Estruturas de Árvores Disjuntas Obtidas através dos Conjuntos Recíprocos Completos e Não Recíprocos e seu uso na Determinação das p-Medianas de uma Rede.

Eduardo Andrade Veloso

Universidade Estadual da Paraíba – UEPB.
Centro de Ciências e Tecnologia – CCT.
Departamento de Matemática, Estatística e Computação – DMEC.
58.109.763 Campus de Bodocongó, Campina Grande, PB.
E-mail: veloso_ba@uol.com.br

1 – Resumo.

Dada uma rede com N vértices, o problema das p – medianas consiste em determinar *p* facilidades (medianas) de modo que a soma de todas as distâncias mínimas dos (*N* – *p*) pontos de demanda às *p* facilidades seja minimizada. Neste trabalho é mostrado que, quando o número médio de alocações por vértice mediana, calculado por ((*N* – *p*)ú *p*) é menor ou igual a 1, os conjuntos recíprocos completos e os conjuntos não recíprocos a eles associados, geram estruturas de árvores disjuntas.Tal fato faz com que seja possível o desenvolvimento de um limite inferior para os *(N – p)* custos associados aos pontos de demanda, bem como de um algoritmo que busque uma solução ótima para o problema geral, determinando, para cada árvore associada a cada conjunto recíproco completo em particular, qual o número ótimo de facilidades. Caso haja *p* conjuntos recíprocos completos com interseção vazia, uma solução ótima para o problema pode ser determina em *p* iterações e, não existindo *p* conjuntos recíprocos completos, no pior caso em $((p_1 - r)\times p_2 n_1^2)$ iterações, onde $p_1=(N-p)$, n_1 é número de vértices da maior árvore e *r* é o número de alocações a cada iteração. Devido a isso, é possível usar o algoritmo em redes de grande porte.

2 – Introdução

Uma rede é um grafo conectado e não orientado $G(V,E)$, com um número não negativo $w(v_1)$, onde w é o peso associado a cada um dos seus $|V| = N$ vértices, e um número $l(e)$, onde l é o comprimento associado a cada um dos seus $|E|$ arcos. Hakimi [6], [7] provou que existe um subconjunto de *p* vértices $V^p \subset V$, que fornece uma solução ótima para o problema das *p – medianas.* Em vista disso, quando nos referirmos na busca de uma solução ótima, procuramos determinar um subconjunto V^p de vértices da rede. Neste trabalho, todos os pesos associados aos vértices serão iguais a um, ou seja, $w(v_1)$ =1. Tal fato dificulta a determinação da solução ótima, já que cada vértice da rede tem a mesma importância para ser facilidade, o que não ocorre nos problemas práticos.

3 – O Modelo matemático para o problema das *p-medianas.*

O problema das *p – medianas*, segundo o modelo de Programação Linear e Inteira Binário de Revelle e Swain [15] é apresentado na forma:

$$
\begin{aligned}
& Min \quad z = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_{ij} \\
& s.a. \\
& \sum_{i=1}^{n} x_{ij} = 1 \qquad j = 1,2,\ldots\ldots\ldots,n \\
& x_{ii} \geq x_{ij} \qquad i,j = 1,2,\ldots\ldots\ldots,n \quad i \neq j \\
& \sum_{i=1}^{n} x_{ii} = p \\
& x_{ij} \in \{0,1\} \qquad i,j = 1,2,\ldots\ldots\ldots,n
\end{aligned}
\tag{1}
$$

Se uma mediana é estabelecida no vértice v_i, então $=1$

custo de alocação do vértice v_j à mediana v_i.

O modelo acima pode ser resolvido via o Algoritmo Simplex, da Programação Linear, aplicando a relaxação nas variáveis x_{ij}, fazendo $x_{ij} \geq 0$. Se todos os valores das variáveis forem inteiros, então uma solução ótima do problema foi encontrada. Do contrário, uma técnica de ramificação e limites pode ser usada, para a obtenção da solução inteira. O grande problema do modelo acima é que requer n^2 variáveis e $n^2 + 1$ restrições.

