Mining Methods for Adaptation Metrics in E-Learning

Igor Shubin [1[0000-0002-1073-023X]], Victoria Skovorodnikova [2[0000-0003-0436-4197]], Andrii Kozyriev [3[0000-0001-6383-5222]], Mariia Pitiukova [4[0000-0002-0978-2622]]

Department of Software Engineering
Kharkiv National University of Radioelectronics,
Kharkiv, Ukraine
igor.shubin@nure.ua, viktoriia.skovorodnikova@nure.ua,
andrii.kozyriev@nure.ua, mariia.pitiukova@nure.ua

Abstract. In this paper reviews methods of adapt implementation, available metrics and methods of determining their suitability for implementation in systems, which analyze the abilities of users in the process of their learning and adjust to the level of their knowledge. The application of neural networks in the creation of adaptive learning systems, the method of cluster analysis as well as the metrics that can be used to create their own methods for analyzing the characteristics of users of such systems in the process of adaptation are analyzed. Using of relational networks in the process of adaptation is also considered in the paper.

Keywords : Adaptation, Metrics, E-learning, Learning Systems, Relational Networks; Algebra of Predicate Operations

1 Introduction

People who want to improve their skills or get new ones are increasingly turning to individual learning. Instead of visiting courses, in the era of electronic technology, people prefer computerized learning systems. With their help, they can not only get the same knowledge as quickly as possible, but they can do it at their own convenient schedule. Among other benefits of this type of training is also the ability of individual choice, what the user will learn, which topics of a particular course he is interested in, and the most important is a large selection of such systems. So, to find the course right now, designed for people with the appropriate level of knowledge and for an acceptable price is not a problem.

However, if the user faces a topic that was not discovered enough to understand it, he has no options other than self-searching information. This can lead to further misunderstanding because of too specific topics explanation in other sources. Web systems offer online consultants and personal mentors to help to find the right information, but their help is mostly invaluable, and students do not often use it. Also, the major problem of such educational systems is their knowledge verification system, which is usually held in the form of test. In case if further training depends on successful passing of the test, the lack of knowledge forces the user to choose a random answer or to search it on the Internet which can cause waste of time and psychological discomfort. This problem is one of the main drawbacks of the knowledge test system.

It would be possible to solve this problem by using a dynamic learning system that would, for example, not stop the learning process on a specific topic until the user passes the control, and carefully followed the user's answers and memorized the question of which subtopics he encountered. Then, using these data, made the following topic, which would include both a new material and an additional problem-specific subtopic.

There is a large number of learning systems that are in demand among modern professionals who are looking for ways to improve their skills or gain new skills in their new industry. For example, Catalan Open University, National Open University "Intuit", etc are a case in point. But they all combine the same problems that were described above. People have different starting levels of knowledge, purpose and approach to learning.

Each of the described systems is integral, linear, intended for a certain audience, educational material is created by the teacher with his vision of the material's presentation. Therefore, when we have a question, we can neither ask it nor find an answer independently in the case of linear submission of information and invariability of the training program.

Of course, programs cannot replace people yet, but we can create a system that would dynamically adjust to the behavior of the student; would provide information about the subtopic with which one had a problem; would provide additional material and "re-frame" the issue as needed.

Such a program should understand the specifics of the learning process of each user.

2 The Solution of the Problem

2.1 Analysis of the Results of Scientific Developments in the Field of Adaptation of Modern Teaching Technologies

Modern leading scientific developments in the field of computerized education are focused on the application of the whole power of modern high technology for the intellectualization of educational systems. Scientists use the modern level of knowledge in the field of artificial intelligence (AI), the attempts to use the already created AI-technologies are intensifying, and also to create new special ones for the educational sphere.

Intelligent automated learning systems (ALS) are currently under development [1].

In the first generation systems of (selective systems) the creation of the training course is to transform the knowledge of the expert (the course author) on the domain and the teaching method into a special form, which has some features of adaptability and is suitable for use in the process of communication with a student.

