

INTEGRAÇÃO DE MODELOS DE LOCALIZAÇÃO A SISTEMAS DE INFORMAÇÕES GEOGRÁFICAS

Luiz Antonio Nogueira Lorena

Laboratório Associado de Computação e Matemática Aplicada – LAC
Instituto Nacional de Pesquisas Espaciais – INPE

Edson Luiz França Senne

Departamento de Matemática – Faculdade de Engenharia de Guaratinguetá – FEG
Universidade Estadual Paulista – UNESP

João Argemiro de Carvalho Paiva

Divisão de Processamento de Imagens – DPI
Instituto Nacional de Pesquisas Espaciais – INPE

Marcos Antonio Pereira

Laboratório Associado de Computação e Matemática Aplicada – LAC
Instituto Nacional de Pesquisas Espaciais – INPE

Resumo

O problema das p -medianas pode ser descrito como o de localizar p facilidades (ou centros) em uma rede minimizando a soma de todas as distâncias de cada ponto de demanda a sua mediana mais próxima. Em alguns casos podem haver restrições na capacidade de atendimento das facilidades (problema de p -medianas capacitado). Modelos de localização de facilidades têm sido propostos como ferramentas de auxílio à decisão, principalmente quando é possível usar Sistemas de Informações Geográficas (SIGs) na coleta e análise dos dados dos problemas. Apresentamos neste trabalho um relato da integração de modelos de p -medianas aos SIGs *ArcView*, da ESRI, e *SPRING*, um sistema desenvolvido no INPE. O código que foi integrado a estes SIGs implementa uma abordagem recente da heurística Lagrangeana/surrogate, onde a viabilização da solução dual é feita através de uma heurística de localização-alocação alternada. O trabalho apresenta alguns testes computacionais usando dados do município de São José dos Campos, com tamanhos variando até o máximo de 3280 vértices e 1141 medianas para o caso não capacitado.

Palavras-chave: Localização, Sistemas de Informações Geográficas, Heurísticas Lagrangeanas.

1. Introdução

Problemas de localização tratam de decisões sobre onde localizar facilidades, considerando clientes que devem ser servidos de forma a otimizar algum critério (DREZNER (1995), DASKIN (1995)). O termo “facilidades” é utilizado para designar fábricas, depósitos, escolas etc., enquanto “clientes” refere-se a depósitos, unidades de vendas, estudantes etc. Em geral, as facilidades podem tanto ser selecionadas como novos centros a serem abertos como também ser escolhidas no subconjunto de vértices existentes. Por isso, tais problemas também são conhecidos como problemas de localização-alocação, devido ao processo de alocação dos vértices de demanda aos centros abertos.

Em certos casos podem existir restrições sobre a capacidade de atendimento de tais centros. Neste tipo de problema, considera-se que cada cliente possui associada uma demanda a ser satisfeita pela facilidade escolhida para atendê-lo. A soma das demandas de todos os clientes atendidos por uma facilidade não deve superar a capacidade de atendimento da mesma. Quando esse tipo de condicionante estiver presente, dizemos tratar-se de um problema de localização capacitado.

As aplicações de problemas de localização de facilidades ocorrem nos setores público e privado. No caso de setores públicos, procura-se maximizar a satisfação dos clientes em detrimento dos custos necessários para o alcance de tal objetivo. Exemplos de aplicações em setores públicos são a localização de escolas, postos de saúde, corpo de bombeiros, ambulâncias, viaturas de polícia, pontos de ônibus, entre outros. No caso do setor privado, onde custos fixos estão envolvidos, as aplicações envolvem, em geral, fábricas, depósitos, torres de transmissão, lojas de franquias etc.

O problema das p -medianas é um problema clássico de localização de facilidades e consiste em localizar p facilidades (medianas) em uma rede, de modo a minimizar a soma das distâncias de cada vértice de demanda à sua mediana mais próxima. As primeiras formulações dos problemas foram apresentadas em HAKIMI (1964) e HAKIMI (1965). O problema é bem conhecido como sendo *NP-hard* (GAREY & JOHNSON (1979)). Vários métodos heurísticos e métodos que exploram uma busca em árvore têm sido desenvolvidos para o problema das p -medianas (TEITZ & BART (1968), JARVINEN & RAJALA (1972), NEEBE (1978),

CHRISTOFIDES & BEASLEY (1982)). O uso combinado de técnicas heurísticas de relaxação lagrangeana e otimização por subgradientes, de um ponto de vista primal-dual, tem se mostrado eficiente na solução do problema (GALVÃO & RAGGI (1989), BEASLEY (1993) e LORENA & SENNE (1999)).

Modelos de localização de facilidades têm sido propostos, há algum tempo, como ferramentas de auxílio à decisão, principalmente quando uma base de dados geograficamente referenciada estiver disponível. Nestes casos, os Sistemas de Informações Geográficas (SIGs) são muito importantes na coleta e análise desses dados (BURROUGH (1986)). Sistemas de Informações Geográficas (FISCHBECK (1994)) integram uma sofisticada interface gráfica a uma base de dados geo-referenciados, constituindo-se em poderosas ferramentas de análise e planejamento espacial. Problemas complexos de localização de facilidades podem ser tratados com SIGs, levando-se em conta várias informações espaciais e, também, sócio-econômicas.

O uso combinado de SIGs e técnicas de Pesquisa Operacional para resolver problemas de localização ainda não está totalmente difundido na comunidade científica internacional. Mas, levando-se em conta sua capacidade de armazenar, exibir e manipular dados espacialmente distribuídos, a integração de algoritmos de localização aos SIGs foi iniciada há alguns anos.

Este trabalho relata a integração de modelos de localização de p -medianas, sem e com restrições de capacidade, aos SIGs *ArcView* da ESRI (*Environmental Systems Research Institute, Inc.*) e *SPRING*, um sistema desenvolvido pela Divisão de Processamento de Imagens (DPI) do INPE. O código integrado aos SIGs implementa uma abordagem recente da heurística lagrangeana/*surrogate*, onde a viabilização da solução dual é feita através de uma heurística de localização-alocação alternada. A integração com o *ArcView* foi feita através de *scripts* escritas em *Avenue* e a integração ao *SPRING* foi realizada criando-se um método que atua na representação vetorial dos modelos de rede, temático e cadastral. O trabalho apresenta alguns testes computacionais usando dados do município de São José dos Campos, com tamanhos variando até o máximo de 3282 vértices e 1141 medianas, para o problema de localização não capacitado.

