Combine Column Generation with GUB to Solve the Steel-Iron Raw Materials Purchasing Lot-Sizing Problem[1)]

GAO Zhen TANG Li-Xin
(*School of Information Science & Engineering, Northeastern University, Shenyang* 110004)
(E-mail: Zhengao2002@yahoo.com.cn)

Abstract The objective of steel-iron raw materials purchasing lot-sizing problem is to identify the purchasing quantity of each item in each period over a given horizon so that the total sum of inventory cost and purchasing cost is minimized under considerations. The lot-sizing problems have been proved to be NP-hard. The existing optimal approaches can only solve small-sized problems. In this paper we formulate the steel-iron raw materials purchasing lot-sizing problem and propose a new approach combining column generation with GUB (the generated upper bound) to solve the problem. The computational results of a practical problem show that the approach is effective and has potential application values in making purchasing decisions.

Key words Lot-sizing, purchasing, column generation, GUB

1 Introduction

The raw materials purchasing optimization problem of large-scale steel-iron enterprises is difficult, since the problem has the following features.

1) Multi-item, as many as thousands;
2) Large-volume, up to tens of millions of tons per year;
3) High-cost, billions of *yuan* per year.

However, the problem is also important, since the raw materials purchasing cost accounts for 60%～70% of the steel-iron production cost.

Optimizing raw materials purchasing, which in mathematics can be viewed as a lot-sizing problem, is to properly determine the purchasing lot size of each item of raw materials in a given horizon to perform the action of reducing cost. It is estimated that if a large-scale steel-iron enterprise could organize the purchasing optimally, millions of yuan of purchasing cost would be saved.

The problem originates from the raw materials purchasing problem of MaSteel. MaSteel is a state-owned large-scale steel-iron enterprise. It purchases 15 million tons of raw materials, which cost 4.5 billion *yuan* each year. The number of items of raw materials is more than 3000. For such a large-scale raw materials purchasing problem how to organize the purchasing is a key problem. This paper proposes a raw materials purchasing lot-sizing model based on the practical problem settings of MaSteel, and provides a solution approach of combining column generation with GUB to solve the problem.

Lot-sizing problems are frequently encountered in industrial production settings, and have become an active research area[1～7]. The general lot-sizing problems have been proved as NP-hard[8]. Up to date, there are no effectively optimal approaches solving this kind of problems[3]. Although, in recent years some near-optimal approaches, such as TS[9], GA[10], the hybrid approach[11], *etc.* have been used, all these approaches are approximate ones and hard to appraise.

Combining column generation with GUB approach has been used for the multi-product

1) Supported by National Natural Science Foundation of P. R. China (70171030, 60274049)
Received October 15, 2002; in revised form July 17, 2003
收稿日期 2002-10-15; 收修改稿件日期 2003-07-17

scheduling problems by Lasdon and Terjung[12], and Dzieinsiki and Gomory[13], respectively. But restricted by the capacities of computers, the approach was not developed further. Now, with the development of computer technique, the column generation-based approaches have been used for many important integer and combinatorial optimization problems[14~18]. In particular, it has been a necessary choice to deal with large-sized problems.

2 Problem description

Raw materials purchasing lot-sizing problem (RMPLP) can be described as follows: given the demand for each item of raw materials each month, the problem is to determine the practical purchasing quantity of each item each month of the year within the horizon such that the total sum of purchasing cost and inventory cost is minimized under considerations. For modeling the problem, we consider two key factors: capacity constraints and cost structure.

2.1 Capacity constraints

1) Purchasing budget: the total expense used to purchase raw materials

2) Maximum inventory capacity: the sum of inventory capacities of all yards

3) Raw materials processing system capacity: convey capacity, sintering capacity, *etc.*

2.2 Cost structure

Because our objective is to minimize the total sum of purchasing and inventory cost, obviously, the analysis on components of the cost is necessary.

