

O USO DE *TABU SEARCH* NA CLUSTERIZAÇÃO DE PONTOS DE DEMANDA: UM MODELO INTEGRANDO NÍVEL DE SERVIÇO E RESTRIÇÕES OPERACIONAIS

Arnaldo Rabello de Aguiar Vallim Filho

Universidade Presbiteriana Mackenzie

Faculdade de Computação e Informática

RESUMO

Este artigo apresenta uma metodologia baseada na meta-heurística *Tabu Search* tendo como foco a solução de um problema logístico de “clusterização”. O modelo foi empregado em um problema real de grande porte que busca a montagem de distritos logísticos de pontos de demanda de carga distribuídos pelos três estados da região sul do Brasil. A pesquisa procurou desenvolver uma modelagem matemática que integrasse requisitos de nível de serviço e restrições operacionais ao processo de clusterização, fato que não é comum na literatura. Normalmente os métodos se preocupam apenas com a montagem de clusters, sem considerações desta natureza. A metodologia aqui proposta considera estes elementos e uma vez que tenham sido devidamente tratados, desenvolve o processo de clusterização, que neste caso, teve por base a metaheurística *Tabu Search* (TS). A metodologia associa um dado nível de serviço e restrições operacionais a um número de clusters necessário para atender a esses requisitos. Uma vez definido este número, é gerada uma clusterização inicial que será entrada para um modelo baseado em TS, que procurará chegar a uma configuração final “otimizada”. Este artigo procura desenvolver e estruturar esta metodologia e avaliar diferentes estratégias e alternativas para seus parâmetros.

ABSTRACT

This paper presents a methodology based on the metaheuristic *Tabu Search* focusing on the solution of a logistic clustering problem. The model was employed on a real-world large-scale logistics districting problem, with demand points spread out over the three states of the Brazil southern region. The research tried to develop a mathematical modeling to integrate service level requirements and operational constraints to the clustering process, which is not common in literature. Usually the methods are concerned only with the clustering process, with no considerations of this nature. The proposal methodology considers these elements and once they had been treated, develops the clustering process, which, in this case, was based on the metaheuristics *Tabu Search* (TS). The methodology associates one given level of service and operational constraints to a number of clusters necessary to satisfy those requirements. Once this number is defined, an initial clustering is generated, which will be an input for a TS based model, that will try to achieve an “optimized” final configuration. This paper intends to develop and to structure this methodology, as well, to evaluate different strategies and alternatives for its parameters.

1. INTRODUÇÃO

Dependendo do porte de uma operação logística o número de pontos de demanda pode chegar a centenas e em muitos casos aos milhares. Devido a esse grande volume, o tratamento individual de cada ponto só realmente ocorre nos processos operacionais finais em que se deve efetivamente definir como será efetuado o atendimento de cada demanda específica. Estudos voltados para planejamento não tratam, em princípio, clientes e/ou fornecedores de forma individual (Ballou, 2006). Em geral, os clientes são agregados em clusters, denominados distritos, que devem ser atendidos por uma instalação logística. A questão da definição desses distritos constitui-se assim, em etapa fundamental para todo o planejamento. O estudo de critérios e métodos para o estabelecimento desses distritos surge, assim, como um importante ramo de pesquisa em Logística. E é nesta linha que segue o desenvolvimento deste

trabalho. Por outro lado, a pesquisa aqui desenvolvida procura tratar aspectos que nem sempre são considerados na solução deste problema. Procura-se aqui, desenvolver uma modelagem matemática que integra requisitos de nível de serviço e restrições operacionais ao processo de “clusterização”, fato que não é comum na literatura. Normalmente os métodos partem diretamente para a montagem de clusters, sem considerações desta natureza. A metodologia aqui proposta considera estes elementos e uma vez que tenham sido devidamente tratados, desenvolve o processo de clusterização, que neste caso, teve por base a metaheurística *Tabu Search* (TS). A metodologia associa um dado nível de serviço e restrições operacionais a um número de clusters necessário para atender a esses requisitos. Definido este número é gerada uma clusterização inicial que será entrada para um modelo baseado em TS, que procurará chegar a uma configuração final “otimizada”. Este artigo procura desenvolver e estruturar esta metodologia e avaliar diferentes estratégias e alternativas para seus parâmetros. Na próxima seção é apresentada uma revisão da literatura disponível sobre o assunto de forma a alicerçar o desenvolvimento metodológico que foi conduzido; a seção seguinte é a fase central da pesquisa, onde é apresentada a metodologia proposta; a seção 4 é dedicada à aplicação prática da metodologia e, finalmente, na seção 5 são apresentadas as conclusões e recomendações do estudo.