4 – Outros Métodos para a determinação das *p – medianas*.

Há vários métodos voltados para a solução dos problemas das *p – medianas,* tais como:

4.1 - Algoritmos de Ramificação e Limites.

O primeiro algoritmo de ramificação e limites para o problema das *p - medianas* foi desenvolvido por Jarvinen, Rajala e Sinervo [10]. Outros algoritmos foram desenvolvidos por El-Shaieb [2], Galvão [4], Khumawala [11], Neebe [14], Veloso [18], Celesi [21]. O principio básico de um algoritmo de ramificações e limites é, através do limite inferior, eliminar implicitamente as soluções viáveis cujas soluções serão piores que a disponível.

4.2 – Métodos Duais baseados na Relaxação Lagrangeana.

A Relaxação Lagrangeana é uma técnica que provou ser eficiente na solução de vários problemas de natureza combinatória. Tal técnica é aplicada e discutida por Geoffrion [5], Held, Wolfe and Crowder [9], Fisher, Nothrup and Shapiro [3], Senne [22]. Aplicações voltadas especificamente para o problema das *p - medianas* podem ser encontradas em Cornuejols et al [1]. Uma visão geral do problema de localização em redes pode obtida em [8].

4.3 – Métodos Heurísticos.

Apesar dos métodos heurísticos não terem a devida atenção dos matemáticos, os mesmos podem ser usados para estimar uma solução viável inicial, que poderá ser usada como limite superior ou inferior para problemas de grande dimensão. Então uma técnica exata poderá ser usada para a determinação da solução ótima do problema, como por exemplo, a técnica de ramificação e limite.

Há vários métodos heurísticos voltados para a solução do problema das *p - medianas*. Dentre eles podemos citar os de Maranzana [13], Kuehn and Hamburger [12], Teitz and Bart [16], Veloso [17,19].

5 – Definições Básicas.

Em Veloso [20] são apresentadas várias definições básicas. Algumas delas serão apresentadas neste trabalho, para um melhor entendimento do problema, já que são fundamentais para isso.

Seja $v_i \in V$. A cada $v_i \in V$ está associado um subconjunto $V_i \subset V,\ i = 1,2,.........,N$, composto por $(\lceil Q \rceil + 1)$ vértices, com o número médio de alocações por vértice facilidade dado por *((N – p) /P)* é menor ou igual a um. Todos eles ordenados em ordem crescente de seus custos de alocação, a partir do vértice v_i.

Definição 1 – Dois conjuntos $V_i \, e \, V_k$, com $|V_i| = |V_k| = (\lceil Q \rceil + 1)$ são definidos como ***não recíprocos*** se:

$$V_i \cap V_k \neq \varnothing \qquad 2$$

com $v_i \in V_k$, porém $v_k \notin V_i$. (ou $v_k \in V_i$, porém $v_i \notin V_k$.)

Definição 2 – Dois conjuntos $V_i \, e \, V_k$, com $|V_i| = |V_k| = (\lceil Q \rceil + 1)$ vértices são definidos como ***recíprocos* completos** se:

$$V_i \cap V_k \neq \varnothing \qquad 3$$

com $v_i \in V_k$, e $v_k \in V_i$

Definição 3 – Uma **árvore** composta por **um conjunto recíproco completo** e **conjuntos não recíprocos** é um grafo conexo que não contém ciclos.

Definição 4 – Uma solução viável V^p é composta por exatamente p vértices estabelecidos como facilidades.

Definição 5 – O custo de alocação dos vértices não medianas v_j, associados à cada conjunto V_i, $i = 1,2,.........,N$ é dado por:

$$S_i = \sum_{v_j \in V_i} d(v_i, v_j) \qquad 4$$

onde $d(v_i, v_j)$ representa a distância mínima do vértice v_i ao vértice v_j.