2. O Problema das p -Medianas

Os problemas das p -medianas considerados neste trabalho podem ser modelados como problemas de programação inteira 0-1. Sem perda de generalidade, vamos considerar que as

medianas serão escolhidas do conjunto de vértices considerado no problema. Assim, para o caso não capacitado, o modelo matemático será:

$$v(P) = \mathbf{Min} \sum_{i=1}^n \sum_{j=1}^n d_{ij} x_{ij} \quad (1)$$

$$(P) \quad \text{sujeito a:} \quad \sum_{j=1}^n x_{ij} = 1, \quad \forall i \in N \quad (2)$$

$$\sum_{j=1}^n x_{jj} = p \quad (3)$$

$$x_{ij} \leq x_{jj}, \quad \forall i, j \in N, i \neq j \quad (4)$$

$$x_{ij} \in \{0, 1\}, \quad \forall i, j \in N \quad (5)$$

onde:

- $N = \{1, \dots, n\}$, sendo n o número de vértices na rede;
- p é o número de facilidades (medianas) a serem localizadas;
- $[d_{ij}]_{n \times n}$ é uma matriz de custo (distância), com $d_{jj} = 0, \forall j \in N$;
- $[x_{ij}]_{n \times n}$ é a matriz de alocação, com:

$$x_{ij} = \begin{cases} 1, & \text{se o vértice } i \text{ é atendido pelo centro } j, i \neq j; \\ 0, & \text{caso contrário.} \end{cases}$$

e

$$x_{jj} = \begin{cases} 1, & \text{se o vértice } j \text{ é um centro;} \\ 0, & \text{caso contrário.} \end{cases}$$

As restrições (2) e (4) garantem que a cada vértice i é alocado a somente um vértice j , que deve ser uma mediana. A restrição (3) determina o número exato de medianas a ser localizado (p), e (5) corresponde às condições de integralidade. Para o caso capacitado (*PMC*), substitui-se as restrições (4) por:

$$\sum_{i=1}^n a_i x_{ij} \leq b_j x_{jj}, \quad \forall j \in N \quad (4')$$

onde:

- a_i é a demanda do vértice i ;
- b_j é a capacidade de atendimento do vértice j , se este for escolhido como centro (mediana).

A técnica heurística usada para resolver (P) de forma aproximada é conhecida como relaxação lagrangeana/*surrogate* e já foi aplicada com sucesso a outros problemas de otimização combinatória (LORENA & LOPES (1994), LORENA & NARCISO (1996), LORENA & SENNE (1996), SENNE & LORENA (1997), LORENA & SENNE (1999) e NARCISO & LORENA (1999)). Uma discussão sobre as relaxações lagrangeana e *surrogate* pode ser encontrada em PARKER & RARDIN (1988). A relaxação lagrangeana/*surrogate* utilizada neste trabalho para o problema de p -medianas não capacitado está descrita em SENNE & LORENA (2000).

3. A Heurística Lagrangeana/*Surrogate*

A relaxação lagrangeana/*surrogate* desenvolvida para resolver de forma aproximada o problema de p -medianas capacitado (PMC) apresenta melhores resultados que a relaxação lagrangeana usual, obtendo limitantes de igual qualidade com menor esforço computacional (NARCISO & LORENA (1999) e LORENA & SENNE (1999)).

Para um dado $\lambda \in \mathfrak{R}_+^n$, a relaxação lagrangeana/*surrogate* de (PMC) é dada por:

$$v(L_t PMC^\lambda) = \mathbf{Min} \sum_{i=1}^n \sum_{j=1}^n (d_{ij} - t\lambda_i)x_{ij} + t \sum_{i=1}^n \lambda_i$$

sujeito às restrições (3), (4') e (5).

Para um dado $t \geq 0$ e $\lambda \in \mathfrak{R}_+^n$, $v(L_t PMC^\lambda) \leq v(PMC^\lambda) \leq v(PMC)$, onde PMC^λ é a relaxação *surrogate* de PMC . O problema ($L_t PMC^\lambda$) pode ser resolvido considerando-se a restrição (3) implicitamente e decompondo-se o problema para o índice j , obtendo-se os seguintes n sub-problemas da mochila:

$$\mathbf{Min} \sum_{i=1}^n (d_{ij} - t\lambda_i)x_{ij}$$

sujeito às restrições (4') e (5). Cada sub-problema é facilmente resolvido fazendo-se:

$$\beta_j = \sum_{i=1}^n [\min\{0, d_{ij} - t\lambda_i\}]$$

e escolhendo-se J como o conjunto de índices dos p menores β_j . Assim, uma solução x_{ij}^λ para o problema (L_tPMC^λ) é dada por:

$$x_{ij}^\lambda = \begin{cases} 1, & \text{se } j \in J \\ 0, & \text{caso contrário} \end{cases}$$

e, para todo $i \neq j$:

$$x_{ij}^\lambda = \begin{cases} 1, & \text{se } j \in J \text{ e } d_{ij} - t\lambda_i < 0 \\ 0, & \text{caso contrário} \end{cases}$$

Neste caso, o valor da solução lagrangeana/*surrogate* é dada por:

$$v(L_tPMC^\lambda) = \sum_{j=1}^n \beta_j x_{ij}^\lambda + t \sum_{i=1}^n \lambda_i$$

Fazendo-se $t = 1$ na relaxação (L_tPMC^λ) tem-se a relaxação lagrangeana usual, usando o multiplicador λ . Para um multiplicador λ fixo, o melhor valor para t pode ser encontrado resolvendo-se o problema dual lagrangeano: $v(D_t^\lambda) = \max_{t \geq 0} v(L_tPMC^\lambda)$.

O melhor valor da relaxação lagrangeana/*surrogate* fornece um limitante melhor do que o obtido pela relaxação lagrangeana usual. Uma solução exata para (D_t^λ) pode ser obtida por uma busca sobre diferentes valores de t (ver MINOUX (1975) e HANDLER & ZANG (1980)). Entretanto, em geral, existe um intervalo $t_0 \leq t \leq t_1$ (com $t_0 = 1$ ou $t_1 = 1$) que também produz limitantes melhores, como pode-se ver pela Figura 1 a seguir:

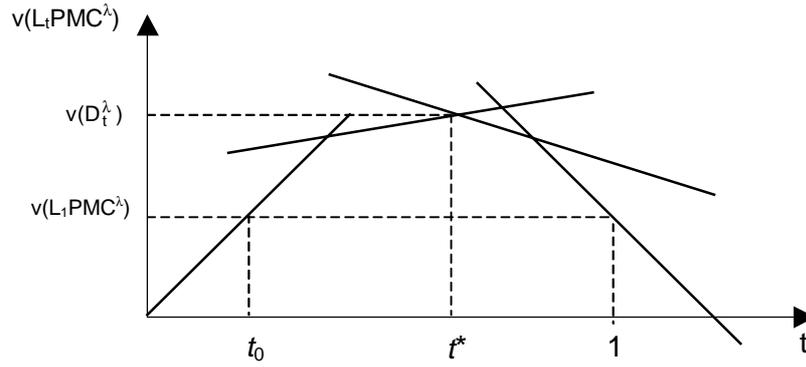


Figura 1: Limitantes lagrangeano/surrogate.

