76	0,66	0,54	0,04	0,00	0,55	0,58	0,61
C_17	0,75	0,74	1,05	1,05	0,54	0,51	0,50	0,48	0,76	0,75	0,97	0,97	0,88	0,83	0,55	0,55	0,00	0,09	0,11
C_18	0,69	0,68	1,08	1,08	0,46	0,44	0,43	0,40	0,75	0,75	0,94	0,94	0,84	0,79	0,56	0,58	0,09	0,00	0,10
C_19	0,71	0,69	1,03	1,04	0,49	0,48	0,45	0,45	0,77	0,76	0,96	0,95	0,85	0,81	0,60	0,61	0,11	0,10	0,00

Figure 4. Matrix of Euclidean distances on the results of testing by means of authorship CTS of experts.

In particular, in Figure 5 we can distinguish a cluster of training entities - 12 people who passed test as "excellent":C3, C49–C50, C27, C32–C39; a cluster of training entities – 38 people with "good" mark ("good" 14 people – C28–C31, C40–C41, C4, C11–C12, C24–C26, C47–C48 and "satisfactory" 24 people – C1–C2, C5–C10, C12–C23, C43–C46). A key result is that an underachieving cluster of training entities has disappeared.

Having examined dendrogram of test results by means of traditional CTS of experts (Figure 6), we can conclude that a cluster of training entities - 11 people who did not make it, is present (C26–C29, C41–C43, C45, C38, C12, C49), and cluster of "excellent" students (in this case cluster consists of seven people– C24, C20, C15, C7, C3, C4, C36) is twice smaller in number compared to the results presented in Figure 5. Cluster of "good" students consists of 32 people ("good" mark received 14 people – C8–C9, C13–C14, C35, C18–C19, C21–C23, C46, C32 and 18 training entities received "satisfactory"– C1–C2, C5–C6, C37, C44, C10–C11, C34, C47, C16–C17, C25, C29–C31, C33, C39).

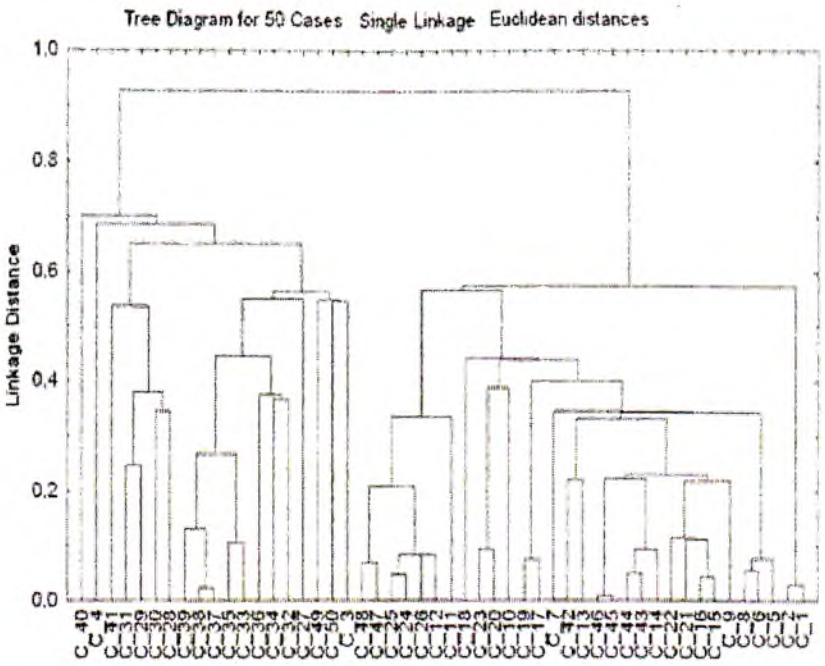

Figure 5. Figure captions are indented 5 spaces and justified. If you are familiar with Word styles, you can insert a field code called Seq figure which automatically numbers your figures.