2. REVISÃO BIBLIOGRÁFICA

A questão da clusterização é um problema antigo e não é exclusivo da Logística, tendo uma abrangência bem ampla, pois, pode ocorrer em diversas áreas de atividade, tais como: ciências sociais, psicologia, agricultura, biologia, produção industrial, gestão empresarial, geometria computacional, mineração de dados e outras (Eaton, 1983; Thorndike, 1978). Mais recentemente com o crescimento das aplicações de computação o assunto ganhou um novo impulso, em particular com o crescimento das áreas de Mineração de Dados e Geometria Computacional.

Do ponto de vista formal, poder-se-ia dizer que “clusterização” é o processo de organizar elementos em grupos cujos membros são similares de alguma forma (Matteucci, 2006). O conceito básico do problema é de que os elementos de um cluster devem possuir um grau de similaridade maior com os outros elementos de seu próprio cluster do que com indivíduos de outros clusters. Assim, pode-se ter o problema de K-Clusterização que ocorre quando K é pré-definido. Já quando K não é previamente conhecido, tem-se o Problema de Clusterização Automática (PCA). O que caracteriza este tipo de problema em mineração de dados, é que nesta área a base de dados costuma ter um porte elevado e o número de atributos de cada elemento é, em geral, alto (Ochi et al., 2005; Oliveira, 2005).

Do ponto de vista de complexidade de algoritmos, os algoritmos de otimização para montagem de clusters se encontram na mais alta classe de complexidade (classe NP-Difícil). Assim, uma pesquisa em todo o espaço de soluções é inviável, dado o número de possibilidades. Desta forma, as soluções são alcançadas através de métodos heurísticos ou por meio de metaheurísticas. Sendo que os principais métodos são do tipo hierárquico ou de particionamento. O tipo hierárquico pode ser de aglomeração ou de divisão. Na estratégia hierárquica de aglomeração (*bottom-up*) pode-se iniciar com cada elemento se constituindo em um cluster, assim, cada cluster que tiver um tamanho maior que 1 terá sido composto por um conjunto de clusters. A união de clusters sempre ocorre com base em alguma medida de

similaridade. Na estratégia hierárquica de divisão (*top-down*), se inicia com um só cluster e procede-se a uma seqüência de divisões, com base em algum critério.

Quando o método é de particionamento, parte-se de um conjunto “p” de clusters e procede-se a um intercâmbio de elementos entre os clusters já formados. A cada rearranjo dos clusters tem-se uma nova configuração. Para se avaliar se a configuração deve ser aceita ou não é preciso uma função objetivo. Esta função é de grande importância, pois é fator decisivo do processo, devendo ser criteriosamente definida, de forma a ser efetivamente representativa do objetivo que se busca alcançar. No que tange especificamente à área Logística as estratégias hierárquica e de particionamento têm sido usadas com freqüência (Ballou, 1991b; Novaes et al., 2000; Zhou et al., 2002; D’Amico et al., 2002; Vallim Fº., 2004).

Em Logística o problema de clusterização, costuma ser identificado pelo termo *districting*. Estando a questão presente principalmente, em problemas de distribuição e também em localização de instalações. Neste caso, não é incomum ter que se trabalhar com centenas de locais candidatos a instalações e milhares de clientes o que leva naturalmente a um processo de agregação desses pontos. Do ponto de vista de modelagem matemática, tais pontos são representados por um conjunto de nós em uma rede logística.