Second generation educational systems, being intelligent hypermedia learning systems (HLS), are not universal in nature, compared to as selective ones and are usually created for a narrow domain (SCHOLAR, WHY, WEST, SOPHIE). These systems are intelligent HLS with a narrow object orientation and are poorly adapted to function in other domains. Some adaptation options are offered by the GUIDON system, which provides rich opportunities for dialogue with the learner giving him an explanation, but has a small selection of training strategies.

The main way of creating systems of the second generation were authoring languages (AL) and authoring systems (AS). The number of known AL items exceeds 100 items and requires a certain level of teacher knowledge in programming.

In the area of intelligent HLS, there is an urgent need creating ASs that generate courses on knowledge of domain, listeners and learning process. The HLS tools can serve as a tool for the accumulation of knowledge and methods of computer technology learning and promote the creation of integrated knowledge bases, which take into account interdisciplinary connections. A generalized structural scheme of modern HLS is presented in Fig. 1.

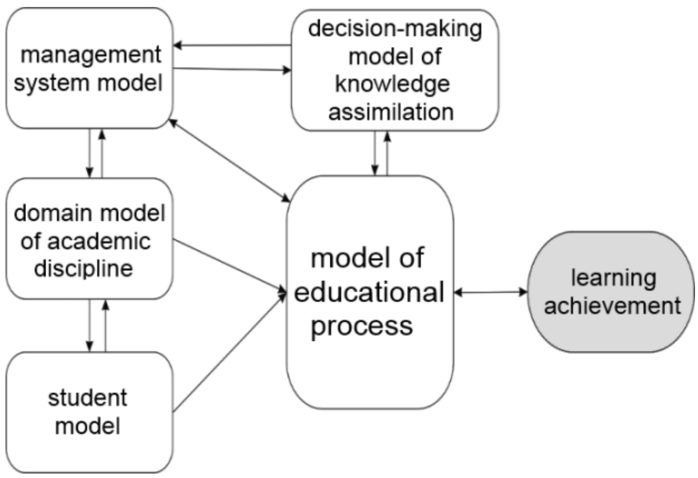

Fig. 1. A generalized structural scheme of modern HLS.

Adaptive and Intelligent Educational Internet Systems (AIEIS) provide an alternative to the traditional "just present it to the Internet" approach in developing educational software. AIEIS try to be more adaptable by building a model of goals, benefits and knowledge for each individual student, using this model during interaction with the student in order to adapt to his needs. They also try to be more intelligent by combining and performing certain activities that are traditionally performed by a human teacher, such as instructing students, or verifying their misunderstanding.

The type of advanced educational Internet systems is most often called adaptive educational or intelligent educational Internet systems. Speaking about adaptive systems, we emphasize that these systems try to be different for different students and groups of students by adding to the account information that is accumulated in an individual or group model of students. Speaking of intelligent systems, we emphasize that such systems use technology from the field of artificial intelligence to provide users with a wider and better support. At the same time, many systems can be classified as both intelligent and adaptive at the same time, but a significant number of systems fall into one of these categories. For example, many intelligent diagnostic systems, including German Tutor and SQL-tutor, are non-adaptive, that is, they will provide the same evaluation in response to the same problem solution, regardless of the past experience of the student with the system. On the other hand, a large number of adaptive hypermedia and adaptive information filtering systems, such as AHA or WEBCOBALT use efficient but very simple technologies that can be classified as "intelligent" hardly. The intersection between adaptive and intelligent systems is still great (Fig. 2). The boundaries between "intelligent" and "non-intelligent" systems are still unclear, and both groups are no doubt the subject of the community's interest in the field of artificial intelligence in education.

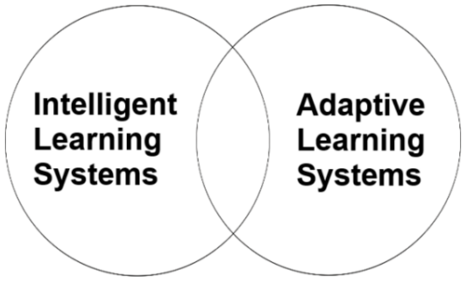

Fig. 2. The relationship between adaptive and intelligent learning systems.

