O Problema da Cobertura de Conjuntos com Facilidades Alocadas a P-Medianas

Rodolfo Ranck Júnior
Instituto Nacional de Pesquisas Espaciais – INPE. Caixa Postal 515 – 12.227-010 – São José dos Campos – SP – Brasil

***Resumo* — Neste trabalho propõe-se resolver um problema de cobertura em que as facilidades a serem localizadas precisam estar alocadas a *p* centros medianos. Trata-se de um problema com dois objetivos em mútua relação dependência. Formula-se este problema como um modelo de programação linear inteira, e desenvolvem-se métodos de solução heurísticos para resolvê-lo. Alguns resultados computacionais são apresentados comparando as soluções exatas com as obtidas através das heurísticas propostas.**

***Palavras-chaves* — Análise Operacional, Problema da Cobertura de Conjuntos com Facilidades Alocadas a P-Medianas, Otimização combinatória.**

I. INTRODUÇÃO

Problemas de localização, neste contexto, tratam de decisões sobre onde localizar facilidades (fábricas, depósitos, antenas, centros de saúde, etc.) em uma rede na qual existem pontos de demanda (clientes, áreas de monitoramento, pacientes, etc.) que necessitam ser atendidos em função de um determinado objetivo a ser otimizado [1,2]. Dentre os possíveis objetivos estão minimizar o número de facilidades atendendo todos os pontos de demanda e minimizar a distância entre pontos de demanda e suas respectivas facilidades, cuja disponibilidade é limitada.

Devido à importância desses problemas para a economia, a diversidade com que ocorrem na prática e a dificuldade em resolvê-los, esses problemas vem sendo estudado por um número crescente de pesquisadores, gerando uma literatura extensa a respeito [3]. Os problemas de localização têm grande valor operacional nas áreas de defesa, quando, por exemplo, há interesse em posicionar radares de vigilância, tropas, redes de sensores, aeronaves, avaliar conflitos de rotas aéreas, localizar aeroportos, entre outros ([1] e [4]).

A seguir, apresenta-se o Problema de Cobertura de Conjuntos (PCC) e o Problema de P-Medianas (PPM), por serem relevantes na formulação do modelo proposto para o problema focado neste trabalho.

O PCC consiste em, dado uma rede, determinar a localização de um número mínimo de facilidades de forma a cobrir (atender) todos os pontos de demanda. Nesse problema, para que um ponto de demanda seja atendido por uma facilidade, a distância entre eles deve ser menor que o raio de cobertura (raio de alcance máximo do serviço) dessa facilidade; nesse caso dizemos que esse ponto de demanda está coberto.

Rodolfo Ranck Júnior. rodolforanck@gmail.com, Tel +55-11-8932-8932. O presente trabalho foi realizado com apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil.

O modelo do PCC ([5]) pode ser apresentado como um problema de programação linear inteira na forma (1-3):

$$\text{Min } z = \sum_{j \in N} y_j \qquad (1)$$

$$\text{s.a.} \quad \sum_{j \in N} a_{ij} y_j \geq 1 \; ; \quad i, j \in N \qquad (2)$$

$$y_j \in \{0,1\}, \; j \in N \qquad (3)$$

em que:

- N é o conjunto dos nós N={1,... n};
- n é o número de nós;
- a_{ij} é um parâmetro que toma o valor 1 se um ponto de demanda i puder ser coberto por uma facilidade alocada na posição (nó) j e 0, caso contrário. Isso é:

$$a_{ij} = \begin{cases} 1, & se \; d_{ij} < d \\ 0, & caso \; contrário \end{cases}$$

- d é o raio de cobertura, a distância máxima que um ponto de demanda pode estar de uma facilidade para ele ser coberto;
- d_{ij} é a distância entre os nós *i* e *j;*
- y_j é uma variável binária que toma o valor 1 se há uma facilidade alocada ao nó *j* e 0, caso contrário.

