The Mathematical Model of the Optimal Choice of a Software Package for an Enterprise Information System

Liudmyla Koliechkina
Algorithms and Databases Department
University of Lodz, Lodz, Poland
Poltava University of Economics and Trade
Poltava, Ukraine
ludaplt1971@gmail.com

Alla Nahirna
Department of Informatics
National University of "Kyiv-Mohyla Academy"
Kyiv, Ukraine
naghirnaalla@ukr.net

***Abstract*—The mathematical definition of the optimal choice problem for software package for an enterprise information system is considered. The problem of the optimal software package choice is a mathematical model of combinatorial optimization on a set of combinations. The algorithm for solving this problem is presented for finding the minimum value of the target function considering additional linear constraints. This method makes it possible to significantly simplify the finding procedure of the optimal solution, since inequalities in the growth of constraints allow us to immediately determine whether a point in the set of combinations will be a support solution or not. In the positive case, the support solution is improved, taking into account the properties of many combinations when directly checking the growth of the target function. The algorithm and its realization are demonstrated by a numerical example. The proposed mathematical model and solution algorithm can be used for a similar class of problems that are modeled by optimization models, where the set of feasible solutions is presented in the many combinations form.**

Keywords—mathematical model, information system, combinatorial set of combinations, optimal solution, growth of the objective function

I. Introduction

The deep penetration of information and communication systems and telecommunication technologies into applied areas is a key element in an enterprise. The concept of information and communication systems includes information systems (hardware and software), telecommunication equipment (subscriber, network) and telecommunication services [1, 2].

In turn, the Information System (IS) is an interconnected set of informational, technical, software, mathematical, organizational, legal, ergonomic, linguistic, technological and other means, intended for the collection, processing, storage and issuing economic information and making managerial decisions [3]. In any information system, processes related to input, processing, storage, withdrawal, and information protection are important. Such processes are carried out using software packages.

Software is an essential component in building an information system in an enterprise. Therefore, the choice of the necessary software is one of the main tasks in the construction of an information system.

When modeling various types of processes and events in different areas, the mathematical models are often used that represent the problems of conditional optimization [4-8].

The models of such problems considered on combinatorial sets deserve special attention. These models are widely used in computer technology, mechanical engineering, logistics, in the complex development of data flow protection systems in enterprises, etc. Nowadays, a lot of scientific works are devoted to the study of combinatorial models [4, 9–12].

A universal method for solving such problems is a complete enumeration of options, which may be applicable to problems of small dimension, but does not give the desired result for large dimensions. Of course, there are other methods for solving such problems, but as a rule, each of them has its own advantages and disadvantages [13-22]. Therefore, it becomes necessary to develop new methods, both exact and approximate, that take into account the specifics of an objective function and constraints of an optimization problem on combinatorial sets.

Based on the consideration of the inequalities in a growth of constraints and objective function, considering properties of a set of combinations, the proposed method can significantly simplify a procedure for finding an optimal solution by determining in a few steps that a considered point of a set of combinations belongs to a reference solution [13].

II. Formal Problem Statement

When building an information system in an enterprise, considerable attention is paid to the choice of software as one of the main components. As a rule, software is defined as a set of programs for implementing a goals, objectives of an information system and normal functioning of both individual and complex hardware [3]. For ease of understanding, this set of programs will be called a software package.

There are many software products on the software market, depending on specifics of an enterprise and features of an information system. But the main task is to effectively select a set of programs that would ensure necessary functioning of an information system in the enterprise, at the same time the acquisition cost should be minimal.

978-x-xxxx-xxxx-x/19/$31.00 ©2019 IEEE

Considering the main tasks of building an information system in an enterprise, we assume that n programs were selected to form a software package. Due to financial capabilities, an enterprise can purchase only k programs from n $(k \le n)$.

For ease of understanding, we assume that 1 is a program with the lowest rating, respectively, n is a program with the highest rating. The main task is to choose a set of programs with the highest rating taking into account the conditions of an enterprise imposed on information system. This set of programs will form a necessary software package. We assume that the average prices for the corresponding programs with a certain rating are known. Then the choice should provide for the minimum cost of acquiring a software package with the highest rating.

