



RESOLVENDO PROBLEMAS DE LOCALIZAÇÃO DE *HUBS* COM
ALOCÇÃO MÚLTIPLA EM MODELAGENS CONTÍNUA E DISCRETA TIPO
P-MEDIANAS ATRAVÉS DA ABORDAGEM DE SUAVIZAÇÃO
HIPERBÓLICA

Claudio Joaquim Martagão Gesteira Neto

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientador: Adilson Elias Xavier

Rio de Janeiro
Dezembro de 2017

RESOLVENDO PROBLEMAS DE LOCALIZAÇÃO DE *HUBS* COM
ALOCAÇÃO MÚLTIPLA EM MODELAGENS CONTÍNUA E DISCRETA TIPO
P-MEDIANAS ATRAVÉS DA ABORDAGEM DE SUAVIZAÇÃO
HIPERBÓLICA

Claudio Joaquim Martagão Gesteira Neto

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

Prof. Adilson Elias Xavier, D.Sc.

Prof. Marcia Helena Costa Fampa, D.Sc.

Prof. Luiz Pinguelli Rosa, D.Sc.

Prof. Amaro Olímpio Pereira Júnior, D.Sc.

Prof. Argimiro Resende Secchi, D.Sc.

Prof. Ricardo Saraiva de Camargo, D.Sc.

Prof. Mario Cesar Mello Massa de Campos, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
DEZEMBRO DE 2017

Gesteira Neto, Claudio Joaquim Martagão

Resolvendo Problemas de Localização de *Hubs* com Alocação Múltipla em Modelagens Contínua e Discreta Tipo p -Medianas através da Abordagem de Suavização Hiperbólica/Claudio Joaquim Martagão Gesteira Neto. – Rio de Janeiro: UFRJ/COPPE, 2017.

X, 84 p.: il.; 29, 7cm.

Orientador: Adilson Elias Xavier

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2017.

Referências Bibliográficas: p. 75 – 75.

1. Hub Location. 2. Multiple Allocation. 3. Min-Sum-Min Problems. 4. Hyperbolic Smoothing. 5. Localização de Hubs. 6. Alocações Múltiplas. 7. Problemas Min-Sum-Min. 8. Suavização Hiperbólica. I. Xavier, Adilson Elias. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

Dedico esta tese a meus falecidos pais, que muito se esforçaram para me propiciar a educação e os ensinamentos que me permitiram seguir na vida com uma pitada de sucesso, alguns amigos, alegrias e decepções na mesma proporção de quase todo mundo e muita paz de espírito.

Agradecimentos

Agradeço a meu orientador, aos demais membros da banca e à minha mulher e meus filhos, que têm sido muito compreensivos em relação aos meus dias de ausência física ou mental, devido aos múltiplos afazeres de trabalho e estudo.

Agradeço a grande ajuda do Prof. Ricardo Camargo, da UFMG, e do seu aluno Rodrigo Carvalho, D.Sc, pelo fornecimento de diversas instâncias de problemas de *hub-and-spoke* que foram fundamentais para o ajuste e validação de versões iniciais da metodologia proposta nesta pesquisa, e também por outras orientações recebidas. Agradeço também a Alexsander Andrade de Melo, estudante do PESC/COPPE, pelo fornecimento de figuras ilustrando diversas topologias de rede.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

RESOLVENDO PROBLEMAS DE LOCALIZAÇÃO DE *HUBS* COM
ALOCAÇÃO MÚLTIPLA EM MODELAGENS CONTÍNUA E DISCRETA TIPO
P-MEDIANAS ATRAVÉS DA ABORDAGEM DE SUAVIZAÇÃO
HIPERBÓLICA

Claudio Joaquim Martagão Gesteira Neto

Dezembro/2017

Orientador: Adilson Elias Xavier

Programa: Engenharia de Sistemas e Computação

As redes *hub-and-spoke* (HS) constituem um conceito importante para o projeto de sistemas de transporte e telecomunicações. Nelas, o tráfego se origina em cada um de diversos pontos distribuídos no espaço, e tem como destino todos os demais pontos. O tráfego flui através de diversos caminhos (*spokes*) a partir dos pontos de origem, se concentrando num conjunto menor de pontos (*hubs*), interconectados através de ligações de baixo custo unitário e grande capacidade, provendo economias de escala, e deles finalmente seguem para seus respectivos destinos. O problema em estudo é o da localização de um determinado número p de hubs, escolhidos no espaço plano contínuo ou, alternativamente, entre os pontos que se quer conectar, para servir como p -medianas. Procura-se encontrar os *hubs* que formam, junto com os pontos de origem e destino, a rede HS mais barata, ao mesmo tempo atribuindo tráfego a cada um desses *hubs*, considerando as demandas de tráfego entre cada par de pontos origem-destino e os respectivos custos de transporte. Na formulação adotada, cada ponto pode receber e enviar fluxos através de mais de um *hub*. A especificação do problema corresponde a uma formulação min-sum-min fortemente não-diferenciável. O método proposto supera essa dificuldade com a estratégia de suavização hiperbólica, que já se provou capaz de resolver com bastante eficiência instâncias grandes de problemas de agrupamento (*clusters*). A solução é obtida, em última análise, ao se resolver uma sequência de subproblemas diferenciáveis de otimização sem restrições, de baixa dimensão. A consistência do método é mostrada através de um conjunto de experimentos computacionais realizado em espaços contínuos e discretos com grandes problemas *hub-and-spoke*, de até mil pontos.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

SOLVING HUB LOCATION PROBLEMS WITH MULTIPLE ASSIGNMENT
THROUGH CONTINUOUS AND DISCRETE MODELING BY THE
HYPERBOLIC SMOOTHING APPROACH

Claudio Joaquim Martagão Gesteira Neto

December/2017

Advisor: Adilson Elias Xavier

Department: Systems Engineering and Computer Science

Hub-and-spoke (HS) networks are an important concept in the design of transportation and telecommunications systems. In those systems, items originate in each of several points scattered in space and, from each of those points, they can be destined to any of the other points. The traffic flows through several routes (spokes) from the points of origin and is concentrated in a smaller set of points (hubs), interconnected through links with a low unit cost and high capacity, capable of providing economies of scale, and finally flowing from these hubs to their respective points of destination. The problem under study is the location of a certain number p of hubs, chosen in a flat continuous space or, alternatively, among the points that have to be connected, to serve as p -medians. It consists in finding hubs that make up, along with the origin-destination points, the cheaper HS network and, at the same time, to assign traffic to each of these hubs, considering the traffic demands between each pair of points and the transportation costs. In this particular formulation, it is assumed that each point can receive and send flows through more than one hub. The problem specification corresponds to a strongly non-differentiable min–sum–min formulation. The proposed method overcomes this difficulty with a hyperbolic smoothing strategy, which has already been proven capable of solving quite efficiently large instances of clustering problems. The solution is obtained, ultimately, by solving a sequence of low-dimensional differentiable optimization subproblems without constraints. The consistency of the method is shown through a set of computational experiments in both continuous and discrete spaces, addressing large hub-and-spoke problems with up to one thousand points.

Sumário

Lista de Figuras	ix
Lista de Tabelas	x
1 Introdução	1
2 Aplicações do Conceito de Hubs e do UMapHMP no Mundo Real	8
3 Variedades de Problemas de Localização de Hubs	14
4 Especificação do Problema de Localização de Hubs a ser Tratado	22
5 Suavização do Problema de Localização de Hubs do Tipo p -Medianas Contínuo	27
6 Resultados Computacionais: Problema de Localização Hubs do Tipo p -Medianas Contínuo Simétrico	38
7 Suavização do Problema de Localização de Hubs do Tipo p -Medianas Discreto	50
8 Resultados Computacionais Problema Localização Hubs Tipo p -Medianas Assimétrico Discreto	57
9 Conclusões	69
Referências Bibliográficas	75
A Os Dados AP200	81

Lista de Figuras

2.1	Representação de hubs no Cérebro	11
2.2	Conjunto de 20 localidades extraídas do AP200, comparando os links diretos entre as localidades com aqueles feitos através de 3 hubs com alocação múltipla, usando a configuração ótima dos hubs no espaço discreto	12
3.1	Outras topologias de rede	16
3.2	Redes subaquáticas	17
4.1	Uma conexão entre a cidade j e a cidade l	24
4.2	O conjunto de conexões entre o ponto j e o ponto l	25
5.1	Parcelas do somatório em (5.4)	29
5.2	Parcelas originais e suavizadas do somatório em (5.4)	31
5.3	Somatório dos termos da equação (5.12)	33
5.4	Idealização da Eliminação de Mínimos Locais	37
6.1	German Towns, tempo médio de execução em função de p e α	43
6.2	Cidades Alemãs, $p = 4$, para diversos valores de α	44
6.3	Cidades Alemãs, $p = 4$ - Trajetória dos hubs: $0 \leq \alpha \leq 1$	45
6.4	Tempo médio de execução do algoritmo de localização de hubs no espaço contínuo em função de p , para quatro instâncias encontradas na literatura	49
9.1	Topologias de rede diversas	74
A.1	Números inteiros de 1 a 40 que geram os pontos do AP200	82
A.2	Distribuição espacial dos pontos da instância AP200, incluindo dez pares de pontos muito próximos, assinalados em cor escura e com quadrados maiores	84

Lista de Tabelas

6.1	Resultados para Cidades Alemãs ($\alpha = 0,5$)	41
6.2	Resultados para Cidades Alemãs, com α variável e $p = 4$	41
6.3	Cidades Alemãs: alocações múltiplas como uma função de α ($p=4$) .	42
6.4	Resultados para a instância das Cidades Alemãs (German Towns): solução contínua pelo algoritmo HSHS, solução discreta mais próxima e solução discreta ótima	45
6.5	Resultados da instância AP200 Simétrico ($\alpha = 0,5$)	46
6.6	Resultados da instância rd400 do TSPLIB ($\alpha = 0,5$)	47
6.7	Resultados da instância d657 do TSPLIB ($\alpha = 0,5$)	47
6.8	Resultados da instância dsj1000 do TSPLIB ($\alpha = 0,5$)	47
8.1	Resultados AP200 - Instâncias (m,p) de $(20,3)$ a $(60,6)$	61
8.2	Resultados AP200 - Instâncias (m,p) de $(60,7)$ a $(110,10)$	62
8.3	Resultados AP200 - Instâncias (m,p) de $(120,3)$ a $(160,10)$	63
8.4	Resultados AP200 - Instâncias (m,p) de $(170,3)$ a $(200,10)$	64
8.5	Comparações de resultados obtidos a partir do AP200 com a literatura	65

Capítulo 1

Introdução

A organização hierárquica dos sistemas de transporte e de telecomunicações pode ser encontrada em várias aplicações no mundo real, tais como a localização de estações de correio ou de centros de comutação telefônica, e desempenha um papel importante no desenvolvimento desses sistemas. Um dos objetivos sempre presentes é a minimização de custos, e, através do uso de várias topologias de agregação de fluxos, consegue-se engendrar economias de escala locais que diminuem os custos totais do sistema. (O’Kelly and Miller, 1994; Klinecicz, 1998; Luna, 2006).

As redes de transporte, telecomunicações e de computadores frequentemente empregam arquiteturas de *hub-and-spoke*, também conhecidas como eixo-raio, para direcionar de forma eficiente os fluxos entre origens e destinos múltiplos. A característica chave desta arquitetura reside no uso de pontos concentradores para transformações e redirecionamentos, aqui chamados de instalações de hubs (*hub facilities* ou simplesmente *hubs*, ou ainda polos ou nós concentradores). A formação de uma rede de hubs permite conectar um grande número de pares de origem-destino (O-D), também chamados de nós terminais, usando um pequeno número de ligações. Fluxos com a mesma origem, mas com destinos diferentes, são consolidados e roteados para os hubs e lá, depois de separados e antes de seguirem para os respectivos destinos, são combinados com outros fluxos de origens diferentes, mas com o mesmo destino.

Além desses redirecionamentos, os hubs podem executar transformações nos fluxos que passam por eles, que variam conforme o tipo de rede. No caso do transporte de produtos ou de correspondências, elas costumam envolver o transbordo, classificação de mercadorias, consolidação de embalagens, re-empacotamento, etc. No transporte de passageiros são feitas conexões, mudanças modais e manipulação e roteamento automático de bagagens. E nas telecomunicações podem ocorrer diversas transformações no sinal transmitido pelas redes, tais como ampliações de sinal,

compressões de dados, multiplexação, mudanças de frequências e de meio físico de transmissão, troca de protocolos, etc.

A criação de hubs ajuda a reduzir os custos de estabelecimento da infraestrutura, a centralizar operações de manuseio e triagem de mercadorias e a alcançar economias de escala nos custos de roteamento através da consolidação de fluxos. Em geral, os problemas de localização de hubs (em inglês, *Hub Location Problems* ou HLP) consistem em planejar a localização de instalações de hubs e de projetar redes de *hub-and-spokes* (eixo-raio) com a preocupação de otimizar um objetivo global, em geral baseado em minimização de custos ou maximização da capacidade de prover serviços.

Os HLPs constituem uma classe desafiadora de problemas NP-complexos (de complexidade não polinomial), envolvendo decisões conjuntas de localização e de projeto de rede. Sua principal dificuldade deriva da interrelação inerente entre dois níveis do processo de decisão. O primeiro nível considera a seleção de um conjunto de nós, ou de pontos no espaço, para localizar as instalações de hubs, enquanto o segundo nível lida com a configuração da rede dos hubs, selecionando os links que serão usados para conectar as origens e os destinos com os hubs, bem como definindo o roteamento dos fluxos através da rede.

Os HLPs são essenciais no planejamento de redes de transporte e telecomunicações. As áreas de aplicação de HLPs nos transportes são diversas. Elas incluem a entrega expressa de encomendas, o frete aéreo, as viagens de passageiros (aéreas, rodoviárias e ferroviárias), os correios, o transporte rodoviário de carga e os sistemas interligados e multimodais de trânsito urbano (ônibus, metrô, barcas, BRTs, etc.). A demanda a ser atendida corresponde a pessoas ou a objetos (por exemplo, pacotes expressos, passageiros, cartas, mercadorias), transportados por veículos (caminhões, trens, aviões, navios) que se movem em suportes físicos específicos, tais como estradas e ferrovias ou através do ar e da água. As instalações dos hubs correspondem a centros de classificação ou terminais de transporte em que um ou mais modos de transporte interagem. São exemplos os portos e aeroportos, rodoviárias, centros de distribuição de mercadorias, etc. A consolidação dos fluxos nos hubs permite que economias de escala incidam sobre os custos de transporte, não apenas com o roteamento de fluxos entre hubs, mas também entre nós O-D e os hubs.

As aplicações do HLP nas telecomunicações surgem na concepção de várias redes de dados distribuídos, onde a demanda corresponde a dados transformados em sinais eletrônicos que são encaminhados através de uma variedade de ligações físicas (cabos de fibra óptica, coaxiais e outros) ou através do ar ou do vácuo fora da atmosfera (canais de satélite, links de microondas, etc.). As instalações dos hubs são componentes de hardware como switches, modems, concentradores, multiplexadores e

roteadores. As economias de escala na transmissão de dados e na utilização da rede, contrapostas aos custos para a configuração das instalações de hubs e links de comunicação, são fatores que motivam o estudo de arquiteturas do tipo *hub-and-spoke*.

Evolução histórica do problema de localização de hubs

O estudo dos HLPs começou com o trabalho pioneiro de O’Kelly (1986a), para modelos de localização em espaços contínuos e O’Kelly (1986b, 1987), para espaços discretos, e desde então evoluiu para variadas áreas de pesquisa. Nas últimas três décadas, a localização de hubs tem sido estudada por pesquisadores de todo o mundo em diferentes disciplinas, tais como a ciência da localização, geografia, planejamento regional, otimização de redes, transportes, telecomunicações e ciência da computação. Existem vários estudos classificatórios e comparativos e levantamentos de pesquisas feitas sobre os HLPs, cada um deles focado em diferentes aspectos desses problemas. As primeiras análises dos HLPs, por O’Kelly e Miller (1994) e Campbell (1994a), contêm esquemas de classificação para os modelos fundamentais e para as estruturas topológicas aplicáveis às redes de hub. Klineciewicz (1998) concentra-se no projeto de redes de hub no contexto das redes de telecomunicações, e Bryan e O’Kelly (1999) apresentam uma pesquisa focada em redes de transporte aéreo. Campbell et al. (2001) fazem um levantamento abrangente de HLPs em que a localização dos hubs é a decisão-chave. Alumur e Kara (2008) fazem uma classificação e uma revisão da crescente literatura sobre os modelos de localização de redes de hubs até a data do seu estudo. Campbell e O’Kelly (2012) fornecem uma visão sobre as motivações iniciais para a análise de HLPs e destacam as direções tomadas pela pesquisa recente. Zanjirani Farahani et al. (2013) reveem os métodos de solução e as aplicações para várias classes de HLPs.

Este estudo se concentra nos problemas em que o acesso local e as redes centrais são projetadas em conjunto, e em que a decisão sobre a localização dos hubs é uma parte fundamental do problema. A adoção de uma arquitetura *hub – and – spoke* proporciona uma simplificação suficiente, em ambos os níveis de rede, para permitir uma visão integrada do problema de projeto dessas redes.

Para determinar a localização dos hubs capaz de produzir os melhores resultados, existem duas abordagens principais, a contínua e a discreta. Historicamente, como será exposto em maior detalhe no capítulo 4 adiante, nos primeiros estudos admitia-se que os hubs poderiam ocupar qualquer posição no plano, sendo sua localização definida pelas suas coordenadas cartesianas. Essa abordagem, entretanto, requer que a procura de uma solução ótima seja feita canonicamente através de

uma formulação que é muito difícil de ser tratada por métodos analíticos tradicionais, porque resulta numa problema matemático definido num espaço não convexo e altamente não diferenciável.

A abordagem discreta, por contraste, limita a escolha de hubs ao próprio conjunto das localidades de origem e destino. Isto restringe a escolha da solução ótima à busca de um subconjunto do conjunto de localidades que se quer interligar, e o número de soluções possíveis passa a ser finito. Em problemas com poucos pontos a serem conectados e poucos hubs, pode até ser viável se achar a solução ótima através do exame de todas as soluções logicamente possíveis, com o uso de um programa de computador relativamente simples.

Essa facilidade levou a que a maioria dos problemas de localização de hubs tratados na literatura dos últimos anos tenha preferido a abordagem discreta. Esta última também tem a característica de só propor a instalação de hubs em locais onde já existem instalações e/ou equipamentos, uma vez que são pontos de origem ou de destino de tráfego, enquanto a abordagem contínua propõe, em geral, a instalação de hubs em locais desocupados. A abordagem discreta tem, por sua vez, um problema sério: quando o número de localidades e o número de hubs aumenta, o número de soluções possíveis cresce muito rapidamente, podendo ocorrer uma explosão combinatória.

Por exemplo, com 10 hubs a serem escolhidos entre 200 localidades, o número de escolhas possíveis de hubs entre as localidades é muito grande, da ordem de 2×10^{16} . E, uma vez escolhida a localização dos hubs, a alocação do tráfego de saída e de chegada tem de ser definida pelo menos para todos os 200 pontos de origem ou destino ou, alternativamente, para cada um dos 40 mil pares de O-D, dependendo do tipo de alocação desejado (simples ou múltipla), como será visto adiante. No primeiro caso, cada decisão envolve uma escolha entre os dez hubs, enquanto no segundo é preciso decidir entre as 100 alternativas possíveis de pares de hubs, o que resulta em um número de escolhas extremamente alto, que é difícil de examinar exaustivamente caso a caso.

As diversas metodologias propostas para resolver a versão discreta do problema diminuem o número de passos a serem seguidos na busca da solução, mas ainda são sujeitas à explosão combinatória, sem conseguir em contrapartida, em muitos casos, a garantia de identificação do ótimo global. Maiores considerações sobre o número de casos de localização-alocação possíveis serão feitas no capítulo 2.

Este estudo examina o problema no espaço contínuo e sugere uma maneira de superar a dificuldade da não diferenciabilidade. É pressuposto que a rede central seja uma rede totalmente interligada, onde cada hub está ligado a cada um dos outros hubs através de um link mais eficiente do que as demais ligações entre eles

e os pontos terminais de origem e de destino. Não são permitidas ligações diretas entre pontos de origem e destino sem passar por pelo menos um hub, embora essa restrição possa ser relaxada com uma variante simples do modelo aqui proposto. Os raios (*spokes*) e o grafo completo da rede central impõem um diâmetro três, o que significa que qualquer demanda pode ser atendida por um máximo de três ligações entre a origem e o destino do tráfego.¹

Em particular, são considerados aqui problemas determinísticos em que os locais dos hubs são escolhidos dentro de uma região *contínua* plana. Este foi o caso do artigo seminal de O’Kelly (1986) sobre a localização de instalações hub que interagem umas com as outras. Algumas importantes contribuições posteriores foram apresentadas por Aykin (1988) e Aykin e Brown (1992). No entanto, depois daqueles estudos iniciais, e através dos inúmeros trabalhos subsequentes, as rotinas computacionais para resolver problemas de hubs de localização contínua não progrediram muito.

A literatura sobre sistemas *hub-and-spoke* tem apenas trinta anos, mas os problemas de localização *min-sum* originaram-se no século XVII, quando Fermat propôs a questão de, dados três pontos num plano, encontrar um ponto mediano no mesmo, tal que a soma das distâncias de cada um dos pontos até o ponto mediano seja minimizada. Alfred Weber, há um século, apresentou o mesmo problema, acrescentando também um peso em cada ponto para considerar as diferentes demandas dos diversos pontos. O problema de Fermat-Weber (uma das várias denominações dadas ao mesmo problema por Wesolowsky (1993), localiza instalações (medianas) em posições *contínuas* no plano euclideano. A ideia por trás do problema foi introduzida em grafos por Hakimi (1964), que definiu a mediana absoluta como o ponto de um grafo que minimiza a soma das distâncias ponderadas entre esse ponto e os vértices do grafo. Ele permitiu que esse ponto ficasse em qualquer lugar ao longo das bordas do gráfico, mas provou que a mediana absoluta ótima está sempre localizada num nó do grafo, proporcionando assim uma representação discreta de um problema contínuo, embora limitado às bordas do grafo. Seguindo o trabalho de Hakimi, Goldman (1969), propôs algoritmos para a localização de um hub único numa rede. Esses estudos se referem a espaços contínuos

Já a literatura sobre *hub-and-spoke* na abordagem discreta é baseada no trabalho pioneiro de O’Kelly (1987), introduzindo um programa quadrático no espaço dos inteiros para a localização das instalações hub. Outras formulações alternativas de programação linear inteira para problemas de localização discreta de hubs foram

¹Como é suposta uma rede de hubs totalmente interligada, não há razão para passar por mais do que dois hubs. No plano, a ligação direta entre dois hubs é sempre mais eficiente do que passando por um terceiro, devido à desigualdade triangular. E, em alguns casos, a ligação mais eficiente passa por somente um hub.

feitas por Campbell (1994-A), Ernst e Krishnamoorthy (1996, 1998), O’Kelly et al. (1996) e Hamacher et al. (2004). De fato, grande parte da literatura sobre hub-and-spoke, assim como os desenvolvimentos computacionais relacionados à mesma, tem seguido uma abordagem *discreta* para a localização dos hubs, selecionando um subconjunto dos nós existentes para exercerem o papel de hubs.

Campbell et al (2002), registrou que ”para cada um dos problemas discretos clássicos fundamentais de localização de instalações (problema da p -mediana, problema das instalações não-capacitado, problema de p -centros e problema de cobertura), problemas de localização análoga de hubs foram formulados e estudados, a saber: o problema de hub p -medianas, o problema dos hubs não-capacitados, o problema dos centros p -hub e o problemas de cobertura de hubs”. Igualmente Campbell (1994-B) faz uma ampla dissertação sobre o tema da localização de hubs. Reese (2005) apresenta uma revisão da problemática de localização de hubs e das diferentes metodologias usadas na sua resolução. Uma pesquisa do estado arte da literatura científica sobre a localização de hubs à época da publicação foi fornecida por Alumur e Kara (2008), que apresentam uma discussão ampla e completa deste tema. As metodologias de resolução para formulações discretas de hubs têm sido melhoradas continuamente. Usando a decomposição de Benders, Contreras et al. (2010), encontraram soluções exatas para o problema de locação de hubs não-capacitado de alocação múltipla, para instâncias com até quinhentos pontos.

A ideia deste trabalho é voltar para as formulações *hub and spoke contínuas* originais, mostrando que as manipulações da técnica de suavização hiperbólica são capazes de abordar problemas maiores do que aqueles da ordem de magnitude alcançada pelas abordagens *discretas*. Em certo sentido, o processo pelo qual isto é alcançado é uma extensão de um esquema de suavização hiperbólica, uma adaptação do método de penalização hiperbólica originalmente introduzido por (Xavier, 1982). Foi provado recentemente que a aplicação da técnica de suavização hiperbólica é bastante eficiente na solução de grandes instâncias de problemas de clustering (Xavier, 2010) e (Xavier e Xavier, 2011). A motivação para resolver versões de grande escala de problemas *hub-and-spoke* contínuos decorre, dentre outras utilizações no mundo real, de aplicações na área de correios e de logística de transporte rodoviário, transporte aéreo, redes de petróleo off-shore e várias aplicações na área de telecomunicações.

Este trabalho está organizado da seguinte maneira. Os diversos problemas práticos do mundo real em que os conceitos e as técnicas de hubs podem ser aplicados são descritos no capítulo 2. O capítulo 3 faz uma panorama resumido das diversas categorias de problemas de *hub-and-spoke* encontrados na literatura técnica. A especificação do problema contínuo de *hub-and-spoke* p -medianas é detalhada no

capítulo 4. A metodologia de suavização hiperbólica é descrita no capítulo 5. Os resultados computacionais ilustrativos do caso contínuo são apresentados no capítulo 6. O capítulo 7 trata da extensão da metodologia exposta no capítulo 5 para lidar também com problemas em espaços discretos. Os resultados computacionais obtidos para alguns problemas discretos são descritos no capítulo 8. O nono e último capítulo é dedicado às conclusões.

Capítulo 2

Aplicações do Conceito de Hubs e do UMA_pHMP no Mundo Real

A ciência da locação lida com diversas situações onde é preciso se determinar no espaço o lugar mais adequado para se colocar um determinado equipamento ou instalação. É presumido que existam diversos pontos distribuídos no espaço (ou diversas áreas que são representadas por pontos) que tenham necessidade de estabelecer certas ligações (links) de algum tipo com outros pontos (ou áreas). Alguns exemplos de tais ligações são rotas de distribuição de produtos de consumo a partir de centros de distribuição, rotas aéreas ligando as cidades de uma região, transferência de dados entre pontos diferentes de uma rede de telecomunicações, estações de metrô atendendo a todos os moradores de uma região, etc.