Francis et al. (1999) apresentam algumas razões para a “clusterização”. O primeiro aspecto a se considerar é o esforço computacional para resolver um modelo em um problema de localização. Conforme já dito a maior parte dos problemas é do tipo NP-Difícil (Daskin, 1995), o que leva pesquisadores a procurar formas para reduzir as dimensões do problema. As aplicações com número muito grande de nós de demanda continuam inviáveis sob o ponto de vista computacional quando os nós são tratados de forma desagregada. Um outro aspecto diz respeito à representação espacial da demanda. Na verdade, em problemas de médio e longo prazo, uma representação espacial do *padrão de demanda*, mais do que a demanda detalhada em si, é que determina os resultados do problema real no futuro. Assim, em termos de resultados a se obter, é melhor tratar com uma demanda clusterizada, do que com cada nó individual.

Sobre o uso de métodos heurísticos ou de metaheurísticas aplicados à Logística há vários artigos na literatura. Ballou (1991b) sugere uma heurística para a montagem de *clusters* que é baseada na proximidade geográfica dos pontos de demanda. A heurística parte da idéia de que em princípio cada ponto se constitui em um *cluster*. Em seguida procura-se para cada ponto o *cluster* mais próximo e destes dois gera-se um novo. Este processo se repete até que um número desejado de *clusters* seja atingido. As coordenadas de um *cluster* corresponderão às coordenadas do centro de gravidade entre seus pontos.

Novaes (1989) também estudou a questão de distribuição urbana e através de uma heurística desenvolve a subdivisão de uma área de estudo criando um conjunto de anéis a partir de quadriculas elementares. Posteriormente, Novaes e Graciolli (1999) propuseram um novo procedimento para criação de *clusters* destinados a uma operação de distribuição de carga a ser efetuada em uma certa região de estudo. No modelo proposto a região é particionada em setores a partir de um depósito central. Os autores utilizam aqui o mesmo conceito utilizado por Novaes (1989) de quadriculas elementares. Novaes et al. (2000) desenvolvem um outro trabalho nessa linha, gerando distritos homogêneos em termos de esforço de distribuição. É desenvolvida uma heurística para geração dos distritos e a solução é melhorada por meio de

um algoritmo genético associado a um método de otimização direta, o chamado método do gradiente (Novaes, 1978).

Zhou et al. (2002) desenvolvem com o apoio de algoritmos genéticos um processo de geração de zonas de atendimento de centros de distribuição. O modelo busca o balanceamento de custo de transporte e nível de serviço e segundo os autores os resultados foram satisfatórios. Outro trabalho nesta linha é o de D'Amico et al. (2002) que apresenta um redesenho de jurisdições policiais em Buffalo, New York. O artigo mostra o uso da metaheurística *simulated annealing*, que partindo de uma solução inicial pesquisa novas soluções que minimizem a heterogeneidade entre as cargas de trabalho dos policiais.

Vallim F^o e Gualda (2003) apresentaram uma heurística de clusterização de pontos de demanda de forma a identificar nesses clusters locais que se constituíssem em “melhores candidatos” a centros de distribuição. Estes candidatos são utilizados como entrada de modelos de otimização para a definição final dos locais. Em um outro artigo Vallim F^o e Gualda (2004) trataram a localização de instalações a partir de clusters com o uso de um modelo de otimização combinado à metaheurística *simulated annealing*. Vallim F^o (2004) combinou e aprimorou estas idéias fazendo uso de *simulated annealing*, para geração dos clusters e identificando a “melhor posição teórica” (MPT) em cada cluster para a localização de um centro de distribuição. Uma pesquisa era feita em torno desta MPT para identificar na rede, nós operacionalmente viáveis para a implantação da instalação.

Outra metaheurística utilizada em clusterização é *Tabu Search* (TS), que foi proposta por Glover (1986) e inicialmente foi aplicada a problemas de programação matemática, expandindo-se posteriormente para diversas áreas de aplicação, inclusive clusterização. Dentre as principais aplicações pode-se citar o problema clássico de partição de grafos que tem sido resolvido por esta técnica (Pirlot, 1996; Glover, 1990). Células de manufatura, que podem ser definidas por clusterização têm sido resolvidas por TS. O problema de roteirização, em que cada rota representa um cluster, é outro problema logístico que também tem sido resolvido por *tabu search* (Crainic e Laporte, 1998).

