

2

Formulação do Problema

2.1

Reconhecimento de cenas por correspondência de grafos

Um problema de reconhecimento estrutural de cena pode ser descrito como a caracterização dos objetos presentes na cena e das relações entre eles. Diversas abordagens deste problema são encontradas na literatura. Tipicamente, o grau de detalhamento necessário para a descrição das características da cena é função do tipo de aplicação em questão. Entretanto, uma estrutura adequada de representação é imprescindível para o processo de reconhecimento [14].

Em geral, deseja-se obter uma interpretação estrutural de uma cena a partir de um modelo esquemático genérico. O reconhecimento de uma imagem pode ser visto como o processo de associação entre regiões (objetos) da imagem e objetos presentes no modelo utilizado. Uma analogia frequentemente utilizada na literatura é a de rotular-se cada região da imagem de forma a associá-la a algum objeto do modelo.

Dentre as aplicações para este problema encontram-se o reconhecimento de estruturas 3D do cérebro humano a partir de imagens obtidas por ressonância magnética, o problema de reconhecimento facial, o reconhecimento de imagens aéreas e a identificação de imagens previamente conhecidas, como estradas e mapas de cidades, entre outras. Por exemplo, em [44, 45] estuda-se este problema com o objetivo de propor a melhor descrição de uma região de Paris, baseando-se no mapa da região. Alguns algoritmos de reconhecimento de cenas são propostos em [21, 23, 31, 33, 49, 50, 53]. A Figura 2.1 apresenta um exemplo de imagem do cérebro humano obtida por ressonância magnética e um modelo retirado de um atlas médico.

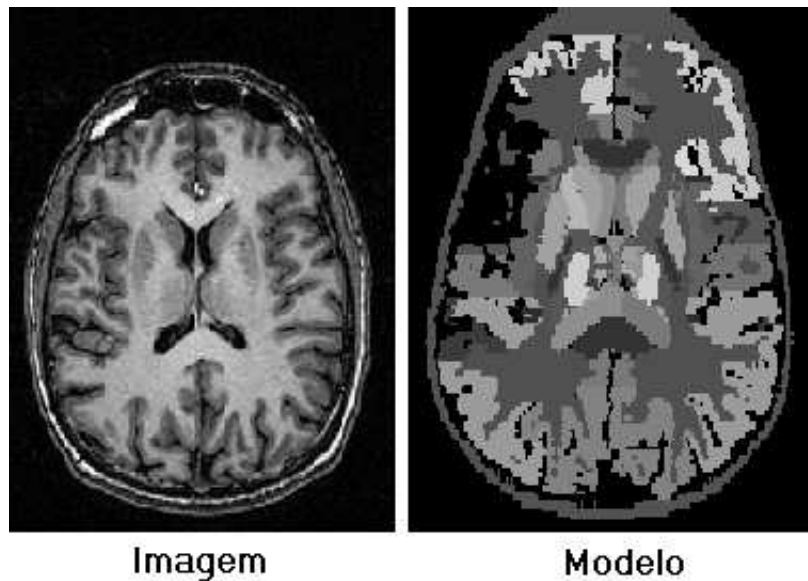


Figura 2.1: Exemplo de imagem obtida por ressonância magnética e um modelo retirado de um atlas médico referentes a um corte do cérebro humano.

Segundo Boeres [14], grafos são estruturas matemáticas adequadas para a descrição estrutural de cenas. Neste caso, problemas de reconhecimento podem ser formulados como problemas de correspondência de grafos (*graph matching*). A representação de cenas por meio de grafos é feita associando-se a cada vértice um objeto. As arestas representam as relações entre tais objetos.

Imagens a serem reconhecidas devem ser processadas por um método de segmentação automática. Neste trabalho, assume-se que é utilizado um método de supersegmentação, ou seja, a cena gerada possui mais objetos que o modelo. Tal escolha se deve ao fato de que parece ser mais fácil associar vários objetos a uma região do modelo, do que identificar contornos das regiões do modelo na imagem, no caso de subsegmentação [14].

Diversos fatores representam complicações para os problemas de reconhecimento. Primeiramente, a segmentação de objetos e eventuais imprecisões nas imagens constituem dificuldades usuais para o problema. Além disso, geralmente o modelo apresenta um aspecto esquemático, em virtude de sua generalidade, o que pode resultar em diferenças em relação à imagem. Uma cena pode conter diferentes objetos similares e um objeto pode apresentar variações na forma em diferentes instanciações de uma cena. A estrutura dos objetos, essencial para o processo, não é única, ou seja, diversos modelos estruturais podem ser extraídos de uma mesma cena.