Quanto à natureza das ligações que são estudadas na ciência da locação, é possível distinguir, dentre outras classificações, dois tipos básicos: no primeiro, cada ponto tem apenas a necessidade de receber, através das ligações que chegam a ele, certos fluxos (de objetos, pessoas, informações, etc.), independentemente de sua origem, ou então precisa enviar um fluxo, cujo destino final não é especificado. Como exemplos de fluxos desse tipo nos dois sentidos, pode-se citar a cadeia de suprimentos que abastece um supermercado (podem vir de qualquer centro de distribuição) e o lixo coletado das residências (seu destino final não é determinado individualmente pelos moradores).

O segundo tipo de ligação ocorre quando o fluxo através das ligações não é associado apenas a uma localidade, e sim a um par específico de localidades de origem e destino (O e D). Nesse caso, cada localidade tem a necessidade de se ligar a diversas outras (possivelmente a todas as outras localidades consideradas), com demandas específicas para cada ligação O-D individual. Isso ocorre, por exemplo, no caso dos possíveis trajetos de cartas mandadas pelo correio, que são enviadas a partir de qualquer endereço de origem ou de coleta para qualquer endereço de

destino que se deseje, ou nas ligações telefônicas, nas quais qualquer número de telefone pode chamar qualquer outro. Nestas situações, a opção de se fazer uma ligação direta entre cada localidade de origem e de destino pode ser muito onerosa, à medida que aumenta o número de localidades a serem conectadas.

O conceito de hubs surgiu naturalmente nessa situação, em que havia necessidade de se ligar diversos pares de localidades distribuídos numa região física de topologia plana, com ligações específicas para cada par. Uma das primeiras aplicações práticas desse conceito ocorreu no início do século vinte, quando se começou a instalar redes telefônicas através de fios ligando as diversas habitações de uma cidade.

A inviabilidade de ligações diretas casa a casa levou à criação de hubs, na forma de estações telefônicas de comutação, em que as residências, e também locais de trabalho, eram conectadas fisicamente a uma única estação situada na sua vizinhança, e estas se conectavam com outras estações através de ligações de maior capacidade. Cada ligação ia assim sucessivamente de sua origem a uma estação, desta para uma segunda estação, e finalmente da segunda estação para o local de destino.

Por exemplo, num pequeno lugarejo com apenas cinco casas, seria possível interligar todos os pares de casas por um cabo telefônico dedicado a cada par de casas, havendo necessidade de apenas dez cabos, o dobro do número de casas. Mas numa cidade média, de mil casas, haveria necessidade de quase 500 mil cabos separados, o que já é bem mais complicado de se instalar, e numa cidade grande com dois milhões de residências e escritórios, seriam necessários dois trilhões de cabos, algo completamente inviável.

Para enfrentar essa dificuldade, utilizou-se o conceito de pontos concentradores, ou simplesmente concentradores (hubs em inglês, abreviação H). Esses são pontos especiais para os quais convergem ligações vindas de muitas localidades de origem ou dirigidas a muitas localidades de destino. Nesses hubs o fluxo que é recebido pode ser transformado de diversas maneiras e depois é redirecionado para outro ponto, que tanto pode ser a localidade final de destino como um outro hub, sendo que este segundo hub poderá transformar o fluxo mais uma vez e o redirecionar novamente. Com isso consegue-se uma diminuição muito grande no número de ligações necessárias, sendo que essa diminuição se torna mais significativa à medida que o número de localidades a serem interligadas aumenta.

Nos exemplos anteriores, com o uso de concentradores: o número mínimo de ligações necessárias para o lugarejo de cinco casas pode cair para a metade (de dez para cinco), usando um único hub independente; na cidade média de mil casas, ao se usar trinta hubs independentes, o número de ligações cai para 1435 (o que corresponde a 0,3 por cento do valor anteriormente estimado); e na cidade grande, caso se use dois mil hubs, o número de links físicos necessários se reduz para cerca

de quatro milhões, resultando na média de apenas duas ligações por localidade de origem/destino.



A comutação das linhas centrais que permitiam que todos se ligassem a todos era feita inicialmente com a ajuda de telefonistas, que se utilizavam de painéis de comutação para conduzir a ligação telefônica por cabos instalados fisicamente, e que conduziam os sinais analógicos de sua origem a seu destino. Depois, esses painéis foram automatizados em diversos níveis, até se chegar aos sistemas celulares atuais, cuja concepção, como é indicado pelo próprio nome, é baseada em células, que nada mais são do que regiões do espaço definidas pela presença predominante de um hub. Nesses sistemas, as transmissões são feitas predominantemente através de ondas hertzianas, cabos coaxiais ou fibras óticas, e as conexões são todas feitas automaticamente por computadores. Tais redes fazem uso de diversos sistemas de hubs de alocação múltipla, embora os terminais finais (os telefones celulares) ainda se liguem, de forma dinâmica, a hubs com alocação simples, conforme varia a localização do aparelho no espaço ao longo do tempo.

Ao chegar aos hubs o fluxo pode ser modificado. Exemplificando, essas modificações, no caso de redes de comunicações, assumem a forma de ampliação do sinal, alterações de faixas de frequência, re-codificação segundo diversos protocolos, mudança de meio físico de transmissão dos dados, etc. No caso de sistemas de transportes de cargas, são feitos reembalagens, separação de cargas segundo o des-

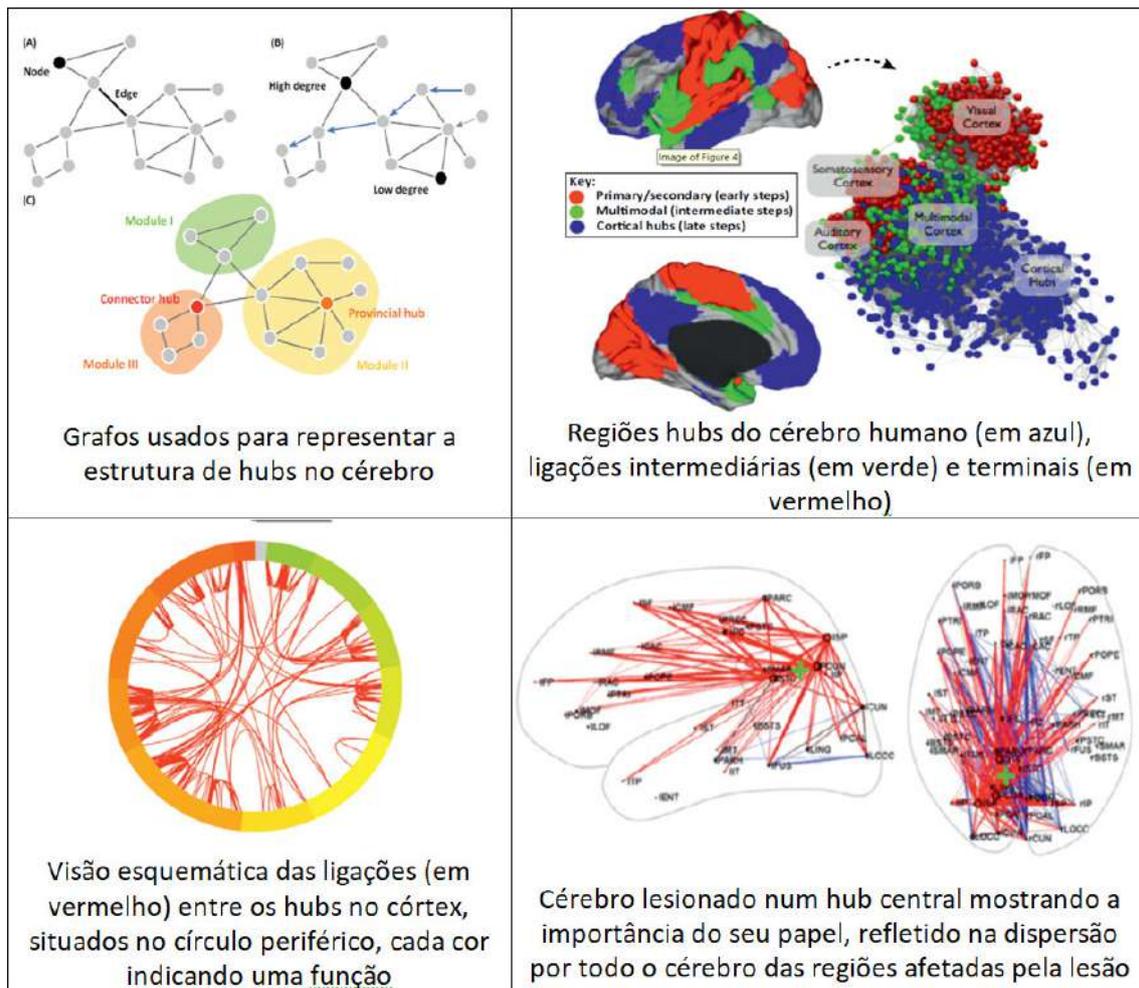
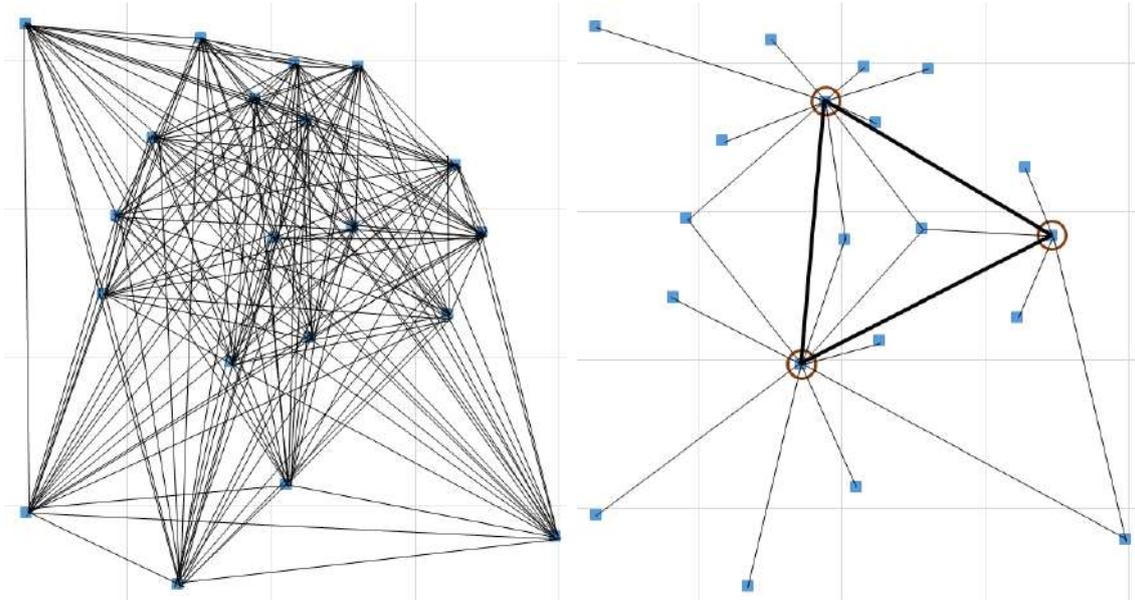


Figura 2.1: Representação de hubs no Cérebro

tino e consolidações de volumes. A criação dos hubs faz com que passe por eles uma quantidade maior de tráfego do que a que sai de cada ponto O ou a que chega a cada ponto D . Eles têm de ser dimensionados, portanto, para tratar de um grande número ou de grandes volumes de carga, pessoas ou dados, conforme o tipo de ligação que esteja sendo tratado. Além disso, as ligações entre os hubs, do tipo H-H, por concentrarem tráfego de diversas origens e destinos, devem ter capacidade maior do que as ligações O-H ou H-D. Isso justifica a utilização de tecnologias mais eficientes (e possivelmente com custo de implantação maior), capazes de prover economias de escala nas ligações. No caso de telecomunicações, trata-se de links com banda mais larga, frequentemente com o uso de satélites. No caso de transporte de carga, usam-se caminhões maiores, trens ou navios, e no caso de transporte aéreo de passageiros, aviões maiores e/ou voando com maior frequência entre os hubs, além de outras transformações na carga. No transporte de passageiros, são feitos traslados, mudanças de modo de transporte, controle de bilhetes, etc.

É interessante se notar que os hubs são utilizados mesmo na natureza, sem interferência humana. Um exemplo notável ocorre no cérebro dos seres humanos e



(a) Rede 20 pontos, links diretos, total 190 arcos (b) Mesma rede 20 pontos, com 3 hubs e 25 arcos
 Figura 2.2: Conjunto de 20 localidades extraídas do AP200, comparando os links diretos entre as localidades com aqueles feitos através de 3 hubs com alocação múltipla, usando a configuração ótima dos hubs no espaço discreto

animais. Está provado pela ciência que a maior parte da atividade cerebral se dá através do estabelecimento de ligações especiais entre diferentes pontos do cérebro, os neurônios. O número de ligações pode ser muito grande, já que existem cerca de 10^{11} neurônios num cérebro humano, sendo que cada um deles tem aproximadamente sete mil conexões com outros neurônios, feitas através de sinapses. Estas são em número que pode chegar a 10^{15} . Seria muito difícil conectar diretamente cada neurônio a todos os outros através de um caminho independente, porque seriam 10^{22} ligações simultâneas num espaço tridimensional limitado. Assim, o cérebro recorre a hubs, com características que permitem que as células se interconectem através de links capazes de tratar ao mesmo tempo de diversas ligações entre pares de neurônios.

A figura 2.1 acima contém imagens extraídas de estudo sobre estrutura de hubs no cérebro humano. (Heuvel, Sporns, 2013). O artigo mostra que eventuais danos nessas estruturas centrais de comunicação entre os neurônios podem ter efeito muito mais intenso e generalizado do que danos nas regiões intermediárias ou periféricas, que tendem a provocar sintomas mais localizados e restritos, ilustrando assim a maior importância relativa das regiões que servem como hubs.

A figura 2.2 mostra um conjunto de 20 localidades formadas a partir do conjunto AP200, descrita no apêndice A e para o qual a localização de hubs é computada no capítulo 8. A primeira parte da figura ilustra como seriam as ligações diretas entre todas as vinte localidades, num total de 190 arcos de ligação. A segunda parte mostra os 25 arcos de ligação necessários para ligar as mesmas vinte localidades através de

três hubs com alocação múltipla, escolhidos entre as 20 localidades. Desses, 22 são do tipo raio, unindo um hub a uma localidade não hub, e 3 unem os hubs entre si. A maioria das 17 localidades não hub está ligada a um só hub, três estão ligadas a dois hubs e uma única está ligada aos três hubs.

Para os parâmetros de custo propostos no conjunto AP200, pode se demonstrar ainda que a rede formada com hubs, além de ser bem mais simples de operar, por ter muito menos ligações, tende a ter também um custo de transferências menor do que a rede completamente interligada da primeira parte da figura. De fato, embora cada um dos trajetos entre localidades não escolhidas como hub seja mais longo no segundo caso do que no primeiro (por se desviar da linha reta entre dois pontos para poder passar pelos hubs), a economia feita no trajeto inter-hubs compensa o trajeto maior, resultando num custo total menor do que o das ligações diretas.

Capítulo 3

Variedades de Problemas de Localização de Hubs

Os problemas de localização e alocação de hubs podem assumir características diversas, existindo grande variedades de problemas, sendo que inúmeros artigos foram publicados nos últimos anos lidando com muitas destas variedades.

Além da diferenciação já exposta no capítulo anterior entre os problemas da procura de hubs num espaço contínuo e da procura no espaço discreto formado pelas próprias localidades a serem ligadas, as diferenças entre os tipos de problemas de *hub and spoke* encontrados na literatura se devem a várias outras características desses problemas, algumas das quais são descritas a seguir:

a) A função objetivo a ser minimizada. As variantes mais comum são o custo total, que corresponde a um problema *min-sum* e o custo máximo, que resulta numa formulação *min-max*.

b) Algumas formulações, sobretudo as tratadas nos primeiros HLPs abordados na literatura, procuram um só hub, mas a maioria considera vários hubs. No caso de múltiplos hubs, a hipótese mais comum é de que no máximo dois hubs devem ser usados para a ligação entre duas localidades, não havendo incentivos ao uso de um terceiro hub (veja nota de rodapé à página 4). Uma exceção é o caso em que existem diversas camadas hierárquicas de hubs, com os hubs nas camadas superiores ligando pontos que exercem o papel de hubs para as camadas inferiores.

c) Ainda quanto ao número de hubs, este pode ser determinado a priori ou pode ser buscado como parte da solução do problema.

d) Comumente é exigido que todas as ligações O-D passem por pelo menos um hub, porém é possível tratar também problemas em que esse requisito é deixado de lado, e se permitem ligações diretas de localidade a localidade, sem passar por

qualquer hub.

e) Os tráfegos entre as localidades e os hubs nas duas direções podem ter o mesmo custo ou não. Pode se considerar que o tráfego localidade-hub tenha um custo, normalmente expresso pelo símbolo χ , e que o custo unitário seja δ no sentido contrário. Entre os hubs, o custo unitário é de hábito expresso por α , sendo $\alpha < \chi, \delta$. Quando o tráfego tem o mesmo custo nos dois sentidos, isto é, $\chi = \delta$, é considerado que ambos tenham um custo unitário, e $0 \leq \alpha \leq 1$, sendo que $(1 - \alpha)$ representa a economia que se pode obter com o tráfego entre hubs.

f) Alocação de localidades aos hubs: a alternativa mais simples é que se exija que cada localidade se ligue somente a um hub, a ser utilizado para todas as suas ligações com as demais localidades, o que é chamado de alocação simples ou única (*single allocation*). Alternativamente, pode ser permitido que ela faça suas ligações, tanto de saída como de entrada, através de mais de um hub, em qualquer número, dependendo da conveniência da ligação com cada uma das outras localidades à qual ela deve ser ligada, o que caracteriza a situação de alocação múltipla (*multiple allocation*). Existe também uma situação intermediária, em que se limita o número de hubs r ao qual uma localidade pode ser ligada, sendo $1 \leq r \leq p$, problema proposto por Yaman (2011) e resolvido por Marti et al. (2014). Naturalmente o caso $r = 1$ corresponde à alocação única e o caso $r = p$ à alocação múltipla, porém o artigo de Marti et al. não trata desses casos extremos, não permitindo a comparação do seu método com os de outros autores.

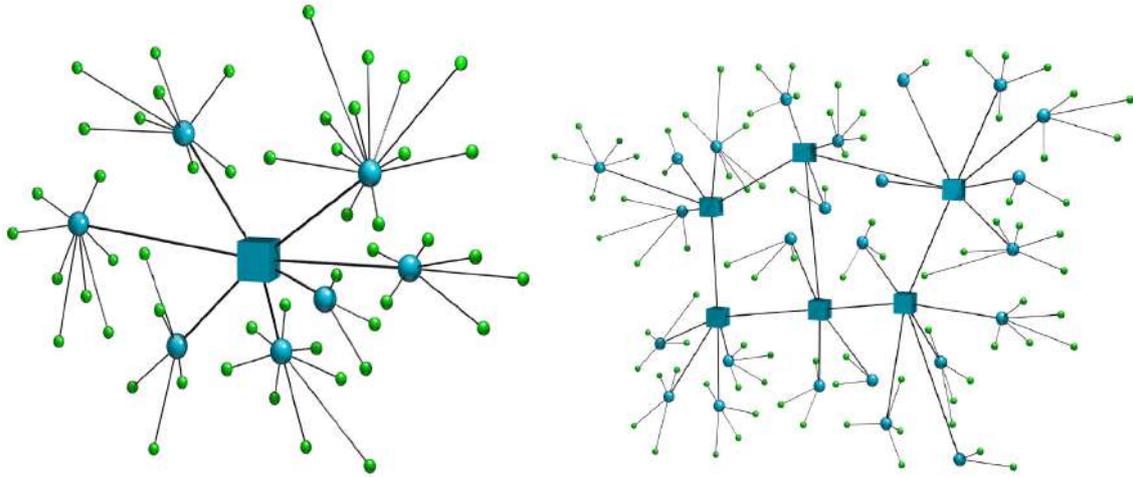
g) As ligações entre os hubs, ou entre os hubs e as localidades, podem ter capacidades máximas de tráfego.

h) Os hubs também podem ter uma capacidade máxima de tráfego. Nessas situações e nas do item anterior, se diz que o problema é capacitado (*capacitated*), por contraste aos problemas em que essas limitações não existem, chamados de não capacitados (*uncapacitated*).

i) Pode haver custo de instalação dos hubs, que são relevantes se o número de hubs não foi definido a priori. Esse custo tem de ser levado em consideração no cálculo da função objetivo a ser minimizada. Ele pode ainda ser fixo ou variável, esse último dependente do fluxo de tráfego através do hub.

j) Pode haver também um custo fixo de instalação das ligações entre as localidades e os hubs, como em O’Kelly et al (2014). Esses links tem sempre um custo variável, proporcional ao tráfego que comporta.

k) A topologia dos hubs também pode variar, sendo o caso mais geral o de grafo completo, em que todos os hubs se ligam a todos os demais, normalmente através de links que têm custo unitário menor que os links entre as localidades e os hubs.



(a) Rede hierárquica em estrela de dois níveis, (b) Rede hierárquica em estrela de dois níveis com do tipo *multicast* múltiplos hubs no segundo nível

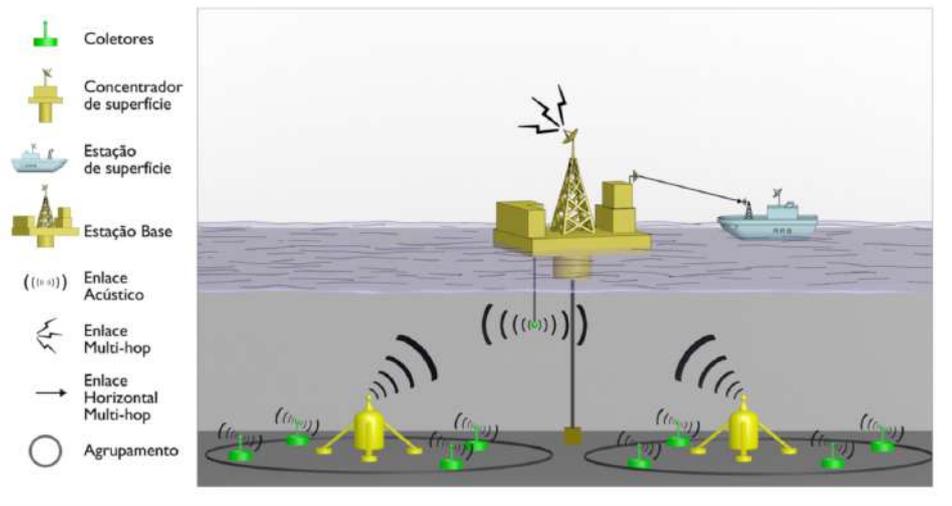
Figura 3.1: Outras topologias de rede

Podem haver diversas configurações de grafos incompletos, como a topologia em anel, as de diversos níveis ou camadas e ainda outras. A figura 3.1, cedida por Barbosa (2017) ilustra algumas das topologias possíveis. Já a figura 3.2, também cedida por Barbosa, mostra a configuração física de alguns tipos de rede usadas na prospecção subaquática de petróleo.

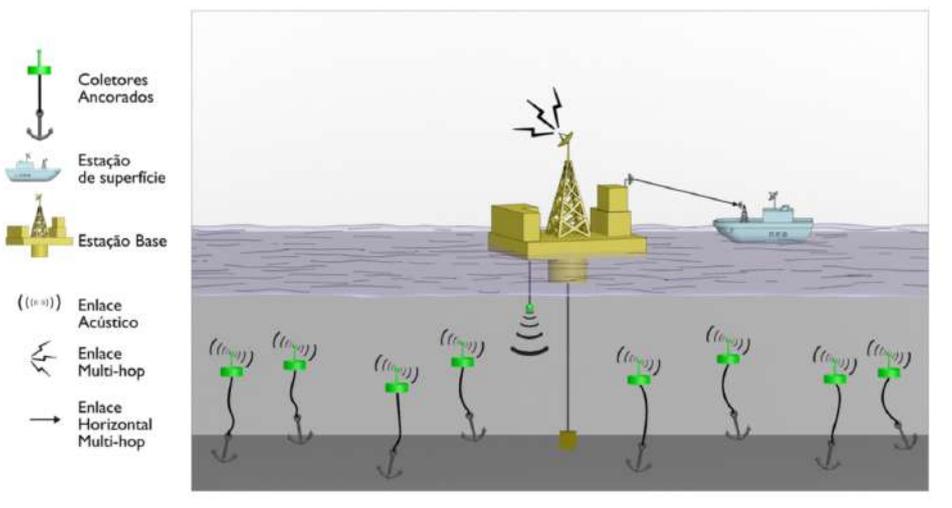
1) O tráfego entre as localidades pode ser determinístico, definido através de uma só medida expressa através de um número real, ou estocástico, em que se presume que ele obedeça a uma dada distribuição de probabilidades.

De acordo com algumas das possibilidades acima, podem ser caracterizados quatro grandes grupos tradicionais de HLPs com múltiplos hubs, conforme a classificação original de Campbell (1994-B):

1. O problema da Mediana p -hub (p -hub Median Location Problem), em que o número de hubs p é definido a priori e se procura localizar esses hubs de tal forma que a soma dos custos de todos os trajetos possíveis entre as localidades seja a menor possível, caracterizando um problema *min-sum*.
2. O problema da alocação de hubs com Custos Fixos (Hub Location Problem with Fixed Costs), em que o número de hubs não é conhecido a priori, e deve ser determinado de forma a minimizar o custo total de escolha dos hubs, considerando os custos fixos de instalação dos mesmos.
3. O problema de Centros p -hubs (p -Hub Center Location Problem), em que o número de hubs é conhecido como no caso 1, mas o objetivo é o de minimizar o custo máximo das diversas ligações O-D ao invés do custo total, caracterizando um problema de tipo *min-max*.



(a) Rede subaquática estacionária 2D



(b) Rede subaquática estacionária 3D

Figura 3.2: Redes subaquáticas

4. O problema de Cobertura de hubs (Hub Covering Location Problem), em que o número de hubs p é determinado endogenamente, levando em consideração custos máximos pré-estabelecidos para as diversas ligações O-D e a necessidade de cobrir toda a demanda.

Através da combinação de diversas outras instâncias das características acima descritas, pode se chegar a um número muito grande de formulações diferentes, que ainda podem ser variadas em casos específicos para atender a situações que ocorrem no mundo real dos transportes ou das telecomunicações. Muitas dessas combinações de características ainda não foram suficientemente exploradas.

Quanto às soluções que são propostas para as diversas formulações possíveis, elas podem recorrer a diversos métodos heurísticos - tais como busca tabu, algoritmos genéticos, simulated annealing, etc. Alguns algoritmos procuram estabelecer o ótimo global, como a Decomposição de Benders. Em grande parte dos problemas de maiores dimensões, não se conhece ainda o ótimo global. Os artigos de Alumur e Kara e de Farahani et al (2013) fornecem resenhas de muitos dos trabalhos publicados nesta área.