Assim, para obter um bom limitante, não é necessário encontrar o melhor valor de t (t^*), sendo suficiente encontrar um valor T tal que $t_0 \leq T \leq t_1$, através de um procedimento de busca heurística (SENNE & LORENA (2000)). Entretanto, se o valor de T se mantiver constante por um número de iterações fixado *a priori*, então esse valor será mantido para todas as relaxações seguintes e o procedimento de busca não será mais executado.

O algoritmo de subgradientes a seguir é usado para resolver o problema das p -medianas. Neste algoritmo, $C = \{j \in N \mid x_{jj} = 1\}$ é o conjunto dos vértices já fixados como medianas.

Dados $\lambda \geq 0$, $\lambda \neq 0$;

Fazer $lb = -\infty$, $ub = +\infty$, $C = \emptyset$;

Repetir

Resolver a relaxação $(L_t PMC^\lambda)$ obtendo x^λ e $v(L_t PMC^\lambda)$;

Obter uma solução viável x^f e seu valor v^f ;

Atualizar $lb = \max \{lb, v(L_t PMC^\lambda)\}$ e $ub = \min \{ub, v^f\}$;

Fixar $x_{jj} = 1$ se $v(L_t PMC^\lambda \mid x_{jj} = 0) \geq ub$, $j \in N - C$;

Atualizar o conjunto C ;

Fazer $g_i^\lambda = 1 - \sum_{j=1}^n x_{ij}^\lambda$, $i \in N$;

Atualizar o tamanho do passo θ ;

Fazer $\lambda_i = \max \{0, \lambda_i + \theta g_i^\lambda\}$, $i \in N$;

Até que (algum teste de parada seja satisfeito).

O valor inicial de λ usado é dado por $\lambda_i = \min_{j \in N} \{d_{ij}\}$, $i \in N$. O tamanho do passo é calculado como:

$$\theta = \frac{\pi(ub - lb)}{\|g^\lambda\|^2}.$$

O controle do parâmetro π é o proposto por HELD & KARP (1971). Inicialmente seu valor é fixado em 2, sendo reduzido à metade em cada iteração sempre que lb mantiver seu valor constante por 30 iterações sucessivas. Os testes de parada usados são os seguintes:

- a) $\pi \leq 0.005$;
- b) $ub - lb < 1$;
- c) $\|g^\lambda\|^2 = 0$
- d) todas as medianas foram fixadas.

3.1. Algoritmo de Busca Local

A solução x^λ não é necessariamente viável, mas o conjunto J identifica os vértices escolhidos como centros que podem ser usados para produzir soluções viáveis para ambos os problemas.

Para o caso não capacitado, os vértices não-medianas são realocados às suas medianas mais próximas, produzindo uma solução viável x^f para (P) dada por:

$$(x_{ij}^\lambda)^f = \begin{cases} 1, & \text{se } j \in J \\ 0, & \text{caso contrário} \end{cases}$$

e para todo $j \neq k$:

$$(x_{ij}^\lambda)^f = \begin{cases} 1, & \text{se } j \in J \text{ e } k = \text{índice de } \min_{j \in J} \{d_{ij}\} \\ 0, & \text{caso contrário} \end{cases}$$

com $v^f = \sum_{i=1}^n [\min_{j \in J} \{d_{ij}\}]$.

No caso capacitado, o processo de obtenção de uma solução factível x^f para o (PMC) utiliza o algoritmo *MTGH* de MARTELLO & TOTH (1990), para resolver de forma aproximada o seguinte problema generalizado de alocação (PGA):

$$\begin{aligned} \mathbf{Max} \quad z &= \sum_{i \in N-J} \sum_{j \in J} p_{ij} x_{ij}^f \\ \text{(PGA) sujeito a:} \quad & \sum_{i \in N-J} a_i x_{ij}^f \leq b_j, \quad \forall j \in J. \\ & \sum_{j \in J} x_{ij}^f = 1, \quad \forall i \in N-J. \\ & x_{ij}^f \in \{0, 1\}, \quad \forall i \in N-J, \forall j \in J. \end{aligned}$$

onde:

- $p_{ij} = -d_{ij}$ é o ganho se o vértice i for atribuído ao centro j , $i \in N-J$ e $j \in J$;
- a_i é a demanda associada ao vértice não-mediana i , $i \in N-J$;
- b_j é a capacidade da mediana j , $j \in J$;
- $x_{ij}^f = \begin{cases} 1, & \text{se o vértice } i \text{ é alocado à mediana } j; \\ 0, & \text{caso contrário.} \end{cases}$

A heurística de localização-alocação, inspirada no trabalhos de COOPER (1963) e TAILLARD (1996), baseia-se na observação que, após a definição de x^f , obtém-se exatamente p aglomerados (*clusters*) $C_m = \{j_m, I_m\}$, onde $j_m \in J$ é o índice do centro relativo ao *cluster* C_m , $m \in \{1, 2, \dots, p\}$, e $I_m = \{i \in N-J \mid x_{ij_m}^f = 1\}$ é o conjunto dos vértices não medianas alocados ao centro j_m . A solução x^f pode ser melhorada procurando-se por uma nova mediana dentro de cada aglomerado, trocando-se a mediana atual por um vértice não mediana e recalculando-se as alocações. Este processo se repete até que não seja mais possível obter melhorias no custo total da alocação. O algoritmo de localização-alocação utilizado neste trabalho, para o caso capacitado, está dado a seguir.

Heurística de Localização-Alocação (HLA)

Sejam:

- I_m o conjunto de índices dos vértices não medianas alocados ao centro m ;

- $Z_m = \sum_{k \in I_m} d_{kj_m}$ o custo do *cluster* $C_m = \{j_m, I_m\}$;
- $D_m = \sum_{k \in I_m \cup \{j_m\}} a_k$ a demanda total do *cluster* C_m ;

Repetir

(A)

Para cada *cluster* C_m , $m = 1, \dots, p$, fazer:

ntrocas = 0;

Se (ntrocas < MAX_TROCAS) então:

Se existe um vértice não mediana $i \in I_m$ tal que:

$$b_i \geq D_m \text{ e}$$

$$Z_{m(i)} < Z_m, \text{ onde } I_{m(i)} = I_m \cup \{j_m\} - \{i\}$$

Então

Intercambiar i com j_m , atualizando o *cluster* C_m ;

ntrocas = ntrocas + 1;

(B)

Seja $J = \{j_1, \dots, j_p\}$ o conjunto atual de índices dos centros.

Resolver o (PGA) considerando o conjunto J , obtendo um novo conjunto de *clusters* C_1, \dots, C_p .

Enquanto (for possível melhorar a solução viável).

Para o caso capacitado, o algoritmo de localização-alocação acima considera ainda, nos pontos assinalados por (A) e (B), a possibilidade de alterações nas alocações dos vértices não medianas às medianas de cada *cluster*, verificando a possibilidade de um vértice n_r pertencente a um *cluster* C_r ser removido deste *cluster* e alocado a outro *cluster* C_s ($r \neq s$), ou então verificando a possibilidade de dois vértices n_u e n_v (pertencentes aos *clusters* C_u e C_v , respectivamente) serem intercambiados, levando-se em conta as demandas dos vértices, as capacidades das respectivas medianas e o custo da solução resultante. Nos testes computacionais foi utilizado MAX_TROCAS = 3.