Let $A=(a_1,a_2,...,a_n)$ be a set of n programs whose rating, respectively, is $a_1=1$, $a_2=2$,..., $a_n=n$. Then $c_1,c_2,...,c_n$ are the average program prices according to the rating. Since it is necessary to form a software package by selecting k programs from n $(n \ge k)$, taking into account their rating, we can look for a solution on a combinatorial set of combinations without repetitions C_n^k.

As far as this package of programs is to ensure the effective operation of an information system in an enterprise as a whole, it is natural to assume that each structural unit of an enterprise has its own requirements for choosing programs. Programs may coincide or differ, which mathematically can be formulated as a system of inequalities $Gx \le (\ge) b$.

It is necessary to form a package in which programs would have the highest rating and satisfy the conditions of structural units of an enterprise. This software package should minimize acquisition costs. When detailing the choice, it is possible to perform additional calculations of the objective function in the range of maximum and minimum average prices of programs taking into account their rating.

III. The Mathematical Model

Given a set $A=(a_1,a_2,...,a_n)$, $(n \in N)$. A combination without repetitions of n elements by k $(n \ge k)$ is called k-elemental subset C of the set A and is denoted by C_n^k. Since the order of writing elements of the set is irrelevant, therefore, as a rule, the elements in each combination are written in ascending order.

Consider a mathematical model of the form:

$$Z(\Phi, C_n^k(A)) : \min\{\, \Phi(a) \mid a \in C_n^k \},\qquad (1)$$

where A is a subset of the combinatorial set of combinations C_n^k, defined by a given system of constraints.

We realize the bijective mapping of the set C_n^k into space R^n by assigning a relevant vector $x \in R^n$ to each element $a \in C_n^k$. The image of a set C_n^k is denoted by $E_k^n \subset R^n$. As a result, we have the problem of combinatorial optimization in the Euclidean formulation (the problem of Euclidean combinatorial optimization)

$$Z(F, E_k^n) : \min\{\, F(x) \mid x \in D \subset E_k^n \}\qquad (2)$$

$$D = \{x \in E_k^n \subset R^n \mid Gx \le (\ge) b\},\qquad (3)$$

where $G - m \times n$ matrix, $b \in R^m$, while $\Phi(a) = F(x)$ for $a \in C_n^k$, $x \in E_k^n$.

Next, we consider linear objective functions of the form

$$F(x) = \sum_{j=1}^{n} c_j x_j .$$

Additional linear constraints form a multifaceted set $D \subset R^n$.

IV. Algorithm

The algorithm for solving the formulated problem consists of three steps, which ensure that the minimum of the objective function is found with additional restrictions on the set of combinations.

Step 1. Finding the first reference solution.

According to Definition 1, the elements of a set of combinations are written in ascending order, so the minimum element of a set of combinations $(x_1,x_2,...,x_{n-1},x_n)$, where $(x_1 < x_2 < ... < x_{n-1} < x_n)$, must be taken as the starting point. Next, check the constraints (2).

If the constraints are satisfied, go to step 2. Otherwise, select the next point in ascending order and repeat step 1. Notably, the number of elements in the set of combinations is a finite set (3).

Step 2. Formation of the initial search conditions for the optimal solution.

A point of a set of combinations that satisfies all constraints (2) will be the first reference solution. For further search for the optimal solution, the initial conditions for finding the optimal solution are formed:

$$f(x_1,x_2,...,x_n) = b,$$

$$\begin{cases} \Delta g_1(x_1,x_2,...,x_n) \le (\ge) \Delta b_1, & \Delta b_1 = b_1 - g_1(x_1,x_2,...,x_n), \\ \Delta g_2(x_1,x_2,...,x_n) \le (\ge) \Delta b_2, & \Delta b_2 = b_2 - g_2(x_1,x_2,...,x_n), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots, \\ \Delta g_n(x_1,x_2,...,x_n) \le (\ge) \Delta b_n, & \Delta b_n = b_n - g_n(x_1,x_2,...,x_n). \end{cases}\qquad (4)$$

Step 3. Improving the first reference solution.

