Algorithm for Finding Guaranteed Solution in Knapsack Problem

Knyaz Mammadov[1], Nazim Mammadov[2]
Cybernetics Institute of ANAS, Baku, Azerbaijan
[1]*Mamedov_knyaz@yahoo.com*, [2]*Nazim_mammedov@mail.ru*

Abstract— **In the paper, an algorithm for finding the *guaranteed* suboptimal solution of the 0-1 variable knapsack problem is given. A program on this algorithm was composed, comprehensive and comparative calculating experiments were done.**

Keywords— Knapsack problem; suboptimal solution; guaranteed solution; guaranteed suboptimal solution; computing experiments

I. Introduction

Consider the following 0-1 variable knapsack problem:

$$\sum_{j=1}^{n} c_j x_j \to \max, \tag{1.1}$$

$$\sum_{j=1}^{n} a_j x_j \le b, \tag{1.2}$$

$$x_j = 0 \cup 1;\ j = \overline{1,n} \tag{1.3}$$

Without loss of generality, suppose that $c_j > 0,\ a_j > 0,\ j = \overline{1,n},\ b > 0$ are integers, and the conditions $\frac{c_1}{a_1} \ge \frac{c_2}{a_2} \ge \frac{c_3}{a_3} \ge \ldots \ge \frac{c_n}{a_n}$ are satisfied. As a problem (1.1)-(1.3) belongs to NP – complete class, there is not methods possessing polynomial time complexity for finding its optimal solution [1,2].But different algorithms were worked out for finding its suboptimal (approximate) solution [2, 3, 4 and etc.]. A great majority of these methods are based on the giving a unique value to the variables that correspond to the greatest of the ratios $\frac{c_j}{a_j}$. Some of the approximate solution principle are resulted in stopping of the calculating process at certain step while calculating the problem by the "branching and boundaries" methods [5, 6 and etc.].

Note that in the papers [7, 8] the notions of guaranteed suboptimal solution are given and a method for finding this solution is elaborated. This method increases the number b by certain quantity and then divides it by the dichotomy principle. But in this paper, we elaborate another method for finding the guaranteed suboptimal solution of problem (1.1)-(1.3). The method in the paper [7] is an iterative process. But in this paper , the method is based on reduction of the stated problem to ordinary knapsack problem by means of different transformations.

II. Problem Statement

Let the optimal solution $X^* = \left(x_1^*, x_2^*, \ldots x_n^*\right)$ or the suboptimal solution $X^s = \left(x_1^s, x_2^s, \ldots x_n^s\right)$ of problem (1.1)-(1.3) be found by one of the known methods. This time the maximum value of function (1.1) is $f^* = \sum_{j=1}^{n} c_j x_j^*$ or at least $f^s = \sum_{j=1}^{n} c_j x_j^s$.

Assume that the values of f^* or f^s must be increased to some value (for example , increase $p\%$).It is clear that then we must increase the number b in relation (1.1). Then naturally there arises some a questions : How much must the number b be increased that the known quantity f^* or f^s may increase by certain Δ. Here in special case we can take $\Delta = \left[f^* \cdot \frac{p}{100} \right]$, i.e. at least $p\%$ increase of function (1.1) should be guaranteed. In other words, to the known number b we should add such a known minimal integer δ that in the appropriate knapsack problem the value of function (1.1) be not less than $f^* + \Delta$ or $f^s + \Delta$. We again note that f^* or f^s and Δ are the given numbers, the minimal value of the quantity δ should be found. Thus we get the following mathematical model :

$$\delta \to \min, \tag{2.1}$$

$$\sum_{j=1}^{n} a_j x_j \le b + \delta, \tag{2.2}$$

$$\sum_{j=1}^{n} c_j x_j \ge f^* + \Delta, \tag{2.3}$$

$$x_j = 0 \cup 1;\ j = \overline{1,n} \tag{2.4}$$

Here, δ must be a positive integer. Otherwise if $\delta = 0$ then is the solution $X^* = \left(x_1^*, x_2^*, \ldots x_n^*\right)$ gives to the function (1.1) maximum value equal to f^*, condition (2.3) may not be satisfied.