Mais recentemente Furuta et. al (2005) fazem uso de uma técnica muito usada em geometria computacional, o diagrama de Voronoi, para resolver um problema de *districting* para alocação de ambulâncias. Galvão et. Al (2006) apresentaram um artigo em que resolvem um problema de *districting* fazendo uso do mesmo diagrama de Voronoi. Fazem isto, partindo de uma “clusterização” na forma de anéis subdivididos em setores e vão relaxando as fronteiras iniciais dos clusters. O processo é iterativo e vai transcorrendo até que uma convergência seja atingida. Como se vê há uma literatura diversificada sobre heurísticas e meta heurísticas aplicadas à clusterização em Logística. Esta pesquisa caminha nesta mesma linha, procurando apresentar uma contribuição à área, agora integrando outros elementos ao processo.

3. DESENVOLVIMENTO DA MODELAGEM

Esta seção apresenta inicialmente a metodologia de incorporação de nível de serviço logístico e de restrições operacionais ao processo de clusterização. Em seguida é apresentada a metaheurística escolhida para ser utilizada neste trabalho.

3.1 Nível de serviço e restrições operacionais

A metodologia aqui proposta considera que o processo de clusterização deve ser desenvolvido de forma que um Nível de Serviço logístico (NS) previamente estabelecido seja atendido e que eventuais Restrições Operacionais sejam contempladas.

Na presente metodologia propõe-se que NS seja definido em termos de um raio máximo de atendimento, definido por meio de um tempo máximo admissível para se atingir um ponto de demanda a partir de uma instalação logística, e como restrição operacional, será adotada uma área média desejada por instalação. Com isto é possível se chegar a um número necessário de clusters para a área de estudo. O processo completo se desenvolve em três passos:

Passo 1: Determinação da Distância Média de Atendimento com base em NS

Define-se um NS baseado no Tempo de Viagem (T_α) da instalação logística ao ponto de demanda. O tempo T_α será o tempo máximo admissível de viagem, a um nível de confiança definido de $1-\alpha$. Assim, para um tempo T a probabilidade de $T \leq T_\alpha$ pode ser expressa por:

$$P(T \leq T_\alpha) = 1 - \alpha \quad (3.1)$$

É possível agora se calcular a distância (D_α) correspondente a T_α utilizando-se uma velocidade de cruzeiro (V):

$$D_\alpha = T_\alpha \cdot V \quad (3.2)$$

Com D_α pode-se determinar a distância média de viagem (D_m), pois uma vez que se assuma $D \sim N(D_m, S_d)$, tem-se:

$$D_\alpha = D_m + S_d \cdot Z_\alpha \quad (3.3)$$

sendo: S_d = desvio padrão das distâncias;

Z_α = valor crítico da distribuição Normal ao nível $1-\alpha$

Sendo C_v o coeficiente de variação da distância, tem-se: $S_d = C_v \cdot D_m$, (3.4)

Logo: $D_\alpha = D_m + C_v \cdot D_m \cdot Z_\alpha$ (3.5)

Portanto: $D_m = D_\alpha / (1 + C_v \cdot Z_\alpha)$ (3.6)

Passo 2: Restrições Operacionais

Neste segundo passo, estabelece-se uma Área Média por instalação logística ($A_{\text{média}}$) baseada em restrição operacional. Assim, o volume médio de armazenagem por instalação ($V_{\text{médio}}$) será dado por:

$$V_{\text{médio}} = A_{\text{média}} \cdot h \cdot \eta_A \quad (3.7)$$

onde : $A_{\text{média}}$ = área média da instalação;

h = pé direito ;

η_A = índice de aproveitamento de área construída;

O volume médio útil por instalação ($V_{\text{médio}}$) é uma parcela do volume disponível para armazenagem, que em geral não é totalmente utilizado. Assim, tem-se:

$$V_{\text{UTIL}} = V_{\text{médio}} \cdot \eta_V \quad (3.8)$$

onde : η_V = índice de aproveitamento do volume de armazenagem.

