Application for Diploma Works Management

Tsvetanka Georgieva-Trifonova
"St. Cyril and St. Methodius" University of Veliko Tarnovo, Bulgaria
cv.georgieva@uni-vt.bg

***Abstract*—In this paper, an information client/server system for the management of data and its extraction from a database containing information for diploma works of students is proposed. The developed system provides users the possibility of accessing information about different characteristics of the diploma works, according to their specific interests. The architecture of the client/server system is described as well as the services offered. The author presents the structure of the created database that stores the necessary information. A client application ADP (access data project) is realized, providing the possibility for insertion, updating and searching, as well as a client application is fulfilled with Java for discovering the constraint-based association rules.**

I. Introduction

The aim of the presented paper is to represent a client/server system for keeping information on diploma works of students. The realized client/server system provides users the possibility of extracting information about the developed diploma works. It allows students and teachers quick access by means of a convenient interface to the data and the files, related with the diploma works of students, graduated the bachelor or the master degree of some of the specialities in department of "Mathematics and informatics" in "St. Cyril and St. Methodius" University of Veliko Tarnovo after the year 2002. The other substantial advantage of using a similar system [18, 19, 20, 21] is that providing the electronic sources permits the students better and more complete ways to represent their developments. Besides the files with the content of the diploma works, the corresponding presentation from their defence, the multimedia files, the programs, the databases, the program source code and the others can be added.

The storage of the data about the diploma works of the students in a database makes suitable conditions for data mining [2, 12], i.e. analyzing the accumulated data with the purpose of extracting the previously unknown and potentially useful information. That motivates the utilization of a program for discovering the constraint-based association rules.

In the presented paper, the architecture of the developed client/server system is described and the services, included in its realization. The structure of the created database is represented, keeping the necessary information. A client application ADP (*access data project*) [16] is proposed, designed for insertion, editing and searching the data. Besides this a Java program is applied [5, 6], providing the possibility for discovering the constraint-based association rules.

The rest of the paper is organized as follows. Section 2 contains a description of the architecture and the interface of the client/server system. In Section 3, we represent the entity-relationship model (*ER model*) of the database that stores information about the diploma works of the students. We also produce the relational tables obtained after the transformation of the created ER model into relational. These relations are realized by using the database management system Microsoft SQL Server. Section 4 consists of the description of the client applications for updating the data and data mining.

II. Overview of the Client/Server System Architecture

The architecture of the developed client/server system is created on the base of the two-layer information model (fig. 1).

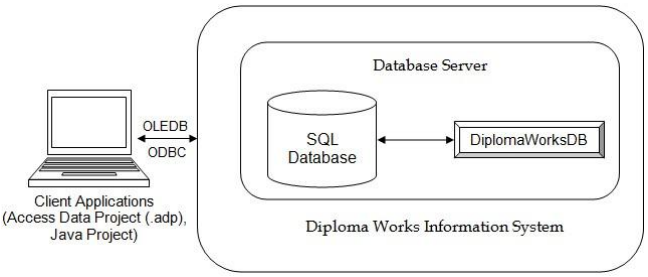

Figure 1. The architecture of the system

The layer for data processing is realized by using the database management system. For the present system we use Microsoft SQL Server, which allows efficient storage of large databases and provides functionality for accessing the data [3, 9, 10, 11, 13, 14, 15].

The client part consists of an ADP application, providing a convenient interface for insertion, updating and searching the data, as well as a Java program for mining the constraint-based association rules.

A. SQL Server database for data storage

The `DiplomaWorksDB` database serves for storage and processing the data for the diploma works, the students and their leaders. Information on the student's faculty number, student's names, the scientific and/or educational degree, the speciality, the form of training, the topic and the annotation of the diploma work, the student's leader, the reviewers, the date of the defence of the diploma work, the

mark obtained by the student for the diploma work is maintained.

For each diploma work a possibility for insertion of an additional information is provided – files (*.doc*, *.pdf*) with its content; presentation (*.ppt*, *.pdf*) of the student for the defence of the diploma work; application realized by the student (such as a database, a program of C, Java, etc.) and others.

