

ICEST 2021

II International Conference on Economic and Social Trends for Sustainability of Modern Society

**COST-EFFECTIVENESS OF USING A MULTIDIMENSIONAL
MODEL FOR SCHEDULING**

O. N. Lopateeva (a)*

*Corresponding author

(a) Reshetnev Siberian State University of Science and Technology 31, Krasnoyarsky Rabochy Av., Krasnoyarsk,
Russia, olga-lopateeva@rambler.ru

Abstract

This paper discusses the use of a multidimensional model and a method of organizing data storage in the form of a hypercube for solving all functional tasks of the "University Schedule" system. The program is ultimately a concrete formulation of an abstract algorithm based on concrete representations and data structures, which will help speed up the process of dispatchers' activities in drawing up a training schedule. This article discusses the use of the multidimensional model "hypercube" for the preparation of training sessions, shows its data structure in the form of ER-diagrams. It also describes the economic calculation with the result from the implementation of the "University Schedule" system based on a multidimensional data structure model. As you know, to create any software package, it is not enough to develop an abstract mathematical solution algorithm. It is equally important to determine the composition of the data describing the real situation and to structure this data appropriately. The purpose of this article is to increase the efficiency of the methodological department by designing an automated information system for scheduling based on a multidimensional model.

2357-1330 © 2021 Published by European Publisher.

Keywords: Information system, multidimensional model, curriculum, hypercube, relational DBMS



This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial 4.0 Unported License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

The information system "University Schedule" is a set of organizational, technical, software and information tools, united into a single system for the purpose of collecting, storing, processing and issuing the necessary information designed to perform the functions of managing the process of drawing up the curriculum of the university.

As you know, to create any software package, it is not enough to develop an abstract mathematical solution algorithm. It is equally important to determine the composition of the data describing the real situation and to structure this data appropriately.

Razumnikov (2013) notes that the effectiveness of any project is characterized by the ratio of costs and benefits in relation to the interests of its participants. When deciding on the implementation of a scheduling system, as a rule, the main criterion is not the return on investment in information technology, but the efficiency of the systems in the framework of increasing labor productivity. Nevertheless, the analysis of the financial side of the project is also not unimportant. Since the educational institution is not a commercial organization, does not sell anything, it is impractical to talk about profit. In this case, we can only talk about reducing costs, since the load on the dispatcher decreases, and his salary can be reduced, or the load can be replaced by other activities.

2. Problem Statement

The task of forming a curriculum in an educational institution belongs to the list of applied problems, the solution of which is to find a partition of a set of elements - resources, each of which has a numerical characteristic - a size, into a given number of non-overlapping subsets, whose sizes are least different or closest to each other.

One of the tasks of this class is the formation of the "schedule hypercube" database.

To form the curriculum, information of four types is needed: conditionally permanent information - "reference books", conditionally permanent information - "local algorithms for including blocks of classes in the curriculum", operational - summary information - "calendar training schedules", operational information - "operational data for the formation of the training schedule".

"Reference books" contain information entered into a computer that characterizes various objects (entities) participating in solving a range of functional tasks associated with the formation of an educational schedule. Such entities are: faculties, departments, specialties, groups, audiences, teachers, disciplines, pairs (class), types of activities.

"Calendar training schedules" store information that can be called operational - final because the composition of this information is constantly changing in the process of executing the algorithm for the formation of calendar training schedules. In the case of "manual" scheduling, information about the calendar training schedules is contained in three magazines: teacher downloads, group downloads and classroom downloads. When using electronic computers, all this information must be stored on machine media (external or internal) in the form of some kind of data structures.

"Local algorithms for the inclusion of blocks of classes in the curriculum" are programs that are stored in a computer for further use in the formation of a calendar curriculum for a particular educational institution, taking into account the requirements of sanitary norms and rules.