Definição 6 – O nível de um vértice associado a um conjunto não recíproco dentro da árvore é igual ao número de vértices associados a conjuntos não recíprocos para se atingir o conjunto recíproco.

Definição 7 – A profundidade de uma árvore é definida como sendo o vértice associado ao conjunto não recíproco de maior nível.

6 – Teoremas sobre Árvores e os Conjuntos Não Recíprocos.

Sejam os $v_j \in V_i, V_i \subset V$ ordenados em ordem crescente de seus custos de alocação a partir de v_i, com $|V_i| = |V_k| = (\lceil Q \rceil + 1)$. Sob essas condições, os seguintes teoremas são válidos.

Teorema 1 – Haverá pelo menos um conjunto V_i composto unicamente de conjuntos recíprocos completos.

Prova: Seja $S_m = \min\{S_i\}$, calculado de acordo com (4). Suponha $v_k \in V_m$ e $v_m \notin V_k$. Então existe $v_j \in V_k$ tal que $d(v_k, v_j) < d(v_k, v_m)$, fazendo com que $S_k < S_m$, o que é uma contradição, já que $S_m = \min\{S_i\}$.

Teorema 2 – Não pode existir uma árvore composta unicamente de conjuntos não recíprocos.

Prova: O teorema 1 garante a existência de pelo menos um conjunto recíproco completo. Logo, havendo apenas um conjunto recíproco completo, a árvore será gerada a partir dele, em função dos conjuntos não recíprocos. Caso haja k conjuntos recíprocos completos, então haverá k estruturas de árvores, como provado no teorema 3.

Teorema 3 – Os conjuntos não recíprocos geram uma estrutura de árvore dentro da rede, a partir de um conjunto recíproco completo.

Prova: Sejam V_i e V_k dois conjuntos recíprocos. Seja V_j um conjunto não recíproco com relação à V_k. Então v_j terá como menor custo de alocação v_k. Suponha que V_m é não recíproco com relação à V_j, então o vértice v_m será alocado a v_j. Como um vértice não pode ser alocado a mais de um outro vértice em função dos custos de alocação, os conjuntos não recíprocos geram uma estrutura de árvore disjunta, já que, pelo exposto acima, os conjuntos não recíprocos evitarão o surgimento de ciclos.

Teorema 4 – Havendo k conjuntos recíprocos em uma rede, então para as k árvores existentes, tais propriedades serão válidas:

$$T_i \cap T_k = \{\} \qquad i \neq k$$
$$\bigcup_{i=1}^{k} T_i = V \qquad 5$$

Prova: De acordo com o teorema 3, cada vértice da rede estará associado ou a um conjunto recíproco completo ou não recíproco e que cada vértice não poderá será alocado a mais de um vértice. Logo, as interseções das árvores serão vazias. Teremos então que a união dessas árvores nos fornecerá o conjunto de vértice da rede que é V.

Teorema 5 – Todos os custos de alocação associados aos conjuntos não recíprocos estarão ordenados em ordem crescente de seus custos de alocação a partir de um conjunto recíproco.

Prova: Seja uma árvore composta de um conjunto recíproco completo e não recíprocos. Foi provado que o conjunto recíproco completo representa o par de vértices com menor custo de alocação. Consideremos os conjuntos não recíprocos V_j, V_m, sendo V_i o conjunto recíproco completo juntamente com V_k. Se V_j é não recíproco com relação à V_k é porque existe um vértice associado a um conjunto recíproco ou não, cujo custo de alocação é menor com relação à V_k. Se V_m é não recíproco com relação à V_j é porque existe um vértice associado a um conjunto não recíproco, no caso V_k, que fornece um custo de alocação com relação à v_j menor que o de v_m. Diante do exposto, os vértices associados aos conjuntos não recíprocos estarão ordenados em ordem crescente de seus custos de alocação a partir do conjunto recíproco.

5 – Determinação do Custo de uma Solução Viável e dos Limites Inferiores.