O volume médio mensal atendido por instalação (V_{CD}) será este volume útil multiplicado pelo giro mensal de estoque (G):

$$V_{CD} = V_{\text{UTIL}} \cdot G \quad (3.9)$$

Passo 4: Número de Clusters

Determina-se inicialmente o momento médio de transporte por instalação (TKM_{CD}) através de:

$$TKM_{CD} = V_{CD} \cdot D_m \quad (3.10)$$

O Momento de Transporte é definido como sendo a multiplicação do peso ou volume movimentado no percurso pela distância percorrida.

A partir de TKM_{CD} , tem-se o número necessário de instalações (N_{CD}) através de:

$$N_{CD} = TKM_{\text{Total}} / TKM_{CD} \quad (3.11)$$

onde: TKM_{tot} = momento de transporte total da área de estudo;

Este momento de transporte total da área de estudo deve ser computado com base na demanda dos pontos a atender e na posição destes pontos.

O número necessário de clusters (N_{clusters}) será igual ao número necessário de instalações. A hipótese adotada é de uma instalação para cada cluster. Portanto, tem-se:

$$N_{\text{clusters}} = N_{CD} \quad (3.12)$$

Este é o número de clusters que atende ao NS desejado e às restrições operacionais impostas.

Na próxima etapa desenvolve-se um processo de clusterização para gerar este número de clusters.

3.2 Metaheurística *tabu search*

A metaheurística *Tabu Search* (TS) foi utilizada na montagem dos clusters. Esta metaheurística desenvolve todo o seu procedimento em torno de três princípios básicos (Viana, 1998): a chamada Lista Tabu, que guarda as últimas soluções; o mecanismo de aceitação ou não de uma solução com base na lista tabu e a alternância entre estratégias de diversificação e intensificação de soluções

O TS parte de uma configuração inicial para o sistema, que pode ter sido definida por algum outro método ou pode ter sido simplesmente adotada para dar início ao processo. Deve-se também definir uma função objetivo que se deseja otimizar. A partir da configuração inicial procede-se a Perturbações nesta configuração de forma a se desenvolver uma pesquisa de outras soluções em toda a sua vizinhança ou em uma sub-vizinhança, se esta for muito ampla. Uma vez pesquisada a vizinhança, deve-se identificar a melhor solução dentre todas testadas. De forma a se evitar situações cíclicas, diversificando a pesquisa de soluções, TS introduz o conceito de Lista Tabu. Mantêm-se nesta lista as últimas soluções encontradas, e a melhor solução de uma vizinhança só é aceita se não pertencer à lista. O TS oferece ainda, um outro mecanismo que desconsidera a lista tabu para aqueles casos em que se atinge uma “solução suficientemente boa”. Esta “solução suficientemente boa” é definida com base na chamada função de Aspiração ou Nível de Aspiração. Esta é uma função que define quando uma solução deve ser aceita, independentemente do fato desta estar ou não na lista tabu. Para que o

TS possa ser implementado, algumas decisões precisam ser tomadas: a especificação da estrutura da vizinhança; a escolha dos atributos a registrar na lista tabu; o tamanho da Lista Tabu; a escolha de um critério de aspiração; a escolha de um critério de parada (número total de iterações); a definição de uma função objetivo.

A literatura oferece orientações sobre estes parâmetros, mas a sua calibração é feita por meio da experimentação. Sobre o algoritmo do TS a figura 3.1 apresenta a idéia básica da sua estrutura.