4. Integração a Sistemas de Informações Geográficas

O algoritmo de p -medianas descrito na seção anterior foi integrado a dois Sistemas de Informações Geográficas: *ArcView* (ESRI (1996)) e *SPRING* (SPRING (1998)). Os estudos foram realizados considerando uma base de dados contendo informações de algumas regiões da cidade de São José dos Campos, SP.

As subseções a seguir contém alguns detalhes das etapas das duas integrações desenvolvidas neste trabalho.

4.1. Integração ao *ArcView*

O algoritmo para a solução dos problemas das p -medianas capacitado e não capacitado foi implementado através de programas escritos em linguagem C e compilados com MS Visual C++. Os dados necessários aos programas foram obtidos a partir da base de dados existente nos mapas temáticos sobre os quais foram feitos os estudos. Através de *scripts* escritas na linguagem *Avenue*, disponível no *ArcView*, esses dados foram organizados em arquivos texto para serem passados como entrada aos respectivos programas.

Em ambos os casos, a distância entre os vértices foi calculada a partir da escala do mapa no qual estão inseridos os pontos considerados. Os valores resultantes representam a distância direta linear entre os pontos ou a distância sobre os arcos (ruas e avenidas) que compõem o mapa. Neste modelo de solução do problema das p -medianas, a distância entre os pontos foi o único parâmetro de custo considerado.

Para a visualização da solução, utilizou-se a função *Spider*, disponível no *ArcView*, que foi modificada para se adequar às necessidades da integração. Esta função verifica as distâncias entre os vértices de demanda, contidos em um tema, e os pontos relativos aos centros ofertantes, contidos em outro tema, e faz a ligação dos vértices de demanda aos centros selecionados para atendimento.

4.1.1. Caso não capacitado

Foram desenvolvidos dois módulos para resolver o problema de p -medianas sem restrições de capacidade. O primeiro módulo utiliza como entrada os dados um arquivo texto gerado por uma *script*. Este arquivo contém no seu primeiro registro o número de vértices e o número de medianas a serem considerados. Os registros seguintes formam uma lista de coordenadas X-Y de todos os vértices do tema de pontos escolhido para estudo. Como

resultado, o programa gera um arquivo tipo texto contendo a matriz simétrica de distâncias entre cada par de vértices.

O segundo módulo contém a implementação da heurística lagrangeana/*surrogate*. O arquivo de distâncias gerado pelo módulo anterior contém os dados necessários ao programa que, após o processamento, fornece a solução para o problema no formato de um arquivo tipo texto, contendo uma lista de triplas formadas por: vértice de demanda, seu centro correspondente e a distância entre eles. O último registro deste arquivo indica o *status* da solução encontrada: "Ótima" (gap fechado por limites) ou "Não-Ótima".

As Figuras 2 e 3 a seguir mostram alguns resultados obtidos utilizando dados do centro da cidade de São José dos Campos. Os polígonos de fundo correspondem ao tema escolhido para representar as quadras do centro da cidade. Os pontos sobrepostos são os vértices de demanda considerados.

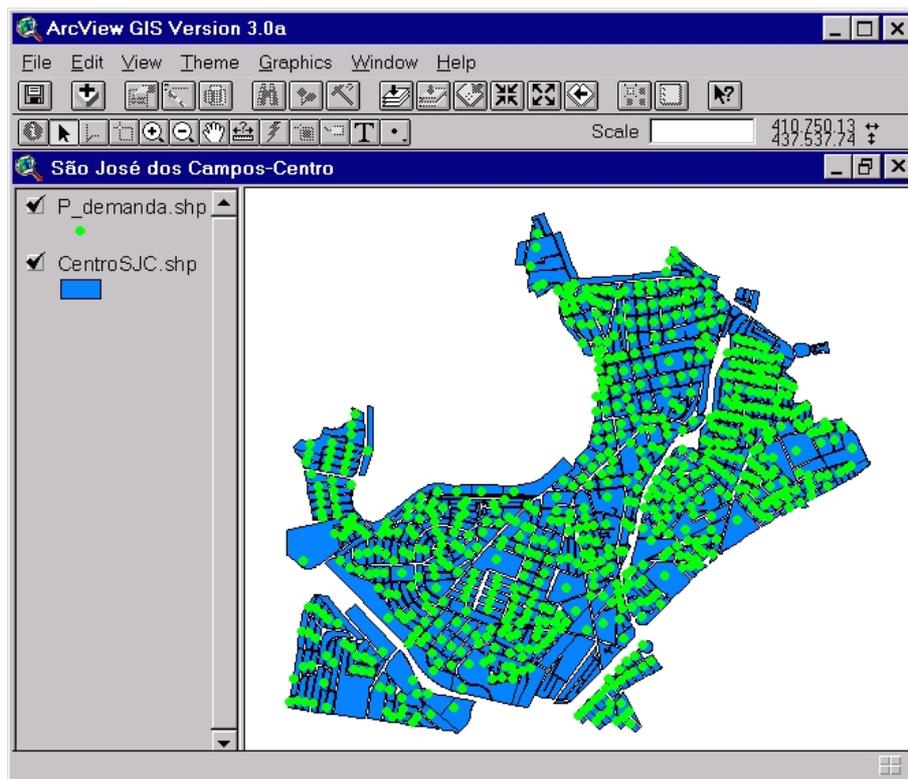


Figura 2: Pontos de demanda no *ArcView*.

A *script* desenvolvida para esta integração utiliza a informação do arquivo contendo a solução do problema e cria dois novos temas. Um tema de pontos representa as medianas encontradas pelo algoritmo, e um tema de linhas representa a alocação centro-vértice:

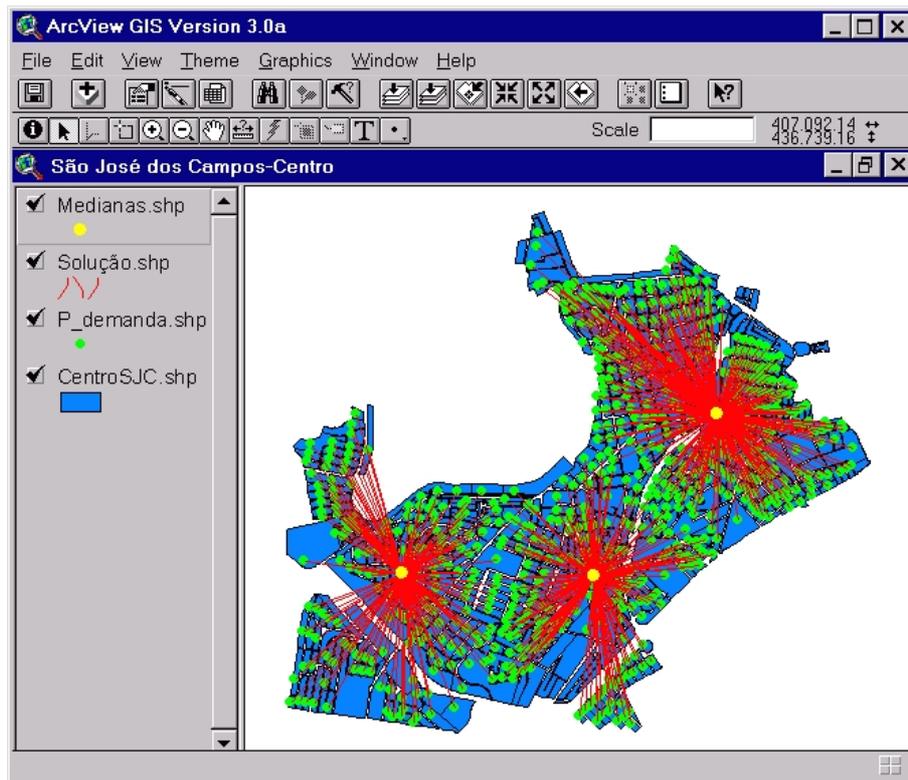


Figura 3: Visualização da solução de problema de p -medianas não capacitado.

4.1.2. Caso capacitado

Para resolução do problema de p -medianas capacitado, apenas um módulo foi desenvolvido, contendo todas as rotinas do procedimento heurístico. Parâmetros específicos na linha de comando permitem selecionar se trata-se de distâncias lineares ou calculadas sobre o tema de rede.

Os valores de demanda considerados neste trabalho foram extraídos dos temas disponíveis, baseado no número de imóveis existentes em cada quadra, onde quadras com número nulo de imóveis recebiam um valor de demanda igual a 1. A partir dessa informação, a *script* calcula a demanda total como sendo a soma da demanda de todos os vértices do tema de pontos selecionado. Este valor é então dividido pelo número de medianas a serem localizadas, definindo assim a capacidade de cada centro de atendimento. Para efeitos de estudo, este valor foi multiplicado por um fator $t > 1$, permitindo modelar cenários com escassez ou excesso na capacidade de atendimento das p facilidades escolhidas. Por questões de factibilidade, as capacidades individuais de atendimento de cada centro (supostas iguais) não devem ser inferiores ao maior valor de demanda individual observado no tema escolhido.

O programa desenvolvido para resolver o problema de p -medianas capacitado utiliza como entrada de dados um arquivo tipo texto gerado pela *script*, que contém no seu primeiro registro o número de vértices e o número de centros a serem considerados. Dependendo do tipo de distância considerada (linear ou na rede), os registros seguintes formam uma lista ordenada de informações que possibilitam ao programa encontrar a solução, apresentada como um arquivo texto, contendo a alocação centro-vértice e a respectiva distância. Um número ao final do arquivo indica o tipo da solução obtida: “Ótima” ou “Não-ótima”.

Nas Figuras 4 e 5 tem-se a visualização da solução de um problema contendo 31 vértices, dos quais foram selecionados 3 para a instalação de facilidades. No primeiro estudo considerou-se distâncias lineares e no segundo foram utilizadas distâncias calculadas sobre a rede que representa um subconjunto das ruas que compõem o centro da cidade de São José dos Campos. Como pode-se observar, existem diferenças entre as soluções dos dois estudos.

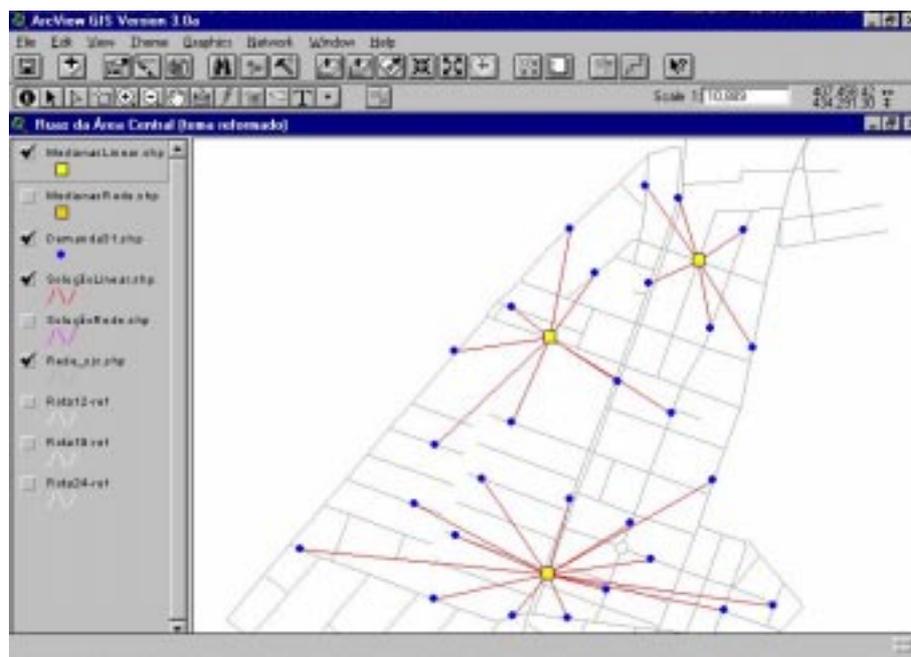


Figura 4: Solução do problema de p -medianas capacitado, distâncias lineares

Observa-se que a formação de agrupamentos permite considerar a possibilidade de roteamento dos vértices associados a cada centro. Neste exemplo, considerou-se que a mediana (em destaque) seria o ponto de origem e de destino da rota. Para o cálculo das rotas foi utilizado o módulo *Network Analyst*, que integra o pacote de módulos opcionais do *ArcView*.