Para a obtenção de uma solução viável, os custos associados aos conjuntos recíprocos completos e não recíprocos são ordenados em ordem crescentes de seus custos de alocação, evitando-se a duplicidade dos custos com relação aos conjuntos recíprocos completos.

5.1 – **Determinação do Limite Inferior**.

A determinação do Limite Inferior é feita de acordo com o teorema abaixo:

Teorema 6 – A soma dos $\overline{p} = (N - p)$ menores custos de alocação, associados aos conjuntos recíprocos completos ou não recíprocos, calculados de acordo com (4), será um limitante inferior para o custo de qualquer solução viável constituída por p vértices facilidades.

Prova: Como deverá ocorrer no máximo $\overline{p}$ alocações dos vértices não facilidades aos p vértices estabelecidos como facilidades, então a soma dos $\overline{p}$ menores S_i é o menor valor possível para qualquer solução viável, já que os mesmos estão ordenados em ordem crescente de seus custos de alocação.

Teorema 7 – Um limitante superior nos custos de alocação para qualquer árvore em particular será dado por:

$$S_{max} = Max(S_{lim}) \qquad 6$$

onde $S_{lim} = \max(S_i)$, para $i = 1,2,\ldots\ldots,\overline{p}$.

Prova. De acordo com o teorema 5, os $\overline{p} = (N - p)$ menores custos de alocação associados a conjuntos recíprocos completos e não recíprocos é um limitante inferior nos custos de alocação. Logo, (6) estabelece para cada árvore em particular um limitante nos custos de alocação, determinando com isto quais vértices por árvore podem ser candidatos a facilidades e quais vértices não facilidades serão a eles alocados.

5.2 – **Determinação da solução viável.**

5.2.1 – Ocorrência de p conjuntos recíprocos completos com interseção vazia.

Teorema 8 – Se houver p conjuntos constituídos unicamente de conjuntos recíprocos completos com interseção vazia, então a solução viável composta pelos p vértices associados a esses conjuntos constituirão uma solução ótima, já que.

$$\cup V_i = V \qquad 7$$

para $v_i \in V_p$, conjunto constituído por p vértices estabelecidos como facilidades.

Prova: Como todos os S_i, $i = 1,2,\ldots\ldots p$ associados aos conjuntos recíprocos completos estão ordenados em crescentes de seus custos de alocação, então existindo p interseção vazia (7) será satisfeita, fazendo que ocorram os $(N-p)$ menores custos de alocações possíveis.

5.2.2 – Não Ocorrência de p conjuntos recíprocos completos.

A não ocorrência de p conjuntos recíprocos completos fará com que seja necessária a determinação de mais de um vértice mediana por árvore. Como cada árvore representa os menores custos de alocação para os vértices que as compõem, então deve - se determinar qual o número ótimo de facilidade por árvore, para obtermos uma solução viável que minimize os custos de alocação. Isso implicará em perdas de custos que compõem o limite inferior, fazendo com que novos vértices associados a conjuntos não recíprocos sejam incluídos como componentes do limite inferior, elevando com isso seu valor.

Teorema 9 – Não ocorrendo p conjuntos recíprocos completos com interseção vazia, então o custo de uma solução viável será dada por:

$$C_p = \sum_{v_i \in V_p} S_i + \sum \Delta S_i^{j+} - \sum \Delta S_k^{j-} \qquad 8$$

onde $\sum \Delta S_i^{j+}$ representa o acréscimo nos custos de alocação devido aos conjuntos não recíprocos, $\sum \Delta S_k^{j-}$ representa as perdas devido a interseção não vazia entre os vértices alocados aos vértices candidatos à facilidade ou a perda de alocação entre conjuntos não recíprocos devido à partição da árvore, já que cada vértice não facilidade será alocada à facilidade mais próxima.

Prova: Dada uma rede com N vértices e p facilidades a serem determinadas, haverá sempre *(N-p)* pontos de demanda que devem ser alocados às p facilidades. Então ocorrendo perdas de alocação devido às interseções não vazias ou à partição da rede, $\sum \Delta S_i^{j+}$ representa os acréscimos nas alocações para compensar as perdas, representadas por $\sum \Delta S_k^{j-}$ para obtermos exatamente *(N-p)* alocações.

