Improvement of system for distance learning based on dialogue by appliance of statistical analysis

Bozidar Kovacic, Igor Jugo and Vanja Slavuj
Department of Informatics
University of Rijeka
Omladinska 14, Rijeka, 51000, Croatia
Phone: (385) 51-345-064 Fax: (385) 51- 345 207 E-mail: bkovacic@inf.uniri.hr, ijugo@inf.uniri.hr, vslavuj@inf.uniri.hr

Abstract - Information technology has crucial importance in e-commerce. E-commerce includes many fields of information technology and their practical usage. The users of e-commerce need to receive a thorough and up-to-date education about the new technological solutions in order to maximize the usage of new technologies and services developed using these technologies. The education of e-commerce users is done through a dialogue-based distance learning system. The system is based on the analysis of a learner's feedback, research of the defined structure of the knowledge domain and generation of new learning materials/units. In order to help e-commerce users to better understand and learn basic and advanced e-commerce services, a specialized knowledge domain about e-commerce was defined as part of our distance learning system. A decrease in learning time leads to a more efficient usage of the services offered and increases the productivity of e-commerce. Statistical analysis, based on monitoring every activity of a user working with the learning system, allows for implementation of adaptive components designed to reduce learning time and maintain the quality level of the newly acquired knowledge. Applying such a system for distance learning enables education of many users, regardless of time and space, and reduces learning time.

I. INTRODUCTION

Information technology is of great importance to e-commerce. Dynamic development of information technology leads to improvement of the existing e-business services and to introduction of new ones. The quality of e-business, and especially e-commerce, is influenced by how well users are educated in the existing and new e-business services. Educating users can be done using a distance learning system, a method independent of time and space. Although important, these e-learning advantages do not guarantee successful learning because crucial factors for the learning process outcome are structure and organization of the learning process. Distance learning system based on dialogue can be applied to an e-commerce system as a supporting system to educate users on all aspects of a given e-commerce system as well as surrounding knowledge [1,2]. The system is based on the analysis of a learner's feedback, research of the defined structure of the knowledge domain and generation of new learning materials/units. The system continuously follows user's reactions to the knowledge material being presented, analyses current knowledge of the user and improves it. By reducing time necessary to educate a user in the domain of e-commerce, while maintaining knowledge quality level, we can improve the results of e-commerce. Statistical analysis of user's activity during the process of learning is used to get data important for evaluating current knowledge level. User's activities while searching through the learning domain, testing acquired knowledge and working with content the system is presenting, are analysed. Data obtained in that way is used by the system to guide user's learning process, in order to eliminate flaws in their knowledge and acquire new. While guiding the learning process, individual and overall results of the statistical analysis are taken into consideration. Selected data is used as the basis for developing advanced features of the system aimed at adaptive behaviour of the system towards each user. Statistical analysis also allows an insight into user's current knowledge and learning styles, as well as guiding the process of learning.

Basic principles of Distance learning system based on dialogue are described in Section II. Section III presents improvements to the system by using statistical analysis. Section IV shows the implementation of improvements based on statistical analysis into the current system. Future work is shown in Section V. Conclusion in Section VI is followed by the reference list.

II. BASIC PRINCIPLES OF A DISTANCE LEARNING SYSTEM BASED ON DIALOGUE

Core elements of the distance learning system based on dialogue are:

- Knowledge base
- Delivery system
- Learner (student)
- Evaluation system
- Database system
- Control system.

Control system uses queries to search the knowledge base and select learning materials containing information to be presented to the user via the delivery system. Multimedia learning materials are presented to the learner who generates a response in the form of certain behaviour (answering a question, selecting subsequent learning materials, etc.). Evaluation system analyses user's reaction, sends the results to the control system and stores them in a database used to archive learning processes data [3]. Control system uses results of the analysis of the current learning process, as well as former learning processes, in order to form a new query for selecting content from the knowledge base that allows learning process to continue. In that way, control system analyses the current learning process and guides it, depending on user's reactions and previous learning experiences of other users.

Knowledge base consists of a group of concepts especially defined for a particular problem domain, and

their mutual relations [4]. The structure of concepts is mainly hierarchic, where simpler concepts belong to lower levels of hierarchy and complex concepts belong to higher levels (Fig. 1). Problem domain and relations are defined by the experts in that domain.