The optimal solution of problem (2.1) - (2.4) is such n-dimensional $X=(x_1, x_2,...x_n)$ vector satisfying condition (2.2) - (2.4) that gives minimum value to the quantity δ.

In this paper, for finding the optimal (suboptimal) solution of problem (2.1) – (2.4), a method was worked out, its program was composed and comparative calculating experiments on different-dimensional problems were conducted.

III. Theoretical Ground of the Method

As first,based on the paper [8] give the following notion:

Definition 1. The vector $X=(x_1, x_2,...x_n)$ satisfying conditions (2.2) - (2.4) for the mentioned integer $\delta>0$ is said to be a possible solution of problem (2.1)-(2.4).

Definition 2. The solution $X^t=(x_1^t, x_2^t,...x_n^t)$ giving the least positive value to the quantity δ among the possible solutions of problem (2.1) - (2.4) is said to be a guaranteed solution of knapsack problem (1.1) – (1.3).

Definition 3. The solution $X^s=(x_1^s, x_2^s,...x_n^s)$ giving the possible least positive value to the quantity δ among the possible solutions of problem (2.1) - (2.4) is said to be a guaranteed suboptimal solution of knapsack problem (1.1) – (1.3).

As first we add the known $y\ge 0$ to the left side of inequality (2.2) and denote the quantity δ :

$$\delta=\sum_{j=1}^{n} a_j x_j + y - b$$

Here $0\le y\le b$ and must get integers. According to the problem statement , the quantity δ should be minimized. Then we get the following problem:

$$\delta=\sum_{j=1}^{n} a_j x_j - b + y \to \min \quad (3.1)$$

$$\sum_{j=1}^{n} c_j x_j \ge f^* + \Delta, \quad (3.2)$$

$$x_j = 0 \cup 1;\ j=\overline{1,n} \quad (3.3)$$

$$y \ge 0 \text{ and integer} \quad (3.4)$$

We can solve the obtained problem (3.1)-(3.4) as an integer minimization problem. Equivalently, we reduce this problem to the maximization problem. Because , for conducting experimental comparisons , our program solve maximization problems. To this end , we write problem (3.1) – (3.4) in the following form:

$$-\delta=\sum_{j=1}^{n}(-a_j x_j) + b - y \to \max, \quad (3.5)$$

$$\sum_{j=1}^{n}(-c_j x_j) \le -f^* - \Delta, \quad (3.6)$$

$$x_j = 0 \cup 1;\ j=\overline{1,n}, \quad (3.7)$$

$$y\ge 0,\ 0\le y\le b, \text{ and integer} \quad (3.8)$$

As in problem (3.1) – (3.8) the coefficient became negative numbers, we make the substitution $x_j = 1 - t_j, (j=\overline{1,n})$ and $y=b-z$ and this problem becomes a positive coefficient problem. Here $t_j = 0\cup 1;\ j=\overline{1,n}, z=\overline{0,b}$. As a result we get the following problem:

$$\sum_{j=1}^{n} a_j t_j + z - \sum_{j=1}^{n} a_j \to \max \quad (3.9)$$

$$\sum_{j=1}^{n} c_j t_j \le \sum_{j=1}^{n} c_j - f^* - \Delta \quad (3.10)$$

$$t_j = 0 \cup 1;\ j=\overline{1,n} \quad (3.11)$$

$$0\le z\le b \text{ and integer} \quad (3.12)$$

We see that, all the coefficients of problem (3.9) – (3.12) are integers. On the other hand, as we consider a maximization problem and the variable z participates only in the function (3.9), we accept its maximum value as $z=b$, then in relation (3.9) the sum $z-\sum_{j=1}^{n} a_j$ becomes a certain constant number. Thus we get the knapsack problem (3.9) – (3.11). Having solved this problem by one of the known methods, we get the optimal solution $T^*=(t_1^*, t_2^*,...t_n^*)$. Writing this solution in the substitution $x_j = 1 - t_j$ we obtain the guaranteed solution $X=(x_1, x_2,..., x_n)$.

