



SOLUÇÃO EXACTA DE PROBLEMAS DE CORTE UNIDIMENSIONAL USANDO O MÉTODO DE PARTIÇÃO E AVALIAÇÃO SUCESSIVAS E GERAÇÃO DIFERIDA DE COLUNAS

J.M. Valério de Carvalho
Dept. Produção e Sistemas,
Universidade do Minho, 4709
Braga Codex, Portugal

Resumo

Quando se pretende obter uma solução inteira para o problema de corte unidimensional, depois de se ter resolvido a sua relaxação linear, é frequente recorrer, quer a técnicas de arredondamento de soluções, quer a diversos tipos de heurísticas. Estas dificuldades decorrem do facto de não ser viável enumerar todas as variáveis estruturais do problema, cujo número pode ser da ordem dos milhões.

Neste artigo, apresenta-se uma formulação em que o número de variáveis e restrições é uma função polinomial da largura do stock e do número de pedidos. Para algumas classes de problemas, é possível enumerar todas as variáveis e obter a solução ótima usando o método da partição e avaliação sucessivas. Para instâncias de maiores dimensões, apresenta-se um procedimento que combina a geração diferida de colunas e o método da partição e avaliação sucessivas. Define-se o subproblema e o modo como é modificado durante a fase de partição e avaliação sucessivas.

São apresentados resultados de testes computacionais para diversos problemas de teste.

Palavras-chave: problemas de corte unidimensional, geração diferida de colunas, partição e avaliação sucessivas.

1. Introdução

O problema de corte consiste em determinar a melhor forma de cortar um inventário de matérias primas

em partes de menores dimensões, de forma a satisfazer as encomendas que são colocadas. O número de dimensões depende das

características do problema. Habitualmente, os problemas de corte unidimensional surgem como problemas de corte de rolos de matéria prima, com uma dada largura inicial, em rolos de menor largura, correspondendo aos pedidos colocados. Um arranjo válido resultante do encaixe de um conjunto de pedidos na largura dos rolos designa-se por padrão de corte ou montagem.

O problema em consideração é classificado, de acordo com o sistema de DYCKHOFF (1990) como 1/V/I/R, significando que se trata de um problema unidimensional, em que existe um conjunto de rolos, todos de dimensão idêntica, e um conjunto de pedidos que devem ser integralmente satisfeitos. Os pedidos, apesar de serem em número relativamente pequeno, traduzem-se num elevado número de rolos a fornecer.

O problema de corte pode ser formulado do seguinte modo: considerem-se m clientes com pedidos de b_i rolos de largura $w_i, i = 1, 2, \dots, m$, a serem cortados a partir de uma quantidade ilimitada de rolos com a largura inicial W . Seja x_j uma variável de decisão que designa o número de rolos a serem cortados segundo o padrão de corte j . A matriz A tem informação sobre os padrões de corte, i.e., cada coluna $A_j = (a_{1j}, \dots, a_{ij}, \dots, a_{mj})^T$ define um padrão de corte. O elemento a_{ij} representa o número de pedidos de largura w_i obtidos a partir do padrão de corte j .

O problema de corte é um problema de programação inteira, porque os rolos são unidades que não interessa fraccionar ao longo do comprimento. O objectivo da formulação apresentada é minimizar o número total de rolos necessários à satisfação dos pedidos.

$$\min \sum_{j \in J} x_j \quad (1.1)$$

$$\text{suj. a } \sum_{j \in J} a_{ij} x_j \geq b_i, i = 1, 2, \dots, m \quad (1.2)$$

$$x_j \geq 0 \text{ e inteiro, } \forall j \in J, \quad (1.3)$$

sendo J o conjunto de padrões de corte válidos. Para o padrão de corte ser válido,

$$\sum_{i=1}^m a_{ij} w_i \leq W$$

$$a_{ij} \geq 0 \text{ e inteiro, } \forall j \in J.$$

A solução da relaxação linear deste problema envolve, por si só, algumas dificuldades. Se, para instâncias de dimensões reduzidas, é possível enumerar todos os padrões de corte e determinar a solução óptima, para problemas de dimensões razoáveis, este método revela-se impraticável. De facto, o número de padrões de corte pode ser facilmente da ordem dos milhões. Para tornar a dificuldade resultante da explosão do número de montagens possíveis, GILMORE & GOMORY (1961) introduziram a geração diferida de colunas. Nesta abordagem, que evita a enumeração explícita de todas as montagens possíveis, as colunas são adicionadas ao problema principal (problema primal restrito) tomando em linha de conta a informação dada pelas variáveis duais do problema, que permite avaliar quais as colunas não enumeradas que são atractivas para o problema principal. Em cada iteração, é geralmente adicionada a coluna mais atractiva, cuja selecção envolve a solução de um subproblema, que é um problema de saco de mochila (GILMORE & GOMORY, 1963).

A determinação de soluções inteiras traz dificuldades acrescidas, pelo que as abordagens usadas são normalmente heurísticas. Citam-se, a título de exemplo, o trabalho pioneiro de GILMORE & GOMORY (1963), que sugeriram arredondamentos para cima e para baixo da solução fraccionária como um método expedito de obter uma solução discreta, e o trabalho de STADTLER (1990) num problema da indústria do alumínio. Esta heurística, iterativamente, faz um arredondamento para cima da variável com maior parte fraccionária, fixa a variável nesse valor, e procede à reoptimização do problema. Se forem gerados excessos de produção, algumas variáveis são decrementadas, o que pode

conduzir a procura não satisfeita. Neste caso, as larguras em falta são combinadas usando uma heurística de FFD.

Existem também algumas referências com abordagens exactas para o problema inteiro. Numa situação em que a razão entre a largura dos rolos e a largura média dos pedidos é pequena, sendo o número de padrões possíveis da ordem de duas ou três centenas, GOULIMIS (1990) utilizou a técnica de planos de corte de Chvátal/Gomory, complementada pelo método de partição e avaliação sucessivas quando os planos de corte falham a obtenção da solução inteira. VANCE *et al.* (1994) desenvolveram um procedimento exacto, baseado na geração diferida de variáveis e no método de partição e avaliação sucessivas, mas apenas aplicável à solução de problemas de corte binário. No problema binário, as quantidades pedidas estão limitadas a serem iguais a uma unidade.

