25 aniversario
ARTÍCULOS
*UTOPÍA Y PRAXIS LATINOAMERICANA. AÑO: 25, nº EXTRA 7, 2020, pp. 251-256*
*REVISTA INTERNACIONAL DE FILOSOFÍA Y TEORÍA SOCIAL*
*CESA-FCES-UNIVERSIDAD DEL ZULIA. MARACAIBO-VENEZUELA*
*ISSN 1316-5216 / ISSN-e: 2477-9555*

# Automation of the Creation of the Schedule at the University: Mathematical Model

*Automatización de la creación del cronograma en la Universidad: modelo matemático*

**R.I. KHABIPOV**
https://orcid.org/0000-0002-5140-0392
paladin86@bk.ru
*Kazan Federal University, Kazan, Tatarstan Republic*

Este trabajo está depositado en Zenodo:
**DOI**: http://doi.org/10.5281/zenodo.4009704

### *RESUMEN*

El objetivo del trabajo es describir un modelo matemático de la automatización del proceso de programación en una institución educativa. También se consideran los enfoques para la compilación del horario óptimo. En las tareas estudiadas se utiliza una gran cantidad de información original inicial, que difiere en su composición, y contiene una gran cantidad de requisitos que deben tenerse en cuenta. Por lo tanto, la programación se refiere a la clase de problemas de programación de enteros NP-completos, lo que implica que a medida que aumenta el número de valores de variables dadas, la complejidad de la solución crecerá exponencialmente.

**Palabras clave:** Algoritmos de programación lineal entera, horario de clases, planes educativos, programación.

### *ABSTRACT*

The work aims to describe a mathematical model of the automation of the scheduling process in an educational institution. Also, the approaches to the compilation of the optimal schedule are considered. In the studied tasks, a large amount of initial original information is used, which differs in its composition, and contains a large number of requirements that must be taken into account. Therefore, scheduling refers to the class of NP-complete integer programming problems, which implies that as the number of values of given variables increases, the complexity of the solution will grow exponentially.

**Keywords**: Algorithms of integer linear programming, class schedule, educational plans, scheduling.

Recibido: 03-08-2020● Aceptado: 07-09-2020

## *INTRODUCTION*

The problem of developing methods and algorithms for solving the problem of preparation of the educational schedule engaged for a long time. The first works appeared in the 1970s yet (Conway: 1975, p. 360). the most approximate is the studies. Approximate and heuristic methods for solving this problem have been developing quite intensively in recent years.The solution of problems of the theory of schedules is complicated by the fact that most of them are NP-difficult, and algorithms for their solution, implemented on a computer, must be adapted to obtain an acceptable option for the maximum allowable time.

The formation of the optimal schedule of training sessions at the University for the academic year requires (Conway: 1975, p. 360) employees performing this work, time-consuming, from several days to several weeks of work. The schedule of classes should meet numerous organizational and methodological requirements with different priorities, which may be mutually contradictory or even mutually exclusive. Such requirements can make it very difficult to build a preliminary schedule let alone an optimal schedule.

In the construction of the schedule as the source data are the academic load, study groups, classrooms, teachers, and a list of disciplines. The task of developing an optimal schedule is to distribute the maximum number of classes per day in the available classrooms for the appropriate student groups, taking into account the load and the wishes of teachers (Tanaev & Shkurba: 1975).

## *METHODS*

There are several approaches (Azami et al.: 2018, pp. 3259-3274; Grankov et al.: 2016; Ivanco et al.: 2017, pp. 111-126; Rustauletov: 2020; Tanaev & Shkurba: 1975) to build a schedule related to associated with the use of:

• Classical integer-valued programming methods and algorithms;

• Approximate algorithms of solving with techniques for the coloring of graphs;

• A brute force algorithm, branches and bounds, and approximate methods that are based on genetic algorithms.

Using these algorithms and methods, you can get a solution to the problem of building a schedule that gratifies the specified criteria. However, the time spent on solving this task can be very long. Therefore, to find the optimal solution, the modeling method is used.

At this method, the algorithm works with a predefined list of lessons that should be included in the schedule. The process starts with the first lesson that is not yet reflected in the schedule, then the algorithm fills the appropriate section of the schedule, using the maximum number of lessons from the available list. Then the next lesson is included in the schedule. The process continues until a full schedule is formed or a specified number of repetitions are completed. The task of building an optimal schedule for large universities is quite complex and time-consuming, so the automation of this process is relevant. Also, the larger the University, the relevance of the automation of this process is higher (Ivanov et al.: 2020, pp. 1-9; Gawiejnowicz: 2020, pp. 46-63).

## *RESULTS*

### The problem statement

Based on the list of occupations, list of disciplines, students, the list of the classroom and the approved curriculum necessary to calculate the teaching load of all departments of the University for the current academic year, and then based on the calculated academic load is required to generate a schedule for this period.