Note that when the number of the unknowns in problem (1.1) – (1.3) is rather great, it becomes difficult to find its optimal solution $X^*=(x_1^*, x_2^*,..., x_n^*)$, the optimal solution of problem (3.9) - (3.11) and also the value f^* of function (1.1). Therefore , at first we find the approximate values of $X^s=(x_1^s, x_2^s,...x_n^s)$ and of f^s write $f^*=f^s$ in relation (3.10) and get a new guaranteed suboptimal solution and appropriate value $f^s=\sum_{j=1}^{n} c_j x_j^s$ of the obtained problem.

IV. RESULTS OF CALCULATING EXPERIMENTS

For clarifying quality of the method suggested in the paper and its comparison with the method from [7], we conducted some calculat experiments on various – di mentional problems, The coefficients of this problem were found in the paper [4] and satisfy the following conditions :

$0 < a_j \le 99$, $0 < c_j \le 99$, $j = \overline{1,n}$ or

$0 < a_j \le 999$, $0 < c_j \le 999$, $j = \overline{1,n}$

and are found as

$$b = \left[0.3 \sum_{j=1}^{n} a_j \right].$$

Here , the sign $[z]$ indicates the integer part of the number z . The result of calculating experiments are given in the following table:

Table 1. Finding the quantity $\delta_{\min}$ when $0 < a_j \le 99$, $0 < c_j \le 99$, $n = (100, 200, 500)$

n	100				200				500			
b	1424				2876				7271			
f^s	3033				5836				15797			
$p\%$	1	2	3	5	1	2	3	5	1	2	3	5
$\Delta = \left[f^* \cdot \frac{p}{100} \right]$	30	60	90	151	58	116	175	291	157	315	473	789
$\delta_{\min}$	24	51	90	154	64	118	180	285	147	287	426	727
$\delta_{\min}[7]$	25	51	103	154	65	130	209	313	140	281	530	795
$f^s(\delta_{\min})$	3071	3101	3133	3193	5900	5953	6017	6127	15964	16115	16273	16594
$f^s(\delta_{\min}[7])$	3066	3101	3151	3184	5902	5969	6047	6144	15955	16112	16383	16669

Table 2. Finding the quantity $\delta_{\min}$ when $0 < a_j \le 99$, $0 < c_j \le 99$, $n = (1000, 2000, 5000)$

n	1000				2000				5000			
b	14496				30168				73400			
f^s	32451				63076				156355			
$p\%$	1	2	3	5	1	2	3	5	1	2	3	5
$\Delta = \left[f^* \cdot \frac{p}{100} \right]$	324	649	973	1622	630	1261	1892	3153	1563	3127	4690	7817
$\delta_{\min}$	296	599	900	1514	609	1226	1856	3128	1505	3023	4560	7691
$\delta_{\min}[7]$	330	594	1057	1585	617	1236	2199	3298	1505	3345	5352	8028
$f^s(\delta_{\min})$	32775	33102	33425	34074	63706	64337	64969	66230	157919	159482	161045	1641 73
$f^s(\delta_{\min}[7])$	32815	33100	33594	34146	63714	64346	65312	66398	157919	159812	161845	1645 04

Table 3. Finding the quantity $\delta_{\min}$ when $0 < a_j \le 999$, $0 < c_j \le 999$, $n = (100, 200, 500)$

n	100				200				500			
b	14381				29030				73351			
f^s	30447				58578				158621			
$p\%$	1	2	3	5	1	2	3	5	1	2	3	5
$\Delta = \left[f^* \cdot \frac{p}{100} \right]$	304	608	1586	1586	1586	1586	1757	2928	1586	3172	4758	7931
$\delta_{\min}$	266	541	786	1428	549	1070	1718	2881	1430	2841	4315	7241
$\delta_{\min}[7]$	196	540	1048	2097	661	1190	2116	3174	1378	2841	5348	8022
$f^s(\delta_{\min})$	30802	31104	31368	31977	59191	59763	60420	61575	160259	161836	163380	166651
$f^s(\delta_{\min}[7])$	30802	31063	31548	32515	59324	59780	60736	61872	160216	161836	164592	167438