• Purchasing cost

Definition 1. raw materials purchasing set-up cost is defined as the sum of all cost occurred for completing a raw materials purchasing activity, including loading, unloading, transportation, leasing vehicle, ship cost etc, except for the raw materials cost, i.e., ore price. It is formulated as follows.

$$\textit{Purchasing cost} = \textit{ore price} + \textit{purchasing set-up cost}$$

• Inventory cost

In a large-scale steel-iron enterprise, the total inventory may be millions of tons. The inventory cost and capital holding is significant.

3 Model

Notation:

• Variables

x_{it} the purchasing quantity of item i in period t

I_{it} the inventory of item i in period t

y_{it} 0-1 binary variable, equal to 1, if $x_{it}>0$; 0, otherwise.

• Parameters

N the number of items of raw materials

T the number of periods

d_{it} the demand of item i in period t

p_i the unit price of item i

h_i the unit inventory cost of item i

s_i the purchasing set-up cost of item i

r a proportional constant determined by the production practice

M a sufficiently large positive number

Ω_{water} the set of items of raw materials that are transported by water

I_{max} maximum inventory capability

$C_{\text{water},t}$ the water transportation capacity in period t

C_{budget} total purchasing budget

The RMPLP can be formulated as follows:

$$\text{(RMPLP)} \quad \min \sum_{i=1}^{N}\sum_{t=1}^{T}(p_i x_{it} + s_i y_{it} + h_i I_{it}) \tag{1}$$

$$\text{s.t.} \quad I_{i,t-1} + x_{it} - I_{it} = d_{it}, \quad i = 1,\cdots, N,\ t = 1,\cdots, T \tag{2}$$

$$I_{it} \geqslant rd_{it}, \quad i = 1,\cdots, N,\ t = 1,\cdots,T \tag{3}$$

$$\sum_{i}^{N} I_{it} \leqslant I_{\max}, \quad t = 1,\cdots ,T \tag{4}$$

$$\sum_{i\in \Omega_{\text{water}}} x_{it} \leqslant C_{\text{water},t}, \quad t = 1,\cdots ,T \tag{5}$$

$$\sum_{t=1}^{T}\sum_{i=1}^{N}(p_i x_{it} + s_i y_{it}) \leqslant C_{\text{budget}} \tag{6}$$

$$x_{it} \leqslant My_{it}, \quad i = 1,\cdots,N,\ t = 1,\cdots,T \tag{7}$$

$$y_{it} \in \{0,1\},\ x_{it} > 0,\ I_{it} > 0, \quad i = 1,\cdots,N,\ t = 1,\cdots,T \tag{8}$$

where y_{it} equals 1, if $x_{it}>0$; 0, otherwise. I_{i0}, $i=1,\cdots,N$, $t=1,\cdots,T$ is the initial inventory.

Objective function (1) expresses minimizing the total sum of purchasing cost and inventory cost. Constraint (2) stands for supply-demand balance. Constraint (3) is the requirement of safe inventory. Constraint (4) is the maximum inventory capability. Constraint (5) is the water transportation capacity. Constraint (6) is the purchasing budget. Constraint (7) indicates the single item maximum purchasing amount restriction. Constraint (8) indicates purchasing quantities and inventories are non-negative.

4 Solution approach

4.1 Model transformation

The above model (RMPLP) is a 0−1 mixed integer programming (MIP). We perform the following steps:

Step 1. The redefinition and sequencing of the variables, e.g., if $I_{i0}=a>0$, then let $I'_{i0}=I_{i0}-a$, and redefine it as I_{i0}

Step 2. Upper and lower bound replacing, if $I_{it}\geqslant R$, then let $I'_{it}\geqslant I_{it}-R$, and redefine it as I_{it}, then the lower bound of I_{it} is 0