Este estudo, como será melhor detalhado nos próximos capítulos, propõe soluções para o problema de alocação de hubs em espaços contínuos e discretos, e se concentra apenas nas instâncias abaixo especificadas das diversas características possíveis desses problemas que foram relacionadas no início deste capítulo, sendo adotada aqui a mesma classificação lá utilizada:

- a) Minimização do custo total (*min - sum*)
- b) Múltiplos hubs.
- c) Número de hubs determinados a priori.
- d) Todas as ligações O-D passam por pelo menos um hub.
- e) Custos χ e δ (conforme definições no início deste capítulo) foram considerados em alguns casos, sendo que na versão tratada no capítulo 6 eles eram considerados iguais e unitários, tornando o problema dependente só de α .
- f) Alocações múltiplas.
- g) Ligações sem capacidade máxima.
- h) Hubs sem capacidade máxima.
- i) Não se consideram os custos de instalação dos hubs.
- j) Não se consideram os custos de instalação das ligações entre hubs.
- k) Grafo completo.

1) Tráfego O-D entre cada par de localidades é determinístico e conhecido a priori.

O método exposto neste estudo pode ser adaptado para tratar muitas outras instâncias dentre as relacionadas acima, com menor ou maior dificuldade. Em particular, as restrições dos itens a, b, d, e, f, k acima podem ser contornadas sem maiores dificuldades, mediante modificações adequadas.

A principal vantagem da metodologia HSHS (a ser abordada adiante no capítulo 5) é que ela permite abordar problemas de ordem de grandeza bem maior do que os algoritmos alternativos. Os algoritmos de otimização combinatória que operam somente em espaços discretos apresentam sempre a deficiência da explosão combinatória, dificultando muito o tratamento de problemas com muitas localidades ou muitos hubs.

Para ilustrar essa explosão combinatória na procura de soluções para o problema de localização de hubs, é feita a seguir o cálculo do número de casos possíveis para duas variantes desses problemas, a saber o Problema de Localização de Hubs do Tipo p -Medianas com Alocação Múltipla no caso discreto (conhecido pela sigla UMApHMP - Uncapacitated Multiple Allocation p -Hub Median Problem), que é tratado nos capítulos 6 e 7 desta tese, e o seu equivalente de alocação simples USApHMP (Uncapacitated Single Allocation p -Hub Median Problem), que não é detalhado neste estudo mas que poderá ser abordado por metodologia semelhante à aqui descrita.

Para calcular quantas soluções existem no caso discreto para o problema de escolha de p -hubs medianos (USApHMP e UMApHMP) é conveniente separar esses problemas em suas duas componentes: localização e alocação. A localização consiste na escolha das localidades que servirão como hubs e a alocação se refere à definição de quais rotas serão seguidas para a ligação entre os diversos pares de localidades, passando pelos hubs.

Caso se considere m localidades e se queira escolher p dentre elas para constituir os hubs, resultam $m!/(p!(m-p)!)$ possibilidades de escolha de hubs.

Para cada escolha de conjunto de pontos que servirão como hubs, diversas alocações são possíveis. O número de possibilidades depende do método de alocação - simples ou múltipla. No caso de alocação simples (USApHMP), como cada localidade não escolhida como hub pode se ligar somente a um hub, existem p possibilidades de escolha para cada uma das $m-p$ localidades não-hub. Como tais escolhas são independentes, o número de possibilidades de alocação para cada conjunto de hubs é de p^{m-p} e o número total de soluções de localização-alocação será, portanto, dado pela fórmula: $N_s(m, p) = m!p^{m-p}/(p!(m-p)!)$.

Note-se que esse cálculo comporta a possibilidade de que um dado hub não seja ligado a nenhuma localidade, afora os outros hubs. Isso ocorreria caso nenhuma das localidades não hub tenha vantagem em se ligar ao hub em causa, o que significa que não estaria havendo qualquer aproveitamento das facilidades que esse hub oferece por parte de outras localidades. Entretanto, levando em consideração que, no espaço discreto, cada hub é também um ponto de origem-destino, ele sempre servirá à sua própria demanda e, portanto, mantém sua condição de hub mesmo que não sirva a qualquer outra localidade. Isso provavelmente não ocorreria em soluções ótimas, que não devem comportar hubs "egoístas", e não faz sentido em espaços contínuos, onde os hubs não são necessariamente pontos de demanda de tráfego, o que tornaria o hub "inútil". Mas no espaço discreto essa possibilidade existe e muitas das alocações possíveis têm esta característica.

Por exemplo, para $p = 3$ e $m = 7$, para cada escolha de três hubs, são possíveis $3^4 = 81$ alocações diferentes, considerando as decisões independentes a serem feitas quanto às ligações das quatro localidades não hub. Entretanto, só 36 dessas alocações, menos da metade, não são "egoístas", no sentido de que todos os hubs servem pelo menos a uma outra localidade, afora servirem a si mesmos.

No caso da alocação múltipla (UMApHMP), o número de possibilidades de alocação tem de ser definido para cada possível par O-D, que existem em número de $(m-p) \times (m-p-1)$. As possibilidades de alocação para cada um desses pares são em número de p^2 , e, portanto, existem $N_m(m, p) = m!p^{2(m-p)(m-p-1)} / (p!(m-p)!)$ soluções possíveis para o problema de alocação-localização. Vale aqui também a observação feita acima com relação a hubs "egoístas", e é possível que algum hub não seja escolhido para fazer parte de qualquer um dos inúmeros trajetos O-D possíveis, exceto aqueles em que o próprio hub é a origem ou o destino.

A explosão combinatória fica bem caracterizada, e vale a pena notar adicionalmente que estes problemas têm um número de soluções consideravelmente maior do que alguns outros problemas tradicionais de alocação, tais como o de particionar um conjunto de m pontos em p grupos, categoria que inclui o problema de clusterização (*clustering*). No USApHMP discreto, o conjunto de localidades que se ligam de forma única a cada um dos hubs constitui, junto com o próprio hub, um grupo que se assemelha muito aos grupamentos formados no processo de clustering e, portanto, pode-se imaginar que o USApHMP tenha o mesmo número de soluções possíveis que o problema de clustering. Entretanto, o fato de que não só é necessário dividir as localidades em p grupos como também é preciso definir em cada grupo qual delas atuará como hub, leva a que o número de soluções possíveis aumente bastante. E no UMApHMP, que tem alocações mais complexas e, no caso particular em que $\alpha = 0$ se torna equivalente ao USApHMP, o número de soluções é diversas

ordens de grandeza maior.

Como exemplo, pode-se conferir as seguintes situações:

1. Para poucos hubs e poucas localidades, como na escolha de $p = 2$ hubs entre $m = 7$ localidades ou na clusterização equivalente de 7 pontos em dois grupos, chega-se a 63 soluções possíveis de clustering, 672 casos possíveis de USApHMP e mais de 23 trilhões de casos de UMAPHMP, o que já torna este último um problema de solução relativamente demorada, pelo método trivial de exame de todas as possibilidades existentes, também conhecido como "força bruta". Isso poderia ser feito sem dificuldades para os outros dois problemas, embora o USApHMP tenha um número de casos mais de dez vezes maior do que o clustering.

2. Aumentando um pouco o número de hubs e localidades, como no caso de $p = 6$ e $m = 20$, os casos possíveis de clustering explodem para cerca de 4 trilhões, e o USApHMP cresce ainda mais, indo para a faixa de 3 quadrilhões e tendo 700 vezes mais casos do que o clustering, enquanto o UMAPHMP chega a $6,8 \times 10^{287}$ casos. Esse último número já está muito longe do alcance de qualquer computador na procura "força bruta", embora a busca da solução ótima possa ser muito acelerada com o uso de diversas técnicas que diminuem fortemente o número de casos a avaliar.

3. Quando a dimensão do problema chega à ordem de grandeza das instâncias maiores tratadas hoje em dia, como no caso de $p = 10$ e $m = 100$, enquanto o número de casos de clustering é $2,8 \times 10^{93}$, os problemas de hubs chegam, no caso da alocação única, a ser cerca de seis bilhões de vezes mais numerosos: existem $1,7 \times 10^{103}$ casos. Ambos os valores são consideravelmente maiores do que a estimativa do número total de átomos existentes no Universo. Quanto à alocação múltipla, chega-se a um número que é até difícil de conceber: $1,7 \times 10^{16033}$, o que ilustra a característica não polinomial (NP) do problema de *hub and spoke*.

Capítulo 4

Especificação do Problema de Localização de Hubs a ser Tratado

No âmbito do tópico Problemas de Localização, conhecido na literatura como *Location Sciences*, há um número expressivo de formulações, muitas das quais foram tratadas no capítulo 2. Considerando só as mais consagradas, esse número atinge certamente a ordem de centenas, como pode ser visto, por exemplo, em Farahani e Hekmatfar (2009). Neste capítulo, serão apresentadas quatro formulações pertencentes ao seletivo grupo das mais estudadas e utilizadas para equacionamento de aplicações práticas reais. As duas últimas são sobre localização de hubs segundo os enfoques contínuo e discreto, objeto central deste trabalho. As duas primeiras são úteis para contextualizar a abordagem aqui adotada.

Esses quatro problemas tem um objetivo comum, qual seja, localizar um determinado número p de pontos (descritos também como polos, nós, centróides ou hubs) num espaço euclidiano de duas dimensões, de forma a minimizar uma função específica de custo de transporte.

Para formular esses problemas, procede-se da forma descrita a seguir. Considere-se que $S = \{s_1, \dots, s_m\}$ denota um conjunto de m localidades ou cidades de origem-destino numa região plana. Dependendo do problema, pode existir uma demanda única de tráfego associada a cada localidade, envolvendo a conexão entre essa localidade e qualquer um dos centróides, ou pode haver uma demanda específica de tráfego associada a todos os pares de localidades, tráfego esse que deverá fluir entre uma localidade e outra, passando por pelo menos um centróide. No primeiro caso a demanda associada a cada localidade j será denotada por w_j . No segundo, a demanda entre duas localidades específicas j e l será representada por w_{jl} . Sejam os centróides, as incógnitas, representados por $x_i, i = 1, \dots, p$, onde cada $x_i \in \mathbb{R}^2$. O conjunto desses centróides constitui o espaço de solução e é representado por $X \in \mathbb{R}^{2p}$.

Problema de Alocação Múltipla de Fermat-Weber

O Problema de Alocação Múltipla de Fermat-Weber tem recebido diferentes nomes, tais como Problema de Alocação-Localização Contínuo, ou simplesmente Problema de Fermat-Weber, como discutido por Wesolowsky (1993). É um problema matemático não-diferenciável e não-convexo, com um grande número de minimizadores locais, conforme mostrado por Rubinov (2006). Logo, é um problema de otimização global de difícil solução.

Os problemas de localização *min – sum* originaram-se no século XVII, quando Fermat colocou a questão de: dados três pontos num plano, encontrar nesse plano um ponto mediano de tal modo que a soma das distâncias de cada um dos pontos para o ponto mediano seja minimizada. Alfred Weber, há um século, propôs o mesmo problema para um número qualquer de pontos, adicionando pesos em cada ponto para considerar a demanda dos clientes. O problema de Weber localiza instalações (medianas) em locais *contínuos* no plano euclideo, conforme mostrado por Koopmans e Beckmann (1957). Neste problema, se procura achar o conjunto de medianas que minimize o custo total de transferência das demandas dos diversos pontos entre os mesmos e a mediana mais próxima de cada um deles.

Para formular o Problema de Alocação Múltipla de Fermat-Weber como um problema *min – sum – min*, procede-se da seguinte forma: dada uma localidade $s_j \in S$, calcula-se inicialmente a distância euclidea de s_j para o centróide mais próximo:

$$z_j = \min_{i=1, \dots, p} \|s_j - x_i\|_2. \quad (4.1)$$

O Problema de Alocação Múltipla de Fermat-Weber consiste então na localização de p instalações (hubs ou centróides) de forma a minimizar o custo total, ou seja:

$$\begin{aligned} & \text{minimizar} \quad \sum_{j=1}^m w_j z_j & (4.2) \\ \text{sujeito a:} \quad & z_j = \min_{i=1, \dots, p} \|s_j - x_i\|_2, \quad j = 1, \dots, m. \end{aligned}$$

Problema das p -Medianas Discreto

O Problema de p -Medianas é muito semelhante ao problema de Alocação Múltipla de Fermat-Weber. A única diferença é que cada centróide deve ser coincidente com uma das m localidades ou cidades especificadas, o que é assegurado através do uso de uma família de variáveis booleanas adicionais v que define quais

dos pontos de origem-destino do tráfego serão usados como hub. A formulação é a seguinte:

$$\begin{aligned}
 & \text{minimizar} \quad \sum_{j=1}^m w_j z_j & (4.3) \\
 \text{sujeito a:} \quad & z_j = \min_{i=1, \dots, p} \|s_j - x_i\|_2, \quad j = 1, \dots, m; \\
 & x_i = \sum_{j=1}^m v_{ij} s_j, \quad i = 1, \dots, p; \\
 & \sum_{j=1}^m v_{ij} = 1, \quad i = 1, \dots, p; \\
 & v_{ij} \in \{0, 1\}, \quad i = 1, \dots, p, \quad j = 1, \dots, m.
 \end{aligned}$$

As três últimas restrições do problema 4.3 têm o papel de que cada hub x_i seja coincidente com um das m cidades.

Problema de Localização de Hubs do Tipo p -Medianas Contínuo

O Problema de Localização de Hubs do Tipo p -Medianas Contínuo, conhecido na literatura como *Continuous p -Median Hub Problem*, é uma extensão natural do problema de Alocação Múltipla de Fermat-Weber, considerando a inclusão de dois hubs em cada trajeto entre duas cidades.

Neste trabalho é sempre feita a hipótese de que os hubs constituem uma topologia em que todos são ligados entre si, formando o que na teoria de grafos é denominado uma *clique*.

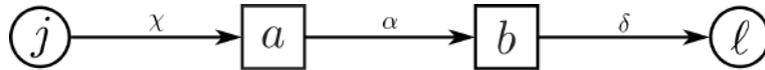


Figura 4.1: Uma conexão entre a cidade j e a cidade l

No Problema de Localização de Hubs do Tipo p -Medianas Contínuo em consideração, as conexões entre cada par de localidades ou cidades j e l , retratadas pela Figura 4.1, têm sempre três partes distintas: a partir da cidade de origem j a um primeiro hub a , de a para um segundo hub b e de b para a cidade de destino l . Os parâmetros $\chi \geq 0$, $\alpha \geq 0$ e $\delta \geq 0$ (vide definições no item e, início do capítulo 3) correspondem aos custos unitários de cada uma das partes do trajeto. A alocação múltipla é permitida, o que significa que qualquer ponto pode ser servido por um ou mais hubs. O primeiro e o segundo hubs podem ser coincidentes (isto é,

$a = b$), o que significa que um hub único seria usado para conectar o ponto de origem j até o ponto de destino l . A Figura 4.2 mostra as p^2 conexões possíveis entre duas cidades. Como ilustrado, os hubs a e b podem ser o mesmo.

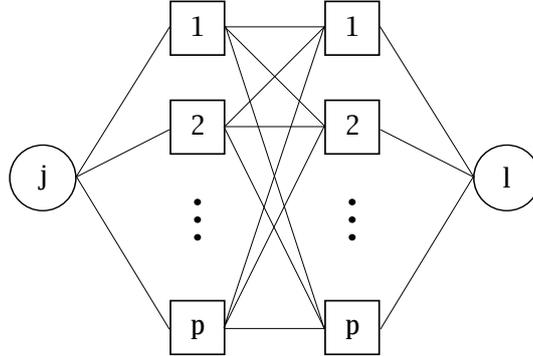


Figura 4.2: O conjunto de conexões entre o ponto j e o ponto l

O custo unitário associado a uma conexão geral (j, a, b, l) é igual à soma das três distâncias euclidianas elementares ponderadas pelos seus correspondentes fatores de custos:

$$z_{jabl} = \chi \|s_j - x_a\|_2 + \alpha \|x_a - x_b\|_2 + \delta \|x_b - s_l\|_2, \quad (4.4)$$

onde $\chi \geq 0$ é o custo de coleta, $\alpha \geq 0$ é o fator de redução provido pela conexão entre hubs, normalmente $0 \leq \alpha \leq 1$, e $\delta \geq 0$ é o custo de entrega.

Como há m^2 pares de cidades origem-destino (lembrando que pode haver tráfego de um ponto de origem-destino para si mesmo), há um total de $m^2 p^2$ possíveis conexões. O custo unitário do ponto de origem j ao ponto de destino l é escolhido como o valor mínimo do custo unitário para todas as conexões possíveis entre os dois pontos:

$$z_{jl} = \min_{a,b=1,\dots,p} z_{jabl}, \quad (4.5)$$

ou

$$z_{jl} = \min_{a,b=1,\dots,p} \{ \chi \|s_j - x_a\|_2 + \alpha \|x_a - x_b\|_2 + \delta \|x_b - s_l\|_2 \}. \quad (4.6)$$

Para o caso particular em que o valor $\alpha = 0$ e para valores $\chi = \delta = 1$, as conexões com valores mínimos serão aquelas formadas pelas ligações da cidade de origem ao hub mais próximo e idem para a cidade de destino, o que leva a soluções idênticas às do problema de Fermat-Weber (4.2).

No caso geral, o problema dos hubs, tratado de forma similar ao de p -medianas contínuo, corresponde a minimizar o custo total de todas as ligações entre os pares

de cidades, considerando o valor do custo unitário, dado para todas as conexões, resultando na formulação:

$$\text{minimizar } \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl} \quad (4.7)$$

$$\text{sujeito a } z_{jl} = \min_{a,b=1,\dots,p} z_{jabl}, \quad j, l = 1, \dots, m; \quad (4.8)$$

$$\text{onde } z_{jabl} = \chi \|s_j - x_a\|_2 + \alpha \|x_a - x_b\|_2 + \delta \|x_b - s_l\|_2.$$

Problema de Localização de Hubs do Tipo p -Medianas Discreto

O problema de localização de hubs do tipo p -Medianas discreto é muito semelhante ao problema contínuo equivalente. Também aqui, a única diferença é que cada centróide deve ser coincidente com uma das m localidades ou cidades especificadas, o que é assegurado através do uso da família de variáveis booleanas adicional v , e com a seguinte formulação:

$$\text{minimizar } \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl} \quad (4.9)$$

$$\text{sujeito a } z_{jl} = \min_{a,b=1,\dots,p} z_{jabl}, \quad j, l = 1, \dots, m;$$

$$\text{onde } z_{jabl} = \chi \|s_j - x_a\|_2 + \alpha \|x_a - x_b\|_2 + \delta \|x_b - s_l\|_2;$$

$$x_i = \sum_{j=1}^m v_{ij} s_j, \quad i = 1, \dots, p;$$

$$\sum_{j=1}^m v_{ij} = 1, \quad i = 1, \dots, p;$$

$$v_{ij} \in \{0, 1\}, \quad i = 1, \dots, p, \quad j = 1, \dots, m.$$

Capítulo 5

Suavização do Problema de Localização de Hubs do Tipo p -Medianas Contínuo

Como já visto, existem diversas formulações para o Problema de Localização de Hubs do Tipo p -Medianas. Além da distinção já abordada entre as versões contínua e discreta do problema, uma das principais características que definem os problemas é o tipo de alocação de hubs que é utilizado. Esta alocação pode ser única ou múltipla. No caso da alocação única, a cada localidade é atribuído um único hub, de tal forma que todas as ligações dessa localidade com as demais passam sempre através desse hub. Já na alocação múltipla é permitido que mais de um hub se ligue a uma dada localidade, resultando que as trajetórias que a unem a outros hubs podem passar por diferentes hubs, dependendo da outra localidade à qual se quer ligar a primeira. O número de hubs que se liga a cada localidade na alocação múltipla pode variar de 1 a p , onde p é o número de hubs desejado.

Existem ainda casos intermediários tratados por alguns autores, como Marti et al (2014), que não serão tratados aqui e nos quais se estabelece um limite (menor que p) ao número de hubs que podem ser conectados a cada ponto O-D. Este estudo está restrito à alocação múltipla, embora uma abordagem bastante semelhante à adotada pudesse também ser utilizada para o caso da alocação única e os intermediários.

Os problemas de localização de hubs também podem ter capacidades máximas atribuídas aos hubs, resultando no problema de hubs com capacidade (*capacitated hubs*). Estuda-se aqui apenas o caso em que não existe essa limitação e se pressupõe que qualquer hub pode assumir a capacidade necessária para lidar com todas as ligações que possam vir a passar por ele (*uncapacitated hubs*).

Finalmente, a topologia das ligações entre os hubs e as localidades pode ter

algumas variações em relação à topologia de grafo completo que é estudada aqui, em que todos os hubs podem se ligar a todas as localidades. Dentre as possibilidades mais comuns de grafos incompletos, costumam aparecer na literatura topologias tais como as de anel ou de árvore, além de diversas outras. Existem também topologias multicast, em que existem diversos níveis de hubs, sendo que cada nível superior encara os hubs do nível imediatamente abaixo como representando as localidades que se pretende ligar. Muitas dessas topologias podem ser tratadas por variantes da metodologia exposta neste estudo, com algumas adaptações, mas neste estudo são tratados apenas os grafos completos de um só nível. Nesse enfoque, o Problema de Localização de Hubs do Tipo p -Medianas Contínuo corresponde a minimizar o custo total entre todos os pares de cidades, considerando o valor do custo unitário, dado para todas as conexões.

$$\begin{aligned}
& \text{minimizar} && \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl} && (5.1) \\
\text{sujeito a} && z_{jl} = \min_{a,b=1,\dots,p} z_{jabl}, && j, l = 1, \dots, m; \\
\text{onde} && z_{jabl} = \chi \|s_j - x_a\|_2 + \alpha \|x_a - x_b\|_2 + \delta \|x_b - s_l\|_2.
\end{aligned}$$

Por sua formulação, este problema *hub-and-spoke* tem uma estrutura chamada na literatura por *min – sum – min*, com as características de ser um problema não-diferenciável e não-convexo, e tendo uma infinidade de mínimos locais, conforme Rubinov (2006). Uma série de transformações serão realizadas a fim de se chegar a uma formulação sucedânea completamente diferenciável.

Primeiro, considerando a sua definição, cada z_{jl} deve necessariamente satisfazer o seguinte conjunto de desigualdades:

$$z_{jl} - z_{jabl} \leq 0, \quad a, b = 1, \dots, p. \quad (5.2)$$

É feito um relaxamento, substituindo as restrições de igualdade originais do problema (5.2) pelas desigualdades acima, obtendo-se o seguinte problema:

$$\begin{aligned}
& \text{minimizar} && \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl} && (5.3) \\
\text{sujeito a} && z_{jl} - z_{jabl} \leq 0, && a, b = 1, \dots, p; \quad j, l = 1, \dots, m.
\end{aligned}$$

As variáveis z_{jl} , na função objetivo do problema (5.1) devem ser não-negativas, uma vez que elas são iguais à soma de normas euclidianas e os parâmetros $\chi, \alpha, \delta \geq 0$. No entanto, pelas restrições de desigualdade dessa formulação, as variáveis z_{jl} não estão mais sujeitas a limites inferiores e, por conseguinte, poderão decrescer indefinidamente. Dessa forma, a função objetivo é ilimitada inferiormente e, portanto, não terá mínimo. A fim de se obter a equivalência desejada com o problema original, deve-se modificar o problema (5.3) para criar um limite inferior adequado para as variáveis z_{jl} . Isso é feito, primeiro, através da introdução da função auxiliar $\psi(y)$ que denota $\max\{0, y\}$. Observa-se que a partir do conjunto de desigualdades em (5.3), segue que para quaisquer j, a, b, l tem-se $\psi(z_{jl} - z_{jabl}) = 0$.

Assim, usando uma forma compactada, são observadas as igualdades:

$$\sum_{a=1}^p \sum_{b=1}^p \psi(z_{jl} - z_{jabl}) = 0, \quad j, l = 1, \dots, m. \quad (5.4)$$

Para um par fixo de localidades diferentes j e l , admitindo uma sequência estritamente crescente das p^2 distâncias ponderadas z_{jabl} , denotadas por $d_1 < d_2 < \dots < d_{p^2}$, a Figura 5.1 ilustra as três primeiras parcelas do somatório (5.4) como função de z_{jl} .

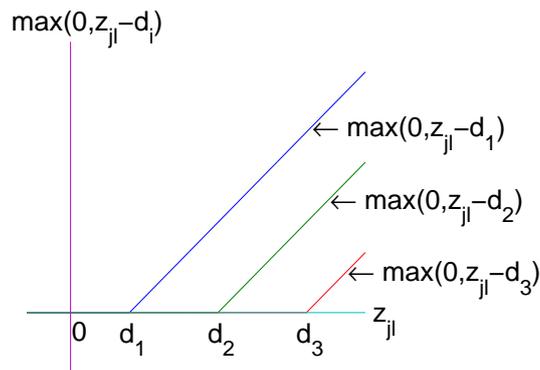


Figura 5.1: Parcelas do somatório em (5.4)

Usando (5.4) no lugar do conjunto de restrições de desigualdade em (5.3), seria obtido um problema equivalente, mantendo a propriedade indesejável de que z_{jl} , $j, l = 1, \dots, m$ ainda não tem um limite inferior. Considerando, no entanto, que a função-objetivo do problema (5.3) irá forçar para baixo cada z_{jl} , $j, l = 1, \dots, m$, pode-se colocar um limite inferior para cada uma dessas variáveis, através da inclusão de uma perturbação $\varepsilon > 0$ em cada restrição (5.4). Assim, se obtém o seguinte problema modificado:

$$\begin{aligned} & \text{minimizar} \quad \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl} & (5.5) \\ \text{sujeito a} \quad & \sum_{a=1}^p \sum_{b=1}^p \psi(z_{jl} - z_{jabl}) \geq \varepsilon, \quad j, l = 1, \dots, m. \end{aligned}$$

Uma vez que o conjunto viável do problema (5.1) é o limite do problema (5.5) acima quando $\varepsilon \rightarrow 0_+$, pode-se considerar a resolução de (5.1) através da solução de uma sequência de problemas como (5.5), com uma sequência decrescente de valores de ε , se aproximando de zero.

De outro lado, a definição da função ψ faz com que o problema (5.5) tenha uma estrutura não-diferenciável extremamente rígida, o que torna sua solução computacional extremamente difícil. Em vista disso, o método numérico que se adota para resolver o problema tem uma abordagem de suavização. A partir dessa perspectiva, define-se a função:

$$\phi(y, \tau) = \left(y + \sqrt{y^2 + \tau^2} \right) / 2, \quad (5.6)$$

para $y \in \mathbb{R}$ e $\tau > 0$.