The basic functions of the database include:

- addition of a new student in the database;
- edition of the data about the students;
- deletion of students from the database;
- browsing the data for the students;
- searching by different criteria.

B. Client application ADP for insertion and searching the data

Microsoft Access allows establishing a connection between the current database and tables from other databases of Microsoft SQL Server and other data sources. ADP is connected with a database of SQL Server and provides an access to the objects created in that database (such as tables, views, stored procedures, triggers, etc.). The data are stored in the database of SQL Server. ADP does not contain any data and tables, but it can be used for easy creation of forms, reports, macros. As a result of that, the end user features opportunity for insertion, editing, and deletion of the data by means of a comfortable interface.

C. Client application for mining the constraint-based association rules

The goal of association rules mining [1] is to find interesting associations or correlation relationships among a large dataset, i.e. to identify the sets of attributes-items, which frequently occur together and then to formulate the rules characterizing these relationships. The constraint-based association rule mining [7] aims to find all rules from given dataset, which satisfy the constraints required from the users. An application is realized with Java for discovering the constraint-based association rules, which in [17] is utilized for the data, obtained after applying the methods of digital processing of signals for analysis of the sounds of the unique Bulgarian bells. This client application is connected with the `DiplomaWorksDB` database with the purpose of performing the association analysis of the data about the diploma works of the students.

III. Modeling of the Data

The model of the `DiplomaWorksDB` database, in keeping with the entity-relationship model (*ER model*), introduced in [4], is shown in Figure 2. The objects of the ER model are depicted as rectangles, their attributes as ellipsis, and the relations as rhombs [8].

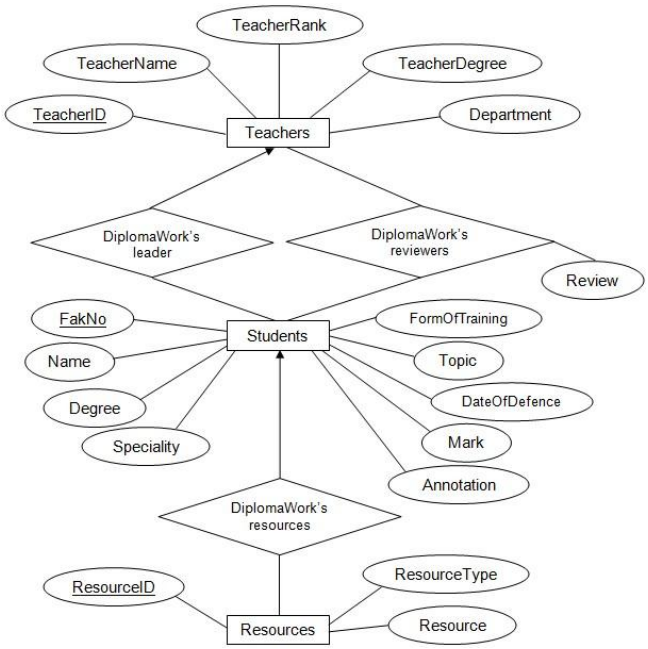

Figure 2. ER diagram of the `DiplomaWorksDB` database

The database is realized by means of the database management system Microsoft SQL Server. The relevant relational tables are shown in Figure 3.

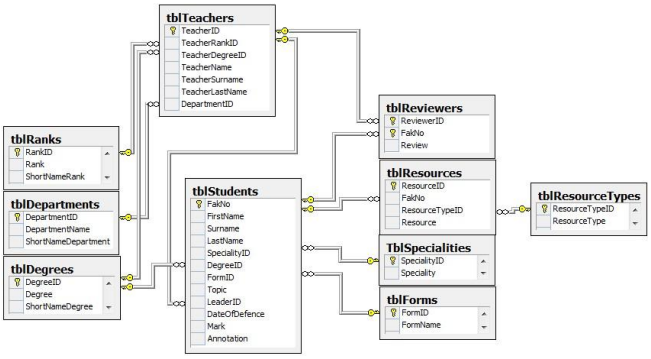

Figure 3. Relational model of the `DiplomaWorksDB` database

The structure of the database is defined to provide the best efficiency of the most frequently used operations – insertion, updating, data searching.

