

# Uma Nova Modelagem Matemática para o Problema da Rotulação Cartográfica de Pontos

Glaydston Mattos Ribeiro  
Instituto Nacional de Pesquisas Espaciais  
E-mail: glaydston@lac.inpe.br

Luiz Antonio Nogueira Lorena  
Instituto Nacional de Pesquisas Espaciais  
E-mail: lorena@lac.inpe.br

## Resumo

O problema da rotulação cartográfica de pontos consiste no processo de rotular pontos em um mapa, procurando a maior visibilidade possível. Este trabalho apresenta uma nova modelagem para esse problema, reduzindo o número de restrições geradas e permitindo encontrar a solução ótima de problemas que até então, não tinham sido relatados na literatura.

**Palavras-chave:** rotulação, modelagem e otimização combinatória.

## 1. Introdução

O termo “rotulação cartográfica de pontos” consiste no processo de inserção de texto de objetos cartográficos (pontos) em um documento cartográfico, também conhecido como “carta” ou mapa (Freeman, 1991). Esse problema é de difícil solução e tem recebido maior atenção por parte dos pesquisadores (Wolff e Strijk, 1996).

O problema da rotulação cartográfica de pontos (PRCP) pode ser tratado como um problema de otimização combinatória. O problema consiste em combinar as posições dos rótulos dos pontos em determinadas posições candidatas, de tal forma que atenda a um conjunto de restrições, minimizando ou maximizando uma função objetivo. Posições candidatas são as possíveis posições que o rótulo de um determinado ponto pode ocupar, definidas através de uma padronização cartográfica (Christensen et al., 1995). A Figura 1 mostra um conjunto de 8 posições candidatas para um ponto, sendo que os números indicam a preferência cartográfica, quanto menor for o número, maior será a preferência.

Ao posicionar os rótulos nas posições candidatas, sobreposições podem aparecer, ocasionando problemas na visualização do mapa. Assim, o PRCP pode ser representado através de um grafo  $G=\{V,A\}$ , no qual  $V=\{v_1, v_2, \dots, v_L\}$  representa o conjunto das posições candidatas (vértices), sendo  $L$  o número

total de vértices; e  $A=\{(v_i, v_j): i, j \in V, i \neq j\}$  o conjunto das arestas que indicam as sobreposições (conflitos). A Figura 2 (b) mostra o grafo que representa o problema mostrado na Figura 2 (a). Verifica-se que a posição candidata  $v_3$  possui um conflito com a posição  $v_6$ , e a posição  $v_4$  com  $v_5$  e  $v_6$ .

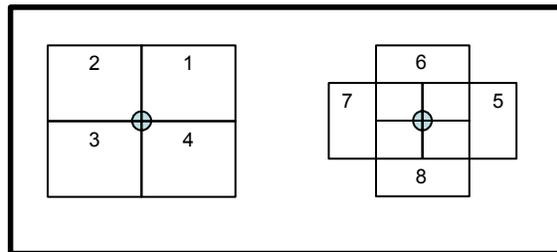


Figura 1. Conjunto de 8 posições candidatas para rotular um ponto (Christensen et al., 1995).

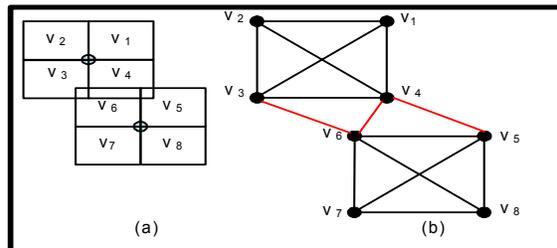


Figura 2. Posições candidatas, conflitos e o grafo relacionado (Yamamoto e Lorena, 2003).

O PRCP pode ser considerado como um problema de Máximo Conjunto Independente de Vértices (Zoraster, 1991; Strijk et al., 2000), como um problema de Maximização do Número de Rótulos Sem Conflito (Christensen et al., 1995 e Yamamoto e Lorena, 2003), e como um problema de Minimização do Número de Conflitos (Ribeiro e Lorena, 2004). As duas últimas abordagens são mais interessantes sob o ponto de vista cartográfico, pois permitem rotular todos os pontos.

O presente trabalho considera o PRCP como um problema de Minimização do Número de Conflitos. Baseada na formulação apresentada por Ribeiro e

Lorena (2004), este trabalho tem como objetivo apresentar uma nova formulação mais eficiente que considera os pontos a serem rotulados. Na próxima seção é apresentada a formulação de Ribeiro e Lorena (2004), seguida pela nova formulação proposta, e por fim, pelos resultados computacionais, as conclusões e os trabalhos futuros.