A função ϕ tem as seguintes propriedades:

$$(a) \quad \phi(y, \tau) > \psi(y), \quad \forall \tau > 0;$$

$$(b) \quad \lim_{\tau \rightarrow 0} \phi(y, \tau) = \psi(y);$$

$$(c) \quad \phi(y, \tau) \text{ é uma função convexa crescente } C^\infty \text{ da variável } y.$$

Por essas propriedades, a função ϕ constitui uma aproximação da função ψ . Adotando as mesmas premissas usadas na Figura 5.1, as três primeiras parcelas do somatório na equação (5.4) e suas aproximações suavizadas correspondentes, dadas por (5.6), são representadas pelas curvas na Figura 5.2.

Usando a função ϕ no lugar da função ψ in (5.5), chega-se ao problema:

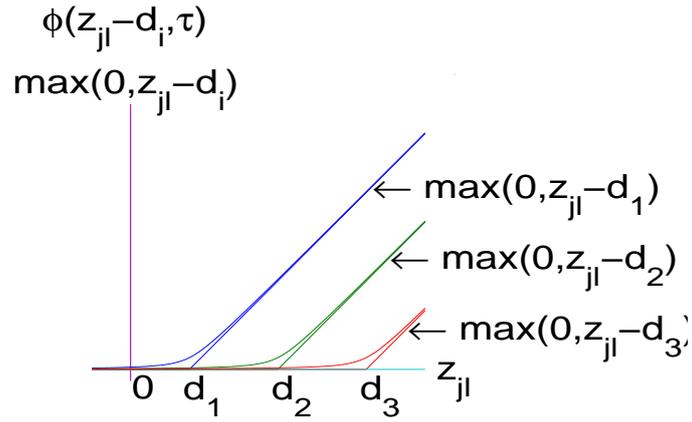


Figura 5.2: Parcelas originais e suavizadas do somatório em (5.4)

$$\begin{aligned} & \text{minimizar} \quad \sum_{j=1}^m \sum_{l=1}^m w_{jl} \tilde{z}_{jl} & (5.7) \\ \text{sujeito a} \quad & \sum_{a=1}^p \sum_{b=1}^p \phi(z_{jl} - z_{jabl}, \tau) \geq \varepsilon, \quad j, l = 1, \dots, m. \end{aligned}$$

Para obter um problema diferenciável, é necessário ainda suavizar as distâncias ponderadas z_{jabl} . Para este fim, é definida a função

$$\theta(u, v, \gamma) = \sqrt{(u^1 - v^1)^2 + (u^2 - v^2)^2 + \gamma^2}, \quad (5.8)$$

onde $u, v \in \mathbb{R}^2$ e $\gamma > 0$.

A função θ tem as seguintes propriedades:

- (a) $\lim_{\gamma \rightarrow 0} \theta(u, v, \gamma) = \|u - v\|_2$;
- (b) θ é uma função C^∞ suavizada da distância euclideana entre os pontos u e v .

Fazendo uso de θ no lugar das distâncias euclidianas, se obtém o seguinte problema completamente diferenciável:

Por essas propriedades, a função θ é uma aproximação

$$\begin{aligned} & \text{minimizar} \quad \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl} & (5.9) \\ \text{sujeito a} \quad & \sum_{a=1}^p \sum_{b=1}^p \phi(z_{jl} - (\chi \theta(s_j, x_a, \gamma) + \\ & \alpha \theta(x_a, x_b, \gamma) + \delta \theta(x_b, s_l, \gamma)), \tau) \geq \varepsilon, \quad j, l = 1, \dots, m. \end{aligned}$$

Agora, as propriedades das funções ϕ e θ permitem buscar uma solução para o problema (5.5) através da resolução de uma sequência de subproblemas como o problema (5.9), produzida através da diminuição dos parâmetros $\gamma \rightarrow 0$, $\tau \rightarrow 0$, $\varepsilon \rightarrow 0$.

O processo de minimização da função objetivo vai trabalhar no sentido de reduzir ao máximo os valores de z_{jl} , $j, l = 1, \dots, m$. Por outro lado, dado qualquer conjunto de hubs x_i , $i = 1, \dots, p$, devido à propriedade (c) da função de suavização hiperbólica ϕ , as restrições do problema (5.9) são funções monotonicamente crescentes de z_{jl} . Conseqüentemente, essas restrições estarão certamente ativas e o problema (5.9) acabará por ser equivalente ao problema:

$$\begin{aligned} & \text{minimizar} \quad \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl} & (5.10) \\ \text{sujeito a} \quad & h_{jl}(z_{jl}, x) = \sum_{a=1}^p \sum_{b=1}^p \phi(z_{jl} - (\chi \theta(s_j, x_a, \gamma) + \\ & \alpha \theta(x_a, x_b, \gamma) + \delta \theta(x_b, s_l, \gamma)), \tau) - \varepsilon = 0, \quad j, l = 1, \dots, m. \end{aligned}$$

Analisando o problema (5.10), vê-se que a dimensão do espaço de domínio das variáveis é igual a $(2p + m^2)$, sendo $2p$ variáveis associadas a x , correspondentes às posições dos hubs, e m^2 variáveis associadas a z , correspondentes aos diferentes pares de cidades. Uma vez que, em geral, é grande o valor do parâmetro m , a cardinalidade do conjunto S das localidades ou cidades s_j , o problema (5.10), em geral, possui um grande número de variáveis.

Analisando novamente o problema (5.10), vê-se que ele tem uma estrutura claramente separável, porque cada variável z_{jl} aparece apenas em uma restrição de igualdade. Ademais, como a derivada parcial de $h_{jl}(z_{jl}, x)$ com respeito a z_{jl} , $j, l = 1, \dots, m$ não é igual a zero, é possível usar o Teorema da Função Implícita para calcular cada componente z_{jl} , $j, l = 1, \dots, m$ como uma função

das variáveis de localização dos hubs $x_i, i = 1, \dots, p$. Desta maneira, se obtém o problema sem restrições:

$$\text{minimizar } f(x) = \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl}(x) \quad (5.11)$$

onde cada $z_{jl}(x)$ resulta do cálculo de um zero de cada equação

$$h_{jl}(z_{jl}, x) = \sum_{a=1}^p \sum_{b=1}^p \phi(z_{jl} - (\chi \theta(s_j, x_a, \gamma) + \alpha \theta(x_a, x_b, \gamma) + \delta \theta(x_b, s_l, \gamma)), \tau) - \varepsilon = 0, \quad j, l = 1, \dots, m. \quad (5.12)$$

Devido à propriedade (c) da função de suavização hiperbólica (5.6), cada termo ϕ acima é estritamente crescente com a variável z_{jl} e, portanto, a equação (5.12) pode possuir no máximo um único zero ou raiz, conforme ilustra a Figura 5.3, que mostra os somatórios das parcelas não diferenciáveis das restrições do problema (5.5) e das correspondentes parcelas diferenciáveis do problema (5.10). Nessa figura, como antes, é assumida uma sequência crescente de distâncias z_{jabl} , denotadas por $d_1 < d_2 < \dots < d_p$.

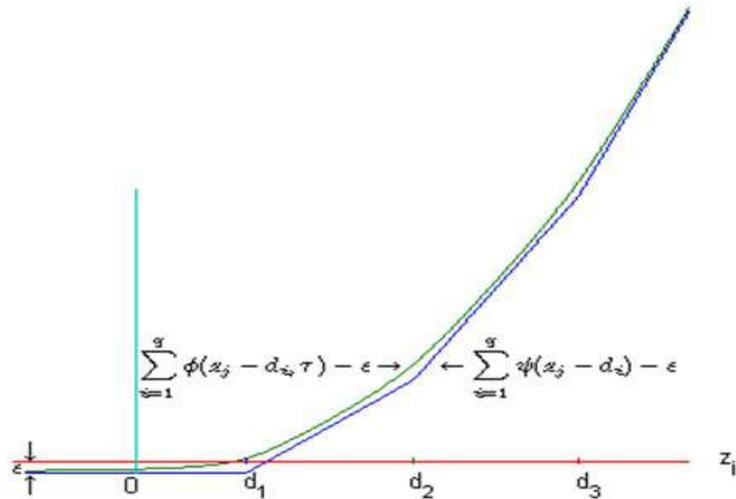


Figura 5.3: Somatório dos termos da equação (5.12)

Novamente, devido ao Teorema da Função Implícita, as funções $z_{jl}(x)$ têm todas as derivadas com relação às variáveis $x_i, i = 1, \dots, p$, e, portanto, é possível

calcular o gradiente da função objetivo do problema (5.11), segundo uma forma absolutamente exata:

$$\nabla f(x) = \sum_{j=1}^m \sum_{l=1}^m w_{jl} \nabla z_{jl}(x), \quad (5.13)$$

onde

$$\nabla z_{jl}(x) = - \nabla h_{jl}(z_{jl}, x) / \frac{\partial h_{jl}(z_{jl}, x)}{\partial z_{jl}}, \quad (5.14)$$

enquanto que $\nabla h_{jl}(z_{jl}, x)$ e $\partial h_{jl}(z_{jl}, x)/\partial z_{jl}$ podem ser obtidas diretamente a partir das equações (5.6), (5.8) e (5.12).

Por partes, em primeiro lugar, usando (5.12), pode-se obter as duas componentes de (5.14):

$$\frac{\partial h_{jl}(z_{jl}, x)}{\partial z_{jl}} = \sum_{a=1}^p \sum_{b=1}^p \phi'(z_{jl} - (\chi \theta(s_j, x_a, \gamma) + \alpha \theta(x_a, x_b, \gamma) + \delta \theta(x_b, s_l, \gamma)), \tau) \quad (5.15)$$

e

$$\nabla h_{jl}(z_{jl}, x) = \sum_{a=1}^p \sum_{b=1}^p \phi'(z_{jl} - (\chi \theta(s_j, x_a, \gamma) + \alpha \theta(x_a, x_b, \gamma) + \delta \theta(x_b, s_l, \gamma)), \tau) (\chi \nabla \theta(s_j, x_a, \gamma) + \alpha \nabla \theta(x_a, x_b, \gamma) + \delta \nabla \theta(x_b, s_l, \gamma)). \quad (5.16)$$

Em seguida, usando (5.6), obtém-se

$$\phi'(y, \tau) = \left(1 + y / \sqrt{y^2 + \tau^2}\right) / 2. \quad (5.17)$$

Finalmente, utilizando (5.8) chega-se a:

$$\nabla \theta(s_j, x_a, \gamma) = \frac{1}{(\sum_{q=1}^2 (x_a^q - s_j^q)^2 + \tau^2)^{\frac{1}{2}}} (0, \dots, 0, x_a^1 - s_j^1, x_a^2 - s_j^2, 0, \dots, 0)^T \quad (5.18)$$

$$\nabla\theta(x_a, x_b, \gamma) = \frac{1}{(\sum_{q=1}^2 (x_b^q - x_a^q)^2 + \tau^2)^{\frac{1}{2}}} (0, \dots, 0, -(x_b^1 - x_a^1), -(x_b^2 - x_a^2), 0, \dots, 0, x_b^1 - x_a^1, x_b^2 - x_a^2, 0, \dots, 0)^T \quad (5.19)$$

$$\nabla\theta(x_b, s_l, \gamma) = \frac{1}{(\sum_{q=1}^2 (s_l^q - x_b^q)^2 + \tau^2)^{\frac{1}{2}}} (0, \dots, 0, s_l^1 - x_b^1, s_l^2 - x_b^2, 0, \dots, 0)^T \quad (5.20)$$

A abordagem acima, baseada essencialmente no uso do Teorema da Função Implícita, emprega o princípio de redução do espaço solução inicial do problema a um espaço de menor dimensão estritamente viável. Dessa forma, entre outros exemplos, se adota a mesma ideia básica de Abadie e Carpentier (1969) para o desenvolvimento do algoritmo geral de gradiente reduzido, destinado à solução do problema de programação não-linear geral sujeito a restrições de igualdade.

Através desse esquema de redução, deve-se enfatizar que o problema (5.11) está definido num espaço que tem somente $2p$ dimensões, correspondentes às variáveis x da localização dos hubs, pois as variáveis z relativas às distâncias entre pares de cidades foram totalmente eliminadas. Ao final, obtém-se um problema de pequeno porte, pois o número de hubs p , nas aplicações do mundo real, é, em geral, pequeno.

Como vantagem adicional, deve ser registrado que, por sua propriedade de completa diferenciabilidade, pode-se resolver o problema (5.11) utilizando qualquer um dos métodos baseados em informações das derivadas de primeira ordem ou de ordens superiores, que, como sobejamente consignado na literatura de programação não-linear, vide por exemplo Luenberger (1984) e Minoux (1986), têm propriedades de robustez e de velocidade superiores.

Em relação à formulação original do problema de hub (5.1), a Figura 5.4 mostra o resultado do conjunto de transformações efetuadas pela metodologia da suavização hiperbólica.

A solução do problema original de hub (5.1), do tipo p -medianas contínuo, pode assim ser obtida usando o algoritmo de Suavização Hiperbólica Hub-and-Spoke (HSHS), descrito abaixo de forma simplificada.

Algorithm 1 Algoritmo HSHS Simplificado

1: Inicialização

- Escolha os valores iniciais: $x^0, \gamma^1, \tau^1, \varepsilon^1$.
- Escolha os valores: $0 < \rho_1 < 1, 0 < \rho_2 < 1, 0 < \rho_3 < 1$;
- Faça $k = 1$.

2: Passo principal Repita até que uma regra de parada seja atingida

- Resolva o problema de otimização irrestrita (5.11) com $\gamma = \gamma^k, \tau = \tau^k$ e $\varepsilon = \varepsilon^k$, a partir do ponto inicial x^{k-1} e seja x^k a solução obtida.
 - Procedimentos de atualização:
 - Faça $\gamma^{k+1} = \rho_1 \gamma^k, \tau^{k+1} = \rho_2 \tau^k, \varepsilon^{k+1} = \rho_3 \varepsilon^k$.
 - Faça $k = k + 1$
-

Assim como em outros métodos de suavização, a solução para o problema de hubs p -medianas é obtida, em teoria, através da solução de uma sequência infinita de problemas de otimização. No algoritmo HSHS, cada problema minimizado é sem restrições e de dimensão reduzida.

Observe-se que o algoritmo faz com que τ e γ se aproximem de zero, fazendo com que as restrições dos subproblemas solucionados no Passo Principal, dadas como em (5.11), tendam àquelas de (5.5). Além disso, o algoritmo faz com que ε se aproxime de zero, de tal forma que, num movimento simultâneo, o problema (5.5) gradualmente se aproxima do problema (5.1).

Adicionalmente às já relatadas vantagens de completa diferenciabilidade e de pequena dimensionalidade dos problemas resolvidos no Passo Principal do Algoritmo HSHS, acima apresentado, deve ser registrado que a utilização da suavização hiperbólica tem o poder, empiricamente observado, de eliminar expressiva quantidade de pequenos mínimos locais, como sugerido pela figura (5.4), deixando grandes regiões ou subespaços com um ponto mínimo dominante. Fazendo uso de uma analogia, essas regiões seriam como grandes bacias hidrográficas. Para o problema de geometria de distâncias é formalmente demonstrada a perfeita convexificação do problema modificado pela suavização hiperbólica, conforme apresentado em Xavier (2003), Souza (2010) e Souza et al (2011).

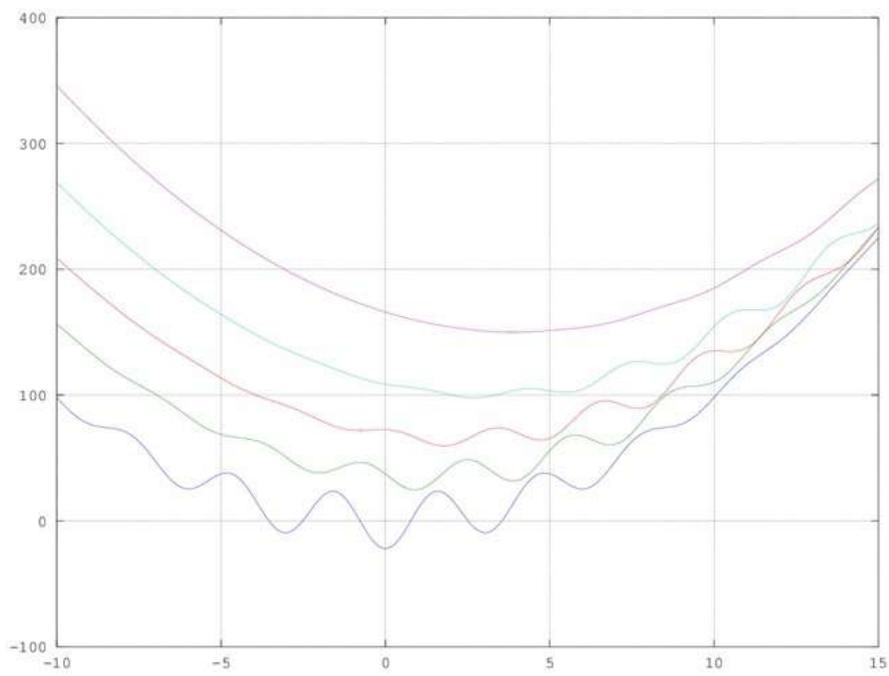


Figura 5.4: Idealização da Eliminação de Mínimos Locais

Capítulo 6

Resultados Computacionais: Problema de Localização Hubs do Tipo p -Medianas Contínuo Simétrico

Os experimentos numéricos deste capítulo foram realizados num PC Intel Celeron com uma CPU de 2,7GHz e RAM de 512MB. Os programas foram codificados com o Compaq Visual Fortran, Versão 6.1. As rotinas de minimização foram feitas através de um algoritmo Quasi-Newton, utilizando a fórmula de atualização BFGS da Biblioteca Harwell, disponível em <http://www.hsl.rl.ac.uk/>.

Neste capítulo será considerado, em particular, o caso simétrico em que $\chi = \delta = 1$, para o qual a formulação assume a forma mais simplificada.

Os resultados computacionais apresentados a seguir foram obtidos a partir de uma primeira implementação do algoritmo, sem qualquer procedimento de poda especial (como será proposto adiante) e onde os hubs iniciais de partida $x_i^0, i = 1, \dots, p$ foram dispostos, de uma forma muito simples, em torno do centro de gravidade do conjunto de cidades, fazendo perturbações aleatórias proporcionais ao desvio padrão deste conjunto, segundo as fórmulas: $\bar{s} = \sum_{j=1}^m s_j / m$, $\sigma = (\sum_{j=1}^m \|s_j - \bar{s}\|_2^2 / m)^{1/2}$, $x_i = \bar{s} + a\sigma$, onde os componentes do vetor a são variáveis aleatórias uniformes no intervalo $[-0,5, +0,5]$.

O valor do primeiro parâmetro de suavização foi tomado sempre como:

$$\tau^1 = 1/100\sigma.$$

As seguintes escolhas foram feitas para os outros parâmetros, depois de uma calibragem:

$$\varepsilon^1 = 4\tau^1, \quad \gamma^1 = \tau^1/100, \quad \rho_1 = 1/4, \quad \rho_2 = 1/4 \quad \text{e} \quad \rho_3 = 1/4.$$

Para ilustrar o desempenho do algoritmo proposto, mostra-se abaixo os resultados obtidos na resolução de cinco problemas teste padrões da literatura. A ideia aqui é usar exemplos que podem ser facilmente reproduzidos para possíveis comparações. As instâncias usadas correspondem aapêndiceos seguintes problemas:

- Cidades Alemãs (German Towns), que usa as duas coordenadas cartesianas de 59 cidades, originalmente apresentado por (Späth, 1980);

- AP200, que é um dos maiores conjuntos padrão de dados de localização de hubs, apresentado originalmente por (Ernst & Krishnamoorthy, 1996) e reproduzido na OR Library (biblioteca de pesquisa operacional) - (Beasley, 1990), acessível através do site:

<http://people.brunel.ac.uk/~mastjjb/jeb/orlib/phubinfo.html>

- rd400, d657 and dsj1000, que são instâncias que usam dados originais para o problema do caixeiro viajante apresentado por (Reinelt, 1991) e conhecido como TSPLIB. O complemento numérico do nome de cada instância corresponde ao seu número de cidades. Os 3 conjuntos de dados estão no site:

<http://www.iwr.uni-heidelberg.de/groups/comopt/software/TSPLIB95/>

Para os experimentos computacionais neste capítulo, considerou-se o caso particular do problema de localização de hubs simétrico, com $\chi = \delta = 1$, com uma matriz de demanda simétrica, $w_{jl} = w_{lj}$ para $j, l = 1, \dots, m$ quando $j \neq l$, e com os elementos da diagonal iguais a zero, $w_{jj} = 0$ para $j = 1, \dots, m$. Dada a propriedade de simetria, basta considerar os termos abaixo da diagonal principal.

Para a instância AP200, considera-se os dados de demanda originais reproduzidos por (Beasley, 1990), mas tendo o valor do elemento w_{jl} como a soma dos elementos originais w_{jl} e w_{lj} . Para as demais instâncias, é suposta uma demanda fixa unitária $w_{jl} = 1$, $j, l = 1, \dots, m$ quando $j \neq l$, para todos os pares de cidades de origem e destino. Assim, o problema (5.11) assume uma formulação mais simplificada para esse caso simétrico:

$$\text{minimizar } f(x) = \sum_{j=1}^{m-1} \sum_{l=j+1}^m w_{jl} z_{jl}(x). \quad (6.1)$$

onde cada $z_{jl}(x)$ resulta do cálculo do zero de cada equação

$$h_{jl}(z_{jl}, x) = \sum_{a=1}^p \sum_{b=1}^p \phi(z_{jl} - (\theta(s_j, x_a, \gamma) + \alpha \theta(x_a, x_b, \gamma) + \theta(x_b, s_l, \gamma)), \tau) - \varepsilon = 0, \quad j = 1, \dots, m-1, \quad l = j+1, \dots, m. \quad (6.2)$$

Para o cálculo do gradiente da função objetivo do problema (6.2), *mutatis mutandis*, pode se usar a expressão (5.13), com idênticas alterações nos índices dos somatórios. A diferença essencial se resume ao problema assimétrico (5.11) ter aproximadamente o dobro de variáveis do problema simétrico (8.1). Por facilidade e por ser mais geral, o problema assimétrico será mantido como referencial até o final deste trabalho.

Os resultados computacionais obtidos são apresentados nas tabelas 6.1 a 6.8. Como o problema (5.11) é um problema de otimização global, usa-se uma estratégia multi-partida com $T = 10$ tentativas diferentes de pontos de partida, escolhidos aleatoriamente para cada instância. Para as tabelas 6.1 e 6.5 a 6.8, o parâmetro de desconto foi definido como $\alpha = 0,5$. A primeira coluna apresenta o número especificado de hubs (p). A segunda coluna apresenta o melhor valor da função objetivo (f_{HSHS}). As três colunas seguintes apresentam o número de ocorrências da melhor solução ($Ocorr$), o erro percentual médio das T soluções (E_{Med}) em relação à melhor solução obtida (f_{HSHS}) e o tempo médio de CPU em segundos (T_{Med}). Definindo f^t como o valor ótimo da função objetivo obtido a partir da tentativa de ponto de partida t , o erro percentual E_{Med} observado em todas as T tentativas é calculado pela expressão:

$$E_{Med} = \frac{100 \sum_{t=1}^T (f^t - f_{HSHS})}{T f_{HSHS}}. \quad (6.3)$$

O esforço necessário para a resolução de (5.11) depende, em primeiro lugar, da dimensão do problema, igual a $2p$. No entanto, cada avaliação da função objetivo envolve o cálculo de $O(m^2)$ distâncias, seja para a formulação completa (5.11) seja para a forma particular mais simplificada (6.2). Uma a uma, cada distância z_{jl} é obtida através do cálculo de um zero da equação (5.12), que tem p^2 termos. Assim, a avaliação da complexidade do cálculo de cada função é igual a $O(m^2 p^2)$. Se admitirmos que o número de iterações feitas na etapa principal do algoritmo HSHS é somente proporcional linearmente à dimensão do espaço de domínio da variável, a complexidade resultante de resolução de cada problema (5.11) no Passo Principal do algoritmo HSHS seria $O(m^2 p^3)$.

Resultados para o Problema Teste German Towns

A Tabela 6.1 apresenta os resultados computacionais obtidos para a instância das cidades alemãs, ou German Towns, de Späth, com 59 cidades. Para os casos de p menores, as ocorrências da melhor solução, dada pela coluna $Ocorr$ são relativamente mais comuns. O fato de que só uma ocorrência atinge a melhor solução para p maiores pode indicar que uma escolha mais criteriosa, não aleatória, dos

pontos de partida ou um número maior de tentativas poderia produzir resultados melhores da função objetivo. Os erros médios apresentados pelas 10 soluções, dados pela coluna E_{Med} , têm valores baixos, que indicam um desempenho consistente do algoritmo.

p	f_{HSHS}	$Ocorr$	E_{Med}	T_{Med}
2	171285	10	0,00	0,53
3	154629	6	0,01	2,48
4	139158	9	0,75	8,47
5	131453	4	2,03	18,23
6	126496	1	0,68	39,30
7	122636	2	0,73	76,28
8	119239	1	0,81	149,52
9	116583	1	1,25	246,99
10	113962	1	1,39	383,56

Tabela 6.1: Resultados para Cidades Alemãs ($\alpha = 0,5$)

α	f_{HSHS}	$Ocorr$	E_{Med}	T_{Med}
0,000	87574	9	0,01	4,72
0,125	102437	10	0,00	7,67
0,250	116102	8	1,83	7,19
0,375	128364	7	2,10	5,72
0,500	139158	9	0,75	8,47
0,625	148391	10	0,00	9,90
0,750	156190	5	0,64	9,39
0,875	161837	7	0,07	7,20
1,000	164388	8	0,28	9,94

Tabela 6.2: Resultados para Cidades Alemãs, com α variável e $p = 4$.

A tabela 6.2 mostra os resultados de nove soluções da instância das Cidades Alemãs associadas a diferentes valores do fator de redução α , a saber $\alpha = 0; 0,125; 0,25; 0,375; 0,5; 0,625; 0,75; 0,875$ e 1. Note-se que, para todos os casos, o número de hubs foi fixado em $p = 4$. As colunas $Ocorr$ e E_{Med} mostram a consistência do algoritmo. A coluna T_{Med} mostra um ligeiro aumento do tempo de CPU quando o valor do parâmetro de desconto α aumenta, embora oscilando muito. A Figura 6.2 ilustra graficamente os resultados correspondentes, mostrando o fenômeno do aumento expressivo no número de ligações entre cidades e hubs quando o fator α aumenta, o que é descrito em detalhe na Tabela 6.3.