Consider the next point of the set of combinations in ascending order after the first reference solution. Find the growth of the constraints (4) and check their fulfillment:

$$\Delta g_i = \Delta g_i^2 - \Delta g_i^1 = \left(x_i^{g_i^2} * c_j + x_j^{g_i^2} * c_i\right) - \left(x_j^{g_i^1} * c_j + x_i^{g_i^1} * c_i\right)\qquad (5)$$

If (5) is not satisfied, then the next point of the set of combinations in ascending order is considered.

Otherwase, find growth of the objective function:

$$\Delta f = \Delta f_2 - \Delta f_1 = \left(x_i^{f_2} * c_j + x_j^{f_2} * c_i\right) - \left(x_j^{f_1} * c_j + x_i^{f_1} * c_i\right)\qquad (6)$$

Since the minimum value of the objective function must be find, a necessary condition for improving the first reference solution is to reduce the growth of the objective function:

$$\Delta f \le 0 \tag{7}$$

If (7) is satisfied, then the optimal solution is found, otherwise, the next point of the set of combinations is considered and verification (5), (7) is carried out.

V. Example

At the software market, five types of programs were chosen to form a software package for the information system of a trade enterprise. These programs have their specific rating A = (1, 2, 3, 4, 5). After a comprehensive assessment of the information system, considering the financial activities of the trade company, it was found that there is a need to purchase three of the five programs. The price of each program, depending on the rating, can vary in the following range: for the program of the 1st rating from \$ 1000 to \$ 2000, 2nd rating from \$ 2500 to \$ 4000, 3rd rating from \$ 4500 to \$ 6000, 4th rating from \$ 6500 -8000 \$, 5th rating from \$ 8500 -20000 \$.

The trade company has 4 departments that are engaged in data processing for three types of goods. Each department has formed its necessary conditions for selection of software package, which are illustrated by the following constraints:

$$\begin{cases} 0{,}9x_1 + 0{,}3x_2 + 0{,}2x_3 \ge 3, \\ 0{,}2x_2 + 0{,}3x_3 \ge 2, \\ 0{,}7x_1 + 0{,}3x_2 + 0{,}8x_3 \ge 6, \\ 0{,}4x_1 + 0{,}2x_2 + 0{,}3x_3 \ge 3. \end{cases}$$

The choice of three out of five possible programs are needed for creation a software package for the enterprise information system, considering the constraints and minimal costs for programs purchase.

The objective function will be: $\min F = c_1x_1 + c_2x_2 + c_3x_3$, where c_1, c_2, c_3 should be considering in the following range: for the program of the 1st rating from \$ 1000 to \$ 2000, 2nd rating from \$ 2500 to \$ 4000, 3rd rating from \$ 4500 to \$ 6000, 4th rating from \$ 6500 -8000 \$, 5th rating from \$ 8500 -20000 \$.

Consider the solution to this problem on a set of combinations $C_5^3(A)$, where $A = (1,2,3,4,5)$, with constraints mentioned above.

Since the programs with the highest rating should be selected, the point of the set of combinations (3, 4, 5) need to be the first one to consider, then the value of the objective function will be in the range:

F (3, 4, 5) = {13500-18000; 26000-32000; 42500-100000}.

Found values of the constraints: $g_1 = 4{,}9 > 3$; $g_2 = 2{,}3 > 2$; $g_3 = 7{,}3 > 6$; $g_4 = 3{,}5 > 3$, constraints are satisfied, then the initial search conditions for the optimal solution:

$$\begin{cases} \Delta g_1 \ge -1{,}9, \\ \Delta g_2 \ge -0{,}3, \\ \Delta g_3 \ge -1{,}3, \\ \Delta g_4 \ge -0{,}5. \end{cases}$$

The point (3, 4, 5) is the first reference solution. Consider the next point of the set of combinations (2, 4, 5) in descending order to improve the first reference solution. Certainly, the growth of the objective function decreases: (2, 4, 5) = {5000-13500 = -8500; 8000-18000 = -10000}. Then we verify the fulfillment of the initial search conditions for the optimal solution:

$$\begin{cases} \Delta g_1 = -0{,}9 \ge -1{,}9, \\ \Delta g_2 = 0 \ge -0{,}3, \\ \Delta g_3 = -0{,}7 \ge -1{,}3, \\ \Delta g_4 = -0{,}4 \ge -0{,}5. \end{cases}$$

Conditions (5) and (7) of the algorithm are satisfied.