Procedimentos que combinam o método da partição e avaliação sucessivas e a geração diferida de colunas foram utilizados noutros problemas, nomeadamente num problema de escalonamento de pessoal e veículos com janelas temporais (DESROSIERS *et al.*, 1984) e no problema de coloração de arcos de um grafo (NEMHAUSER & PARK, 1991). Na generalidade, a dificuldade crucial na implementação de procedimentos que combinam estes dois métodos resulta do facto de o problema perder a sua “estrutura” quando se adicionam as restrições necessárias ao processo de partição e avaliação sucessivas, podendo haver alterações no tipo de subproblema a resolver para determinar as colunas atractivas.

No caso particular do problema de corte, para a formulação apresentada, depois de resolver a raiz da árvore de pesquisa, que corresponde a resolver a relaxação linear do problema, se houver variáveis fraccionárias, é necessário introduzir restrições suplementares para efectuar a partição do domínio. As colunas geradas enquanto se resolve a relaxação linear do problema podem ser suficientes para obter uma solução inteira

com um valor que seja igual ao arredondamento para cima do óptimo da relaxação linear, o que claramente garante que a solução inteira é óptima. No entanto, pode ocorrer que, num dado nodo da árvore de pesquisa, um conjunto de colunas que não foram geradas seja necessário para obter a solução óptima do problema nesse nodo, cujo resultado seria, de outro modo, não-óptimo ou não-válido.

A geração de colunas nos nodos ocorre depois de terem sido introduzidas restrições no problema, a que correspondem variáveis duais suplementares. Nestas circunstâncias, o subproblema é alterado, e a solução óptima do mesmo problema de saco de mochila pode fornecer uma coluna que não deva ser introduzida no problema principal. De facto, pode acontecer que uma particular coluna que foi forçada a tomar o valor nulo no problema principal, pela introdução de uma restrição nesse sentido, venha a ser a coluna mais atractiva gerada pelo subproblema, não fazendo sentido a sua inserção no problema principal. Para resolver esta dificuldade, VANCE *et al.* (1994) introduzem restrições adicionais ao subproblema de saco de mochila, havendo situações em que o subproblema deixa de ter qualquer estrutura e é resolvido como um problema geral de programação inteira.

Neste artigo, exploramos uma formulação alternativa para o problema de corte com um número de variáveis e restrições que é polinomial em função da largura dos rolos e do número de pedidos. Embora esta formulação tenha um maior número de restrições, a sua estrutura permite ultrapassar os problemas surgidos na geração diferida de variáveis nos nodos da árvore de pesquisa. A relaxação linear deste modelo é tão forte como a relaxação da formulação acima apresentada.

Este trabalho constitui uma extensão da aplicação da mesma formulação ao problema de empacotamento unidimensional (VALÉRIO DE CARVALHO, 1995) onde alguns resultados suplementares são apresentados. Nesse problema, esta relaxa-

ção mostrou ser bastante forte, conduzindo a árvores de pesquisa de dimensões tratáveis.

Na Secção 2, introduz-se a formulação e apresentam-se diversos critérios utilizados para a redução do número de variáveis e as inequações válidas que são usadas para reforçar a formulação. Na Secção 3, descreve-se a solução da relaxação linear do problema usando o método de geração diferida de colunas e mostra-se o subpro-

blema a resolver. Na Secção 4, apresenta-se o método de solução que envolve um procedimento de partição e avaliação sucessivas, as regras de partição utilizadas e o modo de articulação deste processo com a geração diferida de colunas. Na Secção 5, descrevem-se alguns detalhes da implementação computacional e apresentam-se resultados relativos à solução de problemas de teste.

2. Formulação

Dados rolos de largura W e pedidos de largura w_1, w_2, \dots, w_m , o problema de determinar um padrão de corte válido pode ser modelado como um problema de determinar um caminho válido num grafo acíclico e orientado com $W + 1$ vértices. Considere-se um grafo $G = (V, A)$ com $V = \{0, 1, 2, \dots, W\}$ e $A = \{(k, l) : 0 \leq k < l \leq W \text{ e } l - k = w_i \text{ para qualquer } i \leq m\}$, significando que existe um arco orientado entre dois vértices se existir um pedido com a largura correspondente. Considerem-se também, sem perda de generalidade, arcos adicionais entre

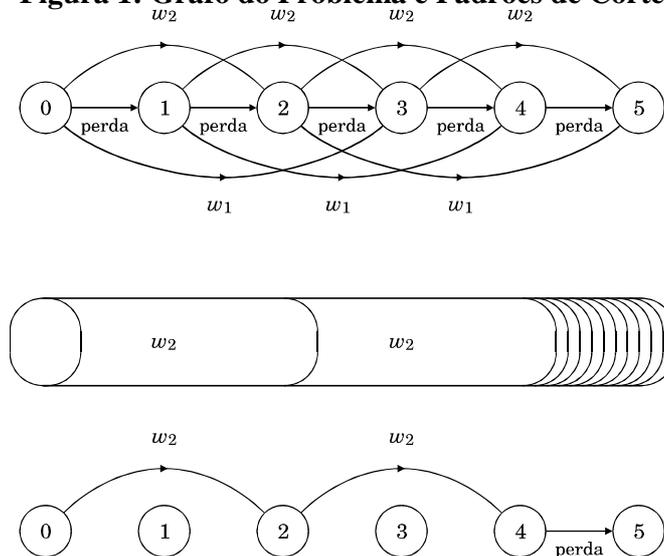
$(k, k + 1), k = 0, 1, \dots, W - 1$, correspondendo a porções não ocupadas do rolo, ou seja, a perdas laterais. Um padrão de corte é válido se e só se existir um caminho entre os vértices 0 e W . Os arcos que constituem o caminho definem o padrão de corte.