A Tabela 6.3 mostra o comportamento das atribuições múltiplas das 59 Cidades Alemãs segundo a variação do parâmetro de desconto. A primeira coluna mostra

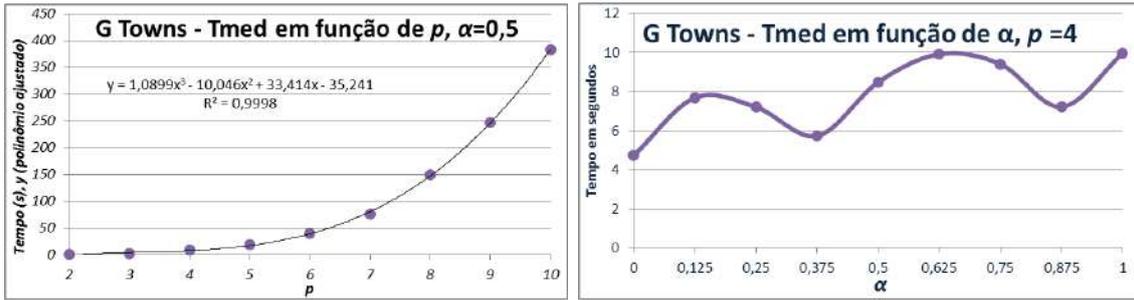
α	1 aloc	2 alocs	3 alocs	4 alocs	Arcos de Acesso
0,000	59	0	0	0	59
0,125	53	6	0	0	65
0,250	51	7	1	0	68
0,375	44	14	0	1	76
0,500	39	17	1	2	84
0,625	35	19	1	4	92
0,750	26	24	1	8	109
0,875	12	28	6	13	138
1,000	1	0	5	53	228

Tabela 6.3: Cidades Alemãs: alocações múltiplas como uma função de α ($p=4$)

o valor de α , as quatro colunas seguintes mostram o número de cidades, respectivamente, com 1, 2, 3 e 4 alocações, e a última coluna mostra o número total de arcos de acesso. Para $\alpha = 0$ o problema resume-se à soma das distâncias mínimas do problema de agrupamento (*clustering*), que é equivalente ao problema de localização de Fermat-Weber (Wesolowsky, 1993), cuja solução é caracterizada por uma única alocação para cada cidade. Não havendo custo de transferência inter-hubs, cada cidade se liga ao hub mais próximo, e as transferências se dão através de um ou dois hubs. Este problema também pode ser resolvido diretamente pela Suavização Hiperbólica, como mostrado por Xavier (2010) e Xavier e Xavier (2011).

O aumento do parâmetro de desconto para 0,125 já faz com que seis cidades sejam servidas por dois hubs. Com o fator α aumentado para 0,25 surge um caso de cidade servida por três hubs e no nível de 0,375 aparece uma servida por quatro. De qualquer modo, quando o valor de α é pequeno, as economias de escala na utilização de ligações inter-hubs inibem as alocações múltiplas. É interessante notar que, entre as 59 cidades, para $\alpha = 0,25$, apenas oito são servidas por mais de um hub, com uma única cidade sendo ligada a três hubs. Para $\alpha = 0,5$, vinte das 59 cidades são atendidas pela alocação múltipla de hubs, com duas dessas cidades sendo ligadas a todos os quatro hubs. À medida que α cresce, a tendência de aumento de alocações múltiplas se mantém, a tal ponto que para $\alpha = 0,875$ apenas 12 cidades são atendidas por um único hub. No outro extremo, para $\alpha = 1$, não há economia de escala em usar interligações com hubs diferentes. Assim na solução ótima, um único hub é usado para qualquer par de cidades origem-destino. Ademais, por esse mesmo efeito, quase todas as cidades são servidas por todos os hubs, como mostra a tabela 6.3, que também indica que o número de arcos é 97% do máximo possível.

A figura 6.1 acima ilustra o fato de que os tempos obtidos para esta instância evoluem de acordo com o número p de uma forma que se aproxima muito de um polinômio do terceiro grau, como esperado. Já a variação dos tempos em função de



(a) tempo versus p

(b) tempo versus α

Figura 6.1: German Towns, tempo médio de execução em função de p e α

α tem um comportamento oscilante, com leve tendência de crescimento à medida que α aumenta, como ilustrado pela mesma figura.

Apesar da alta sensibilidade do número de alocações associado a cada cidade à variação do parâmetro α , como mostrado pela Figura 6.2 e resumido na Tabela 6.3, a localização dos hubs sofre apenas pequenos deslocamentos, como ilustrado pela Figura 6.3. Nos experimentos das Cidades Alemãs, denotando por $x_i^*(\alpha)$, $i = 1, \dots, 4$ a localização ótima dos hubs para o problema (5.11) para diferentes valores do parâmetro α , os deslocamentos entre os locais dos hubs associados aos casos extremos $\alpha = 0$ e $\alpha = 1$, definido por $|x_i^*(1) - x_i^*(0)|$, $i = 1, \dots, 4$, assumem as proporções de 23,9%, 10,9%, 5,7% e 5,7% em relação ao desvio padrão σ , do conjunto das distâncias entre as cidades.

Deve-se registrar o fato de não ter sido possível encontrar quaisquer resultados computacionais para problemas contínuos, de qualquer dimensão, que permitissem a inclusão de uma comparação direta com resultados produzidos pela metodologia proposta da Suavização Hiperbólica, visando uma necessária validação científica.

Não obstante, foram achados dezenas de artigos mostrando resultados numéricos para problemas discretos, que apresentam formulações bem diferentes da contemplada nessa seção, seja na topologia de conexões, na especificação da função objetivo ou no conjunto de restrições.

Dado o clamor de uma validação do desempenho da metodologia proposta foi delineado um experimento alternativo para comprovar de forma irrefutável a qualidade dos resultados computacionais produzidos pelo algoritmo HSHS.

Assim, foi realizada uma comparação dos resultados obtidos para o problema teste German Towns (Cidades Alemãs), no espaço contínuo, com as soluções dos problemas discretos correspondentes. Para cada hub de uma solução contínua, pode-se encontrar a cidade mais próxima do mesmo e, depois, usar esse subconjunto resultante de cidades como uma aproximação para a solução do problema discreto. Na verdade, foram feitas algumas experiências no transcorrer desse estudo que mostram

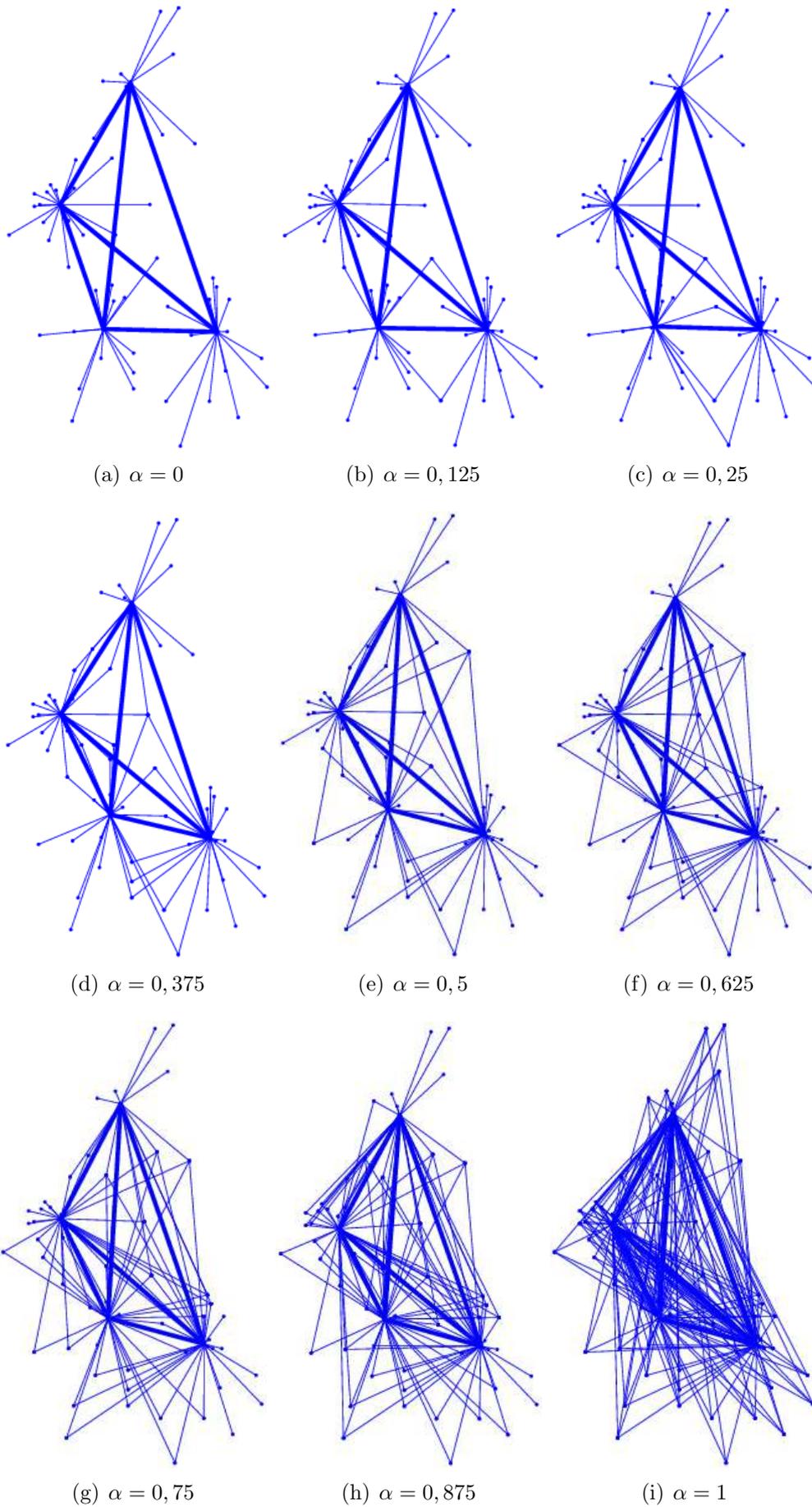


Figura 6.2: Cidades Alemãs, $p = 4$, para diversos valores de α

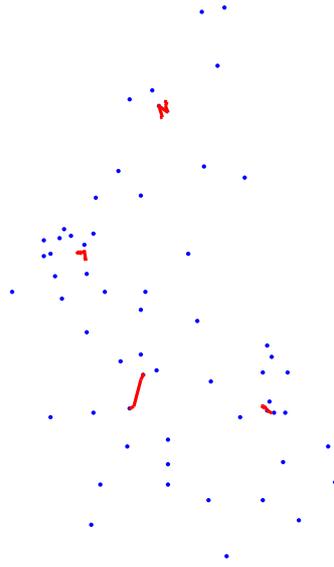


Figura 6.3: Cidades Alemãs, $p = 4$ - Trajetória dos hubs: $0 \leq \alpha \leq 1$

que, em vários casos, este processo nos conduz diretamente ao ótimo global do caso discreto, ou perto dele, como mostrado na tabela 6.4 abaixo.

p	f_{HSHS}	f_{Closer}	$f_{DiscrOpt}$	EC/D	T_{HSHS}	$T_{DiscrOpt}$
2	171285	173156	172968	0.11	0.55	1.29
3	154629	155244	154827	0.27	2.56	49.78
4	139158	139355	139355	0.00	8.48	1147.57
5	131453	131595	131595	0.00	18.06	18853.31

Tabela 6.4: Resultados para a instância das Cidades Alemãs (German Towns): solução contínua pelo algoritmo HSHS, solução discreta mais próxima e solução discreta ótima

A tabela 6.4 exhibe os valores obtidos para a instância das Cidades Alemãs (Späth), que possui 59 cidades. Os valores nas duas primeiras colunas correspondem aos da tabela 6.1 deste estudo. A primeira coluna indica o número de hubs, aqui no intervalo de 2 a 5; A segunda coluna mostra os melhores valores da função objetivo no caso contínuo, obtido através do algoritmo HSHS e retirado deste estudo; a terceira mostra valores equivalentes para o caso discreto que é obtido, como mencionado acima, ao se buscar as cidades mais próximas dos hubs do caso contínuo; a quarta mostra os valores globais ótimos para o caso discreto, obtidos através de uma busca do tipo "força bruta" ou "enumeração completa"; a quinta coluna mostra os desvios percentuais dos valores na coluna 3 a partir do ótimo na coluna 4; finalmente, a sexta e a sétima colunas mostram os tempos necessários para se encontrar as soluções nas colunas 2 e 4, respectivamente. Os valores na coluna 3 exigem um

tempo insignificante para serem computados, depois que se encontra os hubs para o caso contínuo, coluna 2. A quinta coluna mostra quão perto do ótimo global ficou a solução discreta obtida através do método aqui descrito.

Os números falam por si mesmos. Para os casos $p = 4$ e $p = 5$, encontra-se através do método aqui proposto exatamente a melhor solução possível para o caso discreto, gerada pelo procedimento de enumeração completa.

As colunas 6 e 7 mostram que as soluções obtidas através do método proposto neste estudo foram encontradas em tempos muito mais curtos do que a busca pelo mínimo global. Embora seja claro que o método da "força bruta" está longe de ser a maneira mais eficiente de encontrar o mínimo através da utilização de algoritmos discretos, o fato de que a diferença entre a velocidade dos dois métodos aumenta consideravelmente com p é uma indicação da eficiência de processamento do método aqui proposto, especialmente para casos com grande número de cidades. É preciso lembrar que não se fez comparações de resultados para valores maiores de p porque os tempos de processamento para o algoritmo "força bruta" era excessivo. Assim, como já foi dito, acredita-se que a abordagem de Suavização Hiperbólica realmente funciona e pode ser usada para resolver outros problemas de localização, como exemplificado acima, e, em particular, pode resolver problemas de localização de hubs com diversas topologias. Os capítulos 7 e 8 exploram em maior detalhe a utilização do método para resolver problemas discretos.

Resultados para Outros Problemas de Teste em Espaços Contínuos

Em seguida, apresenta-se alguns resultados adicionais para problemas apresentados na literatura. A destacar, o problema AP200, primeiro mostrado a seguir, é a maior instância já tratada na literatura de *hub-and-spoke* contínuo.

p	f_{HSHS}	$Ocorr$	E_{Med}	T_{Med}
2	74900	10	0,00	13,19
3	68593	1	1,06	47,85
4	63239	7	1,47	196,16
5	60742	2	0,45	453,52
6	58469	3	0,32	949,79
7	56547	1	0,61	1910,85
8	54659	1	0,56	3336,10

Tabela 6.5: Resultados da instância AP200 Simétrico ($\alpha = 0,5$)

A Tabela 6.5 apresenta os resultados computacionais obtidos para a instância AP200, um dos maiores conjuntos padrão de dados sobre a localização de hubs. O número máximo de hubs testado foi $p = 8$, quando o tempo de CPU chegou a

p	f_{HSHS}	$Ocorr$	E_{Med}	T_{Med}
2	0,535053E8	9	0,18	33,65
3	0,470813E8	10	0,00	111,54
4	0,434243E8	10	0,00	290,82
5	0,412475E8	1	0,50	750,83
6	0,395146E8	2	0,37	1236,87

Tabela 6.6: Resultados da instância rd400 do TSPLIB ($\alpha = 0,5$)

p	f_{HSHS}	$Ocorr$	E_{Med}	T_{Med}
2	0,349046E9	10	0,00	95,66
3	0,314874E9	10	0,00	326,54
4	0,291443E9	5	0,28	882,16
5	0,273784E9	9	0,10	2132,17
6	0,260834E8	10	0,00	3875,53

Tabela 6.7: Resultados da instância d657 do TSPLIB ($\alpha = 0,5$)

p	f_{HSHS}	$Ocorr$	E_{Med}	T_{Med}
2	0,342083E12	10	0,00	376,66
3	0,285747E12	10	0,00	1296,32
4	0,263992E12	9	0,07	3754,33
5	0,248652E12	4	0,35	8234,88
6	0,240642E12	-	-	17619,41

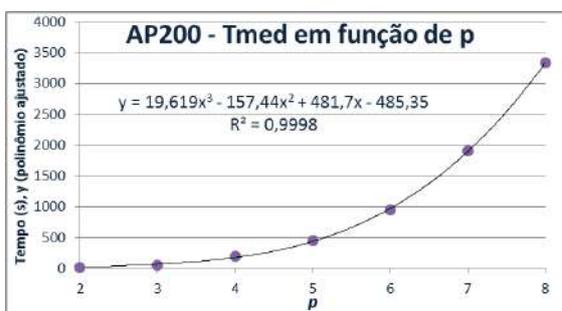
Tabela 6.8: Resultados da instância dsj1000 do TSPLIB ($\alpha = 0,5$)

quase uma hora. Os resultados computacionais produzidos pelo algoritmo HSHS na resolução deste caso têm características semelhantes ao da instância das Cidades Alemãs.

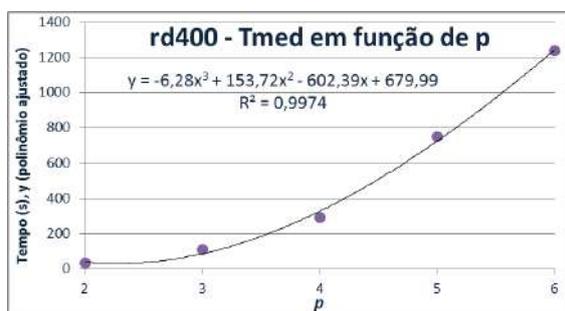
As tabelas 6.6, 6.7 e 6.8 apresentam os resultados computacionais para as instâncias rd400, d657 e dsj1000, provenientes do TSPLIB. Para esses casos maiores, um número máximo $p = 6$ de hubs foi considerado. Para o maior caso de todos, dsj1000, quando $p = 6$, apenas um ponto de partida inicial foi processado, porque o tempo de CPU é muito alto neste caso. Este resultado único mostra o efeito de m , o número de cidades, sobre a complexidade intrínseca do problema de medianas p -hub. Os resultados computacionais produzidos pelo algoritmo HSHS para essas três instâncias têm características semelhantes aos anteriores.

A partir dos resultados apresentados nas tabelas 6.1 a 6.8, vê-se que a primeira implementação do algoritmo HSHS para resolver instâncias grandes de *hub – and – spoke* contínuo confirma a consistência da metodologia proposta. O número expressivo de ocorrências da melhor solução para casos pequenos ($p \leq 4$) mostra o desempenho consistente do algoritmo. Os baixos valores do erro médio das 10 soluções (E_{Med}) em relação à melhor solução obtida mostram a consistência e a estabilidade numérica do algoritmo. Finalmente, os problemas de *hub-and-spoke* p -medianas foram convenientemente resolvidos em tempos de CPU adequados, como consequência tanto da dimensão baixa do problema não-linear (5.11), definido num espaço de dimensão $2p$, e do uso de um algoritmo de minimização que se aproveita da propriedade de diferenciabilidade C^∞ .

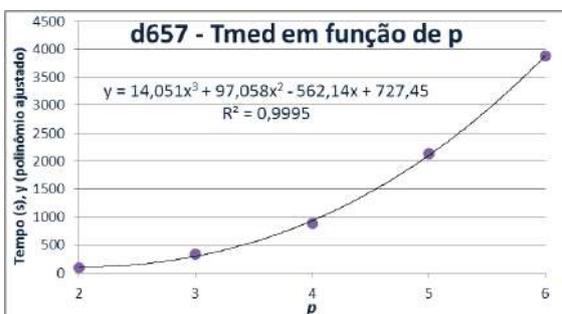
A figura 6.4 mostra a evolução do tempo em função de p nas quatro instâncias e demonstra, mediante o ajuste de curvas polinomiais aos pontos (como já tinha sido feito na figura 6.1 a), que o tempo de execução na prática varia realmente de acordo com p^3 , consoante os valores obtidos para os coeficientes de correlação múltipla R^2 , praticamente iguais a um, comprovando assim o acerto da estimativa de complexidade de $O(m^2p^3)$, feita na página 39.



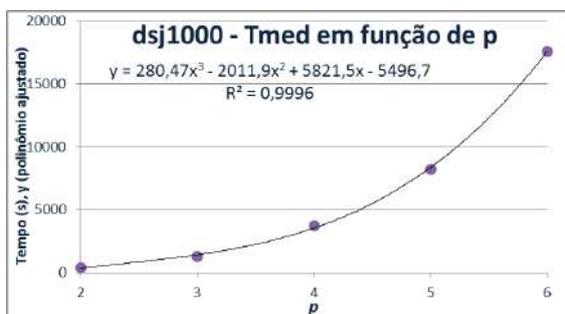
(a) Instância AP200



(b) Instância rd400



(c) Instância d657



(d) Instância dsj1000

Figura 6.4: Tempo médio de execução do algoritmo de localização de hubs no espaço contínuo em função de p , para quatro instâncias encontradas na literatura

Capítulo 7

Suavização do Problema de Localização de Hubs do Tipo p -Medianas Discreto

O Problema de Localização de Hubs do Tipo p -Medianas Discreto é muito similar ao Problema de Localização de Hubs do Tipo p -Medianas Contínuo. A única diferença é que cada centróide deve coincidir com uma das m localidades ou cidades especificadas.

$$\begin{aligned} & \text{minimizar} && \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl} && (7.1) \\ \text{sujeito a} & && z_{jl} = \min_{a,b=1,\dots,p} z_{jabl}, && j, l = 1, \dots, m; \\ \text{onde} & && z_{jabl} = \chi \|s_j - x_a\|_2 + \alpha \|x_a - x_b\|_2 + \delta \|x_b - s_l\|_2; \\ & && x_i = \sum_{j=1}^m v_{ij} s_j, && i = 1, \dots, p; \\ & && \sum_{j=1}^m v_{ij} = 1, && i = 1, \dots, p; \\ & && v_{ij} \in \{0, 1\}, && i = 1, \dots, p, j = 1, \dots, m. \end{aligned}$$

As duas primeiras restrições do problema acima são similares às do problema (5.1). As três últimas têm a finalidade de localizar os p hubs de forma que coincidam com posições de cidades.

Para a resolução do problema 7.1, a literatura tradicionalmente adota formulações baseadas em programação linear inteira, conforme descrito com precisão no survey de Alumur & Kara (2008).

Campbell (1992) foi o primeiro a formular o Problema de Localização de Hubs

do Tipo p -Medianas Discreto como um problema de programação linear inteira, denominando-o como UMAPHMP, nome hoje consagrado pela literatura.

$$\begin{aligned}
& \text{minimizar} && \sum_{j=1}^m \sum_{a=1}^m \sum_{b=1}^m \sum_{l=1}^m w_{jl} y_{jabl} z_{jabl} && (7.2) \\
\text{sujeito a} &&& z_{jabl} = \chi \|s_j - s_a\|_2 + \alpha \|s_a - s_b\|_2 + \delta \|s_b - s_l\|_2; \\
\text{onde} &&& \sum_{a=1}^m Y_{aa} = p \\
&&& Y_{aa} \in \{0, 1\}, \quad \forall a \\
&&& \sum_{a=1}^m \sum_{b=1}^m y_{jabl} = 1, \quad \forall j, l = 1, \dots, m; \\
&&& y_{jabl} \geq 0, \quad \forall j, a, b, l \\
&&& y_{jabl} \leq Y_{aa}, \quad \forall j, a, b, l && (7.3) \\
&&& y_{jabl} \leq Y_{bb}, \quad \forall j, a, b, l && (7.4)
\end{aligned}$$

A função objetivo do problema (7.2), como as demais consideradas neste trabalho, contempla a minimização do custo total de transporte. A primeira restrição calcula o custo de cada caminho $jabl$, estabelecendo, ademais, que todos quatro componentes coincidam com alguma das m cidades especificadas. As duas restrições seguintes garantem a existência de p hubs. As quatro últimas restrições garantem que as variáveis y_{jabl} sejam igualadas a zero ou a um desde que o fluxo entre o par de cidades (jl) seja roteado, ou não roteado, pelo par de hubs (ab) .

Campbell (1994-A) observa que na ausência de restrições sobre as capacidades dos arcos entre as cidades, há uma solução ótima na qual todas variáveis z_{jabl} são iguais a zero ou a um, desde que o fluxo total de cada par origem-destino deve ser roteado via o par de hubs menos custoso. A formulação acima possui $(m^4 + m)$ variáveis, das quais m binárias, com $(2m^4)$ restrições de desigualdade e com $(m^2 + 1)$ restrições de igualdade.

Skorin-Kapov et al (1996) adotam uma formulação linear inteira relaxada do problema 7.2 em que substituem as restrições 7.3 e 7.4 de desigualdade por formas agregadas:

$$\sum_{a=1}^m y_{jabl} = Y_{bb}, \quad \forall j, b, l; \quad (7.5)$$

$$\sum_{b=1}^m y_{jabl} = Y_{aa}, \quad \forall j, a, l; , \quad (7.6)$$

obtendo dessa maneira uma formulação relaxada simplificada com $(m^4 + m)$ variáveis, das quais m binárias, e com $(2m^3 + m^2 + 1)$ restrições de igualdade. Embora seja uma formulação com as restrições de integralidade relaxadas, quase todas as soluções computacionais obtidas para o dataset CAB ¹ produziram soluções inteiras.

Ernst & Krishnamoorthy (1998) propõem uma fórmula baseada em fluxos em que são estabelecidas as seguintes convenções: Z_{ja} como o fluxo do nó j para o hub a ; Y_{ab}^j como o fluxo emanado no nó j roteado pelo par de hubs ab ; X_{bl}^j como a parcela da demanda w_{jl} roteada pelo hub b ; H_a como a variável binária igual a 1 se o nó a for um hub; as demais variáveis são definidas como antes.

$$\text{minimizar } \sum_{j=1}^m \left[\sum_{a=1}^m \chi C_{ja} Z_{ja} + \sum_{a=1}^m \sum_{b=1}^m \alpha C_{ab} Y_{ab}^j \sum_{b=1}^m \sum_{l=1}^m \delta C_{bl} X_{bl}^j \right] \quad (7.7)$$

$$\begin{aligned} \text{sujeito a } & \sum_{a=1}^m H_a = p; \\ & H_a \in \{0, 1\}, \quad \forall a \\ & \sum_{a=1}^m Z_{ja} = \sum_{l=1}^m w_{jl}, \quad \forall j = 1, \dots, m; \\ & \sum_{b=1}^m X_{bl}^j = w_{jl}, \quad \forall j, l = 1, \dots, m; \\ & \sum_{b=1}^m Y_{ab}^j + \sum_{l=1}^m X_{al}^j - \sum_{b=1}^m Y_{ba}^j - Z_{ja} = 0 \quad \forall j, a = 1, \dots, m; \\ & Z_{ja} \leq \sum_{l=1}^m w_{jl} H_a, \quad \forall j, a = 1, \dots, m; \\ & \sum_{j=1}^m X_{bl}^j \leq \sum_{j=1}^m w_{jl} H_b \quad \forall b, l = 1, \dots, m; \\ & Y_{ab}^j, X_{bl}^j, Z_{ja} \geq 0 \quad \forall j, l, a, b = 1, \dots, m; \end{aligned}$$

¹Dados referentes a rotas aéreas nos EUA fornecidas pelo CAB (*Civil Aeronautics Board*) e reproduzidos na OR Library, já referida na página 38.