Then the new search conditions for the optimal solution:

$$\begin{cases} \Delta g_1 \ge -1, \\ \Delta g_2 \ge -0{,}3, \\ \Delta g_3 \ge -0{,}6, \\ \Delta g_4 \ge -0{,}1. \end{cases}$$

The next point of the set of combinations (2, 3, 5): ΔF (2, 3, 5) = {13500-26000=-12500; 18000-32000=-14000}. Verifying the fulfillment of the initial search conditions for the optimal solution:

$$\begin{cases} \Delta g_1 = -0{,}3 \ge -1, \\ \Delta g_2 = -0{,}2 \ge -0{,}3, \\ \Delta g_3 = -0{,}3 \ge -0{,}6, \\ \Delta g_4 = -0{,}2 < -0{,}1. \end{cases}$$

The last inequality is not satisfied; therefore, the point of the set of combinations (2, 3, 5) does not improve the previous solution.

The solution is the point (2, 4, 5), accordingly, the minimum values of the objective function:

F (2, 4, 5) = {5000-8000; 26000-32000; 42500-100000}.

Answer: to build an information system, the trade enterprise has to form a software package that consists of programs with 2nd , 4th and 5th rating. The minimum cost of acquiring them will range from \$ 73500 to \$ 140,000.

To solve the example by the proposed method using a program in the C++ programming language, a computational experiment was carried out, taking into account the increase in the number of sample elements n of the set of combinations with a fixed element $k = 3$. The results of computational experiments are presented in the table 1.

TABLE I.

n	$\left\|C_n^3\right\|$	r	s
5	10	3	4
6	20	6	6
7	35	10	9
8	56	23	15
9	84	37	31
10	120	51	47
11	165	77	56

Table 1 uses the following notation:

n - the number of elements from which a set of combinations are built;

$\left|C_n^3\right|$ - the number of elements of a set of combinations;

r - the number of points that are being considered in the process of finding the optimal solution;

s - the number of steps to find the optimal solution. Analyzing the results of a computational experiment, it should be noted that an increase in the number of sample elements of the set of combinations $\left|C_n^3\right|$ does not lead to a rapid increase in the indicators r and s.

In addition, it should be noted that the proposed method allows for a finite number of steps (s) to find the minimum of the function on a set of combinations.

VI. Conclusion

The trade company has 4 departments that are engaged in data processing for three types of goods. Each department has formed its necessary conditions for selection of software package, which are illustrated by the constraints. The article considers a mathematical model of the optimal choice of a software package for an information system in an enterprise. The optimal software selection problem is presented as a mathematical model of discrete optimization on a combinatorial set of combinations.

Also, a practical task is presented, which is formulated as a mathematical model of the optimal choice of software package for effective work of information system in an enterprise. Using the proposed algorithm, it was determined that the company has to create a software package that consists of the programs with a rating of 2nd , 4th , 5th , with minimal costs ranging from $ 73,500 to $ 140,000.

This method allows to significantly simplify the procedure for finding the optimal solution, since inequalities in the growth of constraints make it possible to immediately determine whether a point in a set of combinations will be a reference solution or not. It is not necessary to make complicated calculations of all constraints and objective function; it is enough to find the growth of constraints and function in case of improving the solution. Further researches are aimed at considering a more complex mathematical model on other combinatorial sets, considering the increase in their dimension and addition of other conditions for solving practical problems in the field of information technology.

References

[1] R. Klempous, J. Nikodem and P. Zolt Baranyi, *Cognitive* Infocommunications, Theory And Applications. Springer Nature, New York, 2018.