The `DiplomaWorksDB` database of SQL Server contains the created views for extracting the data from several related tables, as well as the stored procedures for obtaining the information on the students, defended their diploma works during a specific month and year; the students, whose diploma works' topics comprise a specific string. The stored procedures provide a better performance of the client/server system because they make decreasing the exchange to data between the client and the server. Besides the stored procedures can accept parameters and therefore they can be executed from multiple client applications by applying different input data.

IV. Client Applications for Updating, Searching and Mining the Data

A client application ADP is developed for updating and searching the data about diploma works of students, as well

as a client application realized with Java for mining the constraint-based association rules.

A. ADP client application for insertion, edition and searching the data

Forms for insertion and updating the data are realized. Their purpose is to facilitate actualization of the information. The form for insertion of the data about the students and their diploma works is shown in Figure 4.

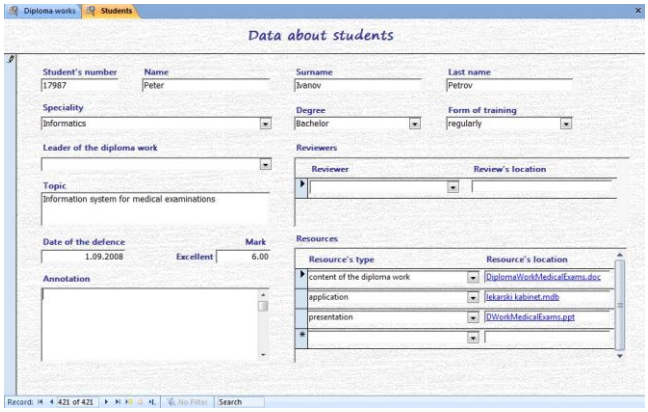

Figure 4. Form for insertion of the data about students

Besides this, the application allows the execution of different queries, which perform finding the specific information, corresponding to the given searching criteria. Each user can fulfil search by filling in text boxes and/or list boxes which correspond to the chosen characteristics of the diploma works of the students, stored in the database. The results from each query are presented in a format convenient for the end user. The forms and the reports are implemented with the record sources – views and stored procedures designed for extracting the data about:

- students from a chosen speciality;
- students with a chosen diploma work's leader;
- students, defended their diploma works during a chosen month and year;
- students, whose diploma works contain in their topics a given string.

B. Client application for discovering the constraint-based association rules

The utilized application allows the user to set constraints for searched rules and finds constraint-based association rules.

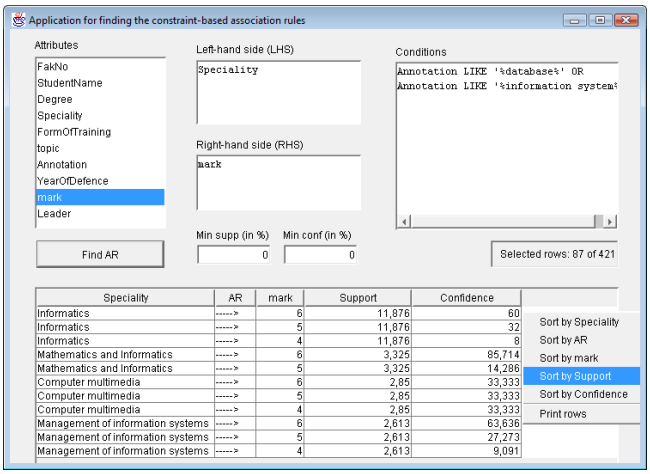

Figure 5. Discovering the constraint-based association rules

Figure 5 shows an example result from the execution of the realized program with given values of the minimal support, minimal confidence and conditions for the values of the attributes. For instance, let the following rule be generated from the database with the diploma works:

Speciality("Informatics") → *Mark*("6")

with values of the support $s = 0.11876$ and the confidence $c = 0.60$. This rule means, that for the students, graduated in the speciality "Informatics", whose diploma works are in the area of the databases and the information systems, one of the most frequent marks from the defence of their diploma works is 6 (with 60.00% confidence) and the students in Informatics with diploma works in databases and information systems represent 11.876% from all students, included in the study.