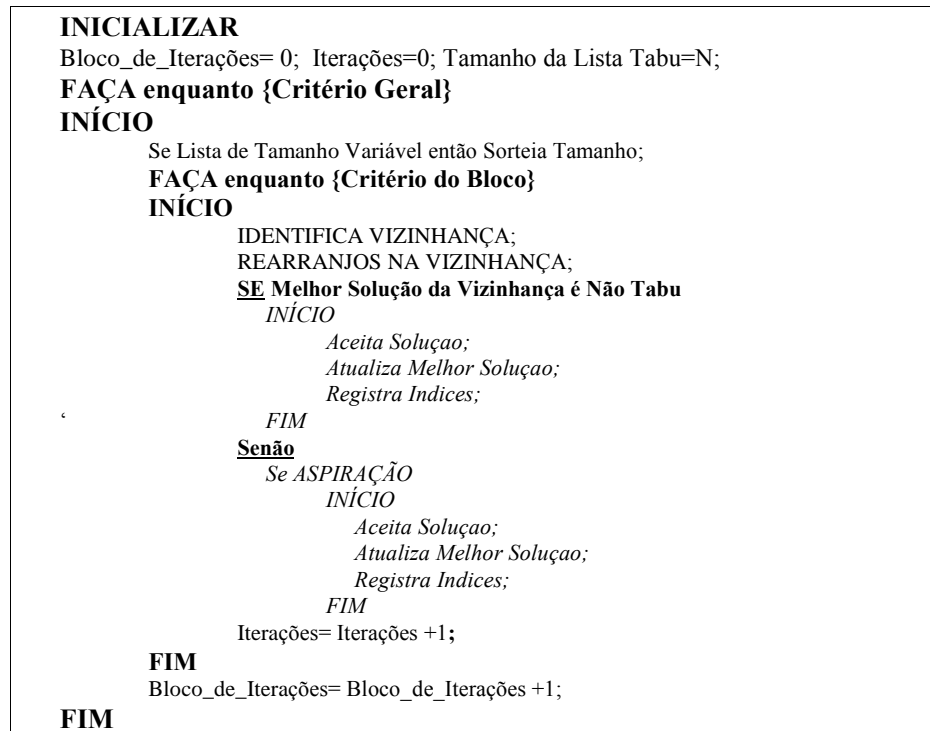


Figura 3.1 – Estrutura Básica do Algoritmo de *Tabu Search*

3.3 *Tabu search* proposto

O modelo aqui proposto parte de uma solução inicial que pode ter sido simplesmente adotada ou então gerada por alguma heurística específica para montagem de clusters. Vallim Fo (2004) propõe uma heurística com tal finalidade. Foi adotada a estratégia de particionamento para a montagem dos clusters, em que a partir da configuração inicial se inicia um processo de migração de pontos entre os clusters de forma a se buscar melhores soluções. Uma solução vizinha de uma dada solução foi definida como sendo aquela mesma configuração de clusters, exceto por um único ponto que foi re-allocado para um cluster adjacente. Para um dado cluster, os clusters adjacentes correspondem aos seus clusters anterior e posterior dentro de um processo de varredura dos pontos no sentido anti-horário. Para isto, os pontos dentro de cada cluster são ordenados segundo o seu ângulo com o eixo horizontal, tendo-se assim, bem definidos o primeiro e último ponto de cada cluster. Somente um destes pontos pode ser transferido para o cluster adjacente. Esta definição de vizinho tem como resultado que a cada iteração todas as soluções da vizinhança serão testadas, fazendo-se com que todos os pontos na fronteira entre dois cluster adjacentes mudem de um cluster para outro.

Sobre a função objetivo adotou-se a mesma proposta por Vallim Fo (2004), que é definida pelo grau de homogeneidade (GH) dos clusters, conforme 3.13:

$$GH = \begin{cases} \left(1 - \frac{s_{TKM}}{TKM_m}\right) \cdot 100, & \text{se } s_{TKM} \leq TKM_m \\ 0, & \text{em caso contrário} \end{cases} \quad (3.13)$$

onde:

s_{TKM} = desvio padrão dos momentos de transporte dos clusters
 TKM_m = média do momento de transporte dos clusters

Foi adotada esta função objetivo, pois o que se busca neste processo de “clusterização” é uma uniformização do nível esperado de produção de transporte em cada cluster. Com isto o planejamento passa a sofrer menos de condições de irregularidades, os recursos podem ser homogeneamente distribuídos e, conseqüentemente, os custos das operações são reduzidos.

O TKM já é uma medida consagrada de produção de transporte pois envolve os dois fatores chave do transporte: peso (ton) e distância (km). Ao se “produzir” transporte está se percorrendo uma distância e movimentando uma certa quantidade de carga. Medir o transporte apenas pela distância percorrida ou apenas pelo peso movimentado é uma métrica incompleta. Assim, a forma adequada de se proceder a essa medida é através de um indicador que relacione as duas variáveis. Como restrição adicional estabeleceu-se um mínimo de pontos que deveria ser mantido em cada cluster, evitando-se assim a ocorrência de clusters muito pequenos.

