1. Introduction

In general the linear integer programming problem has very important real life applications. The general linear integer problem comes in the form of capital budgeting, transportation, traveling salesman, facility location, scheduling, knapsack etc. This model even though it is very easy to model mathematically, has proved to be very difficult to solve. See [1–5] for more on linear integer models.

programing (LP) model. If the linear programming optimal solution is integer then, the optimal solution to the integer problem is available. If the linear programming optimal solution is not integer, then a variable with a fractional value is selected to create two sub-problems such that part of the feasible region is discarded without eliminating any of the feasible integer solutions. The process is repeated on all variables with fractional values until an integer solution is found. In this approach variable sum and additional constraints are generated and added to the original problem before solving. In order to do this the objective bound of knapsack problem is quickly determined. The bound is then used to generate a set of variable sum limits and four additional constraints. From the variable sum limits, initial sub-problems are constructed and solved. The optimal solution is then obtained as the best solution from all the sub-problems in terms of the objective value. The proposed procedure results in sub-problems that have reduced complexity and easier to solve than the original problem in terms of numbers of branch and bound iterations or sub-problems.

Knapsack problem reformulation is not a new idea. The reformulation approaches were once used to solve some knapsack and other problems [8, 9]. The knapsack problem is a special form of the general linear integer problem. There are so many types of knapsack problems. These include the zero-one, multiple, multiple-choice, bounded, unbounded, quadratic, multi-objective, multi-dimensional, collapsing zero-one and set union knapsack problems. The zero-one knapsack [10, 11] problem is one in which the variables assume 0s and 1s only. The reason being that an item can be chosen or not chosen. In other words there is no way it is possible to have fractional amounts or items. This is the easiest class of the knapsack problems and is the only one that can be solved in polynomial by interior point algorithms and in pseudo-polynomial time by dynamic programming approaches. The multiple-choice knapsack problem is a generalization of the ordinary knapsack problem, where the set of items is partitioned into classes. The zero-one choice of taking an item is replaced by the selection of exactly one item out of each class of items.

The multiple knapsack problem [12–14] is a generalization of the standard knapsack problem formed by combining single knapsacks into a group of knapsacks having different capacities. In this case the objective is to assign each item to at most one of the knapsacks in such a way that all capacity constraints are satisfied and that the total profit of all the items put into knapsacks is made maximum. In the bounded knapsack problem [15] there is a knapsack capacity and a set of items, each having a positive integer value, a positive integer weight, and a positive integer limit or bound on its availability. With the bounded knapsack problem the main objective is to select the number of each item type to add to the knapsack in such a way that the total weight is not violated and that the total value is a maximum.

2. Literature review and problem statement

In this case of the unbounded knapsack problem, types of items of different values and volumes are given, then it is required to find the most valuable set of items that fit in a knapsack of fixed volume. The main difference with the bounded knapsack problem is that the number of items of each type is unbounded. A quadratic knapsack [16–19] is a knapsack problem whereby the objective is expressed as a quadratic function subject to a set of linear constraints. The variables in this knapsack problem can either be zero-one or general integers. With the multi-objective knapsack, the objective changes from a single objective into many objectives within the same problem. For example in agriculture, there is an objective to maximize profit and at the same time minimizing transportation costs and maximizing the number of employees. In multi-objective [20, 21] knapsack problems there is the dilemma of dealing with environmental, social, political and or economic concerns. In the multidimensional knapsack problem, several dimensions are considered in the formulation of the problem. The multidimensional knapsack [22–25] problem basically consists of finding a subset of objects that maximizes the total profit while observing some capacity restrictions.

The collapsing zero-one knapsack problem is a type of non-linear knapsack problem in which the knapsack size is a non-increasing function of the number of items included. The set-union knapsack [26] problem is a variation of the zero-one knapsack problem in which each item is a set of elements, each item has a nonnegative value, and each element has a nonnegative weight. The weight of one item is given by the total weight of the elements in the union of the items' sets.

The branch and bound was the first algorithm to be developed in 1960 [6] for these linear integer models. This method was further modified in 1965 to solve the mixed linear integer problem [7]. So many improvements have been done on the branch and bound algorithm in terms of addition of cuts to get the branch and cut algorithm [19, 27–29]. Pricing was introduced within the context of branch and bound to get the branch and price algorithm [30, 31]. The improved versions, branch and cut and branch and price were also combined to get the branch, cut and price [32–34]. In addition to using cuts and pricing within the context of a branch and bound algorithm, preprocessing can reduce the number of sub-problems needed to verify optimality. Even with all these efforts the general linear integer is still very difficult to solve. In fact the general linear integer problem including the knapsack problem is NP hard [10, 18, 19, 26, 35, 36] and there are not aware of any consistent efficient algorithm for these problems. These difficult problems and include the knapsack problem which is a special case with only one constraint.