Problemas de correspondência de grafos podem ser abordados de maneira exata, por isomorfismo, ou através de uma correspondência inexata. No primeiro caso, deseja-se obter um mapeamento bijetivo entre os vértices dos dois grafos comparados, de forma que a estrutura das arestas seja mantida. A segunda abordagem pode ser utilizada em aplicações de reconhecimento que apresentam dificuldades para o uso de isomorfismo.

A abordagem exata é ideal para situações nas quais consegue-se facilmente estabelecer correspondências precisas. Para os casos em que uma relação biunívoca não pode ser estabelecida, uma opção para tratamento exato consiste na utilização de operações de edição dos grafos para transformar o problema em isomorfismo [46, 61, 65]. Definições mais formais de isomorfismo de grafos utilizadas na literatura podem ser encontradas em [14].

Encontra-se mais freqüentemente a modelagem por correspondência inexata de grafos na literatura de reconhecimento de imagens. A principal justificativa são os fatores complicadores descritos anteriormente. Neste caso, o problema a ser tratado é o de se obter um mapeamento não biunívoco entre os grafos comparados, ou seja, a estrutura das arestas não é necessariamente mantida. Este trabalho tem foco na abordagem por correspondência inexata de grafos.

2.2

Abordagens do problema de correspondência de grafos

Boeres [14] realizou um extenso trabalho de revisão bibliográfica e descreveu diversas abordagens de solução do problema de correspondência de grafos. Esta seção apresenta uma breve revisão sobre o tema, baseando-se no trabalho de Boeres.

As maiores dificuldades encontradas na resolução de problemas de correspondência de grafos consistem em encontrar maneiras para se realizar o cálculo das medidas de similaridade entre os grafos e buscar a melhor correspondência de forma eficiente.

Em relação ao cálculo de medidas de similaridade, diversas são as técnicas encontradas na literatura. Operações de edição dos grafos (inserção e remoção de nós e arestas) são utilizadas em algumas dessas técnicas. Em geral, para a realização dos cálculos relativos às operações de edição

é necessária a definição do conceito de distância de um grafo ao outro [23, 61, 63, 64, 67, 68, 69, 70]. Outra possibilidade é a utilização de modelos probabilísticos, os quais podem manipular aspectos de imperfeição da informação [20, 21, 48].

Diversas abordagens, exatas e aproximadas, para resolução do problema de correspondência de grafos são encontradas na literatura e utilizam manipulação de grafos ou descrições relacionais [63]. São apresentados métodos para o cálculo de similaridade entre dois grafos [61, 63, 64] e algoritmos para resolução do problema de isomorfismo [23, 42, 43, 69], entre outros.

Técnicas de inteligência artificial são também utilizadas para adicionar o conceito de aprendizado ao processo de reconhecimento, que pode ser realizado, por exemplo, por um sistema especialista [48].

Modelos probabilísticos são utilizados com o intuito de escapar da convergência prematura para ótimos locais. São propostos na literatura algoritmos genéticos e *simulated annealing* [36, 39, 40]. Outro exemplo de abordagem probabilística foi proposto por Bengoetxea et al. [11], na qual foi desenvolvido um algoritmo de busca que manipula populações de indivíduos (*estimation distribution algorithms*).

Outro tipo de abordagem encontrado na literatura é a utilização da teoria de conjuntos difusos, que oferece uma grande variedade de ferramentas para a modelagem de imprecisões inerentes ao problema de correspondência de grafos. Em [52] são propostos mecanismos de modelagem e manipulação das cenas. A imagem e o modelo são descritos por grafos relacionais de atributos difusos e a correspondência é feita por morfismo difuso de grafos. Perchant [52] também apresenta a teoria que fundamenta o morfismo difuso de grafos. Finalmente, conjuntos difusos também podem ser usados para calcular medidas de similaridades entre os grafos [18, 50, 51, 52, 53, 66].