Exemplo 1.1 A Figura 1 mostra o grafo associado a uma instância com rolos de largura $W=5$ e pedidos de largura 3 e 2. Na mesma Figura, é apresentada a correspondência entre um caminho e um padrão de corte com 2 unidades de largura 2 e com 1 unidade de perda lateral. •

Este tipo de formulação foi já usado para modelar o problema do saco de mochila

como o problema de determinar o caminho mais longo num grafo orientado (PAPADI-

Figura 1: Grafo do Problema e Padrões de Corte



MITRIOU & STEIGLITZ, 1982). Do mesmo modo, pode ser estendido para modelar problemas de corte. Se um caminho corresponde ao fluxo de uma unidade entre os vértices 0 e W, múltiplos caminhos corresponderão a múltiplos padrões de corte.

Pelas propriedades de decomposição de fluxos (ver, por exemplo, AHUJA *et al.*, 1993), todos os fluxos não-negativos podem ser representados por caminhos e por ciclos. O grafo que define o problema é acíclico e, portanto, qualquer fluxo pode ser decomposto em caminhos orientados ligando o único nodo em que existe um excesso de oferta (nodo 0) ao único nodo com déficit (nodo W).

O problema é formulado como um problema de minimização de fluxo entre o

vértice 0 e o vértice W com restrições adicionais que impõem que a soma dos fluxos dos arcos representando cada pedido deve ser maior ou igual à quantidade pedida. Consideram-se variáveis de decisão $x_{ij}, i < j$, correspondentes aos arcos orientados $(i, j) \in A$, acima definidos, que significam o número de pedidos de largura $j-i$ colocados num padrão de corte a uma distância de i unidades do bordo esquerdo do rolo.

Exemplo 1.1 (cont.) A Figura 2 ilustra a posição em que é cortado um rolo de largura 2, correspondente à variável x_{24} . •

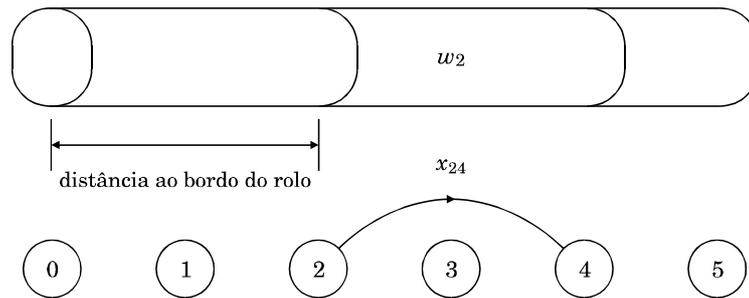


Figura 2: Variáveis do Problema

A formulação do problema é a seguinte:

$$\begin{aligned} \min \quad & z & (2.1) \\ \text{suj. a} \quad & \sum_{(i,j) \in A} x_{ij} - \sum_{(j,k) \in A} x_{jk} = \begin{cases} -z, & \text{se } j = 0 \\ 0, & \text{se } j = 1, 2, \dots, W-1 \\ +z, & \text{se } j = W \end{cases} & (2.2) \\ & \sum_{(k, k+w_i) \in A} x_{k, k+w_i} \geq b_i, i = 1, 2, \dots, m & (2.3) \\ & x_{ij} \geq 0 \text{ e inteiro}, \forall (i, j) \in A & (2.4) \end{aligned}$$

Do ponto de vista prático, para implementar o procedimento de geração diferida de colunas, é vantajoso resolver o modelo apresentado com as restrições de igualdade substituídas por restrições do tipo de maior ou igual.

Existe uma correspondência entre esta formulação de fluxos em arcos e a formulação clássica ((1.1)-(1.3)) de caminhos. A formulação apresentada pode ser vista como um modelo que decompõe os caminhos em fluxos em arcos. Algebricamente, a soma

das colunas correspondentes aos arcos que constituem um caminho produz uma coluna com um elemento -1 na restrição relativa ao vértice 0, um elemento +1 na restrição relativa ao vértice W e elementos não-negativos nas restrições de satisfação de pedidos que correspondem ao número de larguras de cada pedido incluídas no padrão de corte. Estas são as colunas que constituem a formulação clássica.

A matriz que define este problema não possui, na generalidade, qualquer propriedade especial que garanta que as soluções básicas obtidas sejam inteiras, pelo que pode ser necessário, após a solução da relaxação linear do problema, recorrer ao método da partição e avaliação sucessivas. A solução assim obtida possui valores de fluxo inteiro para todos os arcos, podendo, pelas proprie-

dades de decomposição de fluxo acima referidas, ser transformada numa solução inteira para o problema de corte unidimensional.

O limite inferior fornecido pela relaxação linear deste modelo é tão forte como o fornecido pela formulação clássica:

Proposição 1.1 O valor do óptimo da relaxação linear da formulação (2.1)-(2.4) não é menor que o valor do óptimo da formulação clássica (1.1)-(1.3).

Prova: Este facto decorre de a solução óptima do modelo (2.1)-(2.4) ser uma solução válida para a formulação (1.1)-(1.3) que, sendo um problema de minimização, terá um valor óptimo igual ou inferior. •

2.1. Enumeração de colunas e critérios de redução

Usando as variáveis apresentadas, existem muitas soluções alternativas que correspondem exactamente aos mesmos padrões de corte. Entre todas as alternativas, há uma solução em que os pedidos existentes no padrão de corte estão ordenados por ordem decrescente de valor. Se orientarmos a procura para esta solução em particular, é possível reduzir as dimensões do modelo eliminando colunas, sem perder nenhuma solução válida do modelo original.

Considere-se que os pedidos são ordenados por valor decrescente de largura, *i.e.*, para $i < j, w_i > w_j$. O primeiro critério resulta imediatamente das características dos padrões de corte:

Critério 1 Um arco do pedido j , designado por $x_{k,k+w_j}$, apenas pode ter origem num vértice k que seja destino de um outro arco de um pedido i , $x_{k-w_i,k}$, para $w_i \geq w_j$, ou, ainda, a partir do vértice 0, *i.e.*, do bordo do rolo.

Em particular, se o padrão de corte tiver perdas laterais, as variáveis correspondentes serão as últimas em cada padrão de corte,

não podendo ter início a partir do bordo do rolo. Em qualquer solução, o padrão de corte nunca se pode iniciar com uma perda (corresponderia a ocupar um padrão apenas com perdas), devendo ser precedido por um pedido, cuja largura é maior que a largura mínima.