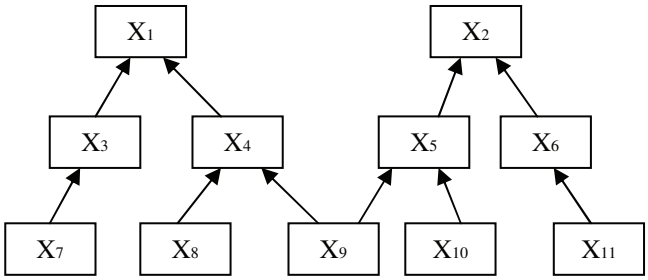

Figure 1. The scheme of a knowledge domain

By selecting a particular domain, users get an overview of all concepts in that domain and select one for learning. The system then selects learning materials for the requested concept from the knowledge base, and searches through the hierarchy of concepts of the selected domain. The system presents to the user multimedia learning materials together with a list of concepts belonging to higher and lower hierarchy levels. The user may guide her learning towards simpler or more complex concepts linked to the current learning item. This represents a way of constant guidance of the user searching through the knowledge domain. The user can test her knowledge, acquired during the learning process, by opting for questions module that belongs to the selected concept. User's answer is then evaluated as either correct or incorrect, and the information about the question, its timestamp and answer accuracy are stored in the database. Analysis of the stored data on current and former learning processes enables system to perform monitoring and make corrections to the current learning process. Besides this systematic approach to studying knowledge domain, which is based on searching through a hierarchy of concepts, the system gives users an insight into their current level of acquired knowledge for a specific domain [5].

III. IMPROVEMENT OF THE SYSTEM BY STATISTICAL ANALYSIS

We can make learning process better and shorten learning time, if we increase the quality of following user progress. This is based on statistical analysis of data collected in the process of learning (choice of concepts, navigation through the concept hierarchy, choice of questions, accuracy of answers, etc.) The results of static analysis depend on the quality of collected data. For the purpose of higher quality data tracking during learning, we have implemented the following modules:

- Administrative module
- Testing module

Administrative module is used for adding new and updating current users. Improvement of the administrative module aims at organising domains to be used for class purposes, as well as groups of users for each domain. Creating new domains or subdomains, and organising users into groups, allows us to analyse data for each domain or subdomain separately, taking into account only targeted groups of users. Administrative module offers the following functions:

- Adding new and editing current learning domains
- Assigning experts to learning domains
- Creating new and editing current groups
- Linking system users to current groups and updating groups.

Testing module is intended for estimating learner knowledge, and the choice of test questions is random. Improvement of the module includes different ways of generating test questions. Testing module allows us to generate questions according to the following criteria:

- Check all the questions that belong to a single concept
- Choose previously incorrectly answered question
- Choose most frequently asked question
- Random.

Checking all the questions that belong to a single concept criterion is based on selecting previously unanswered questions, whether in the current or any other previous learning session of that concept. If there are no previously unanswered questions, one will be chosen randomly. Choosing previously incorrectly answered questions criterion selects questions for which the user has not yet given a correct answer. This is aimed at repetition of what has been learned. If there are no such questions, one will be chosen randomly. Choosing most frequently asked question criterion singles out those questions answered the largest number of times (by all users) in the process of learning. Random criterion selects a question by chance. One criterion is chosen for each domain. Every concept within a domain inherits that domain's criterion for selecting a question, with the possibility of changing that criterion. Selecting a question for the current concept is based on the criterion defined for that learning domain. Then, the system checks whether there is a different criterion set for the current concept. If the defined selection criterion fails to find a question, one is chosen randomly.

By using different selection criteria we change not only the way test questions are selected, but also our approach to testing user's knowledge. Further control over how learning domain is tested is done by defining different criteria for each concept in relation to the domain. Domain expert defines selection criteria for each class domain as well as concepts belonging to it.