As duas primeiras restrições garantem a existência de p hubs. A restrição seguinte garante que todo fluxo saindo de um nó j seja igual à demanda total atendida por esse nó. A restrição seguinte garante que a demanda saindo de um nó j para um nó l seja igual ao total dos correspondentes fluxos passando por todos os hubs b da segunda camada. A restrição seguinte efetua um balanço de massa, pois garante que o fluxo originário do nó j entrando para um hub a , (duas primeiras parcelas na expressão) seja conservado (duas últimas parcelas na expressão). Finalmente, as duas últimas restrições exigem que, para qualquer fluxo passar em um hub, esse deve necessariamente existir.

A formulação de Ernst & Krishnamoorthy é de uma dimensão menor, pois possui um total de somente $(2m^3 + m^2 + m)$ variáveis, das quais m são de natureza binária, todavia ainda possui $(4m^2 + m + 1)$ restrições. Assim, essa formulação tem ainda uma dimensão expressiva.

Como registra Carvalho (2017), para a resolução de problemas de programação inteira, no âmbito de localização de hubs, existem três abordagens clássicas normalmente adotadas: algoritmos de planos de corte, abordagens enumerativas e técnicas de relaxamento e decomposição. De toda sorte, em todas as três formulações clássicas para o problema UMapHMP o número de variáveis e de restrições é altamente explosivo com o número de cidades m , dificultando ou até mesmo impedindo a resolução do problema por qualquer dessas três alternativas.

Resolução do Problema de Localização de Hubs do Tipo p -Medianas Discreto (UMApHMP)

A metodologia proposta neste trabalho para resolução do problema 7.1 é extremamente simples. O primeiro passo é constituído pela relaxação das três últimas restrições de integridade, que obrigam a localização dos hubs a se restringir às posições ocupadas pelas localidades de origem e destino. Dessa forma, chega-se à formulação 5.1, em que os hubs podem ser localizados em todo o plano. A seguir, aplicam-se exatamente os mesmos procedimentos de suavização hiperbólica descritos no capítulo 5. Através do uso do algoritmo HSHS obtém-se uma solução contínua x^* . Para a obtenção de uma solução discreta propõe-se a execução de um procedimento de busca local rápida no entorno da solução contínua x^* , previamente encontrada pela execução do HSHS.

Vale notar que, com frequência, as soluções encontradas no caso contínuo estão extremamente próximas de alguma das localidades. Por exemplo, na instância AP150, quando se procurou experimentalmente, através da HSHS, o melhor conjunto de hubs no espaço contínuo para p variando de 3 a 8 (como será visto no

próximo capítulo), em mais da metade dos casos o conjunto de hubs encontrados estavam praticamente juntos a localidades de origem-destino. Nestes casos, não há alternativa que faça mais sentido do que substituí-los pela localidade vizinha. De toda sorte, é necessário definir um critério de proximidade. Este critério de proximidade pode ser a distância euclidiana entre os hubs no espaço contínuo e as diversas cidades. Com base nessas ideias, foi desenvolvido o algoritmo para resolução de localização de hubs em espaços discretos).

AHSHS - Novo Algoritmo Mais Rápido e Robusto

Trata-se agora de resolver o problema (7.1). Uma ideia natural para resolver esse problema é, primeiro, obter de uma forma rápida um ponto de mínimo contínuo do problema relaxado através do uso do Algoritmo HSHS e, a seguir, efetuar um procedimento de busca local para achar um ponto discreto na vizinhança desse ponto. Com a relaxação das restrições de integridade do problema (7.1), chega-se à formulação:

$$\begin{aligned} \text{minimizar } f(x) &= \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl} & (7.8) \\ \text{sujeito a } z_{jl} &= \min_{a,b=1,\dots,p} z_{jabl}, \quad j, l = 1, \dots, m; \\ z_{jabl} &= \chi \|s_j - x_a\|_2 + \alpha \|x_a - x_b\|_2 + \delta \|x_b - s_l\|_2; \end{aligned}$$

O problema contínuo (7.8) é exatamente igual ao problema (5.10). Dessa forma, pode-se evidentemente usar o algoritmo HSHS para resolvê-lo. Mas o objetivo precípuo agora é achar uma solução discreta, o que constitui a diferença essencial. Para poder dar melhores respostas nesse contexto diferente, foi construído o novo algoritmo AHSHS, mais rápido e robusto.

Como a escolha do ponto inicial tem uma importância determinante na definição da solução final, foi implementado um procedimento para se achar um ponto de partida de boa qualidade. Para tal, foi feita a resolução do problema Multisource Fermat-Weber através do algoritmo HSMFW (*Hyperbolic Smoothing to the Multisource Fermat-Weber problem*) descrito em Xavier, V.L. 2011, considerando o mesmo conjunto de cidades do problema de localização de hubs. A experiência empírica tem demonstrado que o ponto obtido dessa maneira é um bom ponto inicial, como retratado pela Figura 6.3, que mostra que os hubs não se deslocam muito com a variação do parâmetro α .

Levando em consideração que os resultados computacionais obtidos para o problema teste AP200 simétrico, exibidos na Tabela 6.5, mostram tempos computacionais bem superiores àqueles registrados na literatura mais recente: Stanimirovic (2008), Malinovic (2010) e Marti et al (2014), para se obter tempos com menor duração, faz-se necessário implementar modificações no algoritmo HSHS para torná-lo mais competitivo. Com tal finalidade foi feita a inclusão de procedimentos de poda ou *pruning* para reduzir o esforço computacional. No novo algoritmo AHSHS, o procedimento de resolução do problema de otimização irrestrito no Passo Principal é efetuado em três procedimentos, a saber: Procedimento de Aproximação, Determinação dos Caminhos Seletos (*pruning*) e Procedimento de Resolução.

No primeiro procedimento, é executado um número reduzido de iterações do algoritmo de minimização irrestrita na sua forma ortodoxa, em que, para cada avaliação do valor da função objetivo é necessário computar a contribuição de todos os pares de cidades (j, l) , $j, l = 1, \dots, m$, e para cada um desses pares é preciso avaliar p^2 diferentes caminhos $z_{j,a,b,l}$ necessários para calcular $z_{j,l}$ conforme especificado na expressão 5.12. Esse procedimento de aproximação termina quando o número de iterações *iter* do processo iterativo de minimização atinge um número máximo pré-determinado $iter_{MAX}$ ou quando a distância relativa entre duas soluções consecutivas em relação à última solução for menor do que uma especificada tolerância Δx .

$$iter \geq iter_{MAX} \quad (7.9)$$

$$\frac{\|x_{iter} - x_{iter-1}\|_2}{\|x_{iter}\|_2} < \Delta x. \quad (7.10)$$

No segundo procedimento, de Determinação dos Caminhos Seletos, é identificado para cada cidade j o conjunto de hubs a ela fortemente conectados $I(j)$. Um hub i é considerado fortemente conectado a uma cidade j se existe um arco de acesso ligando diretamente j a i , seja na origem ou seja no destino, conforme retratado pelas ligações na Figura 6.2.

No terceiro, Procedimento de Resolução, é executado o algoritmo de minimização irrestrita em que o cálculo das distâncias $z_{j,l}$ somente considera as conexões ativas $I(j)$ e $I(l)$ determinadas no procedimento anterior.

$$h_{jl}(z_{jl}, x) = \sum_{a \in I(j)} \sum_{b \in I(l)} \phi(z_{jl} - (\chi \theta(s_j, x_a, \gamma) + \alpha \theta(x_a, x_b, \gamma) + \delta \theta(x_b, s_l, \gamma)), \tau) - \varepsilon = 0, \quad j, l = 1, \dots, m. \quad (7.11)$$

Efetuando as modificações acima descritas sobre o algoritmo base HSHS destinado à resolução do problema de localização de hub contínuo, chega-se ao algoritmo AHSHS para o caso de localização de hub discreto:

Algorithm 2 Algoritmo AHSHS Simplificado

1: **Inicialização**

- Escolha os valores iniciais: $\gamma^1, \tau^1, \varepsilon^1$.
- Escolha os valores: $0 < \rho_1 < 1, 0 < \rho_2 < 1, 0 < \rho_3 < 1; .$
- Escolha os valores: $iter_{MAX}, \Delta x$.
- Determine o ponto inicial x^0 , como a solução do problema multisource Fermat-Weber resolvido pelo algoritmo HSMFW.
- Faça $k = 1$.

2: **Passo principal**

- **Procedimento de Aproximação:**

- Repita até atingir a condição (7.9) ou (7.10):
 - Resolva o problema de otimização irrestrita (5.11) com $\gamma = \gamma^k, \tau = \tau^k$ e $\varepsilon = \varepsilon^k$, a partir do ponto inicial x^{k-1} e seja x^{kA} a solução aproximada obtida.

- **Procedimento de Determinação dos Caminhos Seletos:**

- Para cada cidade j determine o conjunto de hubs a ela conectados $I(j)$ na solução aproximada obtida x^{kA} .

- **Procedimento de Resolução:** Repita até que uma regra de parada seja atingida:

- Resolva o problema de otimização irrestrita (5.11) simplificado onde o cálculo dos valores $z_{jl}(x)$ somente considera as conexões ativas, conforme equação (7.11), com o uso dos parâmetros $\gamma = \gamma^k, \tau = \tau^k$ e $\varepsilon = \varepsilon^k$, a partir do ponto inicial x^{kA} e seja x^k a solução obtida.

• **Atualização de Parâmetros:**

- Faça $\gamma^{k+1} = \rho_1 \gamma^k, \tau^{k+1} = \rho_2 \tau^k, \varepsilon^{k+1} = \rho_3 \varepsilon^k$.
 - Faça $k = k + 1$
-

■

Capítulo 8

Resultados Computacionais Problema Localização Hubs Tipo p -Medianas Assimétrico Discreto

Deve ser enfatizado que este capítulo se dedica única e exclusivamente à resolução do problema discreto de localização de hubs definido por 7.1, amplamente consagrado na literatura pelo nome UMAPHMP, *Uncapacited Multiple Allocation p -Hub Median Median Problem*, usado, por exemplo, por Campbell et al (2002), Ernst et al (1988-A), Ernst et al (1988-B), Ernst et al (2009) e Contreras et al (2010). A metodologia usada nos experimentos foi a descrita no capítulo anterior, baseada em dois procedimentos independentes: primeiro a resolução do problema irrestrito de programação não linear completamente diferenciável, obtido pelo relaxamento das restrições discretas com a subsequente aplicação da técnica da suavização hiperbólica, via algoritmo AHSMS descrito no capítulo anterior, e, a seguir, o uso de algum tipo de busca local no entorno da solução encontrada no procedimento inicial.

O âmbito dos experimentos computacionais da presente tese se ateu a dois procedimentos de busca local, os mais singelos:

1. Busca do Primeiro Vizinho, ou busca da cidade vizinha mais próxima de cada hub da solução obtida no espaço contínuo. Esta busca resulta num conjunto de hubs no espaço discreto que será designado por HS-1V; e

2. Busca de Mais Vizinhos. Aqui, adota-se uma vizinhança mais larga, incluindo outras cidades vizinhas a cada um dos hubs do espaço contínuo, não se atendo somente à mais próxima, e formando o conjunto de hubs designado por HS-MV. Cada um dos hubs contínuos é substituído por uma cidade vizinha, que passa a atuar como hub discreto. Na vizinhança de cada hub contínuo, designa-se a cidade que irá formar o HS-MV, recaindo a escolha sobre o conjunto de cidades vizinhas que fornece melhores resultados em termos da função objetivo.

Os resultados obtidos com o segundo procedimento de busca são pelos menos tão bons quanto os do primeiro procedimento, sendo frequentemente melhores, e isso é conseguido a um custo relativamente pequeno, em termos do tempo adicional de CPU que é requerido, como será verificado empiricamente adiante.

Os experimentos numéricos foram realizados num PC Intel ACER com uma CPU Intel Core I7 de 2,7GHz e RAM de 512MB. Os programas foram codificados com o Fortran Visual Studio, Versão 9999. Os procedimentos de minimização irrestrita foram feitos através de um algoritmo Quasi-Newton, utilizando a fórmula de atualização BFGS da Biblioteca Harwell, disponível em <http://www.hsl.rl.ac.uk/>.

Os resultados computacionais apresentados a seguir foram obtidos a partir de uma implementação do algoritmo AHSBS com procedimentos análogos aos descritos no capítulo 4 quanto à escolha dos valores dos parâmetros de suavização e de suas atualizações $\varepsilon, \tau, \gamma, \rho_1, \rho_2, \rho_3$. Com respeito à escolha do ponto inicial $x_i^0, i = 1, \dots, p$, ele foi obtido através da resolução do problema Multisource Fermat-Weber usando o algoritmo descrito em Xavier, V.L. 2011. Quanto à escolha dos parâmetros para a regra de parada do procedimento de aproximação, foram definidos $iter_{MAX} = 30$ e $\Delta x = 0,01$.

Para ilustrar o desempenho do estratégia proposta, algoritmo AHSBS + busca local, são mostrados abaixo os resultados obtidos através da utilização de uma única família de problemas teste padrão da literatura:

- Australian Post (AP), que foi apresentado originalmente por (Ernst & Krishnamoorthy, 1996). Essa tipologia oferece problemas teste com número variável de cidades até $m = 200$, sendo AP200 o maior data-set padrão de dados de localização de hubs usado com frequência na literatura e com amplo acesso disponibilizado. Entre outras fontes, esse conjunto é reproduzido na OR Library (biblioteca de pesquisa operacional) - (Beasley, 1990), acessível através do site:

<http://people.brunel.ac.uk/~mastjjb/jeb/orlib/phubinfo.html>

O apêndice A contém informações e comentários adicionais sobre os dados utilizados na tipologia de testes Australian Post.

Nos presentes experimentos foram considerados problemas sem custos fixos de instalação e tomados como fixos aqueles valores mais tradicionalmente usados na literatura: para os custos de coleta: $\chi = 3$, para as economias de escala $\alpha = 0.75$ e para os custos de distribuição $\delta = 2$.

Em relação aos fluxos entre as cidades, diversamente da opção simétrica assumida no capítulo 4, considerou-se a matriz de demandas assimétrica, $w_{jl} \neq w_{lj}$ para $j, l = 1, \dots, m$ quando $j \neq l$, e, com relação aos elementos da diagonal principal

w_{jj} , são usados os valores não nulos para tais fluxos conforme dados obtidos no site administrado por Beasley (1990). Existem alguns estudos que usam os dados da mesma origem, porém impondo $w_{jj} = 0$. Essa alternativa não foi explorada aqui porque a hipótese de fluxo não nulo na diagonal permite mais facilmente comparações com artigos publicados na literatura. Este assunto também é discutido no apêndice A.

Assim, o problema (5.10) assume a seguinte formulação:

$$\text{minimizar } f(x) = \sum_{j=1}^m \sum_{l=1}^m w_{jl} z_{jl}(x). \quad (8.1)$$

O esforço necessário para a resolução de (8.1) depende, em primeiro lugar, da dimensão do problema, igual a $2p$. No entanto, cada avaliação da função objetivo envolve o cálculo de $m^2 p^2$ distâncias elementares z_{jabl} . Uma a uma, cada distância z_{jl} é obtida, no Procedimento de Aproximação do algoritmo AHSBS, através do cálculo do zero da equação (5.12), que tem p^2 termos. Já no Procedimento de Resolução do algoritmo AHSBS, isso é feito através do cálculo do zero da equação (7.11) com número muito menor de termos. A inclusão dessa poda permite um *speed-up* expressivo desse algoritmo em relação à versão sem a poda. No sentido de dar uma medida do ganho, pode-se usar um caso real típico em que $p = 10$. No Procedimento de Aproximação há $p^2 = 100$ termos na equação (5.12), enquanto no Procedimento de Resolução há, em geral, menos de 4 termos na equação (7.11).

Resultados para os Problemas Teste Derivados do Conjunto AP200

Foram obtidos 21 subconjuntos dos dados do AP200, com m variando de 20 a 200 pontos, de dez em dez, e ainda $m = 25$ e $m = 175$. Desses subconjuntos, só foram encontrados artigos na literatura que servissem de comparação para os casos de $m = 20, 25, 40, 50, 100$ e 200 .

Os resultados computacionais obtidos para todos os 21 subconjuntos são apresentados nas tabelas 8.1 a 8.4. Nessas tabelas, as duas primeiras colunas apresentam o número m de localidades considerado, e o número especificado de hubs, p .

A terceira coluna apresenta o melhor valor da função objetivo que se consegue através da Busca do Primeiro Vizinho, designado por f_{HS-1V} . A quarta coluna mostra o tempo médio requerido para se obter a HS-1V, designado por t_{HS-1V} . Este tempo é considerado igual ao tempo necessário para se chegar à solução no

espaço contínuo, uma vez que a determinação adicional da cidade mais próxima de cada hub contínuo, feita no HS-1V, requer um tempo de CPU desprezível.

A quinta coluna contém f_{HS-MV} , os melhores valores da função objetivo que se obtém através da Busca de Mais Vizinhos, e a coluna seguinte, a sexta, mostra o tempo médio t_{HS-MV} necessário para se determinar as cidades que constituirão o HS-MV. Este tempo inclui o tempo de determinação do hub contínuo (presumido igual ao da quarta coluna), acrescido do tempo da Busca de Mais Vizinhos. Nesta definição, os tempos da sexta coluna sempre excedem os da quarta.

As duas colunas finais apresentam comparações entre os resultados obtidos pelos dois procedimentos de busca local, em termos de valores da função objetivo e de tempo de processamento. As comparações são feitas através do cômputo das razões f_{HS-MV} / f_{HS-1V} , que será sempre menor ou igual a 1, e t_{HS-MV} / t_{HS-1V} , sempre maior do que 1.

Como comentado acima, alguns dos resultados obtidos podem ser comparados com valores descritos em artigos publicados nos últimos anos na literatura especializada. Dentre os mesmos, destacam-se os artigos de Stanimirovic (2008), de Malinovic (2010) e de Marti et al (2014), além da pesquisa recente de Carvalho, R. (2017). Em todos esses artigos são propostas formulações capazes de resolver instâncias de *hub – and – spoke* baseadas em dados derivados do conjunto AP200.

A tabela 8.5 mostra os resultados das comparações para esses casos. O número p de hubs varia de 3 a 10, exceto quando $m = 20$ e $m = 25$, porque, para esses valores de m , não existem resultados na literatura para $p = 9$ e $p = 10$. Os resultados transcritos nesta tabela foram os encontrados em Malinovic (2010), e incluem os de Stanimirovic (2008). Não foram usados os resultados mais recentes de Marti et al. (2014) porque, além de não trazerem valores melhores da Função Objetivo (FO) comparativamente a Malinovic (2010), o artigo de Marti et al. transcreve os resultados deste último com alguns erros.

Na tabela 8.5, as primeiras duas colunas, m e p são iguais às correspondentes das tabelas 8.1 a 8.4. Já a terceira e quartas colunas, f_{HS-MV} e t_{HS-MV} , são idênticas às colunas cinco e seis das tabelas 8.1 a 8.4.

A quinta e a sexta colunas da tabela 8.5, f_{lit} e t_{lit} mostram os melhores resultados obtidos na literatura, expressos em termos dos valores da função objetivo e do tempo de CPU utilizado, sendo que na coluna cinco os valores para os quais existe prova de que são os ótimos globais estão destacados em negrito. As duas últimas colunas, de maneira análoga às tabelas 8.1 a 8.4, mostram comparações entre os resultados obtidos pela HS-MV e os da literatura, em termos de função objetivo e de tempo médio de CPU.

m	p	f_{1V}	t_{1V}	f_{MV}	t_{MV}	f_{MV}/f_{1V}	t_{MV}/t_{1V}
20	3	148048	0,05	148048	0,05	1,000	1,08
20	4	135375	0,07	135375	0,07	1,000	1,06
20	5	118935	0,13	118935	0,14	1,000	1,04
20	6	107006	0,21	107006	0,21	1,000	1,02
20	7	97710	0,30	97698	0,31	1,000	1,02
20	8	92109	0,43	92036	0,44	0,999	1,03
20	9	87258	0,67	87235	0,68	1,000	1,02
20	10	81253	0,83	80288	0,85	0,988	1,03
25	3	151081	0,06	151081	0,07	1,000	1,17
25	4	136022	0,11	136013	0,12	1,000	1,05
25	5	120611	0,19	120582	0,20	1,000	1,07
25	6	115746	0,35	114645	0,37	0,990	1,04
25	7	104681	0,49	104681	0,50	1,000	1,02
25	8	97831	0,70	97831	0,71	1,000	1,01
25	9	91973	0,94	91973	0,96	1,000	1,02
25	10	88934	1,48	88767	1,50	0,998	1,01
30	3	152545	0,09	152545	0,09	1,000	1,04
30	4	138036	0,15	138036	0,16	1,000	1,04
30	5	125678	0,27	125678	0,28	1,000	1,04
30	6	117028	0,50	117028	0,50	1,000	1,01
30	7	110836	0,70	110836	0,71	1,000	1,02
30	8	103517	0,94	103517	0,96	1,000	1,02
30	9	99074	1,57	99073	1,58	1,000	1,00
30	10	94685	2,01	94685	2,04	1,000	1,01
40	3	155459	0,17	155459	0,18	1,000	1,06
40	4	140683	0,28	140683	0,29	1,000	1,04
40	5	130385	0,51	130385	0,52	1,000	1,02
40	6	124019	0,99	124019	1,00	1,000	1,01
40	7	116036	1,28	116036	1,31	1,000	1,02
40	8	112999	1,80	109972	2,12	0,973	1,18
40	9	104212	2,86	104212	2,88	1,000	1,01
40	10	101274	3,89	99662	3,93	0,984	1,01
50	3	156015	0,27	156015	0,29	1,000	1,06
50	4	141153	0,49	141153	0,49	1,000	1,00
50	5	129413	0,83	129413	0,85	1,000	1,02
50	6	121672	1,66	121672	1,67	1,000	1,01
50	7	116189	2,22	116189	2,30	1,000	1,04
50	8	109927	2,94	109927	2,97	1,000	1,01
50	9	105291	4,49	104968	4,74	0,997	1,06
50	10	100678	6,79	100678	7,02	1,000	1,03
60	3	157303	0,33	157303	0,35	1,000	1,07
60	4	142171	0,61	142171	0,65	1,000	1,07
60	5	129652	1,17	129652	1,23	1,000	1,05
60	6	121978	2,18	121978	2,22	1,000	1,02

Tabela 8.1: Resultados AP200 - Instâncias (m,p) de $(20,3)$ a $(60,6)$

m	p	f_{1V}	t_{1V}	f_{MV}	t_{MV}	f_{MV}/f_{1V}	t_{MV}/t_{1V}
60	7	116511	2,82	116510	2,89	1,000	1,02
60	8	110237	4,71	110237	4,85	1,000	1,03
60	9	105177	6,29	105177	6,34	1,000	1,01
60	10	101101	8,50	101101	8,65	1,000	1,02
70	3	157658	0,48	157658	0,50	1,000	1,03
70	4	142287	0,88	142287	0,90	1,000	1,02
70	5	132055	1,44	132055	1,46	1,000	1,01
70	6	123601	2,84	123601	2,87	1,000	1,01
70	7	117485	4,02	117485	4,14	1,000	1,03
70	8	112081	5,87	112081	6,84	1,000	1,16
70	9	108179	8,00	108179	8,23	1,000	1,03
70	10	103749	12,18	103749	13,18	1,000	1,08
80	3	158012	0,54	158012	0,56	1,000	1,04
80	4	142613	1,10	142613	1,12	1,000	1,02
80	5	132422	1,90	132422	1,93	1,000	1,01
80	6	125268	3,60	125257	3,68	1,000	1,02
80	7	119105	5,23	119105	5,42	1,000	1,04
80	8	115368	8,23	115368	8,60	1,000	1,05
80	9	109644	11,97	109644	15,81	1,000	1,32
80	10	105652	15,49	105512	16,09	0,999	1,04
90	3	160414	0,66	160399	0,68	1,000	1,04
90	4	142275	1,44	142275	1,49	1,000	1,03
90	5	132422	2,53	132422	3,42	1,000	1,35
90	6	126819	4,54	126819	5,52	1,000	1,21
90	7	119725	6,62	119378	6,84	0,997	1,03
90	8	114160	9,81	114160	10,14	1,000	1,03
90	9	110600	14,37	109674	17,98	0,992	1,25
90	10	105120	20,49	104928	21,92	0,998	1,07
100	3	157870	0,97	157870	1,02	1,000	1,05
100	4	143004	1,90	143004	2,11	1,000	1,11
100	5	133483	3,67	133483	4,11	1,000	1,12
100	6	127389	5,85	127364	6,43	1,000	1,10
100	7	120516	8,44	120423	8,51	0,999	1,01
100	8	115181	13,71	115144	14,52	1,000	1,06
100	9	110751	17,98	110531	18,35	0,998	1,02
100	10	105679	27,62	105679	29,16	1,000	1,06
110	3	157978	1,11	157978	1,15	1,000	1,03
110	4	143110	2,35	143110	2,60	1,000	1,11
110	5	133578	4,13	133578	4,45	1,000	1,08
110	6	126331	7,65	126331	8,20	1,000	1,07
110	7	120643	10,57	120549	11,66	0,999	1,10
110	8	114760	15,76	114681	17,19	0,999	1,09
110	9	110187	21,19	109823	27,49	0,997	1,30
110	10	104968	29,80	104968	31,67	1,000	1,06

Tabela 8.2: Resultados AP200 - Instâncias (m,p) de $(60,7)$ a $(110,10)$

m	p	f_{1V}	t_{1V}	f_{MV}	t_{MV}	f_{MV}/f_{1V}	t_{MV}/t_{1V}
120	3	158021	1,35	158021	1,38	1,000	1,02
120	4	143148	3,11	143148	3,42	1,000	1,10
120	5	133609	4,75	133609	5,37	1,000	1,13
120	6	126339	8,80	126339	9,78	1,000	1,11
120	7	120538	12,19	120538	12,71	1,000	1,04
120	8	114852	17,53	114755	18,67	0,999	1,06
120	9	112417	25,04	112113	27,19	0,997	1,09
120	10	107085	39,32	107085	45,08	1,000	1,15
130	3	158158	1,54	158158	1,57	1,000	1,02
130	4	143287	2,91	143287	3,24	1,000	1,11
130	5	133775	5,47	133775	5,82	1,000	1,06
130	6	127185	10,19	126653	11,38	0,996	1,12
130	7	120988	13,71	120872	14,31	0,999	1,04
130	8	114839	21,57	114839	23,06	1,000	1,07
130	9	110107	28,32	110107	31,00	1,000	1,09
130	10	106294	48,02	106028	53,19	0,997	1,11
140	3	158180	2,12	158180	2,23	1,000	1,05
140	4	143309	3,59	143309	3,95	1,000	1,10
140	5	133734	7,14	133718	8,46	1,000	1,19
140	6	127156	12,12	126645	13,23	0,996	1,09
140	7	120961	14,86	120865	39,40	0,999	2,65
140	8	115297	24,68	115163	26,65	0,999	1,08
140	9	110105	35,49	110105	41,54	1,000	1,17
140	10	106413	54,53	106413	59,49	1,000	1,09
150	3	158437	2,31	158437	2,41	1,000	1,04
150	4	143565	4,35	143565	5,29	1,000	1,22
150	5	134021	8,18	134021	8,74	1,000	1,07
150	6	127038	13,78	126870	17,81	0,999	1,29
150	7	121197	19,31	121085	20,98	0,999	1,09
150	8	115105	27,47	115105	29,12	1,000	1,06
150	9	110245	42,06	110228	44,00	1,000	1,05
150	10	105568	64,90	105568	72,92	1,000	1,12
160	3	158779	2,73	158779	2,88	1,000	1,06
160	4	143935	4,65	143695	10,61	0,998	2,28
160	5	134347	7,64	134342	8,30	1,000	1,09
160	6	127335	15,01	127170	16,98	0,999	1,13
160	7	121474	19,30	121460	20,91	1,000	1,08
160	8	115415	32,58	115415	37,23	1,000	1,14
160	9	110327	48,66	110327	50,61	1,000	1,04
160	10	106246	70,60	106246	81,59	1,000	1,16