Critério 2 Todos os arcos de perda $x_{i,i+1}$ devem ter valores nulos para $i < w_m$.

Seja z_{LP} o valor do óptimo da relaxação linear e $\lceil z_{LP} \rceil$ o menor valor inteiro maior ou igual a z_{LP} . O valor de z_{LP} é um limite inferior para o valor do óptimo inteiro e o espaço de $\lceil z_{LP} \rceil$ rolos não ocupado por pedidos deve ser ocupado por perdas:

Definição 1 O valor mínimo de perda $T_{\min} = \lceil z_{LP} \rceil W - \sum_{i=1}^m w_i b_i$.

Após a solução da relaxação linear, o modelo pode ser reforçado introduzindo uma restrição que força o número de arcos de perda a ser, pelo menos, em quantidade

igual à perda mínima T_{\min} :

Proposição 2.1 A restrição $\sum_{(k,k+1) \in A} x_{k,k+1} \geq T_{\min}$ é uma inequação válida para o problema inteiro.

Pode ser demonstrado que existe uma solução em que todas as restrições de satisfação dos pedidos são respeitadas sem folga (VALÉRIO DE CARVALHO, 1995). Como corolário, a inserção desta inequação válida, que impõe um limite inferior ao valor das perdas laterais, força o valor da solução óptima a ser inteiro.

Na generalidade dos problemas de programação inteira, não é fácil estabelecer um limite superior para a diferença (*gap*) entre os valores do óptimo do problema inteiro e da sua relaxação linear. No caso particular do problema de corte unidimensional, acontece que a generalidade das instâncias possui um *gap* inferior a uma unidade, sendo o valor da função objectivo da solução óptima do problema inteiro igual ao arredondamento para cima do valor do óptimo da respectiva relaxação linear, propriedade que é designada como propriedade de arredondamento para cima (MARCOTTE, 1985,1986). Assim, na generalidade das instâncias, a perda será efectivamente igual a T_{\min} . Se, no entanto, a instância não possuir a propriedade de arredondamento para cima, existindo um *gap* entre os valores do óptimo do problema e da sua relaxação superior a uma unidade, o valor da perda de qualquer solução inteira será da seguinte forma:

Proposição 2.2 Em qualquer solução inteira, a perda $T = T_{\min} + nW$, sendo n um valor inteiro.

Na Secção 4, será apresentado um algoritmo de solução para o problema inteiro de corte unidimensional. Após resolver a relaxação linear do problema, este algoritmo procura, num primeiro passo, obter uma solução inteira de valor igual ao arredonda-

mento para cima do valor do óptimo da relaxação linear. Se for possível encontrar uma solução com esse valor de função objectivo, ela será óptima. Se este valor da perda T_{\min} for pequeno, pode ser possível reforçar a formulação do problema correspondente a este primeiro passo, fixando adicionalmente alguns arcos de perda no valor zero:

Critério 3 Se o valor da perda $T_{\min} < W - w_m$ e a instância possuir a propriedade de arredondamento para cima, todos os arcos de perda $x_{i,i+1}, i = w_m, \dots, W - T_{\min} - 1$ podem ser fixados em 0.

No caso particular da relaxação linear possuir um valor inteiro, se a instância tiver a propriedade de arredondamento para cima, o que garante que o óptimo inteiro será igual a esse valor, o Critério 3 traduz-se na fixação de todos os arcos de perda em zero.

Finalmente, um arco de perda deve ser sempre sucedido por outro arco de perda, pelo que a seguinte restrição é válida:

Proposição 2.3 Para os arcos de perda não fixados em zero, é válida a relação $x_{i,i+1} \leq x_{j,j+1}$, para todos $j > i$.

Outros critérios de redução, que assumem maior relevância em problemas de empacotamento, são referidos em VALÉRIO DE CARVALHO (1995).

Exemplo 2.1 Considere-se o problema de corte de rolos de largura $W=7$ nas medidas $w=(5,3,2)$, nas quantidades $b=(1,3,2)$. A formulação correspondente a este problema é apresentada no Figura 3. O modelo contém uma inequação redundante relativa às perdas que pode ser reforçada depois de resolver a relaxação linear do problema.

É de salientar que o arco x_{46} não é admissível para esta instância particular, porque são apenas requeridas 2 unidades de largura

2. Os pedidos desta largura devem aparecer após pedidos de largura 5 ou 3, ou colocados no início da montagem, a partir do vértice 0. Em nenhum destes casos, os pedidos de largura 2 podem ser colocados a partir do vértice 4. O número de restrições de conservação de fluxo nos vértices pode ser inferior ao valor da largura dos rolos de *stock*. Neste exemplo, não existe nenhuma restrição para o vértice 1.

A solução da relaxação linear produz um

ótimo de valor $z_{LP} = 2.75$. Esta instância possui a propriedade de arredondamento para cima; de facto, a solução óptima é igual a 3. A perda da solução óptima é igual à perda mínima $T_{\min} = 3$. Com base neste resultado, é possível fixar as variáveis $x_{23} = x_{34} = 0$ e reforçar a inequação relativa às perdas, aumentando o termo independente da restrição, que passa a ser $x_{45} + x_{56} + x_{67} \geq 3$. •

	Z	X ₀₅	X ₀₃	X ₃₆	X ₀₂	X ₂₄	X ₃₅	X ₅₇	X ₂₃	X ₃₄	X ₄₅	X ₅₆	X ₆₇	
vértice 0	1	-1	-1		-1									= 0
1														= 0
2					1	-1			-1					= 0
3			1	-1			-1		1	-1				= 0
4						1				1	-1			= 0
5		1					1	-1			1	-1		= 0
6				1								1	-1	= 0
7	-1							1					1	= 0
pedido 1		1												≥ 1
2			1	1										≥ 3
3					1	1	1	1						≥ 2
perdas									1	1	1	1	1	≥ 0

Figura 3: Formulação do Exemplo 2.1

3. Relaxação de Programação Linear

Do ponto de vista prático, para resolver a relaxação linear deste modelo, é vantajoso recorrer à geração diferida de colunas. Para o efeito, pode ser usada a estratégia de inserção de colunas individualmente, ou seja, apenas uma coluna é inserida em cada iteração, procedendo-se depois à reoptimização, ou se insere um conjunto de colunas em cada iteração. Este conjunto de colunas pode, por exemplo, ser o conjunto de larguras que fazem parte da solução de um subproblema de saco de mochila. No trabalho desenvolvido, apenas se implementou o processo de geração individual de colunas, não tendo sido efectuado nenhum esforço de comparação do desempenho dos dois métodos alternativos.