6 – Desenvolvimento do Algoritmo.

De acordo com o Teorema 4, as estruturas de árvores geradas pelos conjuntos recíprocos completos e não recíprocos são disjuntas e cada árvore representa para os vértices que as compõem os menores custos de alocação. Dispondo-se do limitante nos custos de alocação dado por (6), então, em cada árvore, cada vértice candidato à facilidade terá os seus custos de alocações relacionadas aos conjuntos recíprocos completos ou não recíprocos limitadas pelo mesmo. O algoritmo é composto de 2 fases: Na fase I, em cada árvore em particular, cada vértice candidato à facilidade tem o número de vértices a ele alocado com os custos limitados por (6). Havendo perda de alocação, determina-se o novo limite inferior. Na fase II, procura-se determinar o número de facilidades por árvore que minimize os custos de alocação, obedecendo (6). Nesta fase, como poderá ocorrer perda de alocação devido a interseção não vazia entre os vértices alocados aos candidatos à facilidade ou à partição da rede em conjuntos mutuamente disjuntos, um novo limite inferior deve ser determinado, bem como um novo limitante nos custos de alocação.

Fase I – Passos.

- Em cada árvore, para cada vértice candidato à facilidade, determine o número de vértices a ele alocado, os custos de alocação bem como o limite inferior a ele associado, em função das perdas de alocação.
- Selecione como candidato à facilidade o de maior numero de alocações e menor limite inferior.

Fase II – Passos.

- Em cada árvore, verifique a possibilidade de existência de mais um vértice facilidade por árvore, juntamente com o vértice selecionado na fase I. O critério de escolha será o que fornecer o menor limite inferior da fase I ou II.
- Verifique se existem *(N – p)* alocações. Isso ocorrendo, temos uma solução viável. Pare.

- Não existindo *(N – p)* alocações, determine as perdas de alocações por árvore, determine o novo limite inferior, bem como o novo limitante nos custos de alocações.

- Volte a fase I.

7 – Complexidade do Algoritmo.

Caso haja *p* conjuntos recíprocos completos, então a complexidade do algoritmo é linear já que cada árvore terá apenas dois vértices e a solução será ótima. Não ocorrendo *p* conjuntos recíprocos completos, seja então p_2 o número de conjuntos recíprocos completos. Haverá então p_2 árvores. Na fase I do algoritmo, a complexidade será $p_2 \times n_1^2$, onde n_1^2 representa a árvore com maior número de vértices. Havendo *r* alocações após as fases I e II, então um limitante superior no número de operações para se determinar a solução para o problema, com $p_1=(N-p)$ será:

$$((p_1 - r) \times p_2 n_1^2) \qquad 9$$

8 – Resultados Computacionais.

Os resultados computacionais trazem os valores dos limites inferiores, o número de conjuntos recíprocos completos, CRC, bem como o da solução viável, ou seja, *p* vértices estabelecidos como facilidades e *(N – p)* alocações, para diversos valores de *N e p*. A tabela 1 abaixo ilustra alguns resultados computacionais.

Tabela 1

N	p	CRC	LI_1	LI_2	Solução
10	4	3	25	26	26
	5	3	18	18	18
	6	3	12	12	12
20	9	6	166	166	178
	10	6	145	145	157
	11	6	124	124	136
	14	6	70	70	71
30	13	7	184	196	196
	14	7	170	179	179
	15	7	157	162	162
	16	7	144	146	146
	17	7	131	131	131
	23	7	60	60	60
50	25	14	6890	7232	7868
	30	14	4491	4491	5311
	35	14	2608	2608	2858
100	43	16	265	286	347
	50	16	213	217	249
	55	16	178	182	214
200	95	54	413	436	441
	130	54	188	193	195
	140	54	140	140	142

9 – Conclusão.