Tabela 8.3: Resultados AP200 - Instâncias (m,p) de $(120,3)$ a $(160,10)$

m	p	f_{1V}	t_{1V}	f_{MV}	t_{MV}	f_{MV}/f_{1V}	t_{MV}/t_{1V}
170	3	159416	3,00	159416	3,64	1,000	1,21
170	4	143851	4,92	143846	5,22	1,000	1,06
170	5	135635	8,90	135589	10,94	1,000	1,23
170	6	128489	16,82	128384	20,74	0,999	1,23
170	7	122548	23,71	122548	25,87	1,000	1,09
170	8	116647	38,05	116647	39,72	1,000	1,04
170	9	111629	51,99	111629	56,48	1,000	1,09
170	10	107787	80,85	107426	308,30	0,997	3,81
175	3	163266	2,53	162511	4,66	0,995	1,84
175	4	144184	6,66	144051	13,21	0,999	1,98
175	5	136039	10,21	136033	10,90	1,000	1,07
175	6	128971	20,26	128829	21,96	0,999	1,08
175	7	123100	25,47	123099	31,32	1,000	1,23
175	8	117499	39,71	117499	43,38	1,000	1,09
175	9	112616	54,88	112521	60,64	0,999	1,10
175	10	108066	88,16	107978	99,20	0,999	1,13
180	3	159516	2,98	159516	3,98	1,000	1,34
180	4	144303	6,25	144170	13,63	0,999	2,18
180	5	136246	10,71	136150	12,16	0,999	1,14
180	6	129314	19,13	129211	31,01	0,999	1,62
180	7	123298	26,30	123298	29,98	1,000	1,14
180	8	117653	42,22	117653	55,84	1,000	1,32
180	9	112472	57,13	112472	79,41	1,000	1,39
180	10	107770	89,85	107763	110,52	1,000	1,23
190	3	159688	3,46	159688	4,54	1,000	1,31
190	4	144442	6,74	144439	16,71	1,000	2,48
190	5	136668	12,87	136658	14,73	1,000	1,14
190	6	129884	21,95	129720	40,61	0,999	1,85
190	7	123559	32,05	123555	40,22	1,000	1,25
190	8	117769	46,18	117767	49,86	1,000	1,08
190	9	112317	68,05	112317	256,88	1,000	3,77
190	10	108240	102,78	108240	125,69	1,000	1,22
200	3	159725	4,38	159725	5,56	1,000	1,27
200	4	144511	8,04	144508	17,53	1,000	2,18
200	5	136772	14,31	136762	15,53	1,000	1,09
200	6	130255	26,41	129823	35,66	0,997	1,35
200	7	124362	37,15	123918	42,99	0,996	1,16
200	8	117881	52,19	117881	59,49	1,000	1,14
200	9	112607	75,56	112557	92,16	1,000	1,22
200	10	107848	116,86	107847	127,52	1,000	1,09

Tabela 8.4: Resultados AP200 - Instâncias (m,p) de $(170,3)$ a $(200,10)$

m	p	f_{MV}	t_{MV}	f_{lit}	t_{lit}	f_{lit}/f_{MV}	t_{lit}/t_{MV}
20	3	148048	0,05	148048	0,07	1,000	1,30
20	4	135375	0,07	131665	0,12	0,973	1,62
20	5	118935	0,14	118935	0,16	1,000	1,19
20	6	107006	0,21	107006	0,21	1,000	0,98
20	7	97698	0,31	97698	0,25	1,000	0,82
20	8	92036	0,44	91455	0,30	0,994	0,68
25	3	151081	0,07	151081	0,09	1,000	1,29
25	4	136013	0,12	135639	0,15	0,997	1,29
25	5	120582	0,20	120582	0,23	1,000	1,13
25	6	114645	0,37	110836	0,32	0,967	0,88
25	7	104681	0,50	103880	0,42	0,992	0,84
25	8	97831	0,71	97796	0,47	1,000	0,66
40	3	155459	0,18	155459	0,26	1,000	1,44
40	4	140683	0,29	140683	0,39	1,000	1,34
40	5	130385	0,52	130385	0,85	1,000	1,63
40	6	124019	1,00	122171	5,86	0,985	5,85
40	7	116036	1,31	116036	7,12	1,000	5,46
40	8	109972	2,12	109972	8,39	1,000	3,95
40	9	104212	2,88	104212	9,62	1,000	3,35
40	10	99662	3,93	99453	12,48	0,998	3,17
50	3	156015	0,29	156015	0,34	1,000	1,19
50	4	141153	0,49	141153	0,55	1,000	1,12
50	5	129413	0,85	129413	0,79	1,000	0,93
50	6	121672	1,67	121672	8,48	1,000	5,07
50	7	116189	2,30	115912	10,34	0,998	4,50
50	8	109927	2,97	109927	12,51	1,000	4,21
50	9	104968	4,74	104968	15,17	1,000	3,20
50	10	100678	7,02	100509	17,99	0,998	2,56
100	3	157870	1,02	157870	13,32	1,000	13,06
100	4	143004	2,11	143004	18,52	1,000	8,80
100	5	133483	4,11	133483	23,78	1,000	5,78
100	6	127364	6,43	126108	31,27	0,990	4,87
100	7	120423	8,51	120165	41,32	0,998	4,86
100	8	115144	14,52	114296	57,00	0,993	3,93
100	9	110531	18,35	109449	68,74	0,990	3,75
100	10	105679	29,16	104794	87,16	0,992	2,99
200	3	159725	5,56	159725	73,95	1,000	13,30
200	4	144508	17,53	144508	95,51	1,000	5,45
200	5	136762	15,53	136762	156,21	1,000	10,06
200	6	129823	35,66	129556	185,72	0,998	5,21
200	7	123918	42,99	123609	226,30	0,998	5,26
200	8	117881	59,49	117710	284,98	0,999	4,79
200	9	112557	92,16	112374	366,92	0,998	3,98
200	10	107847	127,52	107847	432,69	1,000	3,39

Tabela 8.5: Comparações de resultados obtidos a partir do AP200 com a literatura

Comentários sobre os Resultados dos Problemas Derivados do AP200

As tabelas 8.1 a 8.4 mostram todas os casos calculados, inclusive os 44 que têm correspondentes na literatura e que aparecem também na tabela 8.5. Como já comentado, para cada caso, são calculados um conjunto de hubs no espaço contínuo, que será designado por AHSHS e, a partir dos mesmos, dois conjuntos de hubs no espaço discreto:

–o conjunto correspondente à substituição de cada hub de AHSHS pelo seu vizinho mais próximo no espaço discreto, abreviado por HS-1V (hub-spoke primeiro vizinho);

–o conjunto de hubs obtido através de uma busca de melhores valores da Função Objetivo (FO) realizada no entorno de cada um dos hubs que formam a solução contínua AHSHS, não se restringindo unicamente ao vizinho mais próximo de cada hub, designados por HS-MV (mais vizinhos).

Essas comparações são sintetizadas a seguir:

Quanto aos ganhos obtidos em termos de FO ao se passar de HS-1V para HS-MV, as tabelas 8.1 a 8.4 mostram na coluna 7 que, em 68 do total de 168 casos processados, existe uma melhoria quando se parte para a busca de mais vizinhos, relativamente ao primeiro vizinho. Isso ocorre, portanto, em cerca de 40% dos casos. Logo, a busca de vizinhança traz melhores resultados de uma maneira geral e deve ser preferida. O custo computacional correspondente é relativamente pequeno, uma vez que os aumentos de tempo de CPU, informados na coluna 8 das tabelas 8.1 a 8.4 não é grande, em geral, sendo o aumento médio de 33%.

Pode-se observar que a eficácia da busca de mais vizinhos aumenta com o número de localidades e com o número de hubs. De fato, se dividirmos os 21 valores de m em três faixas de igual tamanho, cada uma com sete valores de m (20-70, 80-140 e 150-200), vê-se que os casos em que a busca por vizinhança é melhor do que a busca do primeiro vizinho ocorrem em proporções crescentes, à medida que cresce o número de localidades, com as proporções assumindo os valores de 23%, 39% e 59%, respectivamente, para as três faixas especificadas. Com relação aos 8 valores de p , divididos em duas faixas: hubs menores (p de 3 a 6) e maiores (p de 7 a 10), as proporções em que a busca de mais vizinhos predomina sobre o primeiro vizinho são também crescentes: 35% e 46%, respectivamente. Para o valor mínimo de p testado, que é $p = 3$, verifica-se que HS-MV melhora os resultados em relação a HS-1V em apenas 10% dos casos, enquanto que, para o valor máximo, $p = 10$, isso ocorre em 48% dos casos testadas.

Conclui-se que a busca na vizinhança é mais eficaz para valores altos de m e de p . Isso ocorre sobretudo porque, para valores baixos de m e de p , a busca mais simples HS-1V já é capaz de chegar com frequência ao ótimo global, não havendo possibilidade de se conseguir melhorias com o HS-MV. Esse fato é comprovado pela tabela 8.5. Por exemplo, para o menor conjunto de localidades testado, ($m = 20$), e hubs menores (p de 3 a 6), o HS-1V já chega ao ótimo global em 3 dos 4 casos.

O tempo adicional requerido para a busca de mais vizinhos é em geral menor do que o requerido pelo algoritmo de procura do AHSHS (que também determina HS-1V), sendo maior só em 7 dos 168 casos analisados. Em 13 casos, o tempo da busca de mais vizinhos foi bem pequeno (menor que 0,01s); em metade dos casos essa busca representou um acréscimo de menos de 7% ao tempo do AHSHS, e em apenas 18 casos esse acréscimo passou de 30%. Logo, pode-se concluir que a busca na vizinhança não afeta significativamente o tempo total de CPU na grande maioria dos casos.

No que tange à comparação com a literatura, nos 44 casos em que havia resultados comparáveis publicados, relacionados na Tabela 8.5, o algoritmo HS-MV chegou a igualar os melhores resultados já obtidos em 25 casos, 57% do total. Nos demais 19 casos não se atingiu os melhores resultados da literatura, mas o HS-MV chegou bem perto desses resultados, em metade dos casos com desvios menores do que 0,3% na função objetivo. Só houve três casos de desvios acima de 1% em relação ao melhor da literatura, sendo o maior deles de 3,4%.

Os tempos de CPU requeridos para obter tais resultados, como é mostrado na última coluna da Tabela 8.5, foram, na grande maioria dos casos, consideravelmente menores do que o relatado na literatura. De fato, houve apenas seis casos dentre os 44 analisados em que o resultado da literatura foi obtido em tempos menores do que o HS-MV, e um caso de empate. Nos demais 37 casos, o HS-MV obteve resultados com mais rapidez, sendo que em 26 deles a redução foi para menos da metade. Esse ganho de velocidade é mais acentuado para valores maiores de m . Por exemplo, para os oito casos de maior tamanho, com $m = 200$, os tempos médios do HS-MV são, em média, 22% dos tempos relacionados na literatura, enquanto ao mesmo tempo se consegue chegar ao melhor resultado já obtido na literatura em metade desses casos, sendo de apenas 0,25% o desvio máximo da função objetivo na outra metade dos casos.

Isso indica que o HS-MV tende a ter um desempenho que se destaca das alternativas existentes, sobretudo quando m e p são grandes, sendo de se esperar um ganho adicional de qualidade e eficiência à medida que esses valores aumentem ainda mais, mantendo ao mesmo tempo a robustez das soluções encontradas.

Como os tempos do HS-MV crescem de uma maneira bem menos explosiva do

que qualquer solução conhecida em espaços discretos, também é de se esperar que o método seja capaz de produzir bons resultados para instâncias de */, hub – and – spoke/*, de dimensões maiores do que aquelas até hoje tentadas na literatura.

Capítulo 9

Conclusões

Preâmbulo

Os problemas de localização de hubs visam determinar os melhores locais para situar instalações que sirvam como polos de distribuição, atendendo às demandas de transporte ou de transferência originadas em algumas localidades específicas e destinadas a outras localidades, de uma forma que minimize o custo de transferência ou de transporte entre tais localidades, e, simultaneamente, visam definir quais instalações serão utilizadas por cada uma delas. São, portanto, problemas que abordam o binômio localização-alocação. Para estimar o custo total, se considera a soma dos custos de todas as transferências, calculado através da multiplicação da demanda específica de cada trecho pelo custo unitário, presumido proporcional à distância percorrida entre os locais de origem e de destino, passando pelas instalações ou hubs a serem estabelecidos.

O Problema de Hubs no Espaço Contínuo

Os sistemas de *hub – and – spoke* são projetados para explorar as economias de escala atingíveis através do uso compartilhado de ligações de alta eficiência entre os polos de distribuição (hubs). Como alternativa à abordagem discreta padrão, que consiste na seleção de hubs a partir de um subconjunto das localidades existentes, esse enfoque explora a possibilidade de localizar os hubs no espaço plano contínuo. Dada a demanda de tráfego entre cada par origem-destino e os respectivos custos de transporte, o problema consiste em encontrar a rede *hub – and – spoke* mais barata, efetuando o roteamento de cada fluxo origem-destino através do conjunto de hubs e atribuindo fluxos a cada um deles.

Este trabalho propõe um novo método para a solução dos problemas de alocação múltipla contínua de hubs p -medianas. Usando a técnica de suavização hiperbólica, o problema foi inicialmente reformulado, através de uma abordagem de aproximação,

como sendo um problema de otimização com restrições completamente diferenciável. Usando o Teorema da Função Implícita, o problema foi novamente reformulado como um problema de otimização sem restrições de baixa dimensão, no espaço euclidiano \mathbb{R}^{2p} , que tem somente $2p$ dimensões. Em seguida, foram apresentados os passos básicos do algoritmo HSHS para resolver o problema original, topologia *clique*, caso simétrico, de *hub – and – spoke* contínuo.

A abordagem adotada para resolução desse problema de hub tem como origem histórica uma extensão natural de trabalhos anteriores sobre análise de agrupamentos (clustering) usando a técnica de suavização hiperbólica, ambos problemas não-diferenciáveis da tipologia *min – sum – min*. Tem sido reiteradamente provado que a aplicação dessa técnica é eficaz, robusta e eficiente na resolução de grandes instâncias dos problemas de agrupamento baseados na minimização da soma dos quadrados MSSC - *Minimum Sum-of-Squares Clustering*: Xavier (2010) e Xavier e Xavier (2011). Um desempenho idêntico foi também observado em relação a um problema de cobertura de duas dimensões (Xavier e Oliveira, 2005), e também para um problema de cobertura de corpos sólidos (Venceslau, Lubke e Xavier, 2015), que são problemas não-diferenciáveis da tipologia *min – max – min*. A aplicação da metodologia para o problema de hub contínuo aqui estudado, que tem características similares ao problema MSSC, resultou em um desempenho exitoso similar às experiências anteriores para os problemas acima relacionadas.

Deve ser ressaltado que problema de *hub – and – spoke* de alocação múltipla de p -medianas no espaço contínuo é um problema de otimização global com uma grande quantidade de mínimos locais. Em consequência, o algoritmo HSHS não garante que se atinja a idealizada solução ótima global. De qualquer forma, este trabalho mostra como uma implementação do algoritmo proposto é capaz de produzir soluções consistentes para instâncias de grande porte, encontradas em muitas aplicações do mundo real. Acredita-se que esta metodologia possa ser adequada para as necessidades de várias aplicações relevantes.

Comentários sobre o Desempenho do Método no Espaço Contínuo Simétrico

A descrição dos resultados obtidos nos experimentos computacionais pelo algoritmo HSHS para a especificação mais tradicional do problema de hub, topologia *clique*, caso simétrico, é apresentada no capítulo 6.

O método proposto foi aplicado a problemas teste da literatura de vários tamanhos, com diversas configurações da distribuição espacial das localidades e para

diversos números de hubs. Conseguiu-se obter soluções satisfatórias para a localização dos hubs em espaços contínuos, em particular em tempos de CPU muito baixos. Foram resolvidas instâncias de grande porte, como o dsj1000 da biblioteca TSPLIB com 1000 cidades e com até 6 hubs. Essa instância tem 499500 pares de fluxos origem-destino, de tamanho inaudito na bibliografia do tópico *hub location*. Entretanto, dado o ineditismo das dimensões dos quatro maiores problemas teste, não foi possível se encontrar exemplos na literatura técnica que permitissem uma comparação direta dos resultados obtidos pelo HSHS com aqueles de métodos alternativos. Para suprir essa lacuna, foi feito um experimento computacional usando um algoritmo tipo força bruta para a instância German Towns, com a enumeração completa de todas combinações das cidades, nos casos $p = 2, 3, 4, 5$ hubs. Os valores da função objetivo, avaliados nas cidades vizinhas mais próximas às soluções contínuas, ficaram perto dos mínimos globais discretos obtidos via força bruta, com desvios relativos de $(0, 11\%, 0, 27\%, 0, 0\%, 0, 0\%)$ respectivamente.

O Problema de Hubs no Espaço Discreto

Considera-se nessa seção a Resolução do Problema de Localização de Hubs do Tipo p -Medianas Discreto, consagrado na literatura com o nome: UMAPHMP.

Para se obter uma solução no espaço discreto, usou-se a articulação de dois procedimentos independentes. Procurou-se primeiro o melhor hub no espaço contínuo, via resolução do problema 7.8. A seguir, substituiu-se os pontos que formam essa solução, através de um procedimento de busca local nessa vizinhança, por localidades ou cidades do conjunto que se quer atender. No segundo procedimento, foram adotadas duas estratégias singelas para a substituição dos pontos no espaço contínuo por localidades do espaço discreto.

A primeira foi uma busca local, a mais simples possível, na qual se substitui cada ponto do hub contínuo pela localidade mais próxima do mesmo. Essa determinação é feita de maneira muito rápida, porque o próprio algoritmo de procura da melhor solução contínua requer o cálculo das distâncias de cada hub a todas as localidades. Dessa forma, resta somente determinar qual dessas distâncias é a menor e verificar a qual localidade corresponde.

A segunda estratégia de busca local adotada, descrita em maiores detalhes no capítulo 7, consiste em procurar a melhor localidade discreta nas vizinhanças dos pontos que formam a solução contínua, não se limitando apenas à localidade mais próxima. Essa estratégia fornece melhores resultados, na forma de valores menores para a Função Objetivo, embora requiera um processamento adicional. Esse cálculo

adicional requer um tempo de CPU que depende do tamanho da vizinhança definida para a procura. Vizinhanças maiores implicam em mais alternativas de hubs discretos e incluem as vizinhanças menores como subconjunto. Por isso, fornecem resultados pelo menos tão bons quanto as menores. Entretanto, consomem mais tempo, por requererem o exame de um número maior de alternativas combinatórias. Optou-se por vizinhanças de tamanho igual ao dobro da distância média entre as localidades, o que forneceu resultados de boa qualidade, chegando ao melhor resultado já registrado na literatura em metade dos casos examinados. Com o aumento do tamanho de tais vizinhanças pode-se chegar a resultados ainda melhores, e o aumento do tempo de CPU não chega a se comparar com o aumento que ocorre na solução do problema discreto por métodos exclusivamente discretos, quando se lida com grandes instâncias (grande número de localidades e de hubs), devido à explosão combinatória que afeta esses últimos. Esse fenômeno é explicitado com clareza pelo número de variáveis das formulações 7.1, 7.2 e 7.5.

O problema contínuo 7.8 resolvido no presente estudo tem uma dimensão muito menor do que as apresentadas por esses trabalhos clássicos, restando entretanto o enfrentamento da natureza discreta da solução. A formulação UMAPHMP tem sido tema de centenas de publicações durante os últimos vinte e cinco anos. A incorporação permanente e acumulada de aprimoramentos consecutivos nas metodologias de sua resolução tem produzidos algoritmos com desempenhos cada vez melhores. Assim, para se obter resultados computacionais de qualidade *vis-à-vis* aqueles publicados mais recentemente na literatura, foram necessárias implementações para tornar o algoritmo HSHS mais rápido e mais preciso, dando origem ao novo algoritmo AHSBS, *Accelerated Hyperbolic Smoothing Hub Spoke*.

O algoritmo AHSBS tem dois novos procedimentos em relação ao anterior HSHS: escolha de ponto inicial de melhor qualidade e inclusão de procedimento de poda ou *pruning* para aumentar a velocidade de execução. O ponto inicial é produzido pela resolução do problema de Fermat-Weber, também pela metodologia de suavização hiperbólica, através do algoritmo HSMFW (*Hyperbolic Smoothing to the Multisource Fermat-Weber problem*) descrito em Xavier, V.L. 2011. A racionalidade dessa escolha se fundamenta na ocorrência de pequenos deslocamentos das posições dos hubs em resposta à variação de α , como mostra a Figura 6.3 para o caso simétrico, o que é também válido para o caso assimétrico.

O procedimento de *pruning* seleciona um seletivo subconjunto de conexões mais relevantes para as avaliações dos fluxos entre as cidades origem-destino, o que reduz substancialmente o volume de trabalho computacional. O cálculo da função objetivo do problema (5.11) exige o cálculo dos zeros de m^2 equações (5.12), uma equação para cada par de cidades, cada uma com p^2 termos. Esta é a tarefa computacio-

nal mais relevante associada ao algoritmo HSHS. O procedimento de poda usa um esquema de partição semelhante ao apresentado por (Xavier e Xavier, 2011), conseguindo com isso uma forte diminuição do número de termos, destarte com grande redução nos tempos de processamento. Nos experimentos computacionais exibidos no capítulo 8, a inclusão do procedimento de *pruning* engendrou um *speed-up* na faixa de cinco a vinte vezes mais rápido.

Comentários sobre o Desempenho do Método no Espaço Discreto

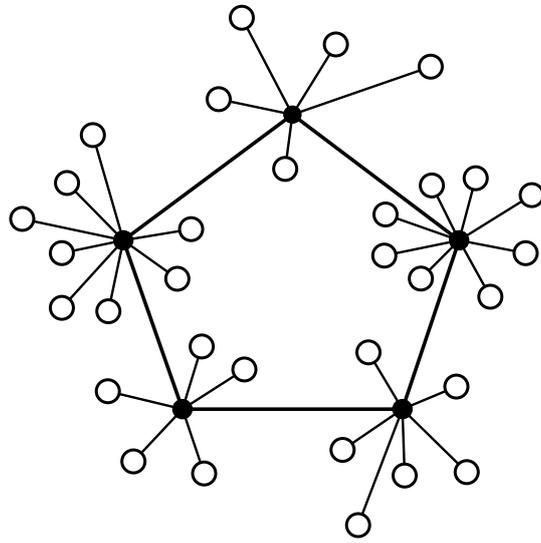
O desempenho da proposta de uso do algoritmo AHSHS articulado ao procedimento de busca na vizinhança da solução contínua foi satisfatório, quer segundo o critério de acurácia, quer segundo o de eficiência. No conjunto de casos com valores de ótimos globais conhecidos, para as instâncias AP20, AP25, AP40, AP50, AP100 e AP200, foram atingidas 25 soluções globais de um total de 44. Nas demais soluções, o desvio relativo do valor da função objetivo da solução encontrada foi da ordem de 1% do valor do mínimo global. De outro lado, os tempos de CPU de processamento foram substancialmente menores, da ordem de 20 % dos valores registrados na literatura pelos melhores algoritmos.

Melhorias e Aprimoramentos

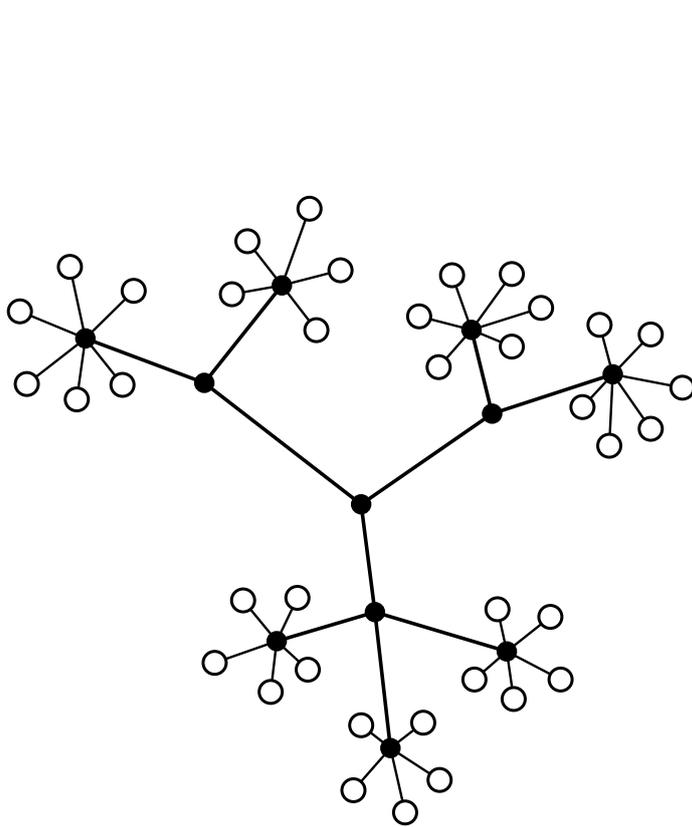
Deve se esclarecer que todos os resultados computacionais apresentados no capítulo 8 foram produzidos por uma implementação preliminar da articulação do algoritmo AHSHS com procedimento de busca local. Nesse contexto, sempre há possibilidades de melhoramento, em particular, nos procedimentos de busca local, que foram de extrema singeleza. A utilização de heurísticas desenvolvidas *ad hoc* para a formulação UMApHMP, como a de Carvalho (2017), ou de meta-heurísticas de uso geral como VNS ou ILS, deve melhorar a eficácia do esquema proposto.