Especificamente, o único método testado consiste em seleccionar a coluna que tem o custo reduzido mais atractivo. O modelo permite uma avaliação de todas as colunas permitidas no problema primal restrito.

Designem-se por $u_j, j = 0, 1, \dots, W$, as variáveis duais associadas às restrições de conservação de fluxo e por $v_k, k = 1, 2, \dots, m$, as variáveis duais associadas às restrições de satisfação dos pedidos.

Proposição 3.1 Uma variável é atractiva se:

$$-u_i + u_j + v_k > 0$$

A determinação da variável mais atracti-

va pode ser feita em tempo linear em função do número de colunas permitidas para a instância do problema.

Prova: Seja B a base corrente do problema primal restrito e $c_B B^{-1}$ a respectiva solução dual. Seja A_{col} a coluna correspondente à variável x_{ij} . A coluna A_{col} é atractiva se o seu custo reduzido $c_B B^{-1} A_{col} - c_{col}$ for maior que zero. Nesta formulação, todos os $c_{col} = 0$ e cada coluna possui apenas três elementos diferentes de zero, nomeadamente um -1 na posição i , um +1 na posição j e um

+1 na posição k , correspondendo à largura $w_k = j - i$. A avaliação de uma coluna envolve apenas a soma de três parcelas, o que pode ser feito em tempo constante. •

Embora a formulação tenha um número de restrições que é mais elevado, o número de elementos diferentes de zero em cada coluna é muito pequeno; este factor é reconhecidamente importante na carga computacional de solução de problemas de programação linear.

4. Método de Solução

A estratégia adoptada para resolver o problema de corte consiste em converter o problema de optimização numa sequência de problemas de decisão, nos quais se pretende determinar se existe alguma solução inteira com um valor de função objectivo igual ao melhor limite inferior conhecido, que será sempre um valor inteiro, designado por Z_{LB} .

O primeiro problema de decisão a resolver consiste em determinar se existe alguma solução inteira com valor de função objectivo igual a $\lceil Z_{LP} \rceil$, i.e., $Z_{LB} = \lceil Z_{LP} \rceil$. Sabendo-se que a maior parte das instâncias goza da propriedade de arredondamento para cima, a solução deste problema de decisão fornece, na grande maioria das vezes, a solução óptima.

No entanto, se não existir nenhuma solução inteira de valor $\lceil Z_{LP} \rceil$, é tentada a solução do problema de decisão com valor inteiro de função objectivo imediatamente superior, i.e., $Z_{LB} = Z_{LB} + 1$. Para o efeito, o problema é reformulado, incrementando o valor do lado direito da restrição relativa às perdas laterais de W unidades, perda que deve necessariamente existir numa solução que tem um rolo suplementar.

O processo é repetido até ser encontrada a primeira solução inteira que será uma solução óptima. A utilização deste método de solução só é justificável em problemas em que o valor esperado do *gap* é minúsculo, não sendo razoável o seu uso em problemas gerais de programação inteira.

4.1 Partição e Avaliação Sucessivas

A solução de cada problema de decisão traduz-se num procedimento de partição e avaliação sucessivas. A partição do domínio é efectuada sobre cada variável individualmente inserindo restrições do tipo:

$$x_{ij} \geq \lceil x_{ij} \rceil \quad \text{e} \quad x_{ij} \leq \lfloor x_{ij} \rfloor$$

Em cada vértice da árvore de pesquisa, é seleccionada a variável x_{ij} com menor valor

de i , sendo em caso de empate escolhida a que corresponder à maior largura, i.e., $\max\{j - i\}$. O algoritmo faz uma pesquisa em profundidade (*depth first search*), sendo as restrições que forcem a variável a subir para o menor inteiro que lhe é superior introduzidas em primeiro lugar.

No que diz respeito à avaliação dos vértices da árvore de pesquisa, dado que os problemas a resolver são problemas de

decisão em que se procura obter uma solução com valor Z_{LB} , qualquer vértice da árvore de pesquisa com um valor de função objectivo estritamente maior é abandonado. Este abandono só pode ser efectuado depois de garantir que a inserção de novas colunas não pode melhorar o valor da função objectivo. Isto acontece quando já não existe mais nenhuma coluna que seja atractiva e o valor da função objectivo do problema primal restrito é estritamente superior a Z_{LB} .

Em síntese, após a inserção de uma restrição de partição, o problema é reoptimizado, podendo acontecer um dos seguintes casos:

- (i) a solução é inteira, o que implica que é óptima;
- (ii) a solução é fraccionária e de valor igual a Z_{LB} , tornando-se necessário introduzir restrições de partição adicionais;
- (iii) a solução tem um valor estritamente maior que Z_{LB} , situação em que se de-

ve recorrer à geração diferida de colunas, processo que pode conduzir a uma das situações (i) ou (ii), ou à conclusão que o vértice não pode produzir nenhuma solução de valor Z_{LB} , devendo ser abandonado. A Figura 4 apresenta um esquema de diagrama de blocos do algoritmo de partição e avaliação sucessivas/geração diferida de colunas.

O factor crucial deste algoritmo é a decomposição dos padrões de corte em variáveis independentes, que permite conciliar, de uma forma simples, a geração diferida de colunas e o método da partição e avaliação sucessivas. Nesta formulação, apenas são candidatas a entrar no problema principal as colunas que dele ainda não fazem parte, pelo que não existe nenhuma incompatibilidade entre as restrições impostas as variáveis do problema principal e a aceitação de uma variável julgada atractiva e indicada pelo subproblema.