4. APLICAÇÃO PRÁTICA

A aplicação da metodologia proposta foi baseada em um caso real com duzentos e vinte e nove pontos de demanda (Vallim Fo, 2004) distribuídos pelos três estados do sul do país, movimentando à época do levantamento um volume em torno de 45.000 m³ por mês com uma frota de cerca de cem veículos. Nesta operação a frota que atendia aos pontos estava rodando cerca de 800.000 quilômetros mensais, com algo próximo de 1.100 viagens.

O processo se inicia pela incorporação de nível de serviço e restrições operacionais ao processo de definição do número de clusters. Conforme visto na seção 3.1, em seu passo 1 este processo requer a definição do tempo máximo admissível desde a instalação logística até um ponto qualquer de demanda, para um nível de confiança desejado. Foi adotado um tempo máximo (T_α) de quatro horas a um nível de confiança ($1 - \alpha$) de 95%. Adotando-se uma velocidade de cruzeiro de 60 km/h obtém-se a distância correspondente $D_\alpha = 240$ km.

A partir de uma varredura dos pontos da região Cv foi estimado em 1,032 e tendo-se $Z_\alpha = 1,645$, utiliza-se a equação 3.6 para se obter a distância média:

$$D_m = 240 / (1 + 1,032 \cdot 1,645) = 89,0 \text{ km}$$

Aplicando-se o passo 2 foi estabelecida uma área média desejada ($A_{\text{média}}$) de 2.000m² e um pé direito (h) de 7,5m. O aproveitamento de área construída (η_A) na atual operação é de 55% e com o uso da equação 3.7 obtém-se o volume médio: $V_{\text{médio}} = 2000 \cdot 7,5 \cdot 0,55 = 8.250 \text{ m}^3$

O volume útil de armazenagem é obtido aplicando-se a $V_{\text{médio}}$ um índice de aproveitamento. Não havia dados disponíveis sobre este coeficiente. Foi adotado então o valor de 76%, que é a

média para o setor, conforme pesquisa efetuada pelo Centro de Estudos Logísticos (CEL/COPPEAD, 2001). Com isto, tem-se: $V_{\text{UTIL}} = 8250.0,76 = 6.270\text{m}^3$

O volume mensal atendido por instalação (V_{CD}) será este valor multiplicado pelo giro médio mensal. O giro atual da empresa em que foi baseada a aplicação prática é de 1,05. Desta forma, tem-se: $V_{\text{CD}} = 6270. 1,05 = 6.584\text{m}^3$. No passo 3 determina-se o momento de transporte por instalação, a partir do volume atendido e da distância média aos pontos de demanda, tem-se:

$$\text{TKM}_{\text{CD}} = 6584. 89,0 = 585.869 \text{ m}^3.\text{Km}$$

O número de instalações é obtido através da equação 3.11, em que se divide o momento de transporte total da área de estudo ($\text{TKM}_{\text{Total}}$) pelo momento de transporte por instalação (TKM_{CD}). A partir da varredura dos pontos da área de estudo obteve-se:

$$\text{TKM}_{\text{Total}} = 3.790.481 \text{ m}^3.\text{Km}.$$

Assim, tem-se:

$$N_{\text{CD}} = 3.790.481 / 585.869 = 6,47$$

Adotou-se então o número de sete instalações e mantém-se este mesmo valor para o número de clusters ($N_{\text{clusters}} = 7$). O processo de clusterização deverá buscar, assim, a formação de sete clusters. Uma vez definido este número, inicia-se a aplicação de TS para a montagem dos clusters. A figura 4.1a mostra a evolução de TS ao longo das iterações, para o caso de lista tabu com tamanho fixo. Note-se que GH varia durante todo o tempo dentro de padrões bastante elevados. Seu valor mais baixo foi de 93,3% (somente a solução inicial teve valor menor). Outro aspecto a salientar é o processo de convergência que se observa, com as soluções de forma *quasi* cíclica sempre convergindo para a solução “ótima”.