2.3

Descrição e formulação do problema

Na representação do problema de correspondência inexata de grafos (PCIG) são utilizados dois grafos não direcionados: $G_M = (V_M, E_M)$ representa o modelo e $G_I = (V_I, E_I)$ a imagem supersegmentada a ser

reconhecida. Portanto, G_I possui mais vértices que G_M . Uma solução para este problema representa uma correspondência entre os vértices dos grafos G_I e G_M . Para cada região do modelo, representada por um vértice de G_M , procura-se as regiões correspondentes na imagem, representadas por vértices de G_I .

Na Figura 2.2 é apresentado um exemplo de solução para um pequeno exemplo fictício do PCIG. As associações entre vértices são representadas pelos índices no grafo G_I e as associações entre as arestas pelos diferentes padrões de cor e espessura, sendo que as arestas são associadas em função dos vértices. No exemplo, o vértice $1 \in V_I$ está associado ao vértice $a \in V_M$ e o vértice $3 \in V_I$ ao vértice $b \in V_M$, o que implica na associação da aresta $(1, 3) \in E_I$ à aresta $(a, b) \in E_M$. As arestas tracejadas no grafo G_I não foram associadas a arestas de G_M .

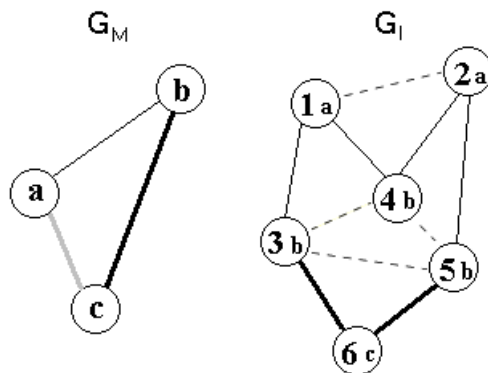


Figura 2.2: Exemplo de solução para o PCIG.

A solução ótima para o problema deveria representar a correspondência na qual a estrutura do grafo modelo fosse mapeada na estrutura do grafo imagem. Dados os grafos G_M e G_I , uma solução para o problema pode ser representada pelo grafo bipartido $G' = (V_M \cup V_I, E')$, onde E' representa as associações entre vértices, ligando cada um dos vértices de V_I a um vértice de V_M .

Uma boa alternativa para a representação de uma solução é pelas listas de adjacências de G' . Para cada vértice $i \in V_M$, define-se o conjunto $A_S(i)$ dos vértices de V_I associados a i numa dada solução S , tal que $A_S(i) = \{j \in V_I | (i, j) \in E'\}$. Para o exemplo da Figura 2.2, $A_S(a) = \{1, 2\}$. A notação $A_S^{-1}(j) = \{i \in V_M | (i, j) \in E'\}$ representa o conjunto unitário que contém o vértice de V_M associado ao vértice j de V_I . Tem-se que $|A_S^{-1}(j)| = 1, \forall j \in V_I$. No exemplo, $A_S^{-1}(6) = c$.

O espaço de busca para este problema é dado por todas as possibili-

dades de associações de vértices. As associações entre arestas são implicitamente derivadas das associações entre vértices, de acordo com a seguinte regra: uma aresta (a, b) de G_M é associada a todas as arestas (a', b') de G_I tais que

$$(A_S^{-1}(a') = a \wedge A_S^{-1}(b') = b) \vee (A_S^{-1}(b') = a \wedge A_S^{-1}(a') = b). \quad (2-1)$$

A condição 2-1 tem como objetivo dar prioridade às associações que preservem a estrutura de G_M em G_I [14], sendo baseada em estudos sobre o morfismo difuso de grafos citado anteriormente [52].

Além dos grafos representando o modelo e a imagem a ser reconhecida, uma instância do PCIG apresenta duas matrizes de similaridade entre os vértices e as arestas. A matriz S^v com $|V_M|$ linhas e $|V_I|$ colunas representa os valores de similaridade entre cada par de vértices (i, j) tais que $i \in V_M$ e $j \in V_I$. A matriz S^a possui $|E_M|$ linhas e $|E_I|$ colunas e representa os valores de similaridade entre cada par de arestas $((i, i'), (j, j'))$ tais que $(i, i') \in E_M$ e $(j, j') \in E_I$. A forma como são calculados os valores de similaridade entre vértices e arestas foge ao escopo deste trabalho.