Na formulação anterior, a função objetivo (1) minimiza o número de facilidades alocadas. A restrição (2) impõe que um ponto de demanda deve ser coberto por pelo menos uma facilidade. A restrição (3) impõe que as variáveis de decisão só podem assumir os valores 0 ou 1.

A localização de *p* medianas é um problema bem conhecido de otimização combinatória. O objetivo é, dado um grafo, localizar (definir) *p* nós denominados medianas, de forma a minimizar a soma das distâncias de cada nó até sua mediana mais próxima. As *p* medianas também podem ser entendidas como facilidades, mas, para não confundir com as facilidades do PCC, neste trabalho elas serão chamadas apenas de medianas.

As primeiras formulações de programação matemática para o PPM podem ser encontradas em [6] e [7]. No caso em que as medianas são não capacitadas, o PPM pode ser modelado como um problema de programação linear inteira conforme (4-8).

$$\text{Min } z = \sum_{k\in N}\sum_{j\in N} d_{kj}x_{kj} \quad (4)$$

$$\text{s.a.} \quad \sum_{k\in N} x_{kj} = 1, \quad j \in N \quad (5)$$

$$\sum_{k\in N} x_{kk} = p \quad (6)$$

$$x_{kj} \le x_{kk}; \quad k,j \in N \quad (7)$$

$$x_{kj} \in \{0,1\}; \quad k,j \in N \quad (8)$$

em que:

d_{kj} é a distância entre os nós *j* e *k;*

x_{kj} é uma variável binária, com $x_{kj} = 1$ se o nó *j* está alocado à mediana *k* ou $x_{kj} = 0$ caso não esteja*;* x_{kk}*=1* se o nó *k* é uma mediana e x_{kk}=0, caso contrário;

p é um número inteiro e positivo de medianas que se deseja localizar.

Na formulação anterior a função objetivo (4) minimiza a distância total entre os pontos de demanda e suas respectivas medianas. A restrição (5) impõe que todo nó *j* deve estar alocado a apenas um nó *k*. A restrição (6) impõe que o número total de medianas seja *p*. A restrição (7) impõe que um nó *j* só pode estar alocado ao nó *k* se neste nó *k* houver uma mediana. A restrição (8) impõe que as variáveis de decisão só podem assumir os valores 0 ou 1.

Os problemas de otimização PCC e PPM, são NP-difíceis (detalhes sobre complexidade podem ser encontrados em [8]), e na maior parte das vezes é computacionalmente difícil resolvê-los utilizando as formulações (1-3) e (4-8), respectivamente. Por este motivo, desde o surgimento dos primeiros modelos de programação matemática para o PCC e para o PPM, nas décadas de 60 e 70, heurísticas têm sido desenvolvidas para esses problemas; veja, por exemplo, [9]-[14] para o caso do PCC e [15]-[18] para o caso do PPM.

No Problema da Cobertura de Conjuntos com Facilidades alocadas a P-Medianas (PCC-PPM), foco deste trabalho, deseja-se localizar o menor número de facilidades (que possuem um raio de cobertura limitado) de modo a atender pontos de demanda espacialmente distribuídos e ainda alocar essas facilidades à *p* medianas com o menor custo. Considera-se o caso não capacitado, isso é, nem as facilidades, nem as *p* medianas possuem um limite de capacidade para o número de nós que podem atender/alocar.

Em contextos práticos, o PCC-PPM pode surgir sempre que for preciso alocar facilidades em uma região e também *p* pontos para atender a eventuais necessidades dessas facilidades. Considere um exemplo em que se deseja localizar antenas para o fornecimento de um serviço em uma cidade e *p* centros para abrigar equipes que realizam a manutenção nessas antenas (fisicamente e no local). Deseja-se primeiramente que a demanda pelo serviço seja atendida com um número mínimo de antenas e depois que os *p* centros de manutenção estejam, ao mesmo tempo, o mais próximo possível de todas as antenas, isso é, que esses *p* centros sejam *p* medianas

Neste trabalho, considera-se que o principal objetivo é minimizar o número de facilidades cobrindo todos os pontos de demanda. Assim, considera-se como objetivo localizar um número mínimo de facilidades alocando-as à *p* medianas com o menor custo (não necessariamente o mínimo). Esse problema é diferente de resolvermos o primeiro objetivo (PCC) e depois resolvermos o segundo (PPM) para as facilidades localizadas, isso por que pode haver diversas maneiras distintas de localizar um número mínimo de facilidades e, para cada uma delas, uma maneira de alocar as facilidades às medianas com o menor custo.