## 2. Formulação baseada nas arestas do grafo de conflito

A formulação descrita abaixo foi proposta por Ribeiro e Lorena (2004) e tem como objetivo, minimizar o número de conflitos em um grafo de conflitos.

$$v = \text{Min} \sum_{i=1}^N \sum_{j=1}^{P_i} \left( w_{i,j} x_{i,j} + \sum_{(k,t) \in S_{i,j}} y_{i,j,k,t} \right) \quad (1)$$

Sujeito a:

$$\sum_{j=1}^{P_i} x_{i,j} = 1 \quad \forall i = 1 \dots N \quad (2)$$

$$x_{i,j} + x_{k,t} - y_{i,j,k,t} \leq 1 \quad \forall i = 1 \dots N \quad (3)$$

$$\forall j = 1 \dots P_i$$

$$(k,t) \in S_{i,j}$$

$$x_{i,j} \text{ e } y_{i,j,k,t} \in \{0,1\} \quad \forall i = 1 \dots N \quad (4)$$

$$\forall j = 1 \dots P_i$$

$$(k,t) \in S_{i,j}$$

Sendo:

- $N$  o número de pontos a serem rotulados;
- $P_i$  o número de posições candidatas do ponto  $i$ ;
- $x_{i,j}$  a posição candidata  $j$  do ponto  $i$ ;
- $w_{i,j}$  a preferência cartográfica associada a posição candidata  $x_{i,j}$ ;
- $S_{i,j}$  um conjunto de pares de índices  $(k,t)$  de posições candidatas  $x_{k,t}$  conflitantes com  $x_{i,j}$  tal que  $k > i$ ;
- $y_{i,j,k,t}$  uma variável de conflito que representa a sobreposição das posições candidatas  $x_{i,j}$  e  $x_{k,t}$ .

A restrição (2) garante que para cada ponto  $i$ , seu rótulo deve estar em uma das suas posições candidatas  $P_i$ . A restrição (3) é a que considera as sobreposições no grafo de conflitos. Se duas posições candidatas conflitantes  $x_{i,j}$  e  $x_{k,t}$  recebem os seus respectivos rótulos, a variável de conflito  $y_{i,j,k,t}$  receberá o valor 1 penalizando assim, a função objetivo. O modelo procura então, encontrar a solução que apresenta o menor número de conflitos possíveis. A restrição (4) é uma restrição que indica que todas as variáveis são binárias.

Como normalmente o grafo de conflitos é esparsos, o número de restrições de conflito é muito grande o que torna o processo de solução demorado. Essa formulação foi aplicada por Ribeiro e Lorena

(2004) a problemas testes propostos por Yamamoto e Lorena (2003), no entanto para problemas com 750 e 1000 pontos, o CPLEX (ILOG, 2001) não conseguiu obter as soluções ótimas, motivando assim a busca por uma modelagem mais eficiente.

## 3. Formulação baseada nos pontos a serem rotulados

Dado que para todo ponto  $i$  apenas uma das posições candidatas receberá o rótulo (restrições definidas em (2)), as restrições de conflitos podem ser representadas entre pontos ao invés de posições candidatas. Assim, além das variáveis e conjuntos definidos por Ribeiro e Lorena (2004), seja  $C_{i,j}$  um conjunto com todos os pontos que possuem posições candidatas em conflito com a posição candidata  $x_{i,j}$ , e  $y_{i,j,c}$  uma variável de conflito entre a posição candidata  $x_{i,j}$  e o ponto  $c \in C_{i,j} : c > i$ . Com isso, as restrições (3) podem ser substituídas, sem falta de generalização, pelo seguinte conjunto de restrições:

$$|C_{i,j}| x_{i,j} + \sum_{(k,t) \in S_{i,j}} x_{k,t} - \sum_{c \in C_{i,j}} y_{i,j,c} \leq |C_{i,j}| \quad \forall i = 1 \dots N \quad (5)$$

$$\forall j = 1 \dots P_i$$

Como o conjunto descrito acima utiliza variáveis de sobreposições que indicam conflitos entre posições candidatas e pontos, a função objetivo também deve ser modificada, assim, o PRCP pode ser modelado como:

$$v = \text{Min} \sum_{i=1}^N \sum_{j=1}^{P_i} \left( w_{i,j} x_{i,j} + \sum_{c \in C_{i,j}} y_{i,j,c} \right) \quad (7)$$

Sujeito a:

$$\sum_{j=1}^{P_i} x_{i,j} = 1 \quad \forall i = 1 \dots N \quad (8)$$

$$|C_{i,j}| x_{i,j} + \sum_{(k,t) \in S_{i,j}} x_{k,t} - \sum_{c \in C_{i,j}} y_{i,j,c} \leq |C_{i,j}| \quad \forall i = 1 \dots N \quad (9)$$

$$\forall j = 1 \dots P_i$$

$$x_{i,j} \text{ e } y_{i,j,c} \in \{0,1\} \quad \forall i = 1 \dots N \quad (10)$$

$$\forall j = 1 \dots P_i$$

$$c \in C_{i,j}$$

A Tabela 1 mostra os números médios de restrições geradas por cada uma das modelagens acima. Os problemas utilizados foram os propostos por Yamamoto e Lorena (2003) e estão disponíveis em [www.lac.inpe.br/~lorena](http://www.lac.inpe.br/~lorena). A primeira coluna mostra o número de pontos, seguidos pelo número de problemas, pelo número médio de restrições para a formulação de Ribeiro e Lorena (2004) e pelo número médio de restrições geradas pela nova modelagem. Percebe-se que a nova formulação reduz de forma considerável o número de restrições.