The proposed algorithm has the advantage that it is parallelizable and independent processors can be used. The knapsack problem has so many real life applications. These include home energy management, cognitive radio networks, mining operations use, relay selection in secure cooperative wireless communication, electrical power allocation management, production planning, in selection of renovation actions, waste management, formulation and solution method for tour conducting and optimization of content delivery networks. Network of electricity that intelligently integrates the users’ concerns. In the multidimensional knapsack problem, there is a knapsack capacity and a set of items, each having a positive integer value, a positive integer weight, and a positive integer limit or bound on its availability. With the bounded knapsack problem the main objective is to select the number of each item type to add to the knapsack in such a way that the total weight is not violated and that the total value is a maximum.
Knapsack problem formulation is used in channel and power allocation for cognitive radio (CR) networks. In this formulation it is assumed that the total available spectrum is divided into several bands, each consisting of a group of channels. A centralized base station, enabled by spectrum sensing, is assumed to have the knowledge of all vacant channels, which will be assigned to various CRs according to their requests. In this case the objective of resource allocation is to maximize the sum data rate of all CRs.

An extension of the precedence constrained knapsack problem where the knapsack can be filled in multiple periods has applications in the mining operations. This problem formulation is known in the mining industry as the open-pit mine production scheduling problem. Both exact and heuristics are used in solving the LP relaxation of this problem.

Knapsack formation is used in cooperative jamming. Cooperative jamming schemes support secure wireless communication in the presence of eavesdroppers. Large numbers of cooperative relays provide better secrecy rate while increasing the communication ad synchronization needs associated with cooperative beam forming.

When renovating a building structure there is a need for the construction manager to select the most feasible renovation activities and the order in which they must be done. The main challenge of renovating a building structure is to determine whether to renovate the existing structure or start building a new building. Such a decision requires use of decision tools such as knapsack modeling.

Waste is another challenge that may cause serious environmental damage if not properly managed. Plastic and paper waste is a serious issue in most developing countries. The many problem of waste in developing countries is that there is very low level recycling in these countries. In other words there is a small amount of plastic and paper waste that is recycled and the rest is sent to the landfills or dumbed on the streets. For recycling to be financially profitable there is need to use effective and efficient ways in selecting items to be produced from a lot of items given a limited amount of money and other resources. This is where knapsack modelling is applied to minimize waste management costs.

Land conservation projects require proper managing and planning for the benefits to be seen. For these land projects there is always the dilemma of how to select the most profitable land projects subject to financial constraints. A multiple knapsack formulation is employed in making such decisions and it outperforms other decision making tools such as benefit targeting, cost-effectiveness analysis, and sequential binary integer programming.

The fast growing populations and introduction of healthcare systems have resulted in increased both inpatients and outpatients to public hospitals, particularly those hospitals that provide special and comprehensive health services in large countries such as China and India. The hospitals in these countries have huge numbers of both inpatients and outpatients. The huge numbers of patients result in overcrowding and these overcrowding conditions are a concern for the hospital managers. The obvious question is how to manage these huge numbers of patients effectively given the fact that some patients require less attention than the others, the lengths of some patients are predictable than the others. In other words those who require less clinical care are less likely to stay longer at the hospital. In order to alleviate the challenge of overcrowding, a multi-criteria knapsack model is used for disease selection in the reception or observation ward of the public hospitals.

So many studies have been done in the field of optimization methods. Unfortunately, sometimes it is not possible to directly apply these optimization methods to practical problems. As an example, a tourist deciding on a traveling schedule within a traveling time limit needs to select travel tourist spots from lists so as to be satisfied as far as possible. This is a problem that can’t be solved by the available conventional methods. This problem is formulated as a tour conducting knapsack problem. The formulation and solution method of tour conducting knapsack problem are based on those of traveling salesman problem and knapsack problem.

The knapsack problem has many very important applications in so many areas of business and engineering and it is certainly very necessary to develop efficient solution algorithms for it. Most of the algorithms for the knapsack problem are branch and bound based and in this paper the branch and bound algorithm is improved for the knapsack problem.

The knapsack problem has so many applications [37–45] and there is definitely a need for efficient and consistent algorithms for this problem. Some of the applications are home energy management [37], cognitive radio networks [38], mining operation use [39], relay selection in secure cooperative wireless communication [40], electrical power allocation management [41], production planning [42], in selection of renovation actions [43], waste management [44], formulation and solution method for tour conducting and optimization of content delivery networks [45].