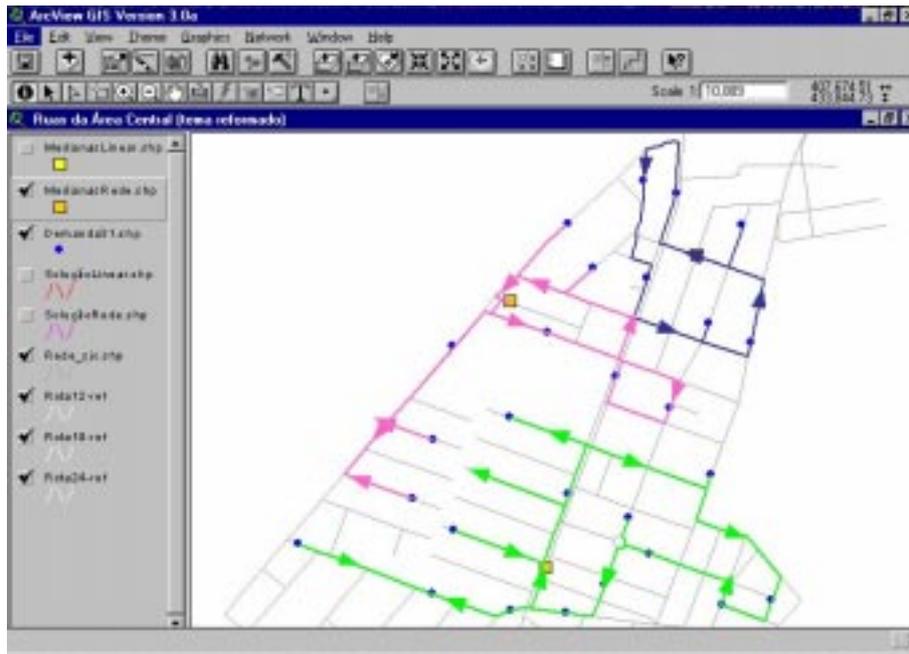


Figura 5: Solução do problema de p -medianas capacitado, com roteamento, considerando distâncias na rede.

4.2. Integração ao *SPRING*

O Sistema de Processamento de Informações Geo-referenciadas (*SPRING* (1998)) é um sistema computacional desenvolvido pela equipe da Divisão de Processamento de Imagens (DPI) do Instituto Nacional de Pesquisas Espaciais. Este sistema objetiva a integração e análise de diferentes tipos de dados espaciais. O modelo de dados do *SPRING* está baseado no paradigma de orientação a objetos (CÂMARA (1995)). Um banco de dados geográfico é composto de planos de informação, de objetos geográficos, e de informações não espaciais. Os planos de informação podem representar informações contínuas no espaço (campos), ou os objetos geográficos individuais. Cada plano de informação pode conter representações espaciais do tipo vetorial ou varredura. A representação vetorial corresponde a linhas, pontos, e polígonos que definem as formas de representação dos objetos espaciais, enquanto a representação de varredura corresponde a uma matriz de pontos com valores em cada célula. Os tipos de dados tratados no *SPRING* são:

- Mapas temáticos: cada informação representa um tema ou classe de informação. Por exemplo as classes de uso do solo de uma região.
- Mapas cadastrais ou mapa de objetos: ao contrário de um mapa temático, cada elemento é um **objeto geográfico**, que possui atributos e que pode estar associado a várias

representações gráficas. Por exemplo, os lotes de uma cidade são elementos do espaço geográfico que possuem atributos (dono, localização, valor venal, IPTU devido, etc.) e que podem ter representações gráficas diferentes (poligonais, lineares, ou pontuais) em mapas de escalas distintas.

- Mapas de redes: correspondem a mapas cadastrais, com a diferença de que geralmente os objetos são representados por elementos lineares ou pontuais. As representações pontuais devem estar localizadas em pontos de intersecção de linhas na rede.
- Modelo numérico de terreno: denota a representação de uma grandeza que varia continuamente no espaço. Comumente associados à altimetria, podem ser utilizados para modelar outros fenômenos de variação contínua (como variáveis geofísicas, geoquímicas e batimetria).
- Imagens: representam dados de sensoriamento remoto ou fotografias aéreas.

O algoritmo para localização das medianas pode ser aplicado no *SPRING* em dados dos modelos temático, cadastral e de redes, da forma descrita a seguir:

- Para uso em um dado temático é necessário que a representação vetorial contenha pontos. As localizações espaciais dos pontos e a distância linear entre os mesmos são utilizados no processo de localização das medianas.
- Para o dado cadastral o procedimento de localização das medianas atua sobre uma determinada categoria de objetos selecionada. Todos os objetos desta categoria que estejam associados a uma representação pontual são utilizados na análise de localização, que usa a distância linear entre os pontos.
- Para o modelo de redes o modo de utilização é similar ao do modelo cadastral, com a diferença de que a distância entre os pontos pode ser escolhida entre linear, ou ser computada a partir da própria rede.

A Figura 6 mostra a interface para execução da função de localização de medianas no *SPRING*. O cálculo das medianas usa a área da informação que está visível no monitor. A partir de um plano de informação ativo, o usuário define o número de medianas a serem calculadas. Se o plano ativo corresponder a um dado temático esta é a única informação necessária a ser fornecida, sendo considerada a distância linear entre os pontos. Para o caso de dados cadastrais ou de redes, a lista de categorias de objeto fica ativa para que seja selecionado um tipo de objeto. A princípio apenas objetos do mesmo tipo entram na análise

de localização, podendo esta restrição não existir em versões futuras. O cálculo da distância entre os pontos corresponde à distância linear para os modelos temático e cadastral, enquanto que para o modelo de redes também está disponível selecionar que a distância seja calculada baseada na própria rede. Para o caso capacitado, a interface apresenta a opção de se associar algum valor de demanda ou peso para os pontos em análise. Este valor pode ser obtido a partir de um atributo do objeto no banco de dados.



Figura 6: Interface de diálogo para localização de medianas no *SPRING*.

As Figuras 7 e 8 mostram os resultados da análise de localização em uma pequena região da cidade de São José dos Campos. Alguns objetos localizados em vértices da rede correspondem à possíveis localizações para instalação de algum tipo de atividade. Dado o número de medianas a se encontrar, o programa gera como resultado na tela, os pontos correspondentes às medianas (representados por círculos) e associa os outros pontos à mediana mais próxima. Pode-se observar que os resultados considerando as distâncias linear e da rede não necessariamente são iguais.

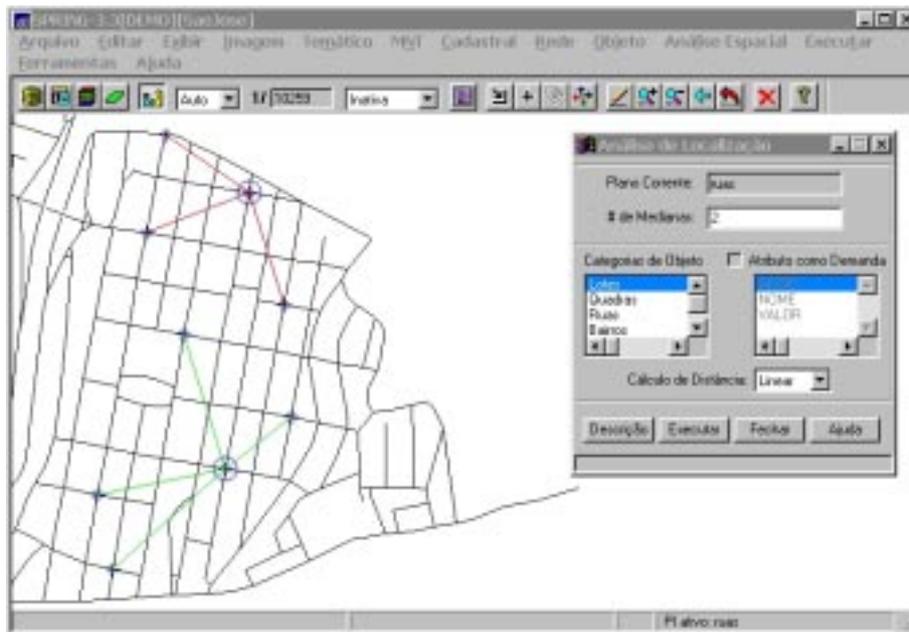


Figura 7: Cálculo de medianas no *SPRING* usando distância linear.