Then, the original problem model can be reformulated as follows:

$$\text{(LSP)} \quad \min \sum_{i=1}^{N}\sum_{t=1}^{T}(p_i x_{it} + s_i y_{it} + h_i I_{it}) \tag{9}$$

$$\text{s.t.} \quad I_{i,t-1} + x_{it} - I_{it} = d_{it}, \quad i = 1,\cdots, N,\ t = 1,\cdots,T \tag{10}$$

$$I_{i0} = I_{iT} = 0, \quad i = 1,\cdots, N \tag{11}$$

$$\sum_{i=1}^{N} l_{ki} x_{it} \leqslant L_{kt}, \quad k = 1,\cdots,K,\ t = 1,\cdots,T \tag{12}$$

$$x_{it} \leqslant My_{it}, \quad i = 1,\cdots,N,\ t = 1,\cdots,T \tag{13}$$

$$y_{it} \in \{0,1\}, \quad i = 1,\cdots,N,\ t = 1,\cdots,T \tag{14}$$

$$x_{it} > 0,\ I_{it} > 0, \quad i = 1,\cdots,N,\ t = 1,\cdots,T \tag{15}$$

where l_{ki} denotes the unit claim on resource k by item i, while L_{kt} denotes the amount of resource k available in period t, according to the inventory capability, water transportation capacity and the purchasing budget in the original problem, respectively. The others are the same as above.

4.2 Solution approach

4.2.1 Column generation principle

The principle of column generation is that for a linear programming (LP) problem, using Dantzig-Wolfe decomposition principle, the original problem is transformed into an equivalent LP problem, called the master problem (MP), which has moderately rows but a huge number of columns. Column generation technology is used to solve the multi-column LP. In the course of solving the MP, not all columns within the constraint matrix are listed explicitly, but only a very small subset of columns, called the restricted master problem (RMP), will be chosen to solve the original problem. In fact, only fewer columns (basic variables) will appear in the optimal solution, while the other nonbasic columns will be ignored because their corresponding variables are equal to zero. The optimality criterion of solution of the RMP is used to check the optimality of MP solution or to generate new entering basis columns to re-optimize until the optimality check is passed.

4.2.2 Column generation reformulation

To use the D-W decomposition, we introduce the definition of dominant schedule.

Definition 2. The apices of the following set are called dominant schedules[6,20]:

$$\begin{cases} I_{i0} = 0 \\ I_{i,t-1} + x_{it} + I_{it} = d_{it} \\ I_T = 0 \\ x_{it}, I_{it} \geqslant 0 \end{cases}, \quad i = 1,\cdots,N,\ t = 1,\cdots,T \tag{16}$$

For each item i, the dominant schedule set is expressed as

$$\Omega_i = \{(x_{i1}^j,\cdots,x_{iT}^j; I_{i1}^j,\cdots,I_{iT}^j),\quad j = 1,\cdots,J(i)$$

Introducing the following notations.

$$d_i^j = \sum_{t=1}^{T}(p_i x_{it}^j + s_i y_{it}^j + h_i I_{it}^j),\quad L_{it}^{kj} = l_{ki} x_{it}^j$$

and by D-W decomposition principle, the feasible solutions of the original problem can be represented as the convex combination of dominant schedules of each item[6, 20].

$$\text{(CG)} \qquad \min \sum_{i=1}^{N}\sum_{j=1}^{J(i)} d_i^j \theta_i^j \tag{17}$$

$$\text{s. t.} \quad \sum_{i=1}^{N}\sum_{j=1}^{J(i)} L_{it}^{kj}\theta_i^j \leqslant L_{kt},\quad k = 1,\cdots,K,\ t = 1,\cdots,T \tag{18}$$

$$\sum_{j=1}^{J(i)} \theta_i^j = 1,\quad i = 1,\cdots,N \tag{19}$$

$$\theta_i^j \geqslant 0,\quad i = 1,\cdots,N,\ j = 1,\cdots,J(i) \tag{20}$$

Feasible solutions of the problem (CG) must be feasible solutions of the problem (LSP). If the optimal solution of the problem (CG) is integral, then it must be the optimal solution of the problem (LSP). In general cases, it is only a lower bound of the original problem, but when $N \gg T$, we can get a very good approximation of the original problem by the following proposition.

Proposition. If $N > T$, then the any basic feasible solution of the problem (CG) includes at least N-T integral values.