Even though there is a lot of effort from researchers to develop an efficient and consistent solution such a method does not exist. The knapsack problem is NP hard [10, 18, 19, 26, 35, 36] and an optimal solution is very difficult to obtain. For example in [10] a parallel algorithm for solving the NP-complete Knapsack Problem was proposed. NP complete is the most difficult subset of the NP hard problems. In [18] it is pointed out that the knapsack problem is an NP-hard optimization problem with so many diverse applications in industrial and management engineering, however, computational complexities associated with this problem still remain in the knapsack problem. In [19] it is also made very clear that the knapsack problem is a well-known NP-hard combinatorial optimisation problem, with many practical applications. Even up to now, approximation methods are still being developed [13, 25, 26, 35, 45] for this problem. The reason for using heuristics is that there are no efficient consistent exact methods for the knapsack problem. In [35] it is clarified that because of the high computational complexity of knapsack problem, three heuristic approaches are proposed. The paper [26] is a recent heuristic which shows that exact efficient approaches for this knapsack problem are not available.

The approximated solution for the knapsack problem is easy to obtain and good for quick decisions but the difference between the approximated solution and the exact solution may be in millions of dollars for large projects such the UN humanitarian projects and the US military operations. There is a need for exact methods for the knapsack problem.

3. The aim and objectives of the study

The aim of the study is to reformulate the knapsack problem given in so that it is easier to solve by branch and bound algorithm. To achieve the set aim the following tasks have been solved:

- to determine the objective bound $Z^*$;
– to use the objective bound to generate the variable sum limits \( \ell_k, \ell_k', \ldots, \ell_k' \) and additional constraints;
– to construct the \( k \) initial parallel sub-problems;
– to illustrate by an example how to reformulate a knapsack;
– to give classes ad examples of difficult knapsack problems.

4. The knapsack linear integer problem

4.1. General form of knapsack problem

The knapsack linear integer problem is a special case of the general integer problem. Even though this integer problem has only one constraint, it is believed to be NP complete and very difficult to solve.

Minimize \( Z = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \)

such that:

\[
\begin{align*}
& a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \geq b, \\
& x_j \text{ is integer.}
\end{align*}
\]  

(1)

4.2. Totally unimodular transportation matrix

The constraints of any linear integer problem can be expressed as (2).

\[ AX = B, \]

(2)

where \( A \) is the transportation coefficient matrix.

**Theorem 1:** Matrix \( A \) is totally unimodular if the determinant of each square submatrix of is \(-1\) or \(+1\).

**Theorem 2:** If matrix \( A \) is totally unimodular, then every vertex solution of (2) is integral.

**Proof of 1 & 2.** Note that every column of \( A \) has exactly two 1’s, thus any column of \( A_k \) has either:

1) two 1’s;
2) only one 1;
3) exactly No. 1.

If \( A_k \) contains a column that has No. 1, then clearly \( \text{det}[A_k] = 0 \) and done for (i). Thus now assume that every column of \( A_k \) contains at least one 1. There are two cases that must be considered here. The first case is where every column of \( A_k \) contains two 1’s. Then one of the 1’s must come from the source rows and the other one must come from the destination rows. Hence subtracting the sum of all source rows from the sum of all destination rows in \( A_k \) will give the zero vector. Thus the row vectors of \( A_k \) are linearly dependent. Hence \( \text{det}[A_k] = 0 \).

What is now left is to consider the case where at least one column of \( A_k \) contains exactly one 1. By expanding \( A_k \) with respect to this column, let’s have \( \text{det}[A_k] = \pm \text{det}[A_{k-1}] \) where the sign depends on the indices of that particular 1. Now the theorem is proved by repeating the argument to matrix \( A_{k-1} \).

Therefore the matrix \( A_k \) is totally unimodular. More on unimodular matrices can be found in [46].

The variable sum inequalities constructed in this chapter have zeros and ones as the only coefficients. Making the coefficient of every linear integer problem unimodular is a very difficult task. In this paper let’s rely on the strategy of introducing new constraints to the knapsack problem with only zeros (0s) and ones (1s) as coefficients. This does not make the knapsack problem unimodular but makes the problem easier to solve than the original form.

4.3. Branch and Bound Algorithm

The branch and bound algorithm in general relies on the usual strategy of first relaxing the integer problem into a linear programing (LP) model. If the linear programming optimal solution is integer then, the optimal solution to the integer problem is available. If the linear programming optimal solution is not integer, then a variable with a fractional value is selected to create two sub-problems such that part of the feasible region is discarded without eliminating any of the feasible integer solutions. The process is repeated on all variables with fractional values until an integer solution is found. The worst case complexity of the branch and bound algorithm on knapsack linear integer models is NP Complete. The number of sub-problems can easily reach levels that are not manageable.

4.4. Variable sum equality

A constraint of the form \( x_1 + x_2 + \ldots + x_n = \ell \), where \( \ell \) is an integer, is called a variable sum equality. Let’s note the coefficients are only ones and this equality is not new and has been used as clique inequality in the general integer programing. Variable sum equalities can be generated for (1). Let \( \ell_i = \text{SINT} \geq \frac{b}{c_i} \) where \( \text{SINT} \) stands for the smallest integer. The objective bound \( Z^0 \) can be found as (3) and can be expressed as (4).

\[
Z^0 = \min \left[ c_1 x_1^0, c_2 x_2^0, \ldots, c_n x_n^0 \right].
\]

(3)

\[
c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \leq Z^0.
\]

(4)

The variable sum bounds \( (\ell_i, \ell_i') \) which are integers can now be determined once the objective bound is known. These two integral bounds satisfy (5).

\[
\ell_i \leq x_1 + x_2 + \ldots + x_n \leq \ell_i'.
\]

(5)

The two variable sum bounds may be found by solving the following two linear programing models (6), (7).