2.2 Analysis of Works from Educational Internet Systems

The purpose of the technology of adaptive navigation support is to help the stu-dent navigate and move in the hyperspace by changing the look of the visible links. The support of adaptive navigation shares the same purpose as the programming of the course – to help the student find the best way through the educational material. At the same time, support for adaptive navigation leaves the student the opportunity to independently choose the next element of knowledge to study, the next task to solve. In the context of WWW, where hypermedia is a basic organizational paradigm, adaptive navigation support is both natural and effective. It was among the three top-notch AIEIS technologies used in systems such as ELM-ART, Interbook, and has become perhaps the most popular technology in AIEIS. KBSHYPERBO, Activemath and ELM-ART demonstrate several options for adaptive annotation (commenting) links. MLTUTOR uses sorting and generating links.

Adaptive information filtering (AIF) is a classical technology from the field of information retrieval. Its purpose is to find several items that are user-friendly, in a large amount of (text) documents. On the Internet, this technology has been used both in the search context and in the context of viewing. Although the mechanisms used in AIF systems are very different from adaptive hypermedia mechanisms, adaptive navigation support techniques are most often used at the interface for AIF for the Internet. There are two fundamentally different types of AIF mechanisms that can be considered as two different AIF technologies – content filtering and compatible filtering.

The first one is based on the contents of the document, while the latter completely ignores the content, trying instead to pick up users who will be interested in the same documents. Modern AIF technology is widely used in machine learning technology, especially for content-based filtering. Being very popular in the field of information systems, AIF has not been used in the educational context of the past. The volume of educational content was relatively small, and the need to direct the user to the most appropriate material was easily supported by adaptive programming (adaptation planning) and adaptive hypermedia. However, the Internet, with its large number of non-indexed open educational resources, has made AIF-technology very attractive for educators. MLTUTOR represents one of the first interesting examples of filtering information based on content in the training.

3 Technologies That Make Adaptation Possible

New research in the field of artificial intelligence gives us more opportunities for designing adaptive learning systems. Thanks to the research of artificial intelligence, today we have close to the biological technology of self-study, analysis and decision-making. To develop an adaptive learning system, you can use neural networks, cluster analysis, or develop your own method, based on one of the existing metrics to analyze the level of knowledge of the user system, or to create their own system entirely.

3.1 Neural Networks

An artificial neural network is a mathematical model, as well as its software and hardware implementation, built on the principle of functioning of biological neural networks – networks of nerve cells of a living body. This concept arose in the study of processes that occur in the brain, and when trying to simulate these processes.

Neural networks show themselves as a very effective tool for predicting human behavior. The best results are achieved by working with a large number of real people who have the ability to correct the prediction of this network if it turned out to be wrong.

Neural networks are now widely used by a variety of online services. They are used not only as entertainment, as with the AutoDraw tool, which tries to guess what the user is drawing now; Deep Dream, working with images and editing them based on their algorithms of recognition; but also, in tools such as Google Translate Google Translate [3].

But, because of the need to constantly gain new experience to improve their algorithms, neural networks are best suited for online services that work with a large number of people. Then, the neural network with each new user will be more successful in guessing the reasons for the wrong answer to specific questions in the control and will better help to get rid of this problem.

Artificial neural networks are not programmed in the usual sense of the word, they are learning. Ability to learn is one of the main advantages of neural networks to traditional algorithms. Technically, learning is to find the coefficients of the relationship between neurons. In the learning process, the neural network is able to detect complex interdependencies between incoming data and output, as well as generalize. For the learning process, you must have an external environment model in which the neural network operates.

There are three general paradigms of learning: "with a teacher", "without a teacher" (self-learning) and mixed.

3.2 Relational Networks

One of the most effective ways to solve the adaptation task is application of relational networks.