A Figura 2.3 mostra um exemplo de correspondência ótima para uma instância pequena do PCIG. As Tabelas 2.1 e 2.2 representam as similaridades entre vértices e arestas para este exemplo. Pode-se notar que os valores de similaridades de arestas não são discriminantes, ou seja, não contribuem de forma relevante para a identificação da estrutura de G_M em G_I .

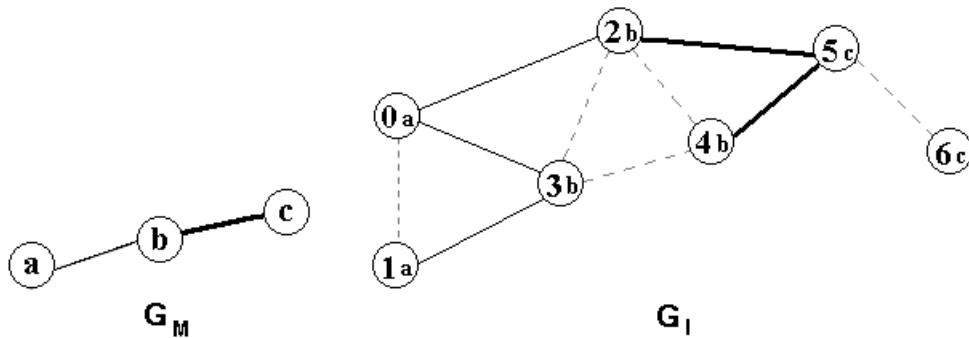


Figura 2.3: Correspondência mais adequada (Exemplo 1).

O segundo exemplo é apresentado na Figura 2.4. Os valores de similaridade mostrados nas Tabelas 2.3 e 2.4 representam um caso em que a matriz de similaridade entre vértices possui informações menos discriminantes que aquela apresentada no Exemplo 1. Neste caso, a matriz

	0	1	2	3	4	5	6
a	0,913	1,000	0,000	0,000	0,000	0,000	0,000
b	0,000	0,000	1,000	1,000	1,000	0,000	0,000
c	0,087	0,000	0,000	0,000	0,000	0,742	0,949

Tabela 2.1: S^v — Similaridades entre vértices de V_M e V_I para o Exemplo 1.

	(0,1)	(0,2)	(0,3)	(1,3)	(2,3)	(2,4)	(2,5)	(3,4)	(4,5)	(5,6)
(a,b)	0,995	1,000	1,000	0,995	1,000	0,995	0,995	0,995	0,995	1,000
(b,c)	0,995	1,000	1,000	0,995	1,000	0,995	0,995	0,995	0,995	1,000

Tabela 2.2: S^a — Similaridades entre arestas de E_M e E_I para o Exemplo 1.

de similaridade de arestas pode tornar-se importante para o processo de correspondência.

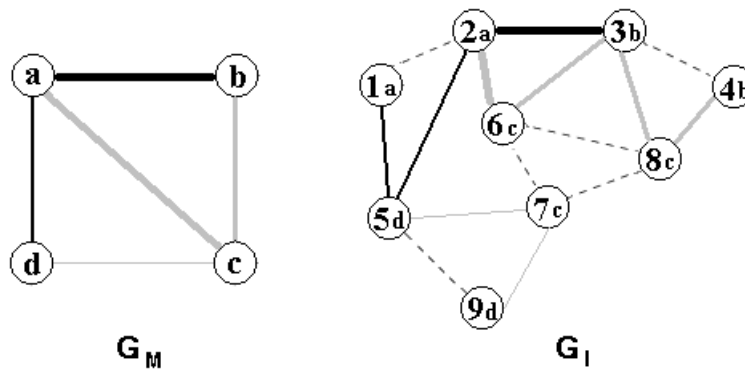


Figura 2.4: Correspondência mais adequada (Exemplo 2).

	1	2	3	4	5	6	7	8	9
a	0,91	1,00	0,75	0,00	0,75	0,85	0,20	0,10	0,00
b	0,00	0,00	1,00	1,00	0,00	0,00	0,00	0,90	0,00
c	1,00	0,00	0,00	0,00	0,00	0,80	1,00	1,00	0,00
d	1,00	0,00	0,00	0,00	1,00	0,00	1,00	0,00	0,70

Tabela 2.3: S^v — Similaridades entre vértices de V_M e V_I para o Exemplo 2.