Maximize \( \ell_i = x_1 + x_2 + \ldots + x_n \). Such that:

\[
\begin{align*}
& a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \geq b, \\
& c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \leq Z^0.
\end{align*}
\]

(6)

where \( x_j \) is integer.

Minimize \( \ell_i = x_1 + x_2 + \ldots + x_n \). Such that:

\[
\begin{align*}
& a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \geq b, \\
& c_1 x_1 + c_2 x_2 + \ldots + c_n x_n \leq Z^0.
\end{align*}
\]

(7)

where \( x_j \) is integer.

The variable sum equality was used recently in [17] to improve the optimality verification process. If there are parallel processors then these can be solved at the same time, otherwise these can be solved as a combined problem given in (8).

Let’s maximize \( \ell_i - \ell_i' \). Such that:

\[
\begin{align*}
& a_1 y_1 + a_2 y_2 + \ldots + a_n y_n \geq b, \\
& c_1 y_1 + c_2 y_2 + \ldots + c_n y_n \leq Z^0,
\end{align*}
\]

(8)

where \( x_j, y_j \geq 0 \) are the unknown variables.

In this case the \( y \) variables are used for the second problem.
4.5. Initial branches

Once the variable sum bounds have been determined then the variable sum constraints can now be constructed as given in (9). From \( \ell_i \leq x_1 + x_2 + \ldots + x_n \leq \ell_j \), let’s have \( k \) equality constraints, i.e.,
\[
\begin{align*}
x_1 + x_2 + x_3 + \ldots + x_n &= \ell_1, \\
x_1 + x_2 + x_3 + \ldots + x_n &= \ell_2, \\
x_1 + x_2 + x_3 + \ldots + x_n &= \ell_3, \\
& \vdots \\
x_1 + x_2 + x_3 + \ldots + x_n &= \ell_k,
\end{align*}
\]
(9)
Each variable sum equality is an initial branch for the branch and bound procedure which imply that the knapsack problem has \( k \) initial branches to be explored. The branches are shown in Fig. 1, the \( k \) initial branches of the proposed in and illustrated in Fig. 1, can be explored independently thus allowing the use of the much needed parallel processors.

4.6. Two additional constraints

Two additional binding constraints can be constructed and added to the original knapsack problem so that the complexity is reduced further. If the variable giving the objective bound is \( x_j \), then an additional variable \( x_{\text{not}} \) can be introduced such that
\[
x_j + x_{\text{not}} = \ell_i, \quad x_{\text{not}} = x_1 + x_2 + \ldots + x_n,
\]
i.e.,
\[
x_1 + x_2 + \ldots + x_n - x_{\text{not}} = 0.
\]
(11)
The variable \( x_j \) is excluded in the sum of variables (11). Let’s note that (11) is obtained by rearranging the variables and that the two constraints (10) and (11) are made up of only (0 s) and (±1 s) as the coefficients. The addition of these two constraints to each branch will significantly reduce the complexity of the problem.

5. Reformulation procedure for the knapsack linear integer problem

5.1. Numerical illustration

Let’s minimize:
\[
Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7,
\]
Such that:
\[
165x_1 + 45x_2 + 33x_3 + 279x_4 + 69x_5 + 6x_6 + 122x_7 \geq 18773,
\]
(12)
where \( x_j \geq 0 \) and integer \( \forall j \).

The branch and bound algorithm takes 1351 sub-problems to verify the optimal solution: \( x_1 = 568, \ x_2 = 5, \ x_3 = 5, \ x_4 = x_5 = x_6 = x_7 = 0 \) & \( Z = 14793 \). This is a very small problem and the 1,351 sub-problems used to verify the optimal solution is too much.

There is definitely a need to preprocess the knapsack linear problem before solving it by the branch and bound method. In this paper there are variable sum constraints and additional constraints and add them to the original problem and then solve,

5.2. Reformulation procedure

Given any knapsack linear integer problem of the form,
\[
\begin{align*}
& \text{maximize or minimize } Z = c_1x_1 + c_2x_2 + \ldots + c_nx_n, \\
& \text{subject to } \begin{cases}
ax_1 + bx_2 + \ldots + nx_n = b, \\
nx_1 + nx_2 + \ldots + nx_n = \ell, \\
x_1 + x_2 + \ldots + x_n - x_{\text{not}} = 0.
\end{cases}
\end{align*}
\]
where \( x_j \) is integer.