V. CONCLUSIONS

In this paper, the automated system is proposed. It explores a client/server based approach to managing the information on the diploma works of the students. The created database contains information about different characteristics of the diploma works and it is realized on Microsoft SQL Server. The interface is developed by means that allow establishing a connection with the database of the ADP project. This gives users the possibility of easily accessing detailed information about the diploma works of the students.

In addition, an application is represented which provides the possibility for finding the constraint-based association rules of the data about the diploma works.

Our future work includes development of an application for mining the constraint-based association rules in the text of the diploma works (*text mining*), which allows performing the association analysis of the different words from their contents.

REFERENCES

[1] R. Agrawal, T. Imielinski, A. Swami, "Mining Association Rules between Sets of Items in Large Databases", *In Proceedings of the ACM SIGMOD International Conference on Management of Data*, Washington, 1993, pp. 207-216.

[2] A. A. Barsegyan, M. S. Kupriyanov, V. V. Stepanenko, I. I. Holod, *Technologies for data analysis: Data Mining, Visual Mining, Text Mining, OLAP*, BHV-Peterburg, 2008 (in Russian).

[3] D. Bieniek, R. Dyess, M. Hotek, J. Loria, A. Machanic, A. Soto, A. Wiernik, *Microsoft SQL Server 2005 Implementation and Maintenance – Training Kit*, Microsoft Press, 2006.

[4] P. Chen, "The Entity-Relationship Model: Toward a Unified View of Data", *ACM Transactions on Database Systems*, Vol.1, No.1, 1976, pp. 9-36.

[5] D. J. Eck, *Introduction to Programming Using Java*, http://math.hws.edu/javanotes, 2006.

[6] B. Eckel, *Thinking in Java*, SoftPress, 2001 (in Bulgarian).

[7] Y. Fu, J. Han, "Meta-rule-guided mining of association rules in relational databases", *In Proceedings of the International Workshop on Integration of Knowledge Discovery with Deductive and Object-Oriented Databases*, 1995, pp. 39-46.

[8] H. Garcia-Molina, J. D. Ullman, J. Widom, *Database Systems: The Complete Book*, Williams, 2002 (in Russian).

[9] Tsv. Georgieva, *Databases Practicum – Part I*, Faber, Veliko Tarnovo, 2006 (in Bulgarian).

[10] M. Gruber, *Mastering SQL*, SoftPress, Sofia, 2001 (in Bulgarian).

[11] F. Houlette, *SQL: A Beginner's Guide*, SoftPress, Sofia, 2001 (in Bulgarian).

[12] M. Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*, John Wiley & Sons, 2003.

[13] D. M. Kroenke, *Database Processing*, Piter, 2003 (in Russian).

[14] Microsoft Corporation, *MCSE Training: Microsoft SQL Server 2000 – Design and Implementing Databases*, SoftPress, Sofia, 2001 (in Bulgarian).

[15] Microsoft Corporation, *Transact SQL*, http://www.microsoft.com/sql.

[16] W. Pearson, *MS Access for the Business Environment: Stored Procedures from the MS Access Client*, http://www.databasejournal.com/features/msaccess/article.php/3363511, 2004.

[17] T. Trifonov, Tsv. Georgieva, "Discovering the Constraint-based Association Rules in an Archive for Unique Bulgarian Bells", *In Proceedings of the International Conference on Information Technology and Computer Science (ITCS'2009)*, Kiev, Ukraine, 2009 (accepted).

[18] HKUST Library, *The Hong Kong University of Science and Technology Electronic Theses*, http://lbxml.ust.hk/th.

[19] UM Graduate School & Fogler Library, *Electronic Theses and Dissertations at the University of Maine*, http://www.library.umaine.edu/theses.

[20] Virginia Tech, *Electronic Theses and Dissertations at Virginia Tech*, http://scholar.lib.vt.edu/theses.

[21] WSU Libraries, *The Washington State University Dissertations Web Site*, http://www.dissertations.wsu.edu.