[2] G. Sallai, “Defining Infocommunications and Relatrd Terms,” Acta Polytehnica Hungarica, vol. 9, no. 6, pp. 5-15, Dec. 2012.

[3] G. Eason, B. Noble, and I.N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529-551, April 1955. (*references*)

[4] R. Stair and G. Reynolds, Fundamentals of of Informations Systems. Cengage Learning, National Geographic Learning. Boston, 2015.

[5] L. Metcalf and W. Casey, Cybersecurity and Applied Mathematics. Syngress, 2016.

[6] V. Bulatov, Mathematical Modeling of Complex Systems. Theory and Applications. Scholars’Press, 2015

[7] V.Ya. Pevnev and M.V. Tsuranov, “Mathematical model of information security,” Information Processing Systems, vol. 3, pp. 62-64, 2010.

[8] M. Y. Ritov, A. P. Gorlov and V. T. Eryomenko, “Mathematical modeling of the assessing process the of security level of the complex organizational and technical systems of industrial enterprises” presented at the 2016 2nd Int. Conf. on Industrial Engineering, Applications and Manufacturing.

[9] A.A. Gerasimov and R. E. Zhuchkov, “The mathematical models of information security mechanisms in personal data automated information systems of the state administration bodies,” Vestnik Voronezhskogo instituta MVD Rossii, vol.4, pp. 282-289, 2014.

[10] L. Koliechkina and O. Pichugina, “Multiobjective Optimization on Permutations with Applications,” presented at the IX International Conference on Optimization and Applications, 2018, paper 61-75.

[11] L. Koliechkina, A. Nahirna and O. Dvirna, “Quadratic Optimization Problem on Permutation Set with Simulation of Applied Tasks,” presented at the Second International Workshop on Computer Modeling and Intelligent Systems, Zaporizhzhia, Ukraine, 2019, pp. 651-663.

[12] S. Yakovlev, O. Kartashov, O. Pichugina and L.Koliechkina, “The Genetic Algorithms in Optimization Problem on Combinatorial Configurations,” in International Conference on Innovations in Engineering, Technology and Sciences, Karnataka, India, 2018, pp. 106-111.

[13] N.V. Semenova, L.N. Kolechkina and A.N. Nagornaya, “On Approach to Solving Vector Problems with Fractionally Linear Functions of the Criteria on the Combinatorial Set of Arrangements,” Journal of Automation and Information Sciences, vol. 42, no. 2, pp. 67-80, Apr. 2010.

[14] A. Nahirna, “Rozv’yzyvannja zadach ymovnoi optumazacii na mnozuni spolychen [Solving the Conditional Optimization Problem on combinations set].” The theory of optimal solutions, pp. 123-127, 2019. (In Ukrainian).

[15] B. Korte and J. Vygen, Combinatorial Optimization: Theory and Algorithms. Springer, New York, 2018.

[16] P.M. Pardalos, D.-Z. Du and R.L. Graham (Eds.), Handbook of combinatorial optimization. Springer, New York, 2013.

[17] K. Bernhard and V. Jens, Combinatorial Optimization, Theory and Algorithms. Springer-Verlag, Berlin, 2012.

[18] R. Karin and A. Saoub, Tour through Graph Theory. Chapman and Hall/CRC, New York, 2017.

[19] V.A. Yemelichev, M.M. Kovalev and M.K. Kravtsov, Polytopes, graphs and optimisation. Cambridge University Press, Cambridge, 1984.

[20] L. Hulianytskyi and I. Riasna, “Formalization and classification of combinatorial optimization problems,” Springer Optimization and its Applications, vol. 130, pp. 239-250, 2017.

[21] S.V. Yakovlev and I.V. Grebennik, “Localization of solutions of some problems of nonlinear integer optimization,” Cybernetics and Systems Analysis, vol. 29, no. 5, pp. 727-734, Oct. 1993.

[22] O.A. Emets and L.N. Koliechkina, “Solution of optimization problems with fractional-linear objective functions and additional linear constraints on permutations,” Cybernetics and Systems Analysis, vol. 40, no. 3, pp. 329-339, Jun. 2004.