A seguir é apresentada a formulação do PCIG como um problema de otimização combinatória, proposta por Boeres [14]. Dados os grafos (G_M e G_I) que representam o modelo e a imagem supersegmentada, bem como as matrizes de similaridade entre vértices e arestas, o problema de correspondência inexata de grafos (PCIG) consiste em determinar a melhor correspondência entre os dois grafos, maximizando a medida de similaridade

	(1,2)	(1,5)	(2,3)	(2,5)	(2,6)	(3,4)	(3,6)	(3,8)
(a,b)	0,99	1,00	1,00	0,99	1,00	0,99	0,99	0,99
(a,c)	0,99	1,00	1,00	0,99	1,00	0,99	0,99	0,99
(a,d)	0,51	0,50	0,50	0,51	0,50	0,51	0,51	0,51
(b,c)	0,99	1,00	1,00	0,99	1,00	0,99	0,99	0,99
(c,d)	0,99	1,00	1,00	0,99	1,00	0,99	0,99	0,99
	(4,8)	(5,7)	(5,9)	(6,7)	(6,8)	(7,8)	(7,9)	
(a,b)	0,99	1,00	0,98	0,99	0,99	0,99	1,00	
(a,c)	0,99	1,00	0,98	0,99	0,99	0,99	1,00	
(a,d)	0,51	0,50	0,52	0,51	0,51	0,51	0,50	
(b,c)	0,99	1,00	0,98	0,99	0,99	0,99	1,00	
(c,d)	0,99	1,00	0,98	0,99	0,99	0,99	1,00	

Tabela 2.4: S^a — Similaridades entre arestas de E_M e E_I para o Exemplo 2.

dada pela função objetivo $f(x)$ dada por

$$f(x) = \frac{\alpha}{|V_M| \cdot |V_I|} \cdot f^v(x) + \frac{1 - \alpha}{|E_M| \cdot |E_I|} \cdot f^a(x), \quad (2-2)$$

com

$$f^v(x) = \sum_{i \in V_M} \sum_{j \in V_I} (1 - |x_{ij} - s^v(i, j)|) \quad (2-3)$$

e

$$f^a(x) = \sum_{(i, i') \in E_M} \sum_{(j, j') \in E_I} (1 - |\max\{x_{ij}x_{i'j'}, x_{i'j}x_{ij'}\} - s^a((i, i'), (j, j'))|), \quad (2-4)$$

sujeito às restrições

$$|A_S^{-1}(j)| = 1, \forall j \in V_I, \quad (2-5)$$

$$G_I^{Ind}(A_S(i)) \text{ é conexo}, \forall i \in V_M, \quad (2-6)$$

$$A_S^{-1}(j) = i \Rightarrow s^v(i, j) > 0, \forall (i, j) \in (V_M \times V_I), \quad (2-7)$$

$$|A_S(i)| \geq 1, \forall i \in V_M, \quad (2-8)$$

onde $\alpha \in [0, 1]$ é um parâmetro utilizado para ponderar cada termo de $f(x)$ (valores de α próximos de 1 indicam um peso maior para a contribuição de vértices). A variável x_{ij} assume o valor 1 se o vértice $i \in V_M$ estiver associado ao vértice $j \in V_I$, isto é, se $A_S^{-1}(j) = i$; 0 em caso contrário. O termo $s^v(i, j)$ é o valor de similaridade entre os vértices $i \in V_M$ e $j \in V_I$ e o

termo $s^a((i, i'), (j, j'))$ é o valor de similaridade entre as arestas $(i, i') \in E_M$ e $(j, j') \in E_I$.

O primeiro termo do lado direito da Equação 2-2 representa a contribuição média dos vértices para a solução. A função $f^v(x)$, definida na Equação 2-3, corresponde à soma das contribuições para cada par de vértices, sendo cada contribuição dada por um termo $t_{ij} = (1 - |x_{ij} - s^v(i, j)|)$. Este é um número real no intervalo $[0, 1]$ e assume valores limites em quatro situações possíveis:

1. $x_{ij} = 0$ e $s^v(i, j) = 0 \Rightarrow t_{ij} = 1$, ou seja, contribuição com valor máximo, uma vez que não se associou vértices com similaridade nula.
2. $x_{ij} = 0$ e $s^v(i, j) = 1 \Rightarrow t_{ij} = 0$, ou seja, contribuição com valor mínimo, pois a não associação de vértices com similaridade alta é penalizada.
3. $x_{ij} = 1$ e $s^v(i, j) = 0 \Rightarrow t_{ij} = 0$, ou seja, contribuição com valor mínimo, pois a associação de vértices com similaridade baixa é penalizada.
4. $x_{ij} = 1$ e $s^v(i, j) = 1 \Rightarrow t_{ij} = 1$, ou seja, contribuição com valor máximo, pois é desejável a associação de vértices com similaridade alta.