II. O MODELO PROPOSTO PARA O PCC-PPM

A seguir, formula-se um problema de programação linear inteira para o PCC-PPM (9-15).

$$\text{Min } z = \sum_{j\in N} y_j + \alpha \sum_{k\in N}\sum_{j\in N} d_{kj}x_{kj} \quad (9)$$

$$\text{s.a.} \quad \sum_{j\in N} a_{ij}y_j \ge 1, \quad i \in N \quad (10)$$

$$\sum_{k\in N} x_{kk} = p \quad (11)$$

$$\sum_{k\in N} x_{kj} = y_j, \quad j \in N \quad (12)$$

$$x_{kj} \le x_{kk}; \quad k,j \in N \quad (13)$$

$$x_{kj} \in \{0,1\}; \quad k,j \in N \quad (14)$$

$$y_j \in \{0,1\}, \quad j \in N \quad (15)$$

em que:

α peso para compatibilizar a grandeza dos dois objetivos de interesse;

e os demais parâmetros e variáveis iguais aos definidos anteriormente.

Na formulação anterior, a função objetivo (9) minimiza o custo com o número de facilidades localizadas e com a alocação dessas facilidades às *p* medianas. A restrição (12) impõe que um nó só pode estar alocado a uma mediana se nele estiver alocada uma facilidade. As outras restrições são descritas conforme anteriormente.

Como se deseja que a solução do problema (9-15) seja mínima para o número de facilidades localizadas, enquanto que o custo com a alocação das facilidades às *p* medianas seja o menor possível, é suficiente impor que o custo em minimizar uma unidade do número de facilidades seja sempre maior que minimizar o custo com a alocação de *p* medianas, isso é:

$$\alpha \sum_{k\in N}\sum_{j\in N} d_{kj}x_{kj} < 1 \quad (16)$$

Um limitante superior para $\sum_{k\in N}\sum_{j\in N} d_{kj}x_{kj}$ é dado por $\sum_{k\in N}\sum_{j\in N} d_{kj}$, a soma de todos os possíveis caminhos entre dois nós em toda rede. Portanto, α em (16) pode ser definido por:

$$0 < \alpha < \frac{1}{\sum_{k\in N}\sum_{j\in N} d_{kj}} \quad (17)$$

Da mesma forma que os problemas PCC e PPM, devido à condição de integralidade das variáveis de decisão, o PCC-PPM pode ser computacionalmente difícil de ser resolvido, principalmente quando a quantidade de pontos do problema é suficientemente grande. Diante desse inconveniente, propõem-se heurísticas para resolvê-lo de maneira aproximada.

III. HEURÍSTICAS PARA RESOLVER O PCC-PPM

Conforme mencionado na introdução, resolver o PCC-PPM é diferente de resolver o PCC correspondente e depois resolver o PPM para as facilidades alocadas. Entretanto, ao resolver o PCC-PPM separadamente, dessa maneira, tem-se uma solução factível (pontos de demanda cobertos por facilidades alocadas a p medianas) e, com sorte, uma boa solução. É nesse princípio que se baseia a Heurística 1 (H1).

Para resolver o PPM quando os pontos de demanda são aqueles onde foram alocadas facilidades (PPM$^{\#}(y)$), basta reescrever sua restrição de alocação (5) conforme (12).

Seja PPM$^{\#}(y)$ o problema de programação linear inteira dado por (4), (6), (7), (8), (12). Os passos a seguir descrevem a heurística H1:

1. Obtenha $y_j,\ j = 1, \dots, n$ resolvendo o PCC;
2. Resolva o PPM(y);
3. A solução aproximada do PCC-PMM é dada por y_j, $j \in N$ e $x_{kj};\ k, j \in N$.