Thinking or conscious human activity is a time-consuming activity of a certain material system. It is very likely that the role of this system (or at least some part of it) is played by the human brain. This system solves some equations. Thinking is the process of solving equations, in the role of the solutions found, i.e. the products of the activity of the system are the thoughts formed by it. If we turn to the external world, we find that everything in it occurs according to the laws of nature, which have the form of equations. Any physical process from the mathematical point of view is the process of solving these equations. For example, the trajectories of the motion of the planets of the solar system are nothing but the result of solving the equations of celestial mechanics.

In turn, the thinking (conscious) system solves the equations of the algebra of predicates. They have the ultimate commonality, since they are able to express any relationship.

Predicates are the basic mathematical tool intended for the formal description of objects of the bionics of the intellect. Language of algebra of predicates is a universal means of the formal description of any mechanisms of human intelligence and machines. Developers who design artificial intelligence tools use predicate algebra for the initial formal description of models. The next stage is the algebra of predicate operations, on which any actions on relations are expressed. Relations express the properties of objects and the connections between them. They are universal means of formal description of any objects [4][5].

What kind of tool does the thinking system solve the equations of the algebra of predicates by? It can only be a mechanism that materializes the algebra of predicate operations, since it is the algebra that can formally describe any actions on arbitrary relations. The mechanism solving the equations of predicate algebra is called the relational network. This name is motivated by the fact that, firstly, the human brain realizes a neural network; secondly, from a psychological point of view, the mechanism of thinking is presented as an associative network; thirdly, from a mathematical point of view, the mechanism of thinking appears as a device for processing relations. The relational network consists of poles and branches connecting the poles. Each pole has its own subject variable xi with the domain of definition A_i $(i = \overline{1, m})$. Pair of poles x_i and x_j ,connecting by the branch $K(x_i, x_j)$, realize a linear logical operator of the first kind

$$\exists x_i \in A_i(P_i(x_i)K_{ij}(x_i, x_j)) = Qj_{j\ max} \quad (1)$$

or the second kind

$$\forall x_i \in A_i(P_i(x_i) \supset K_{ij}(x_i, x_j)) = Qj_{j\ min} \quad (2)$$

Network is called the first kind, if only the first-order operators act in it. Similarly, networks of the second kind are defined. If the network uses operators of both types, the network is called combined. The network searches for the solution of equation

$$K(x_1, x_2, \dots, x_m) = 1 \quad (3)$$

under the constraints imposed on the region of change of variables x_i $(i = \overline{1,m})$ $x_i \in P_i$, where $P_i \subseteq A_i$. If the solution of the equation is sought under more complex constraints

$$L(x_1, x_2, \dots, x_m) = 1 \quad (4)$$

then the network is completed in such a way that it corresponds to equation $K' = 1$, where $K' = K \wedge L$. Building a network that implements a predicate K, predates the binarization of the predicate K, so presenting it in the form

$$K(x_1, x_2, \dots, x_m) = \bigwedge_{\substack{i=1 \\ j=1 \\ i \neq j}}^{m} K_{ij}(x_i, x_j) \quad (5)$$

which can be produced in various ways.

The solution of equation (3) by the network is carried out in cycles. During each clock cycle, all linear logical network operators. In a network of the first kind, after each measure, the intersection of all sets is carried $Q_{j\,\max}$, converging from all sides to each of the poles x_j. In a network of the second kind of set $Q_{j\,\min}$, on the contrary unite. A network of the first kind can form unnecessary solutions, and the second – may not find some of the actual solutions. In the process of solving equation (3) with the increase in the number of the operation cycle of the network of the set $Q_{j\,\max}$ and $Q_{j\,\min}$ converge, always $Q_{j\,\min} \le Q_{j\,\max}$, at some point the approach of sets $Q_{j\,\min}$ and $Q_{j\,\max}$ ceases. If this is achieved simultaneously at all poles, then the process of solving equation (3) ends here. If at all $j = \overline{1,m}$ turns out, that $Q_{j\,\min} =$ $Q_{j\,\max}$, this means, that the network found all the solutions, not missing a single one, and did not include any mistaken solution. If this equality is not achieved at the end of the network, it means that the network has not worked well. This feature can be used in assessing the degree of goodness of the method of network synthesis, in particular – the method of binarization of a predicate K (3). It is important to note that there are such methods of network synthesis that ensure its perfect work in solving any equation of the form (3).