Extensões da Metodologia para Outros Problemas

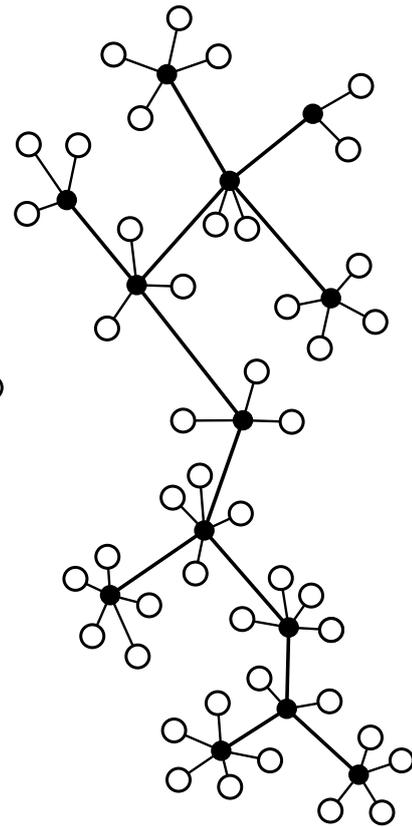
Apesar de ter sido considerada no escopo deste presente trabalho apenas a formulação mais tradicional para uma rede de hubs completamente interligada, conhecida em teoria dos jogos como *clique*, deve ser enfatizado que esta abordagem pode ser usada para resolver uma ampla classe de problemas *hub – and – spoke*. Para topologias diversas daquela retratada pela Figura 4.2, como aquelas ilustradas pelas figuras 3.1 e 9.1, basta modelar o conjunto de conexões origem-destino pertinentes, aplicando procedimentos idênticos aos usados para a topologia *clique*.



(a) Topologia em anel



(b) Topologia em estrela



(c) Topologia em árvore

Figura 9.1: Topologias de rede diversas

Referências Bibliográficas

Abadie, J. e J. Carpentier. 1969. Generalization of the Wolfe Reduced Gradient Method to the Case of Nonlinear Constraints, in *Optimization*, R. Fletcher (editor), Academic Press, London, 37-47.

Alumur, S. e B. Y. Kara 2008. Network hub location problems: The state of the art. *European Journal of Operational Research* **190** 1-21.

Aykin, T. 1988. On the Location of Hub Facilities, *Transportation Science* **22** 155-157.

Aykin, T. e G.F. Brown. 1992. Interacting New Facilities and Location-Allocation Problems, *Transportation Science* **26** 212-222.

Bagirov, A.M. 2008. Modified Global k-means Algorithm for Minimum Sum-of-Squares Clustering Problems, *Pattern Recognition* **41** Issue 10 3192-3199.

Bagirov, A.M. e J. Yearwood 2006. A New Nonsmooth Optimization Algorithm for Minimum Sum-of-Squares Clustering Problems, *European Journal of Operational Research* **170** 578-596

Barbosa, G.V. 2014. Projeto de Redes Hierárquicas Baseado no Mecanismo de Suavização Hiperbólica, D.Sc. Thesis, COPPE/UFRJ, Rio de Janeiro.

Beasley, J.E. 1990. OR-Library: distributing test problems by electronic mail, *Journal of the Operational Research Society* **41** 1069-1072.

Camargo, R.S., G. de Miranda, H.P.L. Luna 2008. Benders Decomposition for the Uncapacitated Multiple Allocation Hub Location Problem, *Computers and Operations Research* **35** 1047-1064.

Campbell, J.F. 1992. Location and allocation for distribution systems with transshipments and transportation economies of scale. *Annals of Operations Research* **40**, 77-99.

- Campbell, J.F. 1994-A. Integer Programming Formulations of Discrete Hub Location Problems, *European Journal of Operational Research* **72** 1-19.
- Campbell, J.F. 1994-B. A survey on network hub location Problem. *Studies on Locational Analysis* **6**, 31-49.
- Campbell, J.F., A.T. Ernst e M. Krishnamoorthy 2002. Hub location problems, in *Facility location: Applications and theory*, Z. Drezner and H. W. Hamacher (editors), 373-407. Springer-Verlag, Berlin.
- Carvalho, R. 2017. Heurísticas paralelas aplicadas a problemas de alocação de concentradores, *Tese de Doutorado em Engenharia Eletrica*, UFMG, Belo Horizonte.
- Chaves, A.M.V. 1997. Resolução do Problema Minimax Via Suavização, M.Sc. Thesis - COPPE - UFRJ, Rio de Janeiro.
- Contreras, I., J-F. Cordeau, G. Laporte 2010. Benders Decomposition for Large-Scale Uncapacitated Hub Location, *Technical Report CIRRELT-2010-26*, Montreal, Canada.
- Cunha, C.B. e M.R. Silva. 2007. A Genetic Algorithm for the Problem of Computing a Hub-and-Spoke Network for a LTL Trucking Company in Brazil, *European Journal of Operational Research* **179** 747-758.
- Demyanov, A. 2005. On the Solution of Min-Sum-Min Problems, *Journal of Global Optimization* **31** 3 437-453.
- Ernst, A.T. e M. Krishnamoorthy 1996. Efficient Algorithms for the Uncapacitated Single Allocation p -Hub Median Problem, *Location Science* **4** 139-154.
- Ernst, A.T. e M. Krishnamoorthy 1998-A. Exact and Heuristic Algorithms for the Uncapacitated Multiple Allocation p -Hub Median Problems, *European Journal of Operational Research* **104** 100-112.
- Ernst, A.T. e M. Krishnamoorthy 1988-B. An exact solution approach based on shortest-paths for p -hub median problems. *Journal on Computing* **10** (2), 149-162.
- Ernst, A.T., H. Hamacher, H. Jiang, M. Krishnamoorthy e G. Woeginger 2009. Uncapacitated single and multiple allocation p -hub center problems. *Computers & Operations Research*, **36**, 2230-2241.

- Farahani, R.Z., M. Hekmatfar, A.B. Arabani e E. Nikbakhsh 2013. Hub location problems: A review of models, classification, solution techniques, and applications, *Computers & Industrial Engineering*, **64** (4), 1096-1109.
- Farahani, R.Z. & Hekmatfar, M. (Editores) 2009. Facility Location: Concepts, Models, Algorithms and case Studies, Optimal Center Location in Simple Networks, Springer-Verlag, Berlin-Heiderbeg.
- Ferreira, R.P.M. e H.P.L. Luna. 2003. Discrete Capacity and Flow Assignment Algorithms with Performance Guarantee, *Computer Communications* **26** 1056-1069.
- Gavish, B. 1991. Topological Design of Telecommunication Networks – Local Access Design Methods , *Annals of Operations Research* **33** 17-71.
- Gavish, B. e K. Altinkemer. 1990. Backbone Network Design Tools with Economic Tradeoffs, *ORSA Journal on Computing* **2** 236-245.
- Gerla, M. e L. Kleinrock. 1977. On the Topological Design of Distributed Computer Networks, *IEEE Transaction on Communications* **25** 48-60.
- Goldman, A.J. 1969. Optimal Center Location in Simple Networks, *Transportation Science* **3** 352-360.
- Hakimi, S.L. 1964. Optimum Location of Switching Centers and the Absolute Centers and Medians of a Graph, *Operations Research* **12** 450-459.
- Hamacher, H.W., M. Labbé, S. Nickel e T. Sonneborn. 2004. Adapting Polyhedral Properties from Facility to Hub Location Problems , *Discrete Applied Mathematics* **145** 104-116.
- Klincewicz, J.G. 1998. Hub Location in Backbone/Tributary Network Design: A Review, *Location Science* **6** 307-335.
- Koopmans, T.C. e M. Beckmann. 1957. Assignment Problems and the Location of Economic Activities, *Econometrica* **25** 53-76.
- Luenberger, D. 1984. Introduction to Linear and Nonlinear Programming, Addison-Wesley (Menlo Park).
- Luna, H.P.L., N. Ziviani e R.M.B. Cabral. 1987. The Telephonic Switching Center Network Problem: Formalization and Computational

- Experience, *Discrete Applied Mathematics* **18** 199-210.
- Luna, H.P.L. 2006. Network Planning Problems in Telecommunications, in *Handbook of Optimization in Telecommunications*, M.G.C. Rezende and P.M. Pardalos (editors), Springer, New York, 213-240.
- Malinovic, M. 2010. A New Evolutionary Based Approach for Solving the Uncapacitated Multiple Allocation p-Hub Median Problem, in *Soft Computing in Industrial Applications*, X. Z. Gao et al. (editors), Advances in Intelligence and Soft Computing **75**, Springer-Verlag, Berlin, 81-88.
- Martí, R., Á. Corberán e J. Peiró. 2014. Scatter search for an uncapacitated p-hub median problem, *Computers and Operations Research* **58**, Elsevier, Amsterdam, 53-66.
- Minoux, M. 1986. Mathematical Programming Theory and Algorithms, John Wiley and Sons, Chichester.
- O’Kelly, M.E. 1986. The Location of Interacting Hub Facilities, *Transportation Science* **20** 92-106.
- O’Kelly, M.E. 1987. A Quadratic Integer Program for the Location of Interacting Hub Facilities, *European Journal of Operational Research* **32** 393-404.
- O’Kelly, M.E. e H.J. Miller. 1994. The Hub Network Design Problem: A Review and Synthesis, *Journal of Transport Geography* **2** 31-40.
- O’Kelly, M.E., D. Bryan, D. Skorin-Kapov, J. Skorin-Kapov 1996. Hub Network Design with Single and Multiple Allocation: A Computational Study, *Location Science* **4** 125-138.
- O’Kelly, M. O., J. F. Campbell, R. S. Camargo e G. Miranda Jr. 2014. Multiple Allocation Hub Location Model with Fixed Arc Costs, *Geographical Analysis, September 2014* 1-24
- Randazzo, C., H.P.L. Luna 2001. A Comparison of Optimal Methods for Local Access Uncapacitated Network Design, *Annals Of Operations Research* **106** 263-286.
- Reese, J. 2005. Methods for Solving the p -Median Problem: An Annotated Bibliography, Trinity University, Mathematics Faculty Research, Paper 28. Available in <http://digitalcommons.trinity.edu/math->

faculty/28/

Reinelt, G. 1991. TSP-LIB: A Traveling Salesman Library, *ORSA Journal on Computing*, 376-384.

Rubinov, A. 2006. Methods for Global Optimization of Nonsmooth Functions with Applications, *Applied and Computational Mathematics* **5** no. 1, 3-15.

Santos, A.B.A. 1997. Problemas de Programação Não- Diferenciável: Uma Metodologia de Suavização, M.Sc. Dissertation - COPPE - UFRJ, Rio de Janeiro.

Silva, M.R. e C.B. Cunha. 2009. New Simple and Efficient Heuristics for the Uncapacitated Single Allocation Hub Location Problem, *Computers and Operations Research* **36** Issue 12, 3152-3165.

Skorin-Kapov, D., J. Skorin-Kapov e M. O'Kelly 1996. Tight linear programming relaxations of uncapacitated p-hub median problems. *European Journal of Operational Research* **94**, 583-593.

Sousa, L.C.F. 2005. Desempenho Computacional do Método de Agrupamento Via Suavização Hiperbólica, M.Sc. thesis - COPPE - UFRJ, Rio de Janeiro.

Souza, M.F. 2010, Suavização hiperbólica aplicada à otimização de geometria molecular, D.Sc. Thesis, COPPE/UFRJ, Rio de Janeiro.

Souza, M.F., A. E. Xavier, L. Lavor, e N. Maculan 2011. Hyperbolic smoothing and penalty techniques applied to molecular structure determination. *Operations Research Letters*, **39**: 461-465.

Späth, H. 1980. Cluster Analysis Algorithms for Data Reduction and Classification, Ellis Horwood, Upper Saddle River, NJ.

Stanimirovic, Z. 2008. An efficient genetic algorithm for the uncapacitated multiple allocation p-hub median problem, *Control and Cybernetics* **37**, No.3, 669-692.

Wesolowsky, G.O. 1993. The Weber Problem: History and Perspectives, *Location Science* **1**, No. 1, pp. 5-23.

Xavier, A.E. 1982. Penalização Hiperbólica: Um Novo Método para Resolução de Problemas de Otimização, M.Sc. Thesis - COPPE - UFRJ,

Rio de Janeiro.

Xavier, A.E. 2003. Convexificação do problema de distância geométrica através da técnica de suavização hiperbólica. *Workshop em Biociências COPPE UFRJ*.

Xavier, A.E., A.A.F. Oliveira. 2005. Optimal Covering of Plane Domains by Circles Via Hyperbolic Smoothing, *Journal of Global Optimization* **31** 493-504.

Xavier, A.E. 2010. The Hyperbolic Smoothing Clustering Method, *Pattern Recognition* **43** 731-737.

Xavier, A.E., V.L. Xavier. 2011. Solving the Minimum Sum-of-Squares Clustering Problem by Hyperbolic Smoothing and Partition into Boundary and Gravitational Regions, *Pattern Recognition* **44** 70-77.

Xavier, V.L. 2011. Resolução do Problema de Agrupamento segundo o Critério de Minização da Soma de Distâncias, M.Sc. thesis - COPPE - UFRJ, Rio de Janeiro.

Yaman, H. 2011. Allocation Strategies in hub networks. *European Journal of Operational Research* 211, No. 3, pp. 442-451.

Apêndice A

Os Dados AP200

Os experimentos computacionais feitos para o caso discreto foram baseados no conjunto de dados AP200, com 200 localidades correspondentes a distritos postais australianos (Australian Post), produzidos por Ernst e Krishnamoorthy para uso em seu estudo de 1994, e colocado à disposição de pesquisadores no repositório OR-Library, acessível em: <http://people.brunel.ac.uk/~mastjjb/jeb/orlib/files/phub2.txt>

As posições são fornecidas como coordenadas nos eixos x e y , variando nos intervalos $x \in [2.882, 57.460]$ e $y \in [0, 55.024]$. A amplitude de variação é de perto de 55 mil nos dois eixos, e os pontos seguem um padrão regular de grade, como pode ser visto na figura A.2 adiante. De fato, dos 200 pontos, 190 tem coordenada x que é um múltiplo de 1441 e coordenada y que é um múltiplo de 1448. Isso permite que as coordenadas possam ser definidas de forma simplificada, como na figura A.1 abaixo, que mostra a posição sequencial das localidades e as coordenadas-índice i_x e i_y , ambas formadas por números inteiros no intervalo $[1, 40]$. Para obter as coordenadas x e y dos pontos do conjunto AP200, basta multiplicar i_x e i_y pelos respectivos fatores de escala 1441 e 1448.

Dez localidades na tabela, as de número: 2, 43, 94, 130, 142, 150, 157, 160, 163 e 177 têm as mesmas coordenadas-índice que a localidade imediatamente anterior, indicando que estão muito próximas. Nesses casos, uma das coordenadas-índice está assinalada em vermelho. As coordenadas dessas localidades podem ser obtidas a partir das coordenadas da localidade anterior, repetindo a coordenada em preto e somando-se 10 à coordenada em vermelho.

Vale notar que esses dez pontos estão muito próximos das localidades imediatamente anteriores, a uma distância de somente 10. Este valor é consideravelmente menor do que a distância mínima entre quaisquer outras duas localidades, que é 1441 na dimensão x e 1448 na dimensão y . Essa situação implica em dificuldade de distinguir duas localidades tão próximas na representação espacial da figura A.2, devido à escala. Para poder assinalar esses dez pares de localidades muito próximas, foram

seq	i_x	i_y																					
1	17	0	26	32	14	51	26	20	76	34	24	101	16	27	126	29	29	151	23	32	176	21	35
2	17	0	27	16	15	52	28	20	77	6	25	102	19	27	127	34	29	152	24	32	177	21	35
3	5	1	28	34	15	53	11	21	78	10	25	103	23	27	128	15	30	153	35	32	178	22	35
4	17	1	29	7	16	54	14	21	79	17	25	104	24	27	129	16	30	154	12	33	179	25	35
5	5	3	30	17	16	55	29	21	80	20	25	105	25	27	130	16	30	155	17	33	180	27	35
6	40	3	31	25	16	56	16	22	81	22	25	106	27	27	131	17	30	156	18	33	181	4	36
7	17	4	32	3	17	57	17	22	82	26	25	107	29	27	132	21	30	157	18	33	182	6	36
8	5	5	33	18	17	58	18	22	83	27	25	108	34	27	133	22	30	158	19	33	183	12	36
9	21	6	34	23	17	59	21	22	84	35	25	109	11	28	134	23	30	159	21	33	184	14	36
10	17	7	35	28	17	60	24	22	85	2	26	110	13	28	135	25	30	160	21	33	185	15	36
11	39	7	36	39	17	61	26	22	86	5	26	111	21	28	136	27	30	161	22	33	186	21	36
12	10	9	37	15	18	62	29	22	87	11	26	112	23	28	137	30	30	162	23	33	187	22	36
13	15	9	38	9	19	63	12	23	88	14	26	113	25	28	138	31	30	163	23	33	188	23	36
14	37	9	39	12	19	64	13	23	89	18	26	114	26	28	139	17	31	164	24	33	189	24	36
15	20	10	40	16	19	65	18	23	90	19	26	115	31	28	140	20	31	165	25	33	190	26	36
16	40	10	41	18	19	66	22	23	91	21	26	116	33	28	141	23	31	166	26	33	191	4	37
17	16	11	42	19	19	67	26	23	92	22	26	117	34	28	142	23	31	167	27	33	192	5	37
18	36	12	43	19	19	68	29	23	93	23	26	118	13	29	143	7	32	168	15	34	193	8	37
19	16	13	44	21	19	69	38	23	94	23	26	119	14	29	144	10	32	169	18	34	194	10	37
20	22	13	45	22	19	70	9	24	95	25	26	120	19	29	145	13	32	170	21	34	195	13	37
21	24	13	46	24	19	71	15	24	96	26	26	121	20	29	146	16	32	171	22	34	196	15	37
22	26	13	47	30	19	72	16	24	97	28	26	122	22	29	147	17	32	172	24	34	197	19	37
23	28	13	48	20	20	73	22	24	98	31	26	123	23	29	148	21	32	173	3	35	198	20	37
24	34	13	49	24	20	74	25	24	99	33	26	124	24	29	149	22	32	174	13	35	199	22	37
25	13	14	50	25	20	75	28	24	100	12	27	125	28	29	150	22	32	175	17	35	200	2	38

Figura A.1: Números inteiros de 1 a 40 que geram os pontos do AP200

empregados símbolos maiores e mais escuros nessa figura. Nesses dez casos, as localidades estão praticamente juntas, o que dá a essas dez localidades um "peso" maior na atribuição de hubs a localidades. Isso porque a escolha de uma das duas localidades vizinhas como sendo um hub permite a redução da distância percorrida pelo tráfego nos dois sentidos entre ambas e todas as demais localidades, devido à minimização da etapa de coleta, agindo as duas quase como se constituíssem um "hub duplo". De fato, na instância AP200, considerando todos os 200 pontos, qualquer que seja o valor de p , há pelo menos uma das localidades que formam esses pares dentre as escolhidas como hubs, embora elas constituam só 10% do total de localidades. Por exemplo, no caso de $p = 3$, dois dos três hubs escolhidos na configuração ótima são de um desses dez pares de localidades; quando $p = 4$, são três entre os quatro e quando $p = 7$ são quatro entre os sete. Esse número só não cresce na mesma proporção que p cresce porque a maioria desses pares de localidades estão concentradas numa área pequena, como ilustrado pela figura A.2, que mostra a distribuição de todas as localidades do AP200 no espaço.

Ernst e Kryshnamoorthy propõem um algoritmo para gerar outras instâncias com $m < 200$, dividindo a área total em sub-áreas menores, e usando uma só localidade para representar cada uma das sub-áreas. A essas localidades sintéticas são atribuídas coordenadas, calculadas a partir das coordenadas das localidades ori-

ginais em cada sub-área, e fluxos de entrada e saída correspondentes à agregação de toda a demanda com destino e origem nas localidades situadas na sub-área. Os autores distribuíram um programa em C para executar tal algoritmo e, através dele, geraram instâncias com $m = 100, 50, 40, 25, 20$ e 10 . Nessas instâncias, desaparecem os pares de localidades muito próximas, sendo interessante notar que alguns desses pares são separados pela divisão em sub-áreas, passando cada uma das duas localidades vizinhas a ser considerada parte de uma sub-área diferente.

Uma das características importantes da instância AP200 e das instâncias derivadas da mesma é que, ao contrário da maioria dos estudos feitos até então, os fluxos de uma localidade para si mesma não são nulos, isto é $w_{ii} > 0$. Os autores justificam tal escolha dizendo que através desse custo é possível considerar o custo de coleta e distribuição de correspondência no próprio distrito postal (*"this allows us to compute the cost of the collection and distribution of mail within the same postal district"*). Este custo é certamente relevante e deveria ser considerado no custo total de operação do sistema de correios. Ou seja, deveriam ser totalizados os custos de coleta e distribuição local de todos distritos e esse total deveria ser somado ao custo das transferências entre os distritos.

Entretanto, em cada distrito, parece intuitivo que este deveria ser um custo associado somente a fatores internos ao distrito e, em princípio, não deveria depender da transferência de correspondência para outros distritos. Assim, em cada distrito, ele deveria equivaler a uma parcela fixa constante, cujo soma pode ser calculada a priori para a totalidade dos distritos, independentemente da escolha das trajetórias a serem usadas para as transferências entre os distritos. Dessa maneira, esses custos não deveriam influenciar a procura das melhores trajetórias para transferências entre distritos.

Entretanto, no estudo de Ernst & Krishnamoorthy em que o AP é proposto e em diversos outros estudos encontrados na literatura e baseados em tais dados (como o de Simanovic, o de Malinovic e o de Marti et al., etc.), é dada uma interpretação diferente: o fluxo w_{jj} de transferência do distrito para ele próprio é tratado da mesma maneira que o fluxo entre dois distritos i e j quaisquer, quando $i \neq j$. Isso significa que são computados os custos de transferência do distrito para o hub e de volta do hub para o distrito, o que significa que toda a correspondência originada no distrito e destinada ao próprio distrito tem de passar primeiro por um hub fora do distrito antes de chegar a seu destino.

Essa interpretação só faz sentido se os distritos não tiverem nenhuma capacidade de separação e classificação da correspondência que recebem, tendo de recorrer sempre a um hub para tanto, mesmo com relação à correspondência interna do distrito. Isso encarece e retarda a distribuição local e tem como consequência o fato de que o custo da mesma passa a depender da distância ao hub mais próximo. A

escolha de um hub mais distante passa a ter impacto no custo da distribuição de correspondência interna ao município, o que parece pouco intuitivo. É possível, entretanto, que os distritos postais prefiram trabalhar dessa maneira, dirigindo a totalidade da correspondência coletada para o hub, inclusive as de destino local, e distribuindo apenas a correspondência recebida do hub, mesmo quando sua origem está no próprio distrito.

A interpretação descrita não é unânime, e muitos autores consideram que $w_{jj} = 0, \forall j, j = 1, \dots, m$. Isso torna mais difícil a comparação dos resultados obtidos por estudos diferentes, sobretudo quando não deixam claro se estão considerando $w_{jj} = 0$ ou não.

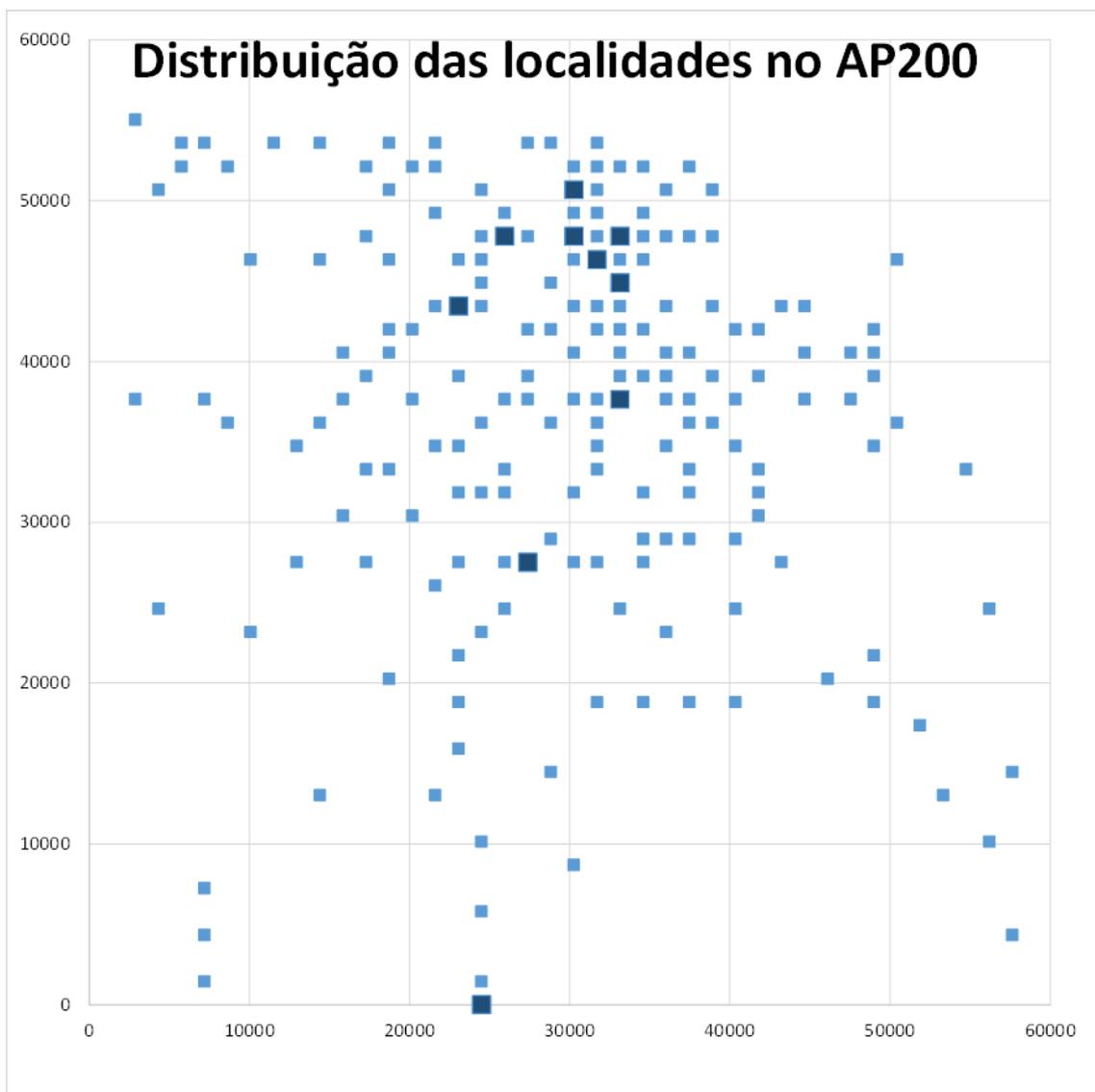


Figura A.2: Distribuição espacial dos pontos da instância AP200, incluindo dez pares de pontos muito próximos, assinalados em cor escura e com quadrados maiores