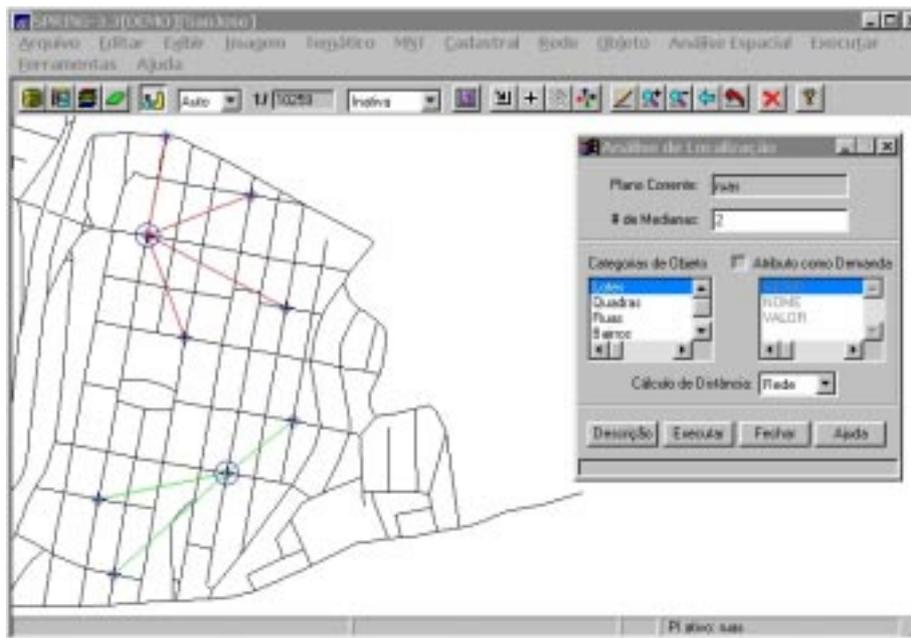


Figura 8: Cálculo de medianas no *SPRING* usando distâncias da própria rede.

5. Testes Computacionais e Resultados

Foram realizados alguns testes computacionais para testar a eficiência do algoritmo de p -medianas apresentado na seção 3. Neste trabalho foi utilizado um microcomputador Pentium MMX 233MHz com 128MB de RAM e os dados correspondem às quadras e ruas da região central da cidade de São José dos Campos.

Os resultados estão apresentados nas Tabelas 1 e 2 a seguir. Nestas tabelas, todos os tempos computacionais excluem o tempo necessário para estabelecer a matriz de distâncias.

As tabelas contém:

- n: número de vértices da rede;
- p: número de medianas;
- LInf: valor da melhor solução dual obtida (limite inferior da solução ótima);
- LSup: valor da melhor solução viável obtida (limite superior da solução ótima);
- Gap: *gap* percentual de dualidade, ou seja, $100\% \times (LSup - LInf)/LSup$;
- Tempo: tempo computacional (em segundos).

Tabela 1 - Resultados dos Testes Computacionais: Caso Não Capacitado

n	p	LInf	LSup	Gap	Tempo
324	5	122518,02	122518,02	0,000	4,72
	10	79250,84	79256,35	0,007	7,30
	20	54467,23	54533,11	0,121	7,33
	50	32094,13	32101,52	0,023	7,65
	108	18719,70	19683,61	4,897	7,84
818	5	604883,69	605855,81	0,160	102,66
	10	382420,75	385371,44	0,766	97,48
	20	251540,45	251717,77	0,070	60,39
	50	146303,64	149251,13	1,975	43,73
	100	97763,44	98992,31	1,241	57,93
	150	75465,67	77440,57	2,550	66,19
	272	47481,36	50086,61	5,201	85,58
3282	5	6381066,50	6381119,00	0,001	1699,88
	10	3911948,00	3914249,75	0,059	1548,43
	20	2342928,75	2350502,50	0,322	1520,00
	50	1288593,00	1308957,25	1,556	1106,45
	100	838007,63	841380,81	0,401	954,24
	500	322401,41	332954,84	3,170	1530,44
	1000	186532,23	194813,50	4,251	1606,07
	1141	164245,19	175905,27	6,629	1526,76

Os resultados da Tabela 1 mostram que os valores para os *gaps* de dualidade são pequenos, demonstrando a efetividade do algoritmo de *p*-medianas para dados reais distribuídos geograficamente.

Os resultados com dados reais para o caso capacitado compõem a Tabela 2. Como pode-se observar, a complexidade do problema reflete-se nos tempos computacionais, sensivelmente mais elevados que os problemas não capacitados de mesma dimensão. Os *gaps* de dualidade também se mantiveram baixos, confirmando a robustez da heurística também para este tipo de problema.

Tabela 2 - Resultados dos Testes Computacionais: Caso Capacitado

n	p	LInf	LSup	Gap	Tempo
100	10	17246,53	17288,99	0,246	187,67
200	15	33225,88	33426,04	0,599	4249,74
300	25	45314,71	45364,30	0,109	4956,45
300	30	40635,80	40635,91	0,000	3403,03
402	30	61843,23	62000,23	0,253	41988,99
402	40	52396,31	53225,30	1,558	9673,38

Os dados utilizados e os resultados obtidos estão disponíveis em <http://www.lac.inpe.br/~lorena/ArsigIndex.html>.

6. Conclusão

Modelos de localização de facilidades têm sido propostos, há algum tempo, como ferramentas de auxílio à decisão, principalmente quando uma base de dados geograficamente referenciada pode ser usada e Sistemas de Informações Geográficas podem ser utilizados para a análise desses dados.

Neste trabalho discutiu-se a integração de modelos de *p*-medianas aos Sistemas de Informações Geográficas *ArcView*, da ESRI, e *SPRING*, em desenvolvimento no INPE. O código integrado a estes SIGs implementa uma abordagem recente da heurística Lagrangeana/*surrogate* que tem se mostrado eficiente em diversas classes de problemas de Otimização Combinatória.

Testes computacionais realizados usando dados do município de São José dos Campos demonstram a efetividade do algoritmo proposto para utilização em Sistemas de Apoio à Decisão usando Sistemas de Informações Geográficas.