An objective bound \( (Z^*) \), variable sum limits \( (\ell_i, \ell_2, \ldots, \ell_k) \) and the two additional constraints as \( x_1 + x_{\text{not}} = \ell_i \) and \( x_{\text{not}} = x_1 + x_2 + \ldots + x_n \) can be determined. The initial \( k \) sub-problems generated are:

\begin{itemize}
\item \text{Initial sub-problem 1.}
\end{itemize}

Let’s minimize:
\[
Z = c_1x_1 + c_2x_2 + \ldots + c_nx_n,
\]
Such that:
\[
\begin{align*}
a_1x_1 + a_2x_2 + \ldots + a_nx_n &\geq b, \\
c_1x_1 + c_2x_2 + \ldots + c_nx_n &\leq Z^*, \\
x_1 + x_2 + \ldots + x_n &= \ell, \\
x_1 + x_2 + \ldots + x_n - x_{\text{not}} &= 0.
\end{align*}
\]
\begin{itemize}
\item \text{Initial sub-problem 2.}
\end{itemize}

Let’s minimize:
\[
Z = c_1x_1 + c_2x_2 + \ldots + c_nx_n.
\]
Such that:
\[
\begin{align*}
a_1x_1 + a_2x_2 + \ldots + a_nx_n &\geq b, \\
c_1x_1 + c_2x_2 + \ldots + c_nx_n &\leq Z^*, \\
x_1 + x_2 + \ldots + x_n &= \ell, \\
x_1 + x_2 + \ldots + x_n - x_{\text{not}} &= 0.
\end{align*}
\]
\begin{itemize}
\item \text{Initial sub-problem k.}
\end{itemize}

Let’s minimize:
\[
Z = c_1x_1 + c_2x_2 + \ldots + c_nx_n.
\]
Such that:
\[
\begin{align*}
a_1x_1 + a_2x_2 + \ldots + a_nx_n &\geq b, \\
c_1x_1 + c_2x_2 + \ldots + c_nx_n &\leq Z^*, \\
x_1 + x_2 + \ldots + x_n &= \ell, \\
x_1 + x_2 + \ldots + x_n - x_{\text{not}} &= 0.
\end{align*}
\]
The \( k \) initial sub-problems can be solved independently and the optimal solution is the best solution (in terms of objective value) from the \( k \) sub-problems.

5. 3. Algorithm

In other words the knapsack linear integer problem is solved using the following steps.

**Step 1:** Determine the objective bound \( Z^0 \).

**Step 2:** Use the objective bound to generate the variable sum limits \( \ell_i, \ell_2, \ldots, \ell_k \) and additional constraints.

**Step 3:** Construct the \( k \) initial sub-problems.

**Step 4:** Solve the \( k \) sub-problems to obtain the optimal solution as the best solution from the \( k \) sub-problems in terms of the objective value.

5. 4. Using the numerical illustration from 5. 1

Let’s minimize:

\[
Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7.
\]

Such that:

\[
1654x_1 + 53x_2 + 3x_3 + 279x_4 + 69x_5 + 6x_6 + 122x_7 \geq 18,773,
\]

where \( x_j \geq 0 \) and integer \( \forall j \).

**Step 1.**

\[
x_j^0 = 289418569682733129154.
\]

It should be a solid formula:

\[
Z^0 = \min \left[ \frac{46,818,15,884,14,794}{20,468,23,751,15,645,21,098} \right] = 14,794. \tag{13}
\]

**Step 2.**

Let’s minimize \( \ell_i = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \).

Such that:

\[
165x_1 + 45x_2 + 33x_3 + 279x_4 + 69x_5 + 6x_6 + 122x_7 \geq 18,773,
\]

\[
162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7 \leq 14,794,
\]

\[
x_j \geq 0.
\]

\[
\ell_i \leq 578.33 \;
\ell_i = 578.
\]

Let’s minimize \( \ell_i = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \).

Such that:

\[
165x_1 + 45x_2 + 33x_3 + 279x_4 + 69x_5 + 6x_6 + 122x_7 \geq 18,773,
\]

\[
162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7 \leq 14,794,
\]

\[
x_j \geq 0.
\]

\[
\ell_i \leq 568.43 \;
\ell_i = 569.
\]

i. e.

\[569 \leq x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 \leq 578.\] \tag{14}

i. e.

\[
\ell_i = 569, \; \ell_2 = 570, \; \ell_3 = 571, \; \ell_4 = 572, \; \ell_5 = 573, \; \ell_6 = 574, \; \ell_7 = 575, \; \ell_8 = 576, \; \ell_9 = 577 & \ell_{10} = 578. \tag{15}
\]

The general two additional constraints are:

\[
x_j + x_k = \ell_i, \forall i = 1,2, \ldots, 9.
\]

\[
x_k = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7.
\]

**Step 3.**

The 9 initial sub-problems are:

1) Sub-problem 1.