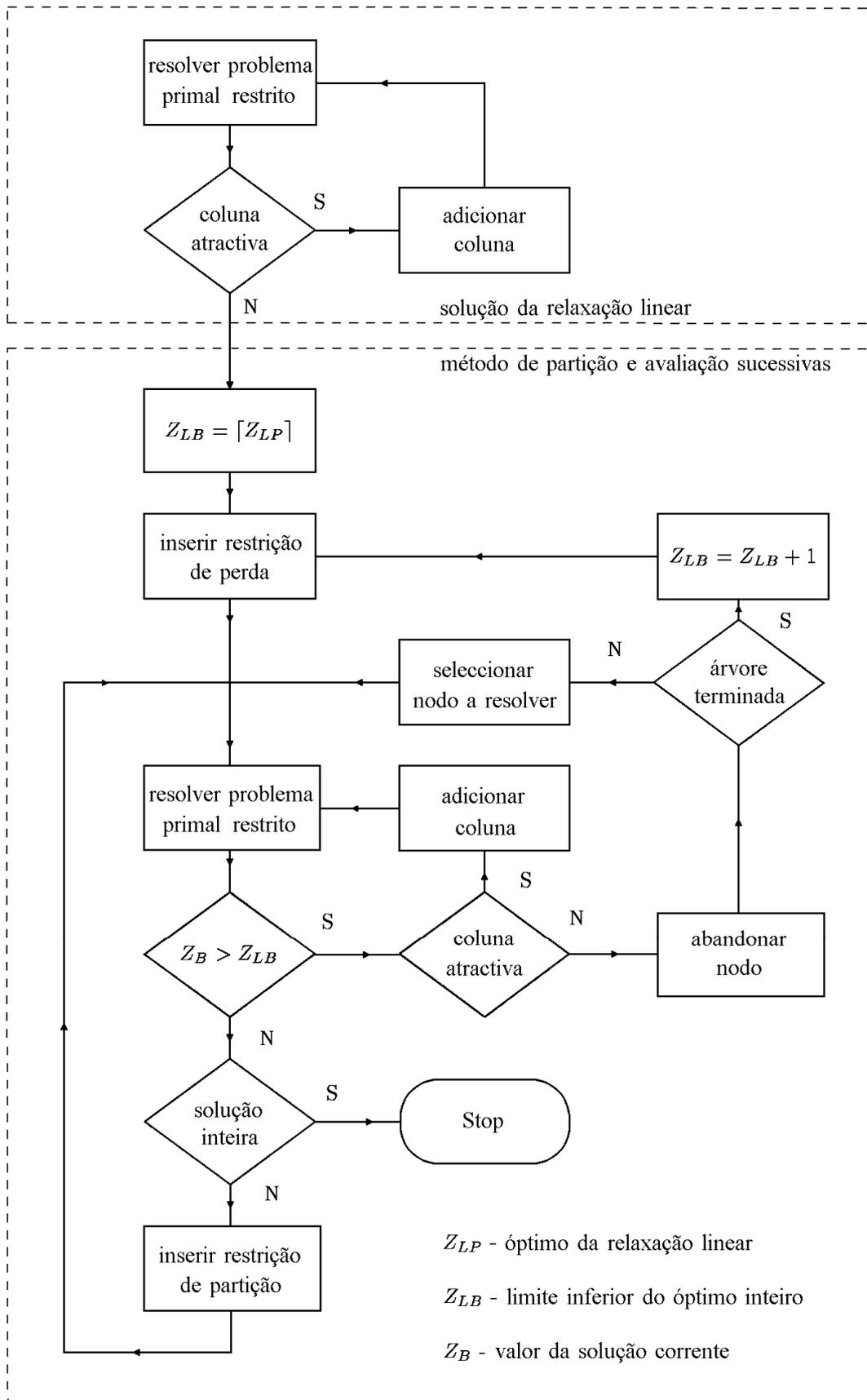


Figura 4: Algoritmo de Partição e Avaliação Sucessivas/ Geração Diferida de Colunas

5. Implementação e Resultados Computacionais

O desempenho do algoritmo foi testado nalguns problemas de teste do conjunto elaborado por WASCHER & GAU (1993). Foram seleccionadas algumas instâncias em que o valor dos pedidos é inteiro. Embora o método possa ser aplicado, sem perda de generalidade, a problemas com pedidos de valor racional, as instâncias resultantes são de maior dimensão, por ser necessário converter esses valores em quantidades inteiras.

O algoritmo faz recurso às rotinas do programa XMP, descrito em MARSTEN (1981), que utiliza o método de simplex revisto e as rotinas LA05 para factorização da inversa. Os limites inferiores e superiores para as variáveis são tratados implicitamente, o que permite uma fácil implementação do método da partição e avaliação sucessivas.

Os tempos referidos nesta secção foram obtidos num microcomputador PC 486 DX a 40 MHz, com 4 Mbytes de memória RAM. Os algoritmos foram codificados em Fortran e compilados usando o compilador FTN77 (SALFORD FTN77, 1995) sob o sistema operativo DBOS, que permite ultrapassar as limitações de utilização da memória disponível, habituais no DOS.

Durante o pré-processamento, são enumeradas todas as colunas permitidas (arcos) para a instância em consideração, procede-se à determinação dos vértices para os quais é necessário considerar restrições de conservação de fluxo e selecciona-se um conjunto inicial de colunas; para cada pedido, são seleccionadas as variáveis $x_{0k}, x_{k,2k}, \dots, x_{(s-1)k,sk}$, sendo $sk \leq W$ e $(s+1)k > W$.

Não foram efectuadas quaisquer operações de remoção de colunas que se tornassem não-atractivas durante o processo, nem efectuadas comparações relativas aos tempos de solução da relaxação linear.

Claramente, para resolver o problema inteiro pode ser utilizado um método que combine as duas formulações. A formulação clássica pode ser usada para obter a solução óptima do problema relaxado, sendo o conjunto óptimo de padrões de corte utilizado para gerar uma solução do modelo cuja formulação aqui foi apresentada. Esta conversão é feita decompondo os padrões de corte e, seguidamente, caso a solução não seja inteira, o método de partição e avaliação sucessivas pode ser usado para determinar a solução inteira.

O primeiro conjunto de problemas possui as características indicadas na Tabela 1. A Tabela apresenta o número de pedidos de cada instância, a largura dos rolos e o número de arcos (variáveis) e de restrições relativos à formulação apresentada.