Proof. Leave out. The reader is referred to [20]. □

Thus, we get the following column generation algorithm outline. In fact, the algorithm is a two-level algorithm. In the high-level, it gets the simplex multiplier or dual prices by solving the corresponding RMP. Let $\boldsymbol{\pi} = (\pi_{11},\cdots,\pi_{KT})$ be the simplex multiplier corresponding to the $K \times T$ resource constraints, and α_i be the simplex multiplier corresponding to ith convexity constraint. While in the low-level problem, it solves the following subproblem by dynamic programming for each i to check the optimality or generate new entering basis columns.

$$\min \quad \beta_i = \sum_{t=1}^{T}\left(\left(p_i - \sum_{k=1}^{K}\pi_{kt}l_{ki}\right)x_{it} + s_{it}Y_{it} + h_{it}I_{it}\right) - \alpha_i \tag{21}$$

s. t. constraint (16) as above.

If all $\beta_i \geqslant 0$, $i=1,\cdots,N$, then the optimal solution has been obtained, or else generate new entering basis columns to reoptimize.

4.2.3 Steps of the column generation algorithm

Step 1. Generate initial solution using lot-for-lot heuristic

Step 2. Solve the RMP, obtaining the simplex multiplier $\boldsymbol{\pi}$ and α_i, $i=1,\cdots,N$

Step 3. Solve the subproblems, checking optimality. If the optimality criterion is met, then stop; else go to next step

Step 4. Select the minimum $\beta_i<0$, generating the entering basis column and return to Step 2

5 Computational experiments

Using the above algorithm, we calculate a group of practical examples from MaSteel. The problem size is 97×12 ($N\times T$). In the problem, for convenience, we only take one kind of resources—water transportation, and test the algorithm behaviors by adjusting its capacity. Clearly, It has no difficulty to consider many kinds of resources by adding T times of couple constraints to the original problem. To protect the business interests of the enterprise, information detail is omitted from this paper, only the computational results are given.

In Table 1, columns 1~3 indicate the problem property. The 4th column is the optimal objective value calculated by the proposed algorithm. The 5th column is the computational results by column generation and branch-and-bound approach[21]. The 6th column is dual gap whose value equals $(V2-V1)/V1\times 100\%$. The 7th column is the total number of columns generated. The 8th column is the total number of nodes fathomed. The 9th column is the running time. The overall algorithm is programmed by VC++ 6.0, running on the Pentium Ⅱ 433 computer.

The differences in the four tested problems are: the capacity constraint of the latter problem is stricter than that of the former. The computational results show that, the looser the capacity of the problem is, the easier it obtains the optimal solution. In addition, the generated columns and running time are also less for those problems. For the different capacities, the optimal solution of the problem is different, in the case where the capacity is abundant (e. g., problem 1), the optimal solution of LP is the solution of the original problem, even no dual gap. Nevertheless, in the case that the capacity is tighter (e. g., problems 2, 3), the cost is enlarged, even no feasible solution (problem 4). A great number of computational experiments show that the performance of the proposed algorithm in this paper is satisfying and the dual gap is within 1% (in the computational results of this group of data, the maximum dual gap is merely 0.03%).

Table 1 Computational results using real world data for two approaches

Problem property			Column generation and GUB Optimal value (V1: *Yuan*)	Column generation and B&B, Optimal value (V2: *Yuan*)	Dual gap(%)	Columns	Nodes	CPU (s)
No.	Size ($N\times T$)	Capacity (10^4 tons)						
1	97×12	200	2301121816	2301121816	0.00	97	1	100
2	97×12	150	2301761315.9974	2301762663	0.01	195	2	194
3	97×12	120	2301802286.7851	2301809458	0.03	504	5	499
4	97×12	110	No feasible solution	—	—	—	—	—

6 Conclusion

The steel-iron raw materials purchasing lot-sizing problem is NP-hard. The approach proposed in this paper can effectively solve this difficult large-sized problem. It has significant potential application value for the enterprise. Lots of computational results show that the proposed algorithm has better performance in solving large-sized problems, and that computational time is less and the dual gap is small. At the same time, we also note that there are some research directions deserving concern in future, such as developing effective heuristics to deal with the tough problems, and so on.