Let’s minimize:

\[
Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7.
\]

Such that:

\[
165x_1 + 45x_2 + 33x_3 + 279x_4 + 69x_5 + 6x_6 + 122x_7 \geq 18,773,
\]

\[
162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7 \leq 14,794,
\]

\[
x_j + x_k + x_1 + x_2 + x_3 + x_5 + x_6 + x_7 = 569,
\]

\[
x_k = 569,
\]

\[
x_k = x_1 + x_2 + x_3 + x_5 + x_6 + x_7.
\]

2) Sub-problem 2.

Let’s minimize:

\[
Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7.
\]

Such that:

\[
165x_1 + 45x_2 + 33x_3 + 279x_4 + 69x_6 + 122x_7 \geq 18,773,
\]

\[
162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7 \leq 14,794,
\]

\[
x_j + x_k + x_1 + x_2 + x_3 + x_5 + x_6 + x_7 = 569,
\]

\[
x_k = x_1 + x_2 + x_3 + x_5 + x_6 + x_7.
\]

3) Sub-problem 3.

Let’s minimize:

\[
Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7.
\]
Such that:

\[ 165x_1 + 45x_2 + 33x_3 + 279x_4 + 
+ 69x_5 + 6x_6 + 122x_7 \geq 18,773, \]

\[ 162x_2 + 38x_3 + 26x_5 + 301x_4 + 
+ 87x_7 + 5x_8 + 137x_9 \leq 14,794, \]

\[ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 571, \]

\[ x_1 + x_2 = 571, \]

\[ x_6 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7. \]

4) Sub-problem 4.
Let’s minimize:

\[ Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7. \]

Such that:

\[ 165x_1 + 45x_2 + 33x_3 + 279x_4 + 
+ 69x_5 + 6x_6 + 122x_7 \geq 18,773, \]

\[ 162x_2 + 38x_3 + 26x_5 + 301x_4 + 
+ 87x_7 + 5x_8 + 137x_9 \leq 14,794, \]

\[ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 572, \]

\[ x_2 + x_3 = 572, \]

\[ x_6 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7. \]

5) Sub-problem 5.
Let’s minimize:

\[ Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7. \]

Such that:

\[ 165x_1 + 45x_2 + 33x_3 + 279x_4 + 
+ 69x_5 + 6x_6 + 122x_7 \geq 18,773, \]

\[ 162x_2 + 38x_3 + 26x_5 + 301x_4 + 
+ 87x_7 + 5x_8 + 137x_9 \leq 14,794, \]

\[ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 573, \]

\[ x_3 + x_4 = 573, \]

\[ x_6 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7. \]

6) Sub-problem 6.
Let’s minimize:

\[ Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7. \]

Such that:

\[ 165x_1 + 45x_2 + 33x_3 + 279x_4 + 
+ 69x_5 + 6x_6 + 122x_7 \geq 18,773, \]

\[ 162x_2 + 38x_3 + 26x_5 + 301x_4 + 
+ 87x_7 + 5x_8 + 137x_9 \leq 14,794, \]

\[ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 574, \]

\[ x_3 + x_4 = 574, \]

\[ x_6 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7. \]

7) Sub-problem 7.
Let’s minimize:

\[ Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7. \]

Such that:

\[ 165x_1 + 45x_2 + 33x_3 + 279x_4 + 
+ 69x_5 + 6x_6 + 122x_7 \geq 18,773, \]

\[ 162x_2 + 38x_3 + 26x_5 + 301x_4 + 
+ 87x_7 + 5x_8 + 137x_9 \leq 14,794, \]

\[ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 575, \]

\[ x_3 + x_4 = 575, \]

\[ x_6 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7. \]

8) Sub-problem 8.
Let’s minimize:

\[ Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7. \]

Such that:

\[ 165x_1 + 45x_2 + 33x_3 + 279x_4 + 
+ 69x_5 + 6x_6 + 122x_7 \geq 18,773, \]

\[ 162x_2 + 38x_3 + 26x_5 + 301x_4 + 
+ 87x_7 + 5x_8 + 137x_9 \leq 14,794, \]

\[ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 576, \]

\[ x_3 + x_4 = 576, \]

\[ x_6 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7. \]

9) Sub-problem 9.
Let’s minimize:

\[ Z = 162x_1 + 38x_2 + 26x_3 + 301x_4 + 87x_5 + 5x_6 + 137x_7. \]

Such that:

\[ 165x_1 + 45x_2 + 33x_3 + 279x_4 + 
+ 69x_5 + 6x_6 + 122x_7 \geq 18,773, \]

\[ 162x_2 + 38x_3 + 26x_5 + 301x_4 + 
+ 87x_7 + 5x_8 + 137x_9 \leq 14,794, \]

\[ x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 = 577, \]

\[ x_3 + x_4 = 577, \]

\[ x_6 = x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7. \]
Such that:

\[
\begin{align*}
165x_1 + 45x_2 + 33x_3 + 279x_4 + \\
+69x_5 + 6x_6 + 122x_7 & \geq 18,773, \\
162x_1 + 38x_2 + 26x_3 + 301x_4 + \\
+87x_5 + 5x_6 + 137x_7 & \leq 14,794, \\
x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 & = 578, \\
x_3 + x_4 & = 578, \\
x_8 = x_1 + x_2 + x_3 + x_5 + x_6 + x_7.
\end{align*}
\]

The initial branches and their corresponding number of sub-problems are given in Table 1.