Improvements of administrative and testing modules enable groups of users to learn groups of concepts within a particular domain or subdomain, by modifying approaches to knowledge testing. Statistical analysis over data collected during learning is now able to provide us with more precise and thorough results of the level of knowledge for all users. Thus, the system improves user tracking while learning concepts within a domain, and better guides the process of learning. Statistical analysis aims at collecting data while:

- Choosing concepts, i.e. navigating within the hierarchy of concepts
- Selecting test questions
- Answering selected questions.

Monitoring how users choose concepts refers to the learning frequency for each user and group the user belongs to. Also, starting time and duration of learning

session for a particular concept, are monitored. Users can see data about the number of learning sessions within a domain. When selecting concepts for learning, the system suggests users to learn concepts that are most frequently chosen by that particular group of users. When selecting test questions, the system registers the question, selection criterion, starting time and time necessary for the user to respond. Monitoring user's answers deals with the evaluation of their answers, number of answers and the overall achievement for a particular concept. The results of statistical analysis are available to users, who are given an overview of their current level of knowledge for a particular domain. Domain experts are able to see the results that refer to the knowledge level of each user and group of users. Using that data, experts get useful feedback about the learning process, so they are able to modify the structure of the knowledge domain (adding new concepts, changing the hierarchy of concepts, introducing new test questions, etc.). By examining the results of the statistical analysis, both users and experts get feedback on user learning process and their level of knowledge. Statistical analysis makes the basis for improvement aimed at better guidance of the learning process of a user or group of users. In that way, it is possible to implement additional functionalities for adaptive system behaviour towards individual users and groups.

IV Implementation of System's Improvement Based on Statistical Analysis

In order to enable data processing for specific groups and knowledge domains, we made some modifications to the application database. Now, database offers the possibility to single out users and experts as well as their grouping and linking to domains. Relational schema of this module currently has the following structure:

domain(domain_id, domain_name, *id_way_question*)
groups(group_id, group_name)
groups_domain(domain_id, group_id)
users(user_id, username,.., *category_id*)
groups_users(group_id, user_id)
categories(category_id, category_name, access)

After building this new model, making the relational schema and creating tables in the database, we have made some modifications to the PHP code of the application. Further, we added new interface for easier addition/deletion of users/domains, following a similar functionality in Moodle (MudRi).

In order to be able to manage the process of knowledge testing, based on previously collected and analysed data, we had to a) collect data and b) introduce different ways of testing. Certain modifications to the database enabled these functionalities, so now the relational schema of this model has a different structure:

question(question_id,..., *type_id*, active, compulsory)
question_type(type_id, type_name)
way_question(way_id, way_name)
concept_domains(...,*way_id*,...)
domains(...,*way_id*,...)
trajectory(trajectory_id, *user_id*, *concept_id*, *domain_id*, access_timestamp)
test_concept(test_id, *concept_id*, *user_id*, *question_id*, correct)
learning_concepts(learning_id, *concept_id*, *user_id*, *domain_id*, date)

Analiza pojmova			
Naziv pojma	Broj posjeta	Broj pitanja	Točnost učenja
Algoritam (kriptografski)	0	7	29%
Algoritam simetričnog šifriranja	4	4	25%
Asimetrični algoritam	0	0	0%
Asimetrični kripto sistem	2	3	0%
Autentifikacija	0	0	0%
Autorizacija	0	0	0%
Certifikat	4	0	0%
Certifikator	1	0	0%
Davatelj internet usluga	0	0	0%
Davatelj usluge elektroničke trgovine	3	1	0%
Dekriptiranje podataka	1	1	0%

Figure 2. Analysis of user data

This relational schema allows precise definition of how questions are generated when testing knowledge on several levels. The selection of questions is defined first by the value of attribute "compulsory". The next level of definition is the way question is generated, as defined in "concept_question" relation. If that value is not defined, there is a default, fall-back setting of the way questions are generated, which is defined at the level of domain. Additional relations allow us to collect a large set of data about the way application is used, which is further used for statistical analysis. In this phase of application development, data about user's trajectories while browsing through learning materials is also collected. Trajectories obtained in this way will be later analysed by a data mining tool. These results will be used to define recommended trajectories, as well as for adjusting knowledge structure. Our final goal is to implement a PHP module that will be able to find user behaviour patterns and suggest domain changes to the expert, who will, in turn, either accept them and make adjustments or decline them. In that way, the system will become adaptive to different user groups and knowledge structures.