Agradecimentos

Os autores agradecem o apoio financeiro da FAPESP - Fundação para o Amparo à Pesquisa no Estado de São Paulo (proc. 96/04585-6). O primeiro, segundo e quarto autores agradecem também o apoio recebido do CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico (procs. 350034/91-5, 302408/88-6 e 380646/99-4, respectivamente).

Referências Bibliográficas

- BEASLEY, J.E.: "Lagrangean Heuristics for Location Problems". *European Journal of Operational Research*, **65**, 383-399, 1993.
- BURROUGH, P.A.: *Principles of Geographical Information Systems for Land Resources Assessment*. Clarendon Press, Oxford, 1986.
- CÂMARA, G.: *Modelos, Linguagens e Arquiteturas para Bancos de Dados Geográficos*. Tese de Doutorado (INPE), São José dos Campos, SP, 1995.
- CHRISTOFIDES, N.; BEASLEY, J.E.: "A tree search algorithm for the p-median problem". *European Journal of Operational Research*, **10**, 196-204, 1982.
- COOPER, L.: "Location-allocation problems". *Operations Research*, **11**, 331-343, 1963.
- DASKIN, M.: *Network and Discrete Location: Models, Algorithms, and Applications*. Wiley Interscience, NY, 1995.
- DREZNER, Z. (ed.): *Facility Location: A Survey of Applications and Methods*. Springer-Verlag, NY, 1995.
- DYER, M.E.: "Calculating surrogate constraints". *Mathematical Programming*, **19**, 255-278, 1980.
- ESRI: *Avenue Customization and Application Development for ArcView*. Environmental Systems Research Institute, Inc., Redlands, CA, 1996.
- FISCHBECK, P.: "GIS: More than a Map". *OR/MS Today*, 42-45, Aug. 1994.
- GALVÃO, R.D.; RAGGI, L.A.: "A method for solving to optimality uncapacitated location problems". *Annals of Operations Research*, **18**, 225-244, 1989.

- GAREY, M.R.; JOHNSON, D.S.: *Computers and intractability: a guide to the theory of NP-completeness*. W. H. Freeman and Co., San Francisco, 1979.
- GLOVER, F.: "Surrogate constraints". *Operations Research*, **16** (4), 741-749, 1968.
- HAKIMI, S.L.: "Optimum location of switching centers and the absolute centers and the medians of a graph". *Operations Research*, **12**, 450-459, 1964.
- HAKIMI, S.L.: "Optimum distribution of switching centers in a communication network and some related graph theoretic problems". *Operations Research*, **13**, 462-475, 1965.
- HANDLER, G.; ZANG, I.: "A dual algorithm for the constrained shortest path problem". *Networks*, **10**, 293-310, 1980.
- HELD, M.; KARP, R.M.: "The traveling-salesman problem and minimum spanning trees: part II". *Mathematical Programming*, **1**, 6-25, 1971.
- JARVINEN, P.J.; RAJALA, J.: "A branch and bound algorithm for seeking the p-median". *Operations Research*, **20**, 173-178, 1972.
- KARWAN, M.L.; RARDIN, R.L.: "Some relationships between Lagrangean and surrogate duality in integer programming". *Mathematical Programming*, **17**, 320-334, 1979.
- LORENA, L.A.N.; LOPES, F.B.: "A surrogate heuristic for set covering problems". *European Journal of Operational Research*, **79**, 138-150, 1994.
- LORENA, L.A.N.; NARCISO, M.G.: "Relaxation Heuristics for Generalized Assignment Problems". *European Journal of Operational Research*, **91**, 600-610, 1996.
- LORENA, L.A.N.; SENNE, E.L.F.: "A Lagrangean/Surrogate Heuristic for Uncapacitated Facility Location Problems". In: *Latin-Iberian-American Congress on Operations Research and System Engineering, 8 / Simpósio Brasileiro de Pesquisa Operacional, 28*. Rio de Janeiro, RJ, Ago. 1996. Anais, p. 854-859 (Vol. 2).
- LORENA, L.A.N.; SENNE, E.L.F.: "Improving traditional subgradient scheme for Lagrangean relaxation: an application to location problems". *International Journal of Mathematical Algorithms*, **1**, 133-151, 1999.
- MARTELLO, S.; TOTH, P.: *Knapsack problem: Algorithms and computer implementations*. John Wiley & Sons, New York, 1990.
- MINOUX, M.: "Plus courts chemins avec contraintes: Algorithmes et applications". *Annals of Telecommunications*, **30**, 383-394, 1975.
- NARCISO, M.G.; LORENA, L.A.N.: "Lagrangean/surrogate relaxation for generalized assignment problems". *European Journal of Operational Research*, **114** (1): 165-177, 1999.

NEEBE, A.W.: "A branch and bound algorithm for the p -median transportation problem".

Journal of the Operational Research Society, **29**, 989-995, 1978.

PARKER, R.G.; RARDIN, R.L.: *Discrete Optimization*, Academic Press, New York, 1988.

SENNE, E.L.F.; LORENA, L.A.N.: "Lagrangean/Surrogate Heuristics for Facility Location Problems". In: *EURO XV - INFORMS XXXIV Joint International Meeting*. Barcelona, Espanha, Jul. 1997. Abstracts, p. 128.

SENNE, E. L. F.; LORENA, L. A. N.: "Lagrangean/surrogate heuristics for p -median problems". In *Computing Tools for Modeling, Optimization and Simulation: Interfaces in Computer Science and Operations Research*, M. Laguna and J. L. Gonzalez-Velarde (eds.), Kluwer Academic Publishers, pp. 115-130, 2000.

SPRING: *Sistema de Processamento de Informações Georeferenciadas*. INPE, São José dos Campos, SP, <http://www.dpi.inpe.br/spring>, 1998.

TAILLARD, E.D.: "*Heuristic methods for large centroid clustering problems*". Technical report IDSIA96-96, IDSIA, 1996.

TEITZ, M.B.; BART, P.: "Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph". *Operations Research*, **16**, 955-961, 1968.

INTEGRATION OF LOCATION MODELS TO GEOGRAPHICAL INFORMATION SYSTEMS

Abstract

The p -median problem is the problem of locating p facilities (medians) on a network minimizing the sum of all distances from each demand point to its nearest facility. Restrictions on the capacity of each facility may be present (capacitated p -median problems). Facility location models have been proposed as decision making tools, mainly when geographic information systems (GIS) can be used to capture, to store and to analyze the data of the problems. In this work we present the integration of a p -median model to both ArcView, a GIS by ESRI (Environmental Systems Research Institute, Inc.), and SPRING, a GIS developed by INPE (National Institute for Space Research). The computer program that has been integrated to these geographic information systems implements a recent approach of Lagrangean/surrogate heuristic which uses a location-allocation heuristics in order to search for the primal feasibility of intermediate dual solutions. The paper presents some

computational tests which have been conducted with data from São José dos Campos city, representing problems with up to 3280 nodes and 1141 medians in the uncapacitated problem.

Key Words: Facility Location, Geographical Information Systems, Lagrangean Heuristics