We will highlight the features of solving the problem of building a schedule:

Since the schedule is traditionally drawn up manually, there is a need to create automated scheduling systems that would be implemented in the management system of the educational process of the University;

To ensure the effectiveness of the educational process, the resulting schedule should be optimized;

For visualization and evaluation of the work in the preparation of the schedule, it is necessary to develop a convenient form for the analysis of its presentation.

The algorithm of formation of the schedule of occupations is realized in two stages:

The primary schedule is formed, based on the initial data, the list of requirements and restrictions;

The resulting primary schedule is optimized, and then its quality is appreciated using the selected criteria.

The main components of the schedule are:

Time interval – the interval of time for the training session, which corresponds to a specific day in the school year; in other words, it is a specific time interval in which the lesson should be held;

Audience – a variety of places possible for classes.

Thus, the initial schedule is considered to be formed if each training session corresponds to a certain time interval and the audience (Tanaev et al.: 2012; Sha et al.: 2004, pp. 101-155).

**The schedule model**

Denote the set of training sessionsE=e_i, i=1...N. Each of them is characterized by such parameters as an academic discipline, audience, student group, and teacher. Training tasks are formed for each discipline and each study group. A set of training tasks can be thought of as a union of two subsetsE=E_T 〖∪E〗_N, where E_T – educational tasks with special requirements for audiences, for example, for work in computer classes, E_N – educational tasks without special requirements for audiences.

We introduce a set of time intervals T - these are time intervals that are uniquely determined by the numbers of the day of the week and the week itself. Thus, the set of time intervals is obtained using the Cartesian product of two-time parameters-week and day, in other words, it is the date-time:$E = E_T \cup E_N$.

Let is denote the set of academic groups and accordingly the set attributes of groups:

$$X = X_{\mathrm{m}},\ \text{m=1...M},\ A_x = \{(a_{x1}, a_{x2}, a_{x3})\},$$

Where a_x1 – number of planned (according to the workload) groups, a_x2 – the number of disciplines already distributed over time intervals in this group, a_x3 – the number of students in this group.

In some cases, groups are divided into subgroups, which is usually due to the capacity of the audience or the specifics of the subject taught. Many groups and subgroups cover the entire contingent of students.

We introduce the following sets:

The set of students

$$Z = X \cup Y, \qquad Z = \{Z_p, p = 1..\mathrm{Q}\},$$

Where X – set of female students, Y – set of male students;

The set of teachers 〖PA〗_k={a_p1,a_p2 }, where a_p1 – number of disciplines taught by this teacher in the schedule, a_p2 – the number of disciplines taught by the teacher and distributed in the schedule.

Denote the set of audienceA=a_k, k =1...K, and the set of attributes for each audienceA_a={(a_a1,a_a2,a_a3 )}, where a_a1 – the total number of disciplines (according to the approved workload), possible to conduct in the classroom for the period under review, a_a2 – the number of disciplines distributed in the schedule in a given audience, a_a3 - the capacity audience.

Thus, a specific training session in the schedule can be represented as a function of the parameters of the discipline, the contingent of students and teachers:

e_i=e(d_j,z_o,p_i ); 〖 d〗_j ϵD, z_o ϵZ, 〖 p〗_i ϵP.

**The imposed restrictions**

The main criterion for the feasibility of the generated schedule is its consistency: this means that there are no situations in the schedule when different lessons that require the same resources are placed in the same time interval of the schedule (audience, students, department, or teacher). Compliance with mandatory restrictions in the formation of the schedule is a sufficient condition for its consistency.

In the case of training lessons, restrictions on teachers will be as follows:

1) No overlays of study groups

∀z_o ϵZ, ∀t_k ϵT, ∀ d_j ϵD
d_j≠d_i↔ω(d_j,z_o,p,a,t_k)≠ ω(d_i,z_o,p,a,t_k);

2) Not combined audiences

∀a_p ϵA, ∀t_k ϵT, ∀ d_i ϵD, ∀ d_j ϵD
d_j≠d_i↔ω(d_j,g,p,a,t_k)≠ ω(d_i,g,p,a,t_k);

3) The lack of combined lessons from the same teacher in one time

∀p_i ϵP, ∀t_k ϵT, ∀ d_i ϵD, ∀ d_j ϵD
d_j≠d_i↔ω(d_j,g,p,a,t_k )≠ ω(d_i,g,p,a,t_k ).

Thus, for each ordered pair (teacher, lesson), there is either a single cycle of classes in which the teacher is "busy" at the specified time, or there is no such cycle, therefore, the teacher is free.