"Operational data for the formation of the curriculum" stores information about working curricula, teaching reviews, workloads of departments and teachers, applications of departments for inclusion in the schedule of lectures, laboratory and practical work with the wishes of teachers about the time of classroom lessons.

In his work Kashirin (2009) suggests using a multidimensional data model for interactive analytical processing. Also, a general concept of using OLAP systems (online analytical processing) for organizing data structures is proposed and formed. Among the advantages, it can be noted that OLAP allows you to work with data in terms of the subject area without knowing the information storage architecture.

Thus, the urgent task is to improve the quality of the organization of the data structure in the "University Schedule" system on the basis of a multidimensional data processing model in order to increase the efficiency of the methodological department.

3. Research Questions

Issues in the field of design and development of scheduling systems in an educational institution were considered in the scientific works of many foreign and domestic researchers: Kashirin (2009), Codd et al. (1993), Zatonsky and Varlamova (2018), Akimkina, (2018), Batishchev (2003), Zykin et al. (2016), Ilyushechkin (2019), Nosov et al. (2020), Ovsyanitskaya (2018), Popov and Lisenkova (2018), Samsonova and Simonov (2018); Khasukhadzhiev and Sibikina (2016) and others.

4. Purpose of the Study

Currently, the established methods and algorithms for the formation of the optimal formation of the curriculum are diverse.

The aim of this work is to improve cost efficiency and scheduling activities based on the application of a multivariate model. The object of the research is the Reshetnev Siberian State University of Science and Technology located in the East Siberian region, in particular in the city of Krasnoyarsk.

5. Research Methods

Currently, there are two somewhat competing, but somewhat complementary approaches to building data warehouses: an approach based on the use of a multidimensional database model (DB), and an approach using a relational database model.

Multidimensional storage model. When using a multidimensional model, data is stored not in the form of flat tables, as in relational databases, but in the form of hypercubes - ordered multidimensional arrays. That is, the multidimensional representation of data is physically implemented here. Of course, this approach requires more memory to store data, and when using it, it is difficult to modify the data structure. However, multidimensional database management systems (DBMS) provide faster search and reading of data compared to relational systems, and also eliminate the need to repeatedly join tables. The average

response time to a complex analytical query when using a multidimensional DBMS is usually 10-100 times less than in the case of a relational DBMS with a normalized structure.

When creating a multidimensional data model “University Schedule”, we use a polycubic data organization scheme. In accordance with this scheme, three hypercubes are defined with the same dimension, but with different dimensions.

The first hypercube *P* is designed to perform various operations with data about teachers. The structure of the hypercube *P* is shown in Figure 1.

The hypercube *P* has three dimensions: days, ribbons, teachers. The cells of the hypercube *P* contain a data structure called a record. For specific measurements (day *d*, tape *l*, teacher *p*), the record contains data about in which classroom, what subject, what type of classes, for which group (or groups), teacher *p* on day *d* with tape *l* conducts classes. The record is filled in during the operation of the optimization algorithm for forming the schedule. Vorobovich and Lopateeva (2006) proposes the following record structure.

```
Type
Element_pdl = RECORD { hypercube cell structure P }
a: Integer; { audience code }
g: Array [1..G] of Integer; { group codes }
z: Integer; { discipline code }
v: Integer; { occupation code }
END;
```