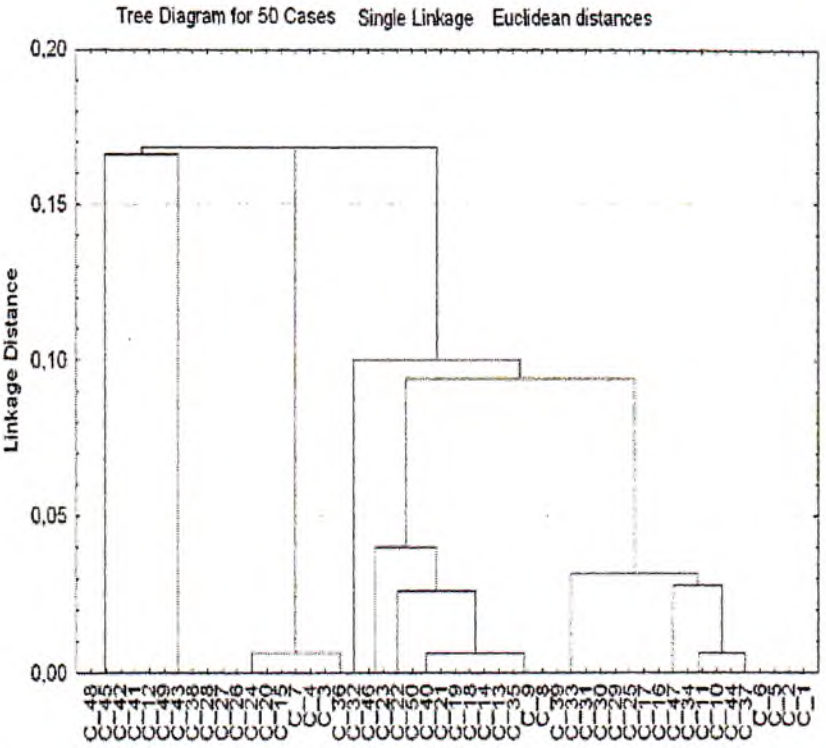

Figure 6. Figure captions are indented 5 spaces and justified. If you are familiar with Word styles, you can insert a field code called Seq figure which automatically numbers your figures.

Thus, the hypothesis is confirmed, because on dendrogram by the test results by means of offered CTS of experts cluster of training entities that made testing for "excellent", changed from 14%, at a traditional approach to 28% by means of CTS of experts, as well as the in CTS of experts cluster of training entities that don't make progress has disappeared.

5. CONCLUSION

In this paper, an experiment to identify the level of quality of content delivery in the author technology of support of information processes of content deliver in CTS of experts which performs individual processes of collecting, analyzing, processing, storage and use of data for personalized content delivery to training entity.

As a result of the experiment, it was found that the delivery of content to the CTS of experts is the best (compared to similar in content courses in which content is delivered without an individual approach) because based on the results the experiment Quality Score of education has changed from 14% in the traditional approach to 28 % by means of CTS of specialists and at the same time in the CTS of experts underachieving cluster of training entities has disappeared, it means that absolute success is achieved.

REFERENCES

[1] Kondratenko, N. R., Troianovska, T. I. and Slobodianiuk, J.,"Research of perception teaching material features by remote course students on cluster analysis base," Visnyk Khmelnytskoho Natsionalnoho Universytetu 4, 249–253 (2009).

[2] Troianovska, T. I., "The processing method of CTS subject preparation individual descriptions data," Visnyk VPI 4, 140–146 (2013).

[3] Troianovska, T. I., "Development of computer subsystem analysis and formation of object-oriented dominant students of distance learning," Visnyk Cherkaskoho Derzhavnoho Tekhnolohichnoho Universytetu, 3, 41–46 (2007).

[4] Horokhovskyi, O. I., Troianovska, T. I. and Trukhmanov, V. O., "Computer program 'Adaptive distance learning'," Certificate of registration of copyright № 26330. - K. State Department of Intellectual Property of Ukraine. - Registration date 22.07.2008.

[5] Mandel, I. D., [Cluster analysis], Finance and Statistics, 176 (1988).

[6] Duran, B. S. and Odell, P. L, [Cluster analysis], Statistics, 128 (1977).

[7] Schlesinger, M. and Glavach, V., [Ten lectures on statistical and structural pattern recognition], Naukova Dumka, Kiev, 67-134 (2004).

[8] Vassilenko, V., Valtchev, S., Teixeira, J. P. and Pavlov. S., "Energy harvesting: an interesting topic for education programs in engineering specialities," Internet, Education, Science (IES-2016), 149-156 (2016).

[9] Shadrach, J. R., [Energy Scavenging for Wireless Sensor Nodes with a Focus on Vibration to Electricity Conversion] Ph.D thesis, University Of California, Berkeley, USA, (2003).

[10] Mateu, L. and Moll, F., "Review of Energy Harvesting Techniques and Applications for Microelectronics", Proc. of SPIE Vol. 5837 VLSI Circuits and Systems II, 359-373 (2005).

[11] T. Starner, and J. Paradiso, "Human generated power for mobile electronics", C.(ed), Low - Power Electronics, CRC Press, 825-894 (2004)

[12] Moll, F. and Rubio, A., "An approach to the analysis of wearable body - powered systems," Proceedings MIXDES 2000, Gdynia,145-161 (2000).

[13] Angrist, S. W., [Direct energy conversion], Allyn and Bacon, Inc. Boston (1971).

[14] Kostishyn, S., Tymchyk, S., Vyrozyb, S., Zlepko, A. and Pavlov, V., "Design features of automated diagnostic systems for family medicine," Proc. of 13th International Conference on Modern Problems of Radio Engineering. Telecommunications and Computer Science (TCSET), 201-216 (2016).

[15] Ławicki, T., „Odszumianie obrazów CT za pomocą transformaty curvelet," Electrical Review 2, 222-224 (2015).

[16] Sawicki D., "Wykorzystanie GPU do określenia obszaru płomienia w wizyjnym systemie diagnostycznym,", Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska 1, 80-85 (2015).