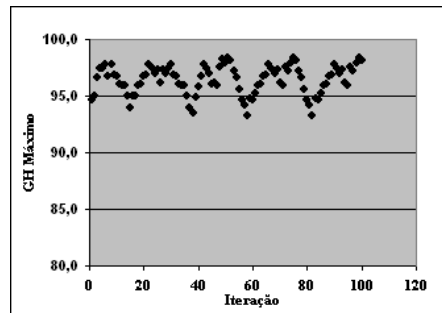


Figura 4.1a – Evolução de GH
Lista Tabu de tamanho Fixo

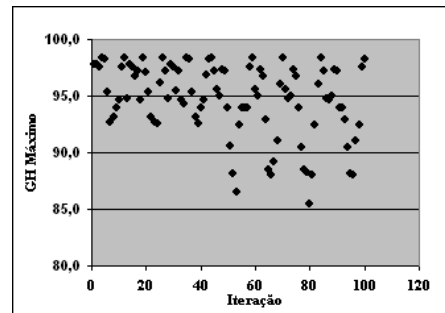


Figura 4.1b – Evolução de GH
Lista Tabu de tamanho Variável

Na figura 4.1a observa-se de modo claro este processo. Percebe-se, inclusive que o número de iterações poderia até ter sido menor. A solução “ótima” foi atingida pela primeira vez na iteração de número 51, voltou a ocorrer na iteração 75 e, finalmente, se repetiu pela última vez na iteração número 98. Uma vez que os tempos de processamento estavam muito baixos, optou-se por manter o número de cem iterações por razões de segurança. Com isto procura-se garantir que a solução “ótima” atingida é efetivamente a melhor solução que o TS pode fornecer. Para o caso de lista tabu com tamanho variável, o experimento foi replicado dezenas de vezes, já que a cada iteração é sorteado o tamanho da lista. Para as várias replicações invariavelmente as soluções convergiram para a solução “ótima”. O desenvolvimento deste processo para uma das replicações é apresentado na figura 4.1b. Pode-se verificar pelas

figuras que o processo agora assume um comportamento mais irregular, quando comparado ao caso de lista tabu de tamanho fixo. Não se verifica mais aquele comportamento *quasi* cíclico que ocorria no caso anterior. Percebe-se agora que há muito mais diversificação das soluções. Esta mudança, entretanto, não significa que aqui o processo do TS foge de outra de suas características básicas que é a intensificação. Na verdade, apesar da grande diversificação, verifica-se ao mesmo tempo um aprofundamento do mecanismo de intensificação, pois as melhores soluções passam a ser visitadas com maior frequência. Para esta replicação apresentada, a solução “ótima” é alcançada pela primeira vez na iteração número 4, voltando a ocorrer nas iterações de número, 12, 19, 26, 35, 44, 59, 70 e 84. Soluções próximas à “ótima” também são visitadas várias vezes.

Quanto à diversificação na busca de soluções, o processo percorre soluções inferiores não tentadas quando a lista tabu é de tamanho fixo. Além disso, amplia a faixa de variação, chegando a um GH de 85,5%, para esta replicação. As figuras 4.2a e 4.2b apresentam uma comparação entre a dispersão dos dois processos, onde pode-se verificar a maior abrangência da lista tabu de tamanho variável. Esta estratégia permite assim, que os fundamentos do TS sejam executados de forma mais intensa. Tem, por outro lado, a desvantagem de que seu tempo de processamento passa a ser mais elevado.

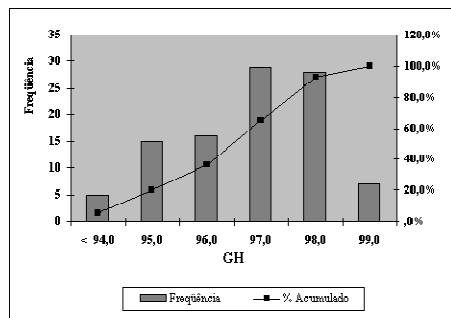


Figura 4.2a – Distribuição de GH
Lista Fixa