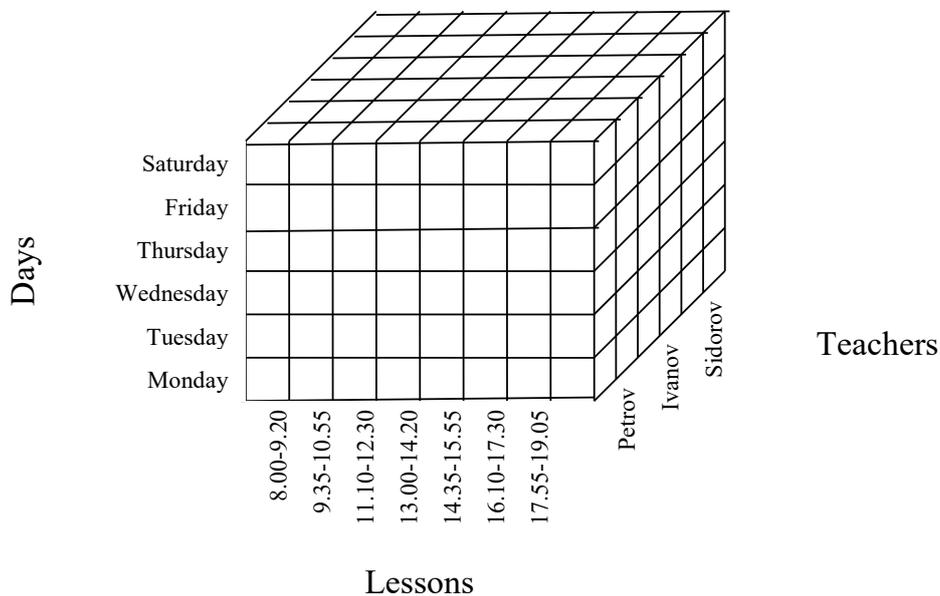


Figure 1. Hypercube "Teachers"

The second hypercube *G* is designed to perform various operations with data on student groups. The structure of the hypercube *G* is shown in Figure 2.

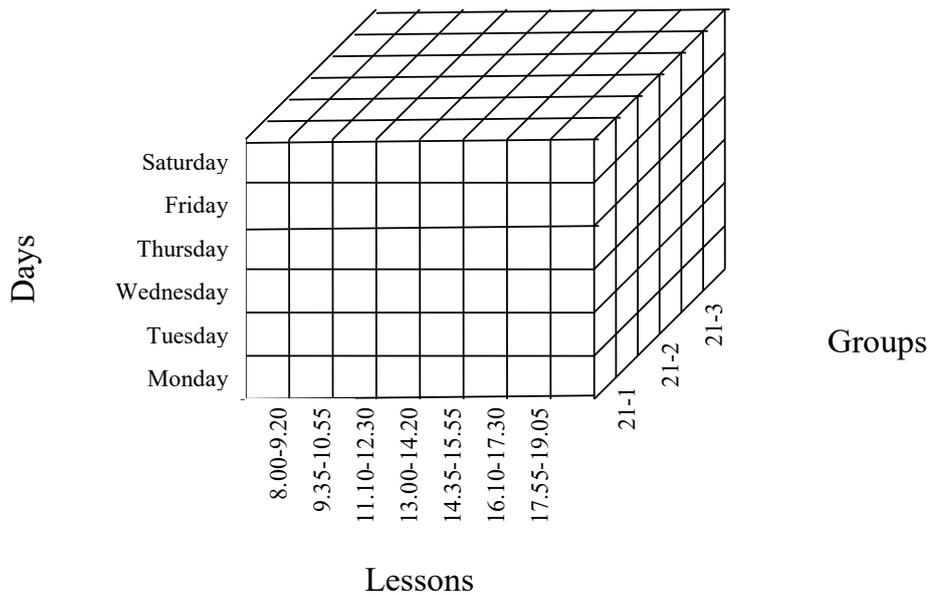


Figure 2. Hypercube "Student groups"

Hypercube G has three dimensions: days, bands, groups. The cells of the hypercube G contain a data structure called a record. For specific measurements (day d , tape l , group g), the record contains data about in which classroom, which subject, what type of class, which teacher group g is in class on day d with tape l . The cells of the hypercube G are filled in during the operation of the optimization algorithm for forming the schedule.

The structure of the record is as follows.

Type

Element_gdl = RECORD { hypercube cell structure G }

a: Array [1...A] of Integer; { audience codes }

p: Array [1...P] of Integer; { teacher codes }

z: Array [1...Z] of Integer; { discipline codes }

v: Integer; { occupation code }

END;

The third hypercube A is designed to perform various operations with data about classrooms. The structure of hypercube A is shown in Figure 3.