Uma vantagem em separar o problema PCC-PPM em dois, dessa maneira, é que seu conjunto de restrições é relaxado facilitando sua resolução computacional aproximada. Além disto, existem diversas heurísticas encontradas na literatura para resolver os problemas PCC e PPM, vide, por exemplo. [9]-[14] e [15]-[18], respectivamente.

Em busca de melhores resultados, seria interessante gerar soluções em que as facilidades estão o mais perto possível uma das outras, pois, com isso, maiores são as chances de encontrarmos p medianas suficientemente próximas de todas as facilidades ao mesmo tempo (o que diminui os custos com a alocação dessas medianas).

Uma maneira de conseguir isso é procurando recobrir os pontos de demanda o maior número de vezes, ou seja, cobrir um ponto de demanda com mais de uma facilidade. Fazendo isso, os raios de cobertura das facilidades tendem a ter uma maior sobreposição, o que implica na maior proximidade entre as facilidades. A variação da heurística H1 que considera esse procedimento é chamada de heurística 2 (H2).

O PCC em (1-3) pode ser modificado para considerar o número de vezes que cada ponto de demanda é recoberto por uma facilidade, reescrevendo a função objetivo (1) como (18).

$$\text{Min } z = \sum_{j \in N} y_j - \beta \sum_{i \in N} \sum_{j \in N} a_{ij} y_j \qquad (18)$$

em que:

β	peso para compatibilizar a grandeza dos dois objetivos de interesse;

O termo $\beta \sum_{i \in N} \sum_{j \in N} a_{ij} y_j$ em (18) representa a soma do número de vezes que cada ponto de demanda i é coberto. Considerando o sinal negativo desse termo na função objetivo, para minimizar (18) é interessante que esse termo seja o maior possível, ou seja, que sejam recobertos o maior número de pontos de demanda. Cobrir os pontos de demanda o maior número de vezes é interessante desde que o número mínimo de facilidades seja mantido. Para manter um mínimo de facilidades é suficiente impor que o custo em minimizar uma unidade do número de facilidades seja sempre maior que maximizar o recobrimento dos pontos de demanda isso é:

$$\beta \sum_{i \in N} \sum_{j \in N} a_{ij} y_j \ < 1 \qquad (19)$$

Um limitante superior para $\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} y_j$ pode ser dado pelo total de pontos de demanda do problema (*totalPts*) multiplicado pelo total de facilidades, isso é, no máximo cada ponto de demanda seria coberto por todas facilidades. Como no máximo pode haver *totalPts* facilidades (uma facilidade para cada ponto de demanda), então, esse limitante superior pode ser dado por $totalPts^2$. Portanto, β em (17) pode ser definido por:

$$0 < \ \beta < \frac{1}{totalPts^2} \qquad (20)$$

Seja PCC$^{\#}$ o problema de programação linear inteira dado por (2), (3), (18), (20). Os passos a seguir descrevem a heurística H2:

1. Obtenha $y_j,\ j = 1, \dots, n$ resolvendo o PCC$^{\#}$;
2. Resolva o PPM$(y)^{\#}$;
3. A solução aproximada do PCC-PMM é dada por y_j, $j \in N$ e $x_{kj};\ k, j \in N$.

IV. TESTES COMPUTACIONAIS

Nesta seção são comparadas as soluções ótimas e as soluções obtidas através das heurísticas H1 e H2 para algumas classes de teste do PCC-PPM.