3.3 Cluster Analysis

Cluster analysis is the task of partitioning a given sample of objects (situations) into a subset called clusters, so that each cluster consists of similar objects, and objects of different clusters differ significantly. The clustering task refers to statistical processing, as well as to a broad class of learning tasks without a teacher.

Cluster analysis is a multidimensional statistical procedure that collects data containing sample selection information and then arranges objects in relatively homogeneous groups-clusters (Q-clustering, or Q-technique, proper cluster analysis).

The main aim of the cluster analysis is to find groups of similar objects in the sample. The best result clustered analysis will show only if it is used in online services, as in this case a large number of users are required for more accurate analysis.

This method can also be used in the local program, which will be provided with the training of the program only on mistakes of its specific user. However, it will complicate the description of behavior to solve problems that it will face, as learn from this program itself can no longer. That is, here it will be necessary to use the teaching paradigm "with the teacher", which can greatly affect the correctness of decision-making, which will not be taken so much an adapted program as the programmer who writes it.

3.4 Metrics

If the use of the previous methods is not an online service, then there will be disadvantages. They can be fixed by writing a method that is suitable for all types of adaptive learning systems: online and local. To do this, you need to define the metrics that will help to develop such a method.

4 Metrics as a Technology for Adaptation

4.1 Numerical Metrics

Numerical metrics characterize the object in numerical form. If you base the analysis of the user's understanding of the topic on numerical metrics, then to achieve its goal, you can use numbers to assess the quality of knowledge of the topic as a whole.

Let P be the user's knowledge of the topic, x is the number of correct answers, y is the number of wrong ones (1). So, we can get the percentage of the topic that the user of the system understands:

$$P = \frac{x}{x+y}\% \quad (6)$$

In shortfall percent, the topic will not be considered passed and will start from the beginning until the test will not be drawn.

Such a method of analysis would be slightly different from the usual statistics collection and would yield the same result as the non-adaptive learning system.

If, for example, using the evaluation system and enter some Mi – an estimate for the answer to the question under subtopic i, then, in addition to the overall assessment for the subject ($\sum_0^i M_i$) we will be able to obtain and evaluate the knowledge of the user in a particular subtopic. This approach is a simple but already working way to implement an adaptive learning system.

It is also numerically possible to calculate how much time a user conducts studying the material and answering questions in control. So, we can get information on how complex a user considers the whole subject.

This method also provides an important part of the information about the user, but it does not make sense to use it separately, because only this information does not provide all the necessary information. For example, it is unknown why reading a theory or answering a question took so much time. Whether it was a problem with understanding a subtopic or just a user distracted from the learning process for other reasons.

4.2 Verbal Metrics

Verbal metrics characterize an object in the form of words, phrases, or just characters. If we base the analysis of the user's understanding of the topic on verbal metrics, then we can achieve this by using words or symbols to display the answer state, and we can indicate with which subtopic the user has a problem.

Let's suppose that each response to the test has the form:

$$Question(id,\ topic,\ theme,\ state) \tag{7}$$

where id is serial number in control, $topic$ – name of the subject being checked, $theme$ – the name of the subtopic to which the question relates, and $state$ – state of the answer (right, wrong).

Thus, we will have more accurate data and we will be able to identify with which subtopic there are problems. This way also provides an opportunity to make a working adaptive learning system.

Verbal one can store more information about learning or passage control, but they just do not work well.

4.3 Graphs

With graphs you can determine the level of knowledge of the user, based on the element, which will stop moving in a row.