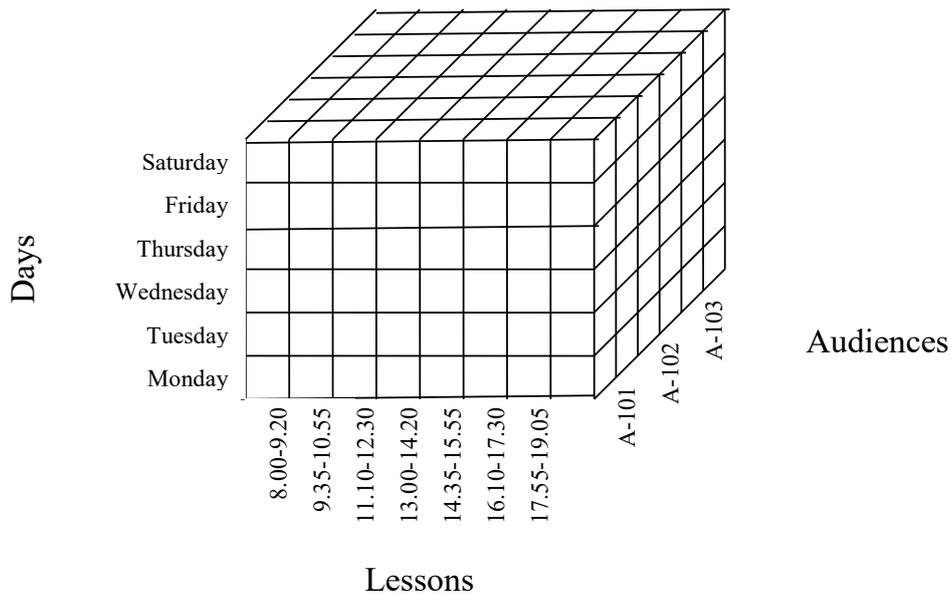


Figure 3. Hypercube "Audience"

Hypercube A has three dimensions: days, tapes, audiences. The cells of hypercube A contain a data structure called a record. For specific measurements (day d , tape l , classroom a), the record contains data about for which group (s), what subject, what type of classes, which teacher is held in classroom a on day d with tape l . The record is filled in during the operation of the optimization algorithm for forming the schedule. The structure of the record is as follows.

Type

```

Element_adl = RECORD { hypercube cell structure A }
    g: Array [1...AG] of Integer; { student group codes }
    p: Integer; { teacher code }
    z: Integer; { discipline code }
    v: Integer; { occupation code }
END;
```

These three hypercubes are analogous to manual repositories of scheduling data - load logs for teachers, student groups, and classrooms.

As noted above, both multidimensional and traditional relational data models can be used to store these three hypercubes. In this case, using the traditional relational data model, the hypercube is emulated by the DBMS at the logical level. Unlike multidimensional relational DBMSs are capable of storing huge amounts of data, but they are inferior in terms of the speed of execution of analytical queries. When using relational DBMSs to organize storage, data is organized in a special way. The most commonly used is the so-called radial pattern. Its other name is "star". This schema uses two types of tables: a fact table (fact table) and multiple lookup tables. In their studies, Koshlich and Gulakov (2019) propose to use schemas of data warehouse structures with radially linked tables (star schema) for storing the hypercubes "Teachers", "Groups", "Audiences", which are shown in Figure 4.