Pelo motivo da função objetivo dos problemas de programação linear inteira do PCC-PPM e das heurísticas terem custos diferentes, após a resolução do PCC-PPM por cada uma das abordagens, todas as soluções são avaliadas através da função (21), apenas a critério de comparação.

$$\text{FN}(\text{x}, \text{y}) = \sum_{j \in N} y_j + \sum_{k \in N} \sum_{j \in N} d_{kj} x_{kj} \qquad (21)$$

Os problemas testes foram gerados de acordo com os parâmetros apresentados na Tabela I. As coordenadas dos nós foram geradas aleatoriamente para o caso bidimensional com a ajuda da função *rand*() do ANSI C. O maior valor possível para as coordenadas é de 99 enquanto o menor é 0. A semente utilizada para inicializar o gerador foi o valor “2009”, a mesma para todos os testes.

TABELA I. PARÂMETROS UTILIZADOS NOS TESTES COMPUTACIONAIS

Classes de Teste (CdT)	Raio de cobertura	Número de nós (pontos de demanda)	Número de medianas (*p*)
1	10	50	2
2	10	150	2
3	30	50	2
4	30	150	2
5	10	50	3
6	10	150	3
7	30	50	3
8	30	150	3

O código foi escrito em linguagem C++ e executado em um computador com processador Intel Centrino Duo T2400, 1,83 Ghz com 1GB de memória RAM do tipo DDR2. O Sistema Operacional utilizado foi o Microsoft Windows XP. Todos os problemas de programação linear inteira foram resolvidos de maneira exata com ajuda do software ILOG CPLEX 11.

Os resultados dos testes computacionais para cada uma das 8 classes estão apresentados na Tabela II. Nessa tabela P^*, P^{H1}e P^{H2} representam, respectivamente, o PCC-PPM resolvido até a otimalidade como (9-15), a heurística H1 e a heurística H2. O tempo de resolução dos problemas de programação linear inteira para P^*, P^{H1}e P^{H2} é dado por t(s).

TABELA II. RESULTADOS DOS TESTES COMPUTACIONAIS

CdT	P^*		P^{H1}		P^{H2}	
	FN(x, y)	t(s)	FN(x, y)	t(s)	FN(x, y)	t(s)
1	702,892	16,062	733,317	0,062	733,317	0,093
2	*	*	1195,79	0,234	1182,52	0,281
3	121.663	57.734	156,821	0,031	137,518	0,031
4	*	*	170,905	0,171	144,143	0,140
5	516,733	5,641	543,67	0,078	543,67	0,046
6	*	*	920,034	0,203	933,874	0,265
7	69,558	25,265	97,744	0,015	75,606	0,015
8	*	*	126,859	0,156	89,911	0,125

* Casos em que após de uma hora de execução o software CPLEX não encontrou a solução ótima.

A partir da Tabela II, observa-se que na maioria dos casos a heurística H2 obtém os melhores resultados entre as heurísticas. Com relação às soluções ótimas encontradas, a menor diferença no valor de $\mathrm{FN}(x, y)$ foi de 26,937 unidades para a heurística 1 na classe de teste 5 e de 6,048 unidades para a heurística 2 na classe de teste 7. Em todos os casos o tempo computacional das heurísticas foi inferior a 0,5 segundos.

Os tempos computacionais para resolver o PCC-PPM até a otimalidade mostraram-se elevados com relação às heurísticas, mesmo para os testes com o menor número de pontos de demanda (classes 1,3,5 e 7). Para as demais, a solução exata não pôde ser obtida num tempo limite de 1 hora.

Observa-se que nos casos em que o raio de cobertura é maior (classes de teste: 3,4,7 e 8), o desempenho da heurística H2 (com relação ao custo $\mathrm{FN}(x, y)$) aumenta em comparação com a heurística H1. Isso acontece, pois quando o raio de cobertura é maior, o número de pontos de demanda que podem ser recobertos também é maior, possibilitando as facilidades estarem mais próximas uma das outras (diminuindo o custo com as medianas).

V. CONCLUSÕES

Neste artigo apresentou-se um modelo de programação linear inteira para o PCC-PPM e, para resolvê-lo aproximadamente, foram propostas heurísticas baseada na relaxação do problema. As heurísticas, em um todo, mostraram ser capazes de gerar soluções factíveis com um baixo tempo computacional para o conjunto de testes avaliado. Em particular, a heurística H2 apresentou, na maioria dos casos, os melhores resultados dentre as heurísticas, com soluções aparentemente próximas ao valor ótimo em um baixo tempo computacional. O conjunto de testes pode ser ampliado para se ter uma avaliação mais robusta do comportamento da heurística proposta.