**Tabela 1.** Número médio de restrições geradas

Número de pontos	Número de problemas	Ribeiro e Lorena (2004)	Nova formulação
25	8	357	96
100	25	202	153
250	25	864	530
500	25	2909	1412
750	25	6181	2481
1000	25	10700	3643

#### 4. Resultados computacionais, conclusões e trabalhos futuros

O resultados mostrados na Tabela 2 são pré-eliminatórios e comprovam a eficiência da nova formulação, sendo LAGCLUS uma relaxação Lagrangeana obtida após a divisão do grafo de conflitos em *clusters*. Os problemas testes foram os propostos por Yamamoto e Lorena (2003), com 4 posições candidatas para cada ponto, sem preferência cartográfica, ou seja,  $w_{ij}=1 \forall i=1...N$  e  $\forall j=1...4$ . Observa-se que, para problemas com até 500 pontos, os resultados são semelhantes com exceção do tempo computacional. Para os problemas com 750 pontos, o CPLEX foi utilizado até encontrar a solução ótima, ou parar por falta de memória (todos os testes foram feitos em um computador com processador Pentium 2,66 GHz e 512 MB de memória). Para esses problemas, apesar da nova formulação apresentar um tempo maior de processamento, o percentual de rótulos livres foi melhor. Isto porque a nova formulação permitiu encontrar as soluções ótimas para 12 dos 25 problemas, enquanto que, com a formulação de Ribeiro e Lorena (2004), o CPLEX só encontrou a solução ótima em 9.

Como trabalhos futuros, deve-se verificar o seu comportamento com relação aos problemas com 1000 pontos, e testar algumas relaxações como a Lagrangeana, Lagrangeana/Surrogate e a LAGCLUS que, como mostrou a Tabela 2, apresentou bons

resultados, com tempo computacional baixo. Espera-se com essas relaxações obter as soluções ótimas para os problemas de 750 e 1000 pontos com um tempo computacional menor do que o utilizado pelo CPLEX.

#### 5. Referências

- [1] Christensen, J., Marks, J., e Shieber, S. (1995) An Empirical Study of Algorithms for Point-Feature Label Placement. *ACM Transactions on Graphics*, v 14, n 3, p 203-232.
- [2] ILOG, (2001), *ILOG Cplex 7.5 Reference Manual*. ©Copyright 2001 by ILOG.
- [3] Freeman, H., (1991), Computer name placement. In: Maguire, D. J.; Goodchild, M. F., Rhind, D. W. eds. *Geographic Information Systems: Principles and Applications*. New York, Longman Scientific, v 1, p 445-456.
- [4] Wolff, A., e Strijk, T., (1996), *A Map Labeling Bibliography*. (On line) <http://i11www.ira.uka.de/map-labeling/bibliography/>.
- [5] Ribeiro, G. M., e Lorena, L. A. N., (2004), Modelagem Matemática e Relaxações Lagrangeana e Lagrangeana/Surrogate para o Problema de Rotulação Cartográfica de Pontos. *SBPO XXXVI* (Aceito para apresentação) – São João Del Rei – MG
- [6] Strijk, T., Verweij, B. e Aardal, K., (2000), *Algorithms for Maximum Independent Set Applied to Map Labeling*. <ftp://ftp.cs.uu.nl/pub/RUU/CStechreps/CS-2000/2000-22.ps.gz>.
- [7] Yamamoto, M., e Lorena, L.A.N., (2003), A Constructive Genetic Approach to Point-Feature Cartographic Label Placement. *In the Fifth Metaheuristics International Conference*, Japão.
- [8] Zoraster, S. (1991), Expert Systems and the Map Label Placement Problem. *Cartographica*, v. 28, n 1, p. 1-9.

**Tabela 2.** Quadro comparativo das formulações e da heurística LAGCLUS (Ribeiro e Lorena, 2004)

Modelos e Heurística	Parâmetros avaliados	Problemas - Número de pontos				
		25	100	250	500	750
Nova formulação	Tempo (s)	0,20	0,02	0,06	0,74	9625,92
	Rótulos em conflitos	4,88	0,00	0,00	1,68	17,84
	Rótulos livres (%)	80,50	100,00	100,00	99,67	97,62
Formulação de Ribeiro e Lorena (2004)	Tempo (s)	1,60	0,02	0,06	3,12	6586,88
	Rótulos em conflitos	4,88	0,00	0,00	1,68	18,40
	Rótulos livres (%)	80,50	100,00	100,00	99,67	97,55
LAGCLUS (Ribeiro e Lorena, 2004)	Tempo (s)	23,88	0,16	2,36	82,72	337,80
	Rótulos em conflitos	5,63	0,00	0,00	3,08	33,56
	Rótulos livres (%)	77,50	100,00	100,00	99,38	95,53