Further, with each test, the system collects data about the question, answer correctness and time necessary for user to answer. Relation "test_concept" contains general data about the number of learning materials (concepts) that have been generated. We added to the application code functionalities and algorithms necessary for question selection, which are based on the preferred way of question delivery (random, incorrectly answered, most frequent, etc.).

Collected data is processed and then shown to users and experts (professors) who manage learning process for a particular domain or group. A user is able to view the following: the number of visits for each unit (concept), the number of answered questions about a particular concept and the success rate of her answers (Figure 2). An expert is able to browse through data for each user (as described above) or view the overall results for the whole domain in order to get a list of most frequently visited concepts, a list of questions and information about whether they were answered correctly or not. This is a good way to notice concepts most interesting to users, as well as those that show the highest level of acquisition or those never visited

by users. The latter can serve as a reliable indicator that it is necessary to introduce certain changes to the knowledge domain.

Figure 3. Visits to concepts - visualization (tag cloud)

In order to visualise numbers that represent visits to each concept, together with concepts for which the highest number of questions has been generated, we have implemented the, so called, tag cloud. More visited concepts are represented by bigger letters, and those less visited by smaller letters (Figure 3). The upcoming versions of this application will contain additional possibilities for monitoring users. These will be used as extra sources of data for statistical analysis and data mining. Further, we will implement live charts based on HTML5/JavaScript platform that will help in visualisation of collected and processed data.

V Future work

Further development of the learning system will aim at improving the process of monitoring user's learning and increasing the level of system adaptivity to user's learning needs. Improving the process of monitoring will be achieved by analysing user's navigation while searching through the learning domain, i.e. by analysing formative paths of learning [6]. System adaptivity aimed at user's learning needs has as its goal a complete and accurate estimate of current knowledge, and higher decision-making quality of how to guide the process of learning [7]. Further, it is important to note the knowledge level of each user and the overall level of acquisition of concepts from a particular learning domain. System adaptivity will enable a more complete search of the knowledge domain as well as better acquisition.

VI Conclusion

The users of e-commerce need to receive a thorough and up-to-date education about the new technological solutions in order to maximize the usage of new technologies and services developed using these technologies. The education of e-commerce users is done through a dialogue-based distance learning system. This paper described the improvement of a distance learning system based on dialogue, by statistical analysis. A decrease in learning time leads to a more efficient usage of the services offered and increases the productivity of e-commerce. Statistical analysis, based on monitoring every activity of a user working with the learning system, allows future implementation of adaptive components designed to reduce learning time and maintain the quality level of the newly acquired knowledge. Applying such a system for distance learning enables education of many users, regardless of time and space, and reduces learning time.

References

B. Kovacic, I. Jugo "Applying a Distance Learning System Based on Dialogue in e-commerce", *Proceedings of the International Convention MIPRO 2009*, section DE, p. 29-32, Rijeka, Croatia, 2009.

B. Kovacic, "Distance learning model based on dialogue", *Proceedings of the International Convention MIPRO 2000*, section CE, p. 29-32, Rijeka, Croatia, 2000.

C. Chou, "Constructing a Computer-Assisted Testing and Evaluation system on the Wide Wide Web – The CATES Experience", *IEEE Trans. on Education*, vol. 43, no. 3, p. 266-272,. 2000.

B. Kovacic, Z. Skocir: *"Formal Model for Distance Learning System Based on Dialogue"*, *Proceedings of the International Conference ICT2001*, vol. 1, p. 231-236, Bukurešt July 2001.

B. Kovacic, "Application of dialogue realized by distance learning system based on dialogue", *Proceedings of the International Convention MIPRO 2003*, section CE, Rijeka, Croatia, 2003.

V. Carchiolo, A. Longheu, M. Malgeri, "*Adaptive Formative Path in a Web-based Learning inviroment*", ???? izvor

E. Sangineto, N. Capuano, M. Gaeta, A. Micarelli, "*Adaptive course generation through learning styles representation*", Journal Universal Access in the Information Society archive Volume 7 Issue 1, March 2008.