<table>
<thead>
<tr>
<th>Sub-problem</th>
<th>Number of sub-problems</th>
<th>Solution</th>
<th>Optimal solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7</td>
<td></td>
<td>14794, x_3 = 569</td>
</tr>
<tr>
<td></td>
<td></td>
<td>x_1 = x_2 = x_4 = x_5 = x_6 = x_7 = 0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>Infeasible</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>Infeasible</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>Infeasible</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>14793, x_3 = 568, x_4 = 5, x_1 = x_2 = x_5 = x_6 = x_7 = 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Best &amp; optimal</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>Infeasible</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>Infeasible</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>Infeasible</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>Infeasible</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>Infeasible</td>
<td></td>
</tr>
</tbody>
</table>

The automated branch and bound algorithm takes only 5 sub-problems to verify the optimal solution: x_3 = 568, x_4 = 5, x_1 = x_2 = x_5 = x_6 = x_7 = 0 & Z = 14793, in the initial parallel problem 5.

### 6. Some difficult classes of knapsack problems

#### 6.1. Knapsack binary linear problems with bizarre behaviour

Let’s maximize:

\[ Z = \sum_{j=1}^{n} x_j \]

or Minimize \( x_n \).

Such that:

\[ 2\sum_{j=1}^{n} x_j \pm x_n = n - 1, \tag{20} \]

where \( x_j = 0 \) or \( 1 \) \( \forall j \) and \( n \) is even.

The complexity of the problem is given in Table 2. The number of sub-problems increases exponentially as \( n \) increases.

### Complexity of the problem as \( n \) increases

<table>
<thead>
<tr>
<th>Value of ( n ) in model</th>
<th>Number of sub-problems created by the branch and bound approach to reach the optimum solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>39</td>
</tr>
<tr>
<td>8</td>
<td>139</td>
</tr>
<tr>
<td>16</td>
<td>25,739</td>
</tr>
<tr>
<td>32</td>
<td>Number of sub-problems exceeds 30,000</td>
</tr>
</tbody>
</table>

This shows that the branch and bound on its own is not a very good approach.

#### 6.2. Second class of bizarre knapsack problems

This is a modification of Class 7.1. Class 7.1 and the general form is given in (21).

Let’s maximize:

\[ Z = \sum_{j=1}^{n-1} x_j \]

or Minimize \( x_n \).

Such that:

\[ 2\sum_{j=1}^{n-1} x_j \pm x_n = n - 1, \tag{21} \]

where \( x_j = 0 \) or \( 1 \) \( \forall j \), \( 3 \leq \kappa \leq n - 1 \), \( \kappa \) is odd and \( n \) is even.

The bizarre behaviour of the branch and bound method is given in Table 3.

### Complexity of the problem as \( n \) increases

<table>
<thead>
<tr>
<th>Value of ( n ) in model</th>
<th>Number of sub-problems created by the branch and bound approach to reach the optimum solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>( \kappa = 3 ), sub-problems=17</td>
</tr>
<tr>
<td>6</td>
<td>( \kappa = 3 ), sub-problems=59</td>
</tr>
<tr>
<td>8</td>
<td>( \kappa = 5 ), sub-problems=59</td>
</tr>
<tr>
<td>16</td>
<td>( \kappa = 3 ), sub-problems=209</td>
</tr>
<tr>
<td>16</td>
<td>( \kappa = 5 ), sub-problems=209</td>
</tr>
<tr>
<td>16</td>
<td>( \kappa = 7 ), sub-problems=209</td>
</tr>
<tr>
<td>16</td>
<td>Sub-problems exceeded 30,000</td>
</tr>
</tbody>
</table>

These are small problems and the branch is not expected to struggle to solve these problems.

#### 6.3. Third class of knapsack problem pure integer case

Changing of variables from binary to pure integer in any difficult knapsack problem automatically increases the complexity of the problem.

Let’s maximize:

\[ Z = \sum_{j=1}^{n} x_j \]

or Minimize \( x_n \).
Mathematics and cybernetics – applied aspects

Such that:

\[ 2 \sum_{j=1}^{n} x_j + \kappa x_n = n - 1, \quad (22) \]

where \( x_j \geq 0, \) integer \( \forall j, \ 1 \leq \kappa \leq n - 1, \) \( \kappa \) is odd and \( n \) is even.