Acknowledgment The research has obtained help from MaSteel Raw Materials Supply & Sales Company.

References

1 Derxl A, Kimms A. Lot-sizing and scheduling—Survey and extensions. *European Journal of Operational Research*, 1997,**99**(2): 221~235

2 Kuik R, Salomon M, Van Wassenhove L N. Batch decisions: structure and models. *European Journal of Operational Research*, 1994,**75**(2): 243~263

3 Maes J, Van Wassenhove L N. Multi-item single-level capacitated dynamic lot-sizing heuristics: A general review. *Journal of the Operational Research Society*, 1988,**39**(11): 991~1004

4 Wang Zhao-Hui, Chen Hao-Xun, Hu Bao-Sheng. Scheduling for batch chemical processes using lagrangian relaxation-based approach. *Acta Automatica Sinica*, 1998,**24**(1): 1~8 (in Chinese)

5 Wagner H M, Whitin T. Dynamic version of the economic lot-size model. *Management Science*, 1958,**5**(1): 89~96

6 Manne A S. Programming of economic lot sizes. *Management Science*, 1958,**4**(2): 115~135

7 Tang Li-Xin, Yang Zi-Hou, Wang Meng-Guang. Research on steelmaking and continuous casting production scheduling. *Acta Automatica Sinica*, 1998,**24**(1): 9~15 (in Chinese)

8 Florian M, Lenstra J K, Rinnooy Kan A H G. Deterministic production planning: Algorithms and complexity. *Management Science*, 1980,**26**(7): 12~20

9 Hindi K S. Solving the CLSP by a tabu search heuristics. *Journal of the Operational Research Society*, 1996,**47**(1): 151~161

10 Dellaert D, Jeunet J, Jonard N. A genetic algorithm to solve the general multi-level lot-sizing problem with time-varying costs. *International Journal of Production Economics*, 2000,**68**(2): 241~257

11 Ödamar L, Barbarosgou G. Hybrid heuristics for the multi-stage capacitated lot-sizing and loading problem. *Journal of the Operational Research Society*, 1999,**50**(8): 810~825

12 Lasdon L S, Terjung R C. An efficient algorithm for multi-item scheduling. *Operations Research*, 1971,**19**(4): 946~969

13 Dzielinski B P, Gomory R E. Optimal programming of lot-sizes, inventory and labor allocations. *Management Science*, 1965,**11**(9): 874~890

14 Desrochers M, Desrosiers J, Soloman M. A new optimization algorithm for the vehicle routing problem with time windows. *Operations Research*, 1992,**40**(2): 342~354

15 Savelsbergh M W P. A branch-and-price algorithm for the generalized assignment problem. *Operations Research*, 1997,**45**(6): 831~841

16 Vance P H, Barnhart C, Johnson E L, Nemhauser G L. Airline crew scheduling: A new formulation and decomposition algorithm. *Operations Research*, 1997,**45**(2): 188~200

17 Barnhart C, Hane C A, Vance P H. Using branch-and-price-and-cut to solve origin-destination integer multi-commodity flow problems. *Operations Research*, 2000,**48**(2): 318~326

18 Barnhart C, Johnson E L, Nemhauser G L, Savelsbergh M W P, Vance P H. Branch-and-price: Column generation for solving huge integer programs. *Operations Research*, 1998,**46**(3): 316~329

19 Yi Ting-Yuan, Liu Kun-Sheng. Industrial Accounting. Beijing: China Railway Press, 1990 (in Chinese)

20 Zhu Dao-Li. Large-scale systems optimization theory and application. Shanghai: Shanghai Jiaotong University Press, 1987 (in Chinese)

21 Gao Zhen, Tang Li-Xin. Branch-and-price algorithm of CLSP. Journal of Northeastern University, 2003, **24**(1): 11~14 (in Chinese)

GAO Zhen He is now a lecturer in the Department of Systems Engineering of School of Information Science & Engineering at Northeastern University. He received his bachelor degree from Inner Mongolia University in 1988, received his master and Ph. D. degree from Northeastern University in 1992 and 2003, respectively. His research interests include logistics management and intelligent optimization.