Na Tabela 2, apresentam-se os dados relativos à respectiva solução, nomeadamente o número de colunas geradas e os tempos de solução. Pode ser verificado que apenas uma pequena fracção das variáveis é utilizada no processo de solução. O significado de cada coluna é o seguinte:

- cols lp - o número de colunas geradas na solução da relaxação linear;
- cols bb - o número de colunas geradas no método de partição e avaliação sucessivas;
- vért bb - o número de vértices explorados no método de partição e avaliação sucessivas;
- t_{pp} - tempo de pré-processamento (segundos);
- t_{lp} - tempo de solução da relaxação linear (segundos);
- t_{bb} - tempo de partição e avaliação sucessivas (segundos);
- t_{tot} - tempo total (segundos);
- Z* - valor do óptimo.

Tabela 1: Características dos Problemas PIE

	pedidos	largura	arcos	restrições
PIE64-1	7	215	501	209
PIE64-2	7	215	501	209
PIE64-3	7	215	707	219
PIE64-4	7	215	707	219
PIE64-5	10	148	376	131
PIE64-6	20	150	268	124
PIE64-7	20	150	639	144
PIE64-8	30	182	250	146
PIE64-9	30	182	250	146
PIE64-10	30	182	1226	186
PIE64-11	30	182	1226	186

Tabela 2: Tempos de Solução dos Problemas PIE

	cols lp	cols bb	vért bb	t _{pp}	t _{lp}	t _{bb}	t _{tot}	z*
PIE64-1	73	0	1	0.11	7.75	0.49	8.35	640
PIE64-2	90	0	4	0.11	10.11	0.88	11.10	245
PIE64-3	131	0	11	0.16	17.25	2.97	20.38	493
PIE64-4	46	0	9	0.17	7.86	1.98	10.00	197
PIE64-5	52	0	4	0.16	3.35	0.44	3.96	239
PIE64-6	40	0	0	0.11	2.31	0.00	2.42	1527
PIE64-7	65	0	43	0.11	5.28	4.18	9.56	1000
PIE64-8	58	42	2	0.11	3.35	3.46	6.92	2845
PIE64-9	58	42	2	0.11	3.35	3.46	6.92	2575
PIE64-10	127	0	45	0.17	14.12	9.73	24.01	1390
PIE64-11	164	0	49	0.16	18.46	10.44	29.06	1242
	82.2	7.6	15.5	0.13	8.47	3.46	12.06	

Na maior parte dos casos, não foi necessário recorrer à geração de colunas adicionais durante o processo de pesquisa em árvore, tendo as colunas geradas durante a solução da relaxação linear sido suficientes para a obtenção da solução inteira. Neste trabalho, não foi implementado nenhum corte do processo de geração diferida de colunas, à maneira de FARLEY (1990), o que poderia ter reduzido o tempo de solução da relaxação linear, conduzindo, possivelmente, a um incremento do número de colunas geradas durante o método de partição e avaliação sucessivas.

O desempenho do algoritmo depende crucialmente da dimensão da largura dos rolos W e, também, do número de pedidos. O conjunto de problemas cujos resultados são apresentados nas Tabelas 3 e 4 correspondem a situações particularmente desfavoráveis, porque a largura dos rolos varia entre 1020 e 1520, embora o número de pedidos seja pequeno, entre 5 e 12. É de salientar que, numa percentagem significativa de instâncias (8 em 50), a solução da relaxação linear é inteira.

Tabela 3: Tempos de Solução dos Problemas WAE

	cols lp	cols bb	vért bb	t _{pp}	t _{lp}	t _{bb}	t _{tot}	z*
WAE84-1	159	0	7	0.55	162.20	55.27	218.02	89
WAE84-2	296	0	20	0.55	271.92	164.40	436.87	382
WAE84-3	27	0	2	0.38	72.58	7.48	80.44	913
WAE84-4	316	0	1	0.27	105.16	2.52	107.96	402
WAE84-5	214	0	2	0.33	111.21	6.98	118.52	410
WAE84-6	211	0	4	0.44	135.66	17.80	153.90	557
WAE84-7	424	0	14	0.60	340.60	109.62	450.82	480
WAE84-8	351	0	2	0.49	240.66	7.15	248.30	315
WAE84-9	275	0	16	0.50	205.82	115.83	322.15	322
WAE84-10	28	0	0	0.38	69.13	0.00	69.51	576
WAE84-11	387	0	0	0.49	240.27	0.00	240.77	428
WAE84-12	97	0	5	0.45	91.48	42.47	134.40	618
WAE84-13	213	0	10	0.49	183.35	46.05	229.89	286
WAE84-14	224	0	4	0.49	196.32	27.97	224.78	317
WAE84-15	310	0	16	0.55	222.25	109.07	331.87	372
WAE84-16	92	0	2	0.38	74.34	6.32	81.04	644
WAE84-17	34	0	11	0.49	89.73	58.30	148.52	510
WAE84-18	60	0	7	0.44	83.96	35.93	120.33	117
WAE84-19	275	0	3	0.44	189.34	18.41	208.19	149
WAE84-20	581	0	22	0.55	523.46	143.68	667.69	160
WAE84-21	284	0	1	0.44	183.90	6.15	190.49	405
WAE84-22	233	3	4	0.55	209.89	67.20	277.64	535
WAE84-23	209	0	18	0.38	124.51	61.98	186.87	243
WAE84-24	610	0	12	0.55	508.85	103.96	613.35	348
WAE84-25	39	0	0	0.38	79.40	0.00	79.78	543
WAE84-26	419	0	13	0.60	328.96	100.27	429.84	370
WAE84-27	358	0	3	0.49	283.90	27.69	312.09	465
WAE84-28	154	2	1	0.38	82.09	11.04	93.52	312
WAE84-29	197	0	15	0.44	142.69	82.97	226.10	470
WAE84-30	419	0	81	0.49	264.95	340.88	606.32	412
WAE84-31	115	0	1	0.33	68.30	2.86	71.48	317
WAE84-32	100	0	7	0.38	101.54	32.03	133.96	586
WAE84-33	213	0	0	0.27	77.42	0.00	77.69	331
WAE84-34	357	0	10	0.49	245.49	53.13	299.12	566
WAE84-35	376	0	10	0.60	309.89	76.32	386.81	604
WAE84-36	335	0	5	0.49	206.15	21.76	228.41	427
WAE84-37	54	0	10	0.44	90.82	52.69	143.96	486
WAE84-38	376	0	22	0.38	176.48	59.67	236.54	243
WAE84-39	370	0	9	0.49	286.04	59.12	345.66	360
WAE84-40	525	0	18	0.55	506.54	117.86	624.94	183
WAE84-41	206	0	21	0.55	194.78	158.74	354.07	261
WAE84-42	176	0	0	0.44	134.73	0.00	135.16	416
WAE84-43	22	0	0	0.49	106.43	0.00	106.92	499
WAE84-44	50	0	1	0.49	92.53	4.45	97.47	651
WAE84-45	281	0	17	0.44	170.22	83.13	253.79	459