Os resultados computacionais mostram que essas estruturas de árvores obtidas através dos conjuntos recíprocos completos e não recíprocos são uma ferramenta eficiente na determinação das *p* medianas de uma rede, dado que a complexidade do algoritmo é polinomial. A diferença entre o valor do limite inferior 2 (LI_2) e da solução viável, deve-se ao fato que nem todas as alocações dos vértices não facilidades aos vértices facilidades estão associadas a vértices adjacentes ou associados a conjuntos não recíprocos, fazendo com o custo da solução viável seja maior, já que haverá vértices intermediários.

Bibliografia.

[1] Cornuejols, G., Fisher, M.L., Nemhauser, G.L., Location of bank accounts to optimize float: An analisys study of exact an approximate algorithms, Man. Sci, 23, 789, (1977).

[2] El-Shaieb, A.M. A new algorithm for location sources among destination, Man. Sic., 20, 221,(1973).

[3] Fisher, M.L., The lagrangian relaxation method for solving integer programming problems, Man, Sci., 27, (1981).

[4] Galvão, R.D., A dual bounded algorithm for the p-median problem, Opers. Res. 28, 1112, (1980).

[5] Geoffrion, A.M., Lagrangian relaxation for integer programming, Math. Prog. Study, 2, 82, (1974).

[6] Hakimi, S.L., Optimum distribution of Switching center in a communication network and some related graph theoretic problems, Opers. Res. 13, 462, (1965).
[7] Hakimi, S.L., Optimal location of switching centers and absolute centers and medians of a graph, Opers. Res. 12, 450, (1964).
[8] Handler, G.Y., Mirchandani, P.B., Location on Networks: Theory and Algorithms, MIT Press, Cambridge, 1979.
[9] Held, p., Wolfe, p., Crowder, H.P., Validation of subgradient optimization, Math. Prog. 6, 62, (1974).
[10] Jarvinen, P., Rajala, J., Sinervo, H., A branch-and-bound algorithm for seeking the p-median, Oper. Res., 20, 173, (1972).
[11] khumawala, B. M., An efficient branch and bound algorithm for warehouse location problem, Mang. Sci., 18, 12, (1972).
[12] Kuehn, A.A., Hamburger, M. J., A heuristic program for location warehouse, Man. Sci., 9, 643, (1963).
[13] Maranzana, F.E., On the location of supply points to minimize transport costs, Operational Res. Quart., 15, 261, (1964).
[14] Neebe, A.W., A branch-and-bound algorithm for the p-median transportation problem, J. Opl .Res. Soc., 29, 979, (1978).
[15] Revelle, C., Swain, R.W., Central facilities location, Geographical Analysis, 2, 30, (1970)
[16] Teitz, M.B., Bart, P., Heuristics methods for estimating the generalized vertex median of a weighted graph, Ope. Res., 16, 955, (1968).
[17] Veloso, E.A., Um algoritmo baseado na interseção de conjuntos para a determinação das p - medianas de uma rede, XXVI SBPO, (1994).
[18] Veloso, E.A., Um algoritmo eficiente para a determinação das p - medianas de uma rede, XXI SBPO E IV CLAIO, (1988).
[19] Veloso, E.A., (1995) – Um novo algoritmo para a determinação das p - medianas de uma rede, II EIMA, (1995).
[20] Veloso, E.A., – Uso de Conjuntos Recíprocos e Não Recíprocos para a Determinação das p – Medianas de uma Rede, XXX CNMAC (2007)
[21] Celesi, A, Righini, G., – A branch-and-price algorithm for the capacitaded *p*-median problem. Working Paper, 2002, Universiity of Milano (2002)
[22] Senne ELF, Lorena LAN – Lagrangean/surrogate heuristics for *p*-median ptoblems. In: Laguna M, Gonzalez-Velarde JL, editors. Computing tools for modeling, optimization and simulation: interfaces in computer science and operations research. Dordrecht: Kluwer Academic Publischers: p. 115-30 (2000).