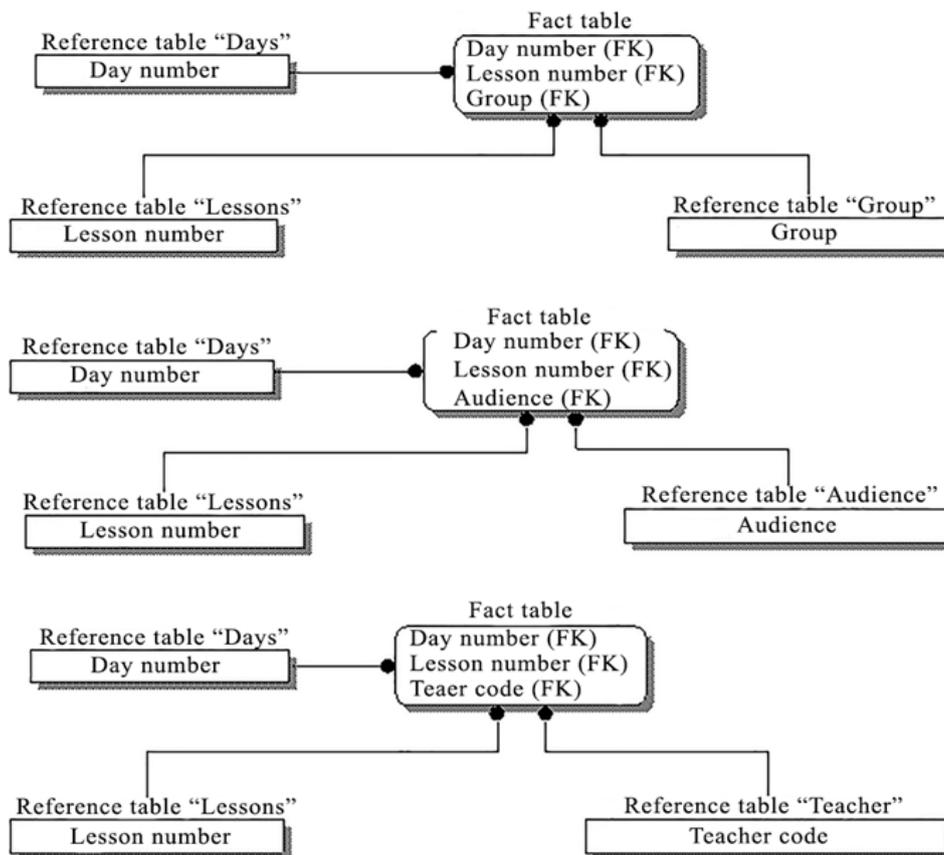


Figure 4. Data type diagram scheduling calendar

6. Findings

In this study, we will assume that a decision was made to reduce the dispatcher's salary, and the cost reduction will be considered as conditional income. Next, you need to make calculations. The costs of implementing the project will conditionally amount to 50,000 rubles. At the moment, all 3 dispatchers of the methodological department are engaged in scheduling. The salary of each of them is conventionally 12,000 rubles.

Automation of the scheduling process will allow to involve only one dispatcher in this process, which means that it is advisable for the remaining two to reduce wages. Since the manual scheduling is a rather laborious process that takes more time, the release of labor time will lead to a decrease in the wages of dispatchers by 3,000 rubles conditionally for each. Thus, we can draw up the following Table 1.

Table 1. Calculation of the conditional income from the implementation of the system

| Staff unit | Standard rate, RUB / month | Rate after sale, RUB / month | Cost reduction, RUB / month | Conditional income, rubles / 5 years |
|--------------|----------------------------|------------------------------|-----------------------------|--------------------------------------|
| Dispatcher 1 | 12.000 | 8.000 | 4.000 | 240.000 |
| Dispatcher 2 | 12.000 | 8.000 | 4.000 | 0 |
| Dispatcher 3 | 12.000 | 12.000 | 0 | |
| | | Total: | 8.000 | 480.000 |

Conditional income for 5 years will be 480,000 rubles at a cost of 50,000 rubles. Next, you need to calculate the efficiency that will be obtained after the implementation of the project. In this case, the main efficiency criterion should be considered the time freed up by dispatchers, previously spent on manual scheduling. The freed up time can be used to resolve other issues of the methodological department, as well as to work with students and the teaching staff of the Institute. At the moment, scheduling for one group for one dispatcher takes about 10 working hours on average, which is more than one working day with an 8-hour working day. On average, about 130 groups study at the Institute. This means that the time for scheduling completely takes about 1,300 hours or about 162 working days, provided that one dispatcher is engaged in this. When dividing the work into 3 dispatchers, scheduling takes more than 50 days – 1/3 of the semester.