Tabela 3 (cont.): Tempos de Solução dos Problemas WAE

	cols lp	cols bb	vért bb	t_{pp}	t_{lp}	t_{bb}	t_{tot}	z^*
WAE84-46	331	0	1	0.49	259.56	6.04	266.10	157
WAE84-47	277	0	3	0.50	232.64	15.50	248.63	397
WAE84-48	16	0	0	0.44	63.79	0.00	64.23	352
WAE84-49	227	2	2	0.28	81.26	25.44	106.98	588
WAE84-50	240	0	0	0.38	150.05	0.00	150.44	312
	242.9	0.1	8.7	0.46	187.46	50.92	238.85	

6. Conclusões

Neste trabalho, foi apresentada uma nova formulação para o problema de corte unidimensional. Esta formulação tem uma estrutura que permite desenvolver um procedimento exacto de solução que combina a geração diferida de colunas e o método de partição e avaliação sucessivas.

Mesmo sem utilizar regras sofisticadas de partição, verificou-se que é possível obter

soluções inteiras em tempos razoáveis. O bom desempenho do algoritmo fica a dever-se ao facto de a formulação ser bastante forte, o que conduz a árvores de pesquisa de dimensões tratáveis. A formulação é particularmente sensível à largura dos rolos; verificou-se que, mesmo para essas instâncias, os tempos de solução são aceitáveis.

Agradecimentos

Este trabalho foi parcialmente suportado pelo Centro de Investigação Algorítmica da Universidade do Minho, e foi executado no âmbito de um projecto do Grupo de

Engenharia Industrial e de Sistemas. Agradecemos as sugestões feitas pelos *referees*, que ajudaram a melhorar a apresentação do material.

Referências Bibliográficas:

- AHUJA, R. ; MAGNANTI, T. & ORLIN, J.: *Network Flows: Theory, Algorithms and Applications*. Prentice Hall, Englewood Cliffs, New Jersey, 1993.
- DESROSIERS, J. ; SOUMIS, F. & DESROCHERS, M.: "Routing with time windows by column generation". *Networks*, 14, p. 545-565, 1984.
- DYCKHOFF, H.: "A typology of cutting and packing problems". *European Journal of Operational Research*, 44, p. 145-159, 1990.
- FARLEY, A.A.: "A note on bounding a class of linear programming problems, including cutting stock problems". *Operations Research*, 38, p.992-993, 1990.
- GILMORE, P.C. & GOMORY, R.E.: "A linear programming approach to the cutting stock problem". *Operations Research*, 9, p. 849-859, 1961.
- GILMORE, P.C. & GOMORY, R.E.: "A linear programming approach to the cutting stock problem - part ii". *Operations Research*, 11, p. 863-888, 1963.
- GOULIMIS, C.: "Optimal solutions to the cutting stock problem". *European Journal of Operational Research*, 44, p. 197-208, 1990.
- MARCOTTE, O.: "The cutting stock problem and integer rounding". *Mathematical Programming*, 33, p. 82-92, 1985.
- MARCOTTE, O.: "An instance of the cutting stock problem for which the rounding property does not hold". *Operations Research Letters*, 4, p.239-243, 1986.
- MARSTEN, R.M.: "The design of the XMP linear programming library". *ACM Transactions on Mathematical Software*, 7, p. 481-497, 1981.
- NEMHAUSER, G. & PARK, S.: "A polyhedral approach to edge coloring". *Operations Research Letters*, 10, p. 315-322, 1991.

- PAPADIMITRIOU, C.H. & STEIGLITZ, K.:** *Combinatorial Optimization: Algorithms and Complexity*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1982.
- SALFORD FTN77.** *Salford Software Ltd*. Adelphi House, Adelphi Road, Salford M3 6EN, UK, 1995.
- STADTLER, H.:** "One-dimensional cutting stock problem in the aluminium industry and its solution". *European Journal of Operational Research*, 44, p. 209-224, 1990.
- VALÉRIO DE CARVALHO, J.M.:** "A model for the one-dimensional cutting stock problem". *working paper*, 1995.
- VANCE, P.; BARNHART, C.; JOHNSON, E.L. & NEMHAUSER, G.:** "Solving binary cutting stock problems by column generation and branch-and-bound". *Computational Optimization and Applications*, 111-130, 1994.
- WASCHER, G. & GAU, T.:** "Two approaches to the cutting stock problem". *Presented at the IFORS's 93 Conference*, Lisbon, July 12, 1993.

EXACT SOLUTION OF ONE-DIMENSIONAL CUTTING STOCK PROBLEMS USING BRANCH-AND-BOUND AND COLUMN GENERATION

Abstract

If an integer solution to the one-dimensional cutting stock problem is required, after solving the linear programming relaxation, one frequently resorts to heuristics based on rounding up and down the continuous solution, or other heuristics similar type. The difficulties arise from the fact that it may not be practically possible to enumerate all the structural variables of the problem, whose number may be in the order of millions, even for instances of moderate size.

In this article we present a formulation with a number of variables and constraints that is polynomial with respect to the width of the stock and the number of orders. For some classes of instances, it is possible to enumerate completely all the variables and to obtain an integer optimal solution using a branch-and-bound method. For larger instances, we present a procedure that combines column generation and branch-and-bound. We define the subproblem, and the way it is modified during the branch-and-bound phase. Computational results are presented for several test problems.

Key-words: cutting stock, cutting problem, column generation.