TANG Li-Xin He is now a professor at the Northeastern University. He received his master and Ph. D. degree from Northeastern University in 1991 and 1995, respectively. His research interests include logistics management, production planning and scheduling, and intelligent optimization algorithms.

列生成与 GUB 相结合求解钢铁原料采购批量问题

高 振 唐立新
（东北大学信息科学与工程学院 沈阳 110004）
（E-mail：Zhengao2002@yahoo.com.cn）

摘 要 钢铁原料采购批量问题的目标是确定各种原料在一定时期（通常为一年）的各个时段（一个月）内的采购量，在满足生产需求的前提下，使总的采购成本和库存费用之和最小. 一般批量问题是 NP-hard，目前只存在有限的、启发式的方法求解小规模问题. 建立了原料采购批量模型并提出一种新的方法——列生成与 GUB（广义上界）相结合方法求解该模型. 一组实际问题测试结果证明了该方法的有效性，同时也表明了该方法的潜在应用价值.

关键词 批量，原料采购，列生成，GUB
中图分类号 TP29

中国自动化学会第 19 届青年学术年会（YAC2004）

征文通知

中国自动化学会第 19 届青年学术年会（YAC2004）将于 2004 年 8 月 16～19 日在美丽的泉城山东济南召开. 本次会议由中国自动化学会、中国自动化学会青年工作委员会主办，山东大学控制科学与工程学院承办. 借此机会，热烈欢迎全国各高等院校教师、科研院所和企事业单位的青年科技工作者及博士生、硕士生积极参加. 会议设有优秀论文奖和优秀应用论文奖.

一、征文范围

1. 广义系统、大系统、非线性系统、混沌系统、系统稳定与镇定；
2. 自适应、预测、变结构控制、H_∞ 控制、优化和鲁棒控制；
3. 智能控制、模糊控制、人工智能与专家系统；
4. 系统建模、辨识和估计；
5. 规划、管理与自制、容错控制和故障诊断系统；
6. 神经网络及应用；
7. 机器人与机器人控制；
8. 离散事件动态系统、调度、决策系统；
9. 混合动力学系统及控制；
10. 计算机视觉、图像处理与模式识别；
11. 自动化仪表与过程控制；
12. 电力系统及其自动化；
13. 电机驱动及运动控制；
14. 传感器与检测技术；
15. 软件工程、并行处理；
16. 计算机集成制造系统；
17. 计算机软硬件技术及其应用；
18. 系统工程理论、方法及应用；
19. 自动化指挥系统；
20. 数据融合与软测量；
21. 单片机控制及其应用技术；
22. 企业改革发展策略及管理决策；
23. 工业过程与生产管理；
24. 其他.

二、征文要求

1. 被录用论文将由山东大学出版社出版《自动化理论技术及应用》（卷 11），论文应具有一定学术或实用价值，未在国内外学术期刊或会议发表过；2. 论文第一作者的年龄一般不超过 40 岁；3. 来稿中英文皆可，请用 Word 或 LaTex 编排，A4 纸打印，一式三份并附软盘（也可通过 E-mail 传送）；4. 格式参考自动化学报；5. 投稿时请注明文章所属的方向标号（见本征文范围）；6. 请说明联系作者的详细通讯地址、电话和电子邮件信箱，并注明是否参加优秀论文评选；7. 因版权等引起的纠纷，作者自负.

三、重要日期

论文截稿日期：2004 年 4 月 10 日； 录用/不录用通知日期：2004 年 4 月 31 日.

四、投稿地址

山东济南 山东大学控制科学与工程学院 YAC2004 组委会收 邮政编码：250061
联系人：王玉振 刘允刚
联系电话：0531-8392625，0531-8392515，0531-8392535； 传真：0531-8392205
E-mail：yzwang@tsinghua.edu.cn（王玉振）；lygfr@263.net（刘允刚）