For example, with the help of trees, one can construct a scheme by which the user, in the wrong answer to a particular subtopic, would fall into a thread with more and more easy questions in this subtopic to make sure that the user has problems with it or not. When passing through different tree leaves (nodes), based on the contents of these nodes, you can dynamically compose material for the next topic, which will be after the control.

The graphs will help not only to form varieties of control that will dynamically change during their passage, which will facilitate understanding of the gaps in the knowledge of the user not only by the program itself, but by the user, will help at the same time to evaluate its successes, since the tree leaves can have different values.

4.4 Mixed Metrics

Of course, none of the metrics described above will be able to independently simulate complex behavior for the correct work of an adaptive learning system. So, by combining all of the above metrics, you can achieve the best, most accurate result of analyzing user behavior when passing control, and thus make the adaptive learning system most user-friendly

For example, combining numerical and verbal metrics into a single mixed, we will have the opportunity to accurately determine the problem area in the understanding of the subject, as well as assess its level of knowledge of the topic as a whole, which, unlike previous methods of definition, based on only one of these metrics, will provide more information to decide whether you need to move on to the next topic, or whether to repeat it again. Also, this method will provide enough information, even in the case of a repetition of the topic, so that the system compiles the next material so that this time the user was able to pass those control points that could not go through it. And when combined with the graph system, this program will adapt to any user and thus provide him with the most necessary information.

Mixed metrics combine all the benefits of each of the metrics used and leave them deficiencies depending on how and for what purpose the metrics were mixed.

5 Implementation and Experimental Results

The results of the research were implemented in educational adaptive systems developed at the Software Engineering Department in KNURE. These systems are based on the formal representation of algorithms based on relational networks.

The purpose of the experiment was to assess the effectiveness of the developed approach to adaptive management of e-learning. Tasks of the experiment:

- To confirm the fundamental feasibility of the proposed model and teaching method using the relational networks;
- Determine the effectiveness of the proposed model with an intelligent adaptive hypermedia system in comparison with other approaches in the organization of training.

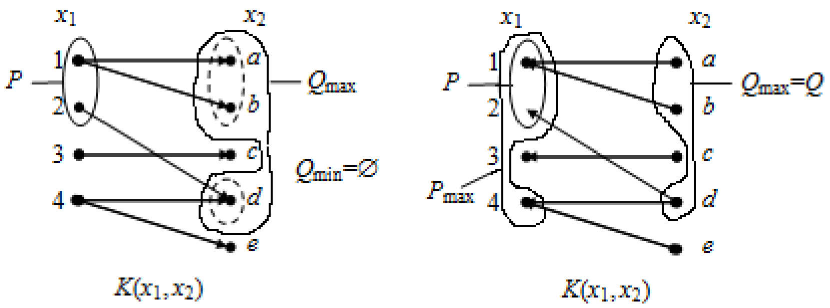

Fig. 3. An example of a schematic representation of linear logical operators in relational networks.

Within the framework of the experiment, various training methods were implemented, and based on the analysis of the results, the effectiveness of the method of conducting adaptive e-learning from relational networks was determined, which by the total rating surpassed the methodology based on the use of other types of metrics.

5.1 The Analysis of the Conceptual Schema of the Database

On the basis of the analysis of the conceptual schema of the database, the systems of the remote workshop are checked; the main features were identified – these are the tasks, users and solutions. Therefore, it was decided to allocate the following clustering objects:

- Students (users of the workshop);
- The task of the workshop;
- Couples "student - task".

In order to determine the set of clustering attributes, the attributes of selected entities were found in the database. For tasks in the database, the following attributes are stored:

1. Task identifier;
2. The limit of processor time and operational task;
3. The minimum percentage of the unique code at which the solution of the problem is considered unique;
4. The expert complexity of the task;
5. The number of users who have solved this task;
6. The number of users attempting to solve this problem;
7. The number of solutions received for this task.