The behaviour of the branch and bound method for \( n = 4, 6, 8 \) and 16 for this class of difficult problems is given in Table 4.

<table>
<thead>
<tr>
<th>Value of ( n ) in model</th>
<th>Number of sub-problems created by the branch and bound approach to reach the optimum solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>( \kappa = 3, ) sub-problems = 23</td>
</tr>
<tr>
<td>6</td>
<td>( \kappa = 3, ) sub-problems = 129</td>
</tr>
<tr>
<td>8</td>
<td>( \kappa = 5, ) sub-problems = 755</td>
</tr>
<tr>
<td>16</td>
<td>Number of the sub-problems exceed 30,000</td>
</tr>
</tbody>
</table>

Changing from binary to general integer means expanding the problem. The number of sub-problems increases and this is expected.

6. 4. Fourth class of hard knapsack problems

Let’s maximize:

\[ Z = \sum_{j=1}^{n} x_j \]

or Minimize \( x_n. \)

Such that:

\[ 2 \sum_{j=1}^{n} x_j + \kappa x_n = \lambda, \quad (23) \]

where \( x_j \geq 0, \) integer \( \forall j, \ 2(n - 1) + k = \lambda, \) \( k, \lambda \geq 0 \) are odd and \( n \) is even.

The standard branch and bound method can’t solve most of these for large values of \( \kappa. \) For example a knapsack problem with the parameters, \( n = 4, \ k = 91, \ \lambda = 97, \) explodes to an unmanageable number of sub-problems.

Let’s minimize: \( Z = x_n. \) Such that:

\[ 2x_1 + 2x_2 + 2x_3 + 91x_4 = 97, \quad (24) \]

where \( x_j \geq 0, \) integer \( \forall j. \)

The branch and bound method requires 7449 sub-problems to verify the optimal solution. For large values of \( \lambda \) the knapsack problems are very difficult to solve by the standard branch and bound algorithm on its own. These numerical illustrations and more on complexity of knapsack problems and other linear integer models are given in [9].

The numerical illustration is given in Section 5. The branch and bound on its own took 1351 sub-problems to verify optimality. The same knapsack problem is reformulated into 10 parallel problems which can be solved independently as given in Section 5. 4. The reformulated same knapsack problem is solved by the branch and bound algorithm. The numbers of iterations required to verify optimality ranges from 3 to 11 for the parallel problems as given in Table 1. Reducing complexity from 1,351 to the worst case of 11 is a very significant improvement. The branch and bound algorithm is a general purpose algorithm for solving the general linear integer problem. Unfortunately this approach on its own has serious weaknesses as presented in Section 6 from Tables 2 to 4.

The reformulated knapsack can be identified by the following features. The new knapsack problem is split into several independent parallel problems. The number of constraints increases from 1 to 5 for each parallel. The 4 new constraints for each parallel problem include an objective bound. Splitting the problem into many parallel problems, increasing the number of constraints from only 1 to 4 and increasing the number of variables by 2 for each split problem are the weaknesses of the proposed approach. The most important feature of the new problem is that it is easier to solve by the branch and bound algorithm than the original single constraint knapsack form.

There is a need to compare the proposed approach with other methods. Again this is a limitation and a shortcoming for this study. What seems to be an obvious weakness is that the reformulation splits the single problem into many but easier problems to solve. The challenge of splitting the problem into parallel independent sub-problem can be alleviated by use of parallel computer processors. There is a need to further reduce the numbers of branch and branch bound iterations needed to solve each sub-problem. The main challenge with this is that the complexity of the general integer problem increases with an increase in the number of variables.

8. Conclusions

1. Determining an objective bound \( (Z_k^0) \) to the knapsack problem. An objective bound which is the initial upper limit to the objective value of the problem. In this study all the \( n \) given variables in the original knapsack problem were used in determining the objective bound.

2. Once the objective bound was determined it became easy to generate the \( k \) variable sum limits \( \ell_1, \ell_2, \ldots, \ell_k \) and the 2 additional constraints. To do this let’s only calculate \( \ell_1 \) and \( \ell_k \) and the rest generated as all the integers between \( \ell_1 \) and \( \ell_k. \) The first additional constraint was easily generated from the objective bound and objective row and the other two constraints were generated from the variable sum limits.

3. The \( k \) parallel initial sub-problems where constructed from the \( k \) variable limits \( \ell_1, \ell_2, \ldots, \ell_k \) and 3 additional constraints. Even though the original knapsack problem looked simpler than the each of the \( k \) parallel sub-problems the truth is that the reformulated parallel problems were easier to solve than the original problem as attested to by the numerical illustration. The available computing power which is in terms parallel processing can be taken advantage of.

4. The numerical illustration was shown in Section 5. Reformulation is the way for a knapsack problem given I this paper.

5. Classes of difficult problems were presented in this study. There is need for more research on knapsack problems as shown from the various applications.
References