Uma análise similar pode ser feita em relação à função $f^a(x)$ apresentada na equação 2-4, que representa a contribuição de arestas na equação 2-2. O termo $\max\{x_{ij}x_{i'j'}, x_{i'j}x_{ij'}\}$ indica se duas arestas estão associadas em função das associações de suas extremidades.

A função $f(x)$ é calculada a partir da avaliação das contribuições de vértices e arestas. Cada uma destas contribuições possui um termo constante que pode ser calculado previamente e equivale ao valor da função objetivo quando nenhuma associação entre os grafos foi realizada. O valor de $f(x)$ pode ser então calculado pela adição do custo incremental relativo a cada associação realizada. O desenvolvimento a seguir mostra como podem ser feitas a inicialização da contribuição de vértices e as posteriores atualizações.

$$\begin{aligned}
 f^v(x) &= \sum_{i \in V_M} \sum_{j \in V_I} (1 - |x_{ij} - s^v(i, j)|) = \\
 &= \sum_{(i,j) \in V_M \times V_I - E'} (1 - |0 - s^v(i, j)|) + \sum_{(i,j) \in E'} (1 - |1 - s^v(i, j)|) = \\
 &= \sum_{(i,j) \in V_M \times V_I} (1 - s^v(i, j)) - \sum_{(i,j) \in E'} (1 - s^v(i, j)) + \sum_{(i,j) \in E'} (1 - |1 - s^v(i, j)|) = \\
 &= \sum_{(i,j) \in V_M \times V_I} (1 - s^v(i, j)) + \sum_{(i,j) \in E'} (2s^v(i, j) - 1)
 \end{aligned}$$

Assim, $f^v(x) = \sum_{(i,j) \in V_M \times V_I} (1 - s^v(i, j)) + \sum_{(i,j) \in E'} (2s^v(i, j) - 1)$, onde $f_c^v = \sum_{(i,j) \in V_M \times V_I} (1 - s^v(i, j))$ é uma constante e $f_{ij}^v = \sum_{(i,j) \in E'} (2s^v(i, j) - 1)$ é o custo incremental relativo à associação de $i \in V_M$ a $j \in V_I$.

O termo relativo à contribuição de arestas pode ser calculado e atualizado de forma análoga.

Com a restrição apresentada na equação 2-5, assume-se que o método de segmentação adotado não gera segmentos da imagem que correspondam simultaneamente a mais de uma região adjacente do modelo, ou seja, os contornos do modelo são preservados. Neste caso, cada vértice de V_I deve se associar a exatamente um vértice de V_M .

Na restrição 2-6, $G_I^{Ind}(A_S(i))$ representa o subgrafo induzido em G_I por $A_S(i)$. Este subgrafo deve ser conexo para todo i em V_M . Esta restrição foi criada com o intuito de evitar uma correspondência incomum em situações reais, pois um método de supersegmentação pode partir um objeto em pedaços, mas não altera as posições desses pedaços.

A restrição 2-7 impede que pares de vértices com similaridades nulas sejam associados.

A última restrição (2-8) garante que todas as regiões do modelo foram mapeadas na imagem, ou seja, cada vértice de V_M é associado a pelo menos um vértice de V_I .

2.4

Discussão

Este capítulo apresentou uma breve revisão bibliográfica sobre o problema de correspondência inexata de grafos. Uma revisão mais detalhada pode ser encontrada em [14]. Também foi descrita sua formulação como um problema de otimização combinatória. A função objetivo apresentada incorpora condições provenientes do morfismo difuso de grafos [52], com o objetivo de assegurar que a estrutura do modelo seja mapeada na imagem. Os algoritmos aproximados propostos no Capítulo 4 utilizam a função objetivo apresentada, respeitando também as restrições impostas na formulação.

Esta formulação apresenta componentes não lineares na função objetivo e na restrição 2-6. O próximo capítulo propõe uma formulação original

como um problema de programação linear inteira que permite uma abordagem exata para resolução do PCIG.