Each user of the system in the database is assigned an identifier, login and password to logon, and also computes data calculated the number of tasks that the user resolved and the number of tasks that the user tried to solve. Each attempt to solve a student problem is recorded in the database, while retaining the following information:

1. Student ID and task number;
2. The date and time when the task is obtained by the verification system;
3. The compiler used;
4. Status of the attempt (correct solution or error code);
5. Characteristics of the correct solution – the time of execution of the program;
6. Request, script, the size of the used memory, the percentage of plagiarism.

After analyzing the attributes of clustering objects stored in the DB and selecting the most significant signs, for each stage of the study, a set of attributes for clusterization was identified.

For the clustering of users checking systems, the following attributes have been selected that will distinguish the student groups by the level of training:

1. The user ID;
2. Relative indicator of student's level of education;
3. The average number of attempts to solve problems;
4. The average complexity of the tasks to be solved;
5. Year of study (1-5, students of the past years are considered as one course "-1").

For clustering of the task checking in the systems the following attributes were selected, which will allow you to highlight groups of tasks in terms of complexity:

1. Task identifier;
2. Relative indicator of the degree of complexity of the task;
3. The number of unusual solutions;
4. The number of partially correct solutions;
5. The average number of attempts to solve the problem.

The following attributes were chosen for clustering the "student-task" pairs, which will determine the appropriate student's task:

- The complexity of the task;
- Relative indicator of the degree of complexity of the task;
- A relative indicator of the student's level of preparation; average complexity of tasks solved by the student;
- The relative indicator with which the task is solved;
- The number of days between the first and the last attempt of the decision.

In the database for each task an expert assessment of the complexity in points (1-200 points) was performed. But an empirical assessment of the complexity of the problem, calculated on the basis of real data on solving this problem by a large number of students, may not always coincide with the expert. Therefore, it is worth noting that in all experiments as an attribute the empirical evaluation of the complexity of the task was used.

6 Conclusions

The article analyzed the types of adaptive learning systems such as selective systems, intelligent hypermedia systems of learning, the use of author languages and authoring systems in the creation of adaptive learning systems, adaptive educational Internet systems (or intelligent educational Internet systems).

In computerized didactic teaching systems are used adapted hypermedia systems, adaptable hypermedia systems and adaptive hypermedia systems).

As a means of adapting educational systems, neural as well as relational networks and cluster analysis methods were considered that would best suit adaptive learning systems on the Internet or collect statistics from their users to develop and improve the work of analysis and decision making of these methods.

In the case of creating the own method on which the learning system's adaptation algorithm would be based, the best metric for analyzing the characteristics and state of its listeners was a mixed metric that combines numerical, verbal metrics and graphs to produce the most accurate result and best behavior of the system itself during adjustment by the user.

References

1. Kinshuk, A.: Intelligent Systems in Distance Learning (2017)
2. Kutsenko, A., Ignatchenko, E.: A Semantic-Statistical Approach to Identifying Key Phrases in the Problem of the Thematic Classification of Tests. In: Vestnik of the National Technical University "KhPI", Kharkiv : NTU "KhPI", vol. 6, pp. 149–154 (2002)
3. A Neural Network for Machine Translation, at Production Scale, https://research.googleblog.com/2016/09/a-neural-network-for-machine.html (2017)
4. Bondarenko, M., Leshchynska, I., Kruglikova, N., Rusakova, N., Shabanov-Kushnarenko, Y.: About relational networks. In: Bionics of Intelligense, vol. 3, KNURE, Kharkiv (2010)
5. Shubin, I., Kyrychenko, I., Goncharov, P., Snisar, S.: Formal representation of knowledge for infocommunication computerized training systems. In: 4th International Scientific-Practical Conference Problems of Infocommunications. Science and Technology, PIC S&T 2017, pp. 287–291 (2017)
6. Witten, I., Eibe, F., Hall, M.: Data Mining – Practical Machine Learning and Techniques. Morgan Kauffmann (2011)
7. Ramaswami, M., Bhaskaran, R.: A Study on Feature Selection Techniques in Educational Data Mining. In: Journal of Computing, vol. 1, Issue 1, 7–11 (2009)