Table 4. Finding the quantity $\delta_{\min}$ when $0 < a_j \le 999$, $0 < c_j \le 999$, $n = (1000, 2000, 5000)$

n	1000				2000				5000		
b	146301				304417				740774		
f^s	325760				633548				1570299		
$p\%$	1	2	3	5	1	2	3	5	1	2	5
$\Delta = \left[f^* \cdot \frac{p}{100} \right]$	3257	6515	9772	16288	6335	12670	19006	31677	1563	3127	4690
$\delta_{\min}$	2978	5961	9046	15135	6121	12285	18591	31317	15102	30316	45680
$\delta_{\min}[7]$	2999	5999	10667	16002	6242	12485	22197	33295	15190	30382	54014
$f^s(\delta_{\min})$	329019	332299	335546	342056	639884	646218	652580	665225	1586017	1601719	1617409
$f^s(\delta_{\min}[7])$	329065	332313	337302	342916	640030	646421	656167	667151	1586079	1601778	1625800

In the tables we accepted the following denotations :

n is the number of unknowns,

b is the rigth side of problem (1.2),

f^s is the value that the suboptimal solution gives to function (1.1) in problem (1.1) – (1.3),

$p\%$ is denotes the increase percent of the quantity f^s

$\Delta = \left[f^* \cdot \frac{p}{100} \right]$ is denotes the increase amount of the quantity f^s,

$\delta_{\min}$ is the minimal mount of the quantity δ in problem (2.1) – (2.4) found by the method in this paper,

$\delta_{\min}[7]$ is the minimal mount of the quantity δ found by the method of the paper [7].

$f^s(\delta_{\min})$ is the value of function (1.1) for the guaranteed suboptimal solution after $\delta_{\min}$ growth of the right side of b in problem (1.1) – (1.3),

$f^s(\delta_{\min}[7])$ is the value of suboptimal solution given to function (1.1) established for the quantity $\delta_{\min}$ found by the method of [7].

V. Conclusions

From the tables it is seen that in 36 from the solved 47 problems the minimal value of the quantity $\delta_{\min}$ given by the method of this paper is less than one in the paper [7]. In 4 problems the minimum values of the quantity $\delta_{\min}$ are equal. Only in [7] problems the result of the paper [7] is best. Thus , the method given in our paper may be considered more qualitative than the method given in [7].

From the tables given on the base of the conducted experiments we see that in the 77% of the solved problems the results of our method is best. On the other hand, as it is seen from the tables , by increasing the coefficients of the problems, the result of the method of this paper growth compared to the result of [7]. This shows that the suggested method is more important from the practical point of view.

References

[1] Garey M.R. and Johnson D.S Computing machines and hardly solved problems.Mir 1982 (in Russian)

[2] Kellerer H. , Pferschy U., Pisinger D. Knapsack problems. Berlin-Heidelberg : Springer-Verlag, 2004, pp.546.

[3] Martello S., Toth P. Knapsack problems. Algorithm and Computers Implementations. J.Willey & Sons: New York Chichester, 1990, pp.296.

[4] Babayev G.A., Mamedov K.Sh., Mektiyev M.Q. Constructions methods of suboptimal solution of multidimensional knapsack problem Zh.V.M & MF 1978, №6,pp. 1443-1453.(in Russian)

[5] Kovalev M.M, Discrete Optimization (integer programming) M.URSA 2003, 191 pp. (in Russian)

[6] Sigal I.Kh., Ivanova.A.P Introduction to applied discrete programming M.FIZMATLIT 2003. (in Russian)

[7] Mammadov N.N. Finding the guaranteed solution in the knapsack problem at the expense of minimal growth of the right side. Proceeding of scientific Conference of pest graduate students of ANAS-Baku,"ELM",2010, pp.93-96 (in Azerbaijan)

[8] Mammadov K.Sh., Mammdov N.N. The notion of guaranteed suboptimal solution for the knapsack problem and the method for its determination. (in Azerbaijan)