7. Conclusion

As a result of the study, it was found that one of the most important problems of the high-quality organization of the educational process in the university is the task of forming a high-quality schedule of training sessions. This task is the main one in the activity of the methodological department. A well-designed schedule should ensure a uniform workload of both student groups and faculty in the next semester.

The introduction of an automated scheduling system based on a multidimensional model will significantly reduce the time for scheduling. The scheduling time for one group depends on the number of disciplines, the severity of the conditions and technical support. It is planned that, on average, scheduling for one group will take 30 minutes, taking into account the relaxed conditions. Then the total time is 65 hours, or about 10 business days. The implementation of the project will reduce the time for scheduling by 16 times, and the cost of wages conditionally by 480,000 rubles over 5 years.

References

- Akimkina, E. E. (2018). Structuring and visualization of indicators in multidimensional data cubes. *Information Technology Bulletin*, 4, 79-87.
- Batishchev, P. S. (2003). Experience in using information technologies in scheduling training sessions. *Secondary vocational education*, 11, 23-30.
- Codd, E. F., Codd, S. B., & Salley, C. T. (1993). *Providing OLAP to User-Analysts: An IT Mandate*. <http://www.arborsoft.com/papers/coddTOC.html>
- Ilyushechkin, V. M. (2019). *Basics of using and designing databases*.
- Kashirin, I. Yu. (2009). Interactive analytical data processing in modern OLAP systems. *Business Informatics*, 2, 12-19.
- Khasukhadzhiev, A. S., & Sibikina, I. V. (2016). A generalized algorithm for scheduling a timetable in a university, taking into account the new requirements of federal state educational standards. *Bulletin of the Astrakhan State Technical University*, 3, 78-86.
- Koshlich, A. D., & Gulakov, K. V. (2019). Modern approaches to building data warehouses. *Sciences of Europe*, 45, 32-37.
- Nosov, A. P., Akhrem, A. A., Rakhmankulov, V. Z., & Yuzhanin, K. V. (2020). Computational complexity analysis of OLAP-hypercube decomposition methods for multidimensional data. *Mathematics and Mathematical Modeling*, 4, 52-64.

- Ovsyanitskaya, L. Yu. (2018). Technology of analysis and visualization of multidimensional data of pedagogical monitoring in higher education. *Modern Information Technology and IT Education*, 14, 793-801.
- Popov, S. G., & Lisenkova, A. A. (2018). Algorithms for dynamic generation of MDX queries to multidimensional OLAP cubes. *SPbSPU Scientific and Technical Bulletin*, 4, 21-35.
- Razumnikov, S. V. (2013). Analysis of existing methods for assessing the effectiveness of information technology for cloud IT services. *Modern problems of science and education*, 3, 124.
- Samsonova, N. V., & Simonov, A. B. (2018). Scheduling in Higher Education. *Mathematical Methods and Software Products E-Management*, 1, 60-69.
- Vorobovich, N. P., & Lopateeva, O. N. (2006). Construction of multidimensional data models for the problems of forming the schedule of classes in universities. *Bulletin of KrasGAU*, 5, 370-378.
- Zatonsky, A. V., & Varlamova, S. A. (2018). Information support for decision-making support on the example of scheduling classes for an educational organization. *Bulletin of the South Ural State University*, 3, 88-106.
- Zykin, S. V., Mosin, S. V., & Poluyanov, A. N. (2016). Technology of separate generation of multidimensional data. *Bulletin of the Don State Technical University*, 2, 114-129.