**The criterion of workload of the training lesson in the schedule**

Each training lesson$E = e_i$, *i*=1...I, can be evaluated, that is, put it by some numerical criterion:

$$V^E = K_X \times K_P \times K_A \times K_Y, \qquad 0 \ll K_X,\ K_P,\ K_A,\ K_Y \ll 1,$$

Where $K_X,\ K_P,\ K_A$ – specific criteria for the workload of groups, teachers, and audience, accounting classes in subgroups, which are calculated by the formula

$$K_i = \frac{a_1^i - a_2^i}{a^T - a_2^i}, \qquad (1)$$

Where $a_1^i$ - number of scheduled classes per period, $a_2^i$ - several distributed audience load criteria, which defined as follows:

- Training lessons that require audiences, such as classes with displays for programming lessons; for these, $K_i$ is calculated by the formula (T'kindt & Billaut: 2006);
- Training lessons that do not require an audience; the private criterion for the training assignments of this subset is calculated as follows:

$$K_A^0 = \frac{a_1^0 - a_2^0}{a^0 - a_2^0},$$

Where $a_1^0$ - number of training lessons without required audiences, $a_2^0$ - the number of lessons included in the schedule of learning tasks without the required audience, $a^0$– the number of time slots of available audiences.

## *DISCUSSION*

The following restrictions are also taken into account when forming the initial schedule and determining the time interval and audience for the lesson:

- Ensuring uniform distribution of classes in each group;
- To ensure the uniform distribution of classes across time intervals.

The first limitation is due to the need to obtain uniform intervals between classes, the second-the rational use of the classroom Fund.

The criterion of the workload of the training session includes private criteria of the workload of the group, the teacher, and the audience and is the final criterion. The described criteria allow us to estimate the average interval between classes for a group or subgroup. Each of the particular criteria is a relative value, the ability to override the value of the particular criterion is provided by changing the number of consecutive resource pairs included in the schedule.

## *CONCLUSION*

We have proposed a mathematical model and a method of creating an initial schedule with the use of workload criteria in the selection of educational tasks. This method is based on the database of the University, which can be automatically generated training tasks for classes, and the workload criteria are used, taking into account the requirements of users and allocating resources, which gives a more optimal result and acceptable performance within 10% than cumbersome methods that require large computational costs.

## *BIBLIOGRAPHY*

AZAMI, A, DEMIRLI, K, & BHUIYAN, N (2018). "Scheduling in aerospace composite manufacturing systems: a two-stage hybrid flow shop problem". The International Journal of Advanced Manufacturing Technology, 95(9-12), pp. 3259-3274.

CONWAY, RV, MAXWELL, VL, MILLER, LV (1975). "Theory of schedules". Moscow: Science. p. 360.

GAWIEJNOWICZ, S (2020). "The classical scheduling theory". In Models and Algorithms of Time-Dependent Scheduling, pp. 47-63. Springer, Berlin, Heidelberg.

GRANKOV, MV, AL-GABRI, VM, & GORLOVA, MY (2016). "Analysis and clustering of the main factors affecting the academic performance of university groups". Engineering Bulletin of the Don, (4).

IVANCO, M, HOU, G, & MICHAELI, J (2017). "Sensitivity analysis method to address user disparities in the analytic hierarchy process". Expert Systems with Applications, 90, pp. 111-126.

IVANOV, D, SOKOLOV, B, & DOLGUI, A (2020). "Introduction to Scheduling in Industry 4.0 and Cloud Manufacturing Systems". In Scheduling in Industry 4.0 and Cloud Manufacturing, pp. 1-9. Springer, Cham.

RUSTAULETOV, B (2020). "The Automation of Course Scheduling in Higher Education Institutions: Mathematical methods and software products". In Proceedings of International Young Scholars Workshop (Vol. 9).

SHA, L, ABDELZAHER, T, CERVIN, A, BAKER, T, BURNS, A, BUTTAZZO, G, ... & MOK, AK (2004). "Real-time scheduling theory: A historical perspective". Real-time systems, 28(2-3), pp. 101-155.

TANAEV, V, GORDON, W, & SHAFRANSKY, YM (2012). “Scheduling theory. Single-stage systems (Vol. 284)”. Springer Science & Business Media.

TANAEV, VS, & SHKURBA, VV (1975). “Introduction to scheduling theory”. Nauk~, Moscow.

T'KINDT, V, & BILLAUT, JC (2006). “Multicriteria scheduling: theory, models and algorithms”. Springer Science & Business Media.

## *BIODATA*

**R.I. KHABIPOV**: Khabipov was born in 1986 in the small town of Bavly, the Republic of Tatarstan of the Russian Federation. studied at the Faculty of Computational Mathematics and Cybernetics of Kazan State University from 2004 to 2009. Since 2010, have been working as a programmer at Kazan Federal University. From 2011-2015, studied in insight on the management of the economy and territorial development. In 2017, entered a postgraduate course in the specialty 05.13.01 system analysis, management, and information processing at the Higher School of Information Technologies and Intelligent Systems of Kazan Federal University.