Nos testes computacionais realizados, todos os problemas de programação linear inteira foram resolvidos de maneira exata. Para problemas de maior porte, esses poderiam ser resolvidos com ajuda de outras heurísticas, por exemplo, aquelas encontradas em [9]-[14] e [15]-[18].

O problema abordado pode ser estendido, considerando variações do problema, por exemplo, no caso que medianas e facilidades são capacitadas.

Esta é uma primeira abordagem para resolução deste problema. Um estudo de novas heurísticas será objeto de desenvolvimentos futuros.

REFERÊNCIAS

[1] M. Daskin, Network and Discrete Location: Models, Algorithms and Applications, Wiley Interscience, New York, EUA, 1995.

[2] Drezner, Z. (ed.) Facility Location: A Survey of Applications and Methods, Springer-Verlag, NY, 1995

[3] Brandeau, M.L.; Chiu, S.S. (1989) “An Overview of Representative Problems in Location Research”. *Management Science*, 35(6), 645-674.

[4] Goldbarg, M.C; Luna, H.P.L. Programacao Linear e Otimizacao Combinatoria: Modelos e Algoritmos. Rio de janeiro : Campus, 2000, v.1. p.639.

[5] Toregas, C.; Swain, R.; ReVelle, C.; Bergman, L. (1971). “The location of emergency service facilities”. *Operations Research*, 19: 1363-1373.

[6] Hakimi, S.L. (1964) “Optimum location of switching centers and the absolute centers and the medians of a graph”. *Operations Research*, 12: 450-459.

[7] Hakimi, S.L. (1965) “Optimum distribution of switching centers in a communication network and some related graph theoretic problems”. *Operations Research*, 13: 462-475.

[8] Garey, M. R. and Johnson, D. S. (1979) Computers and intractability: a guide to the theory of NP-completeness, W. H. Freeman and Co., San Francisco.

[9] Balas, E.; Ho, A. (1980). "Set covering algorithms using cutting planes, heuristics and subgradient optimization: a computational study". *Mathematical Programming S*tudy 12: 37-60

[10] Vasko, F.J.; Wilson, G.R.. (1984). "An efficient heuristic for large set covering problems". *Naval Research Logistics Quarterly*, 31: 163-171.

[11] Beasley, J.E. (1990)."A Lagrangian heuristic for set covering problems". *Naval Research Logistics*, 37: 145-164.

[12] Jacob L. W.; Brusco M. J. (1993). "A simulated annealing based heuristic for the set-covering problem", Working paper, Operations Management and Information Systems Department, Northern Illinois University, Dekalb, IL.

[13] Lorena, L.; Lopes, F.B. (1994). "A surrogate heuristic for set covering problems". *European Journal of Operational Research*, 79: 138-150.

[14] Lorena, L.; Lopes, L. S. (1997). "Genetic algorithms applied to computationally difficult set covering problems". *Journal of the Operational Research Society*, 48: 440-445

[15] Lorena, L.; Senne, E. (2001) Local search heuristics for capacitated p-median problems. Networks and Spatial Economics.

[16] Resende, M.G.C. and Werneck, R.F. (2004) A hybrid heuristic for the p-median problem, J. of Heuristics, vol. 10, pp. 59-88.

[17] Hansen, P. and Mladenovic, N. (1997) Variable neighborhood search for the p-median. Location Science, 5:207-226.

[18] Senne, E. L. F. and Lorena, L. A. N. (2000) Langrangean/surrogate heuristics for p-median problems. In M. Laguna and J. L. Gonz_alez-Velarde, editors, Computing Tools for Modeling, Optimization and Simulation: Interfaces in Computer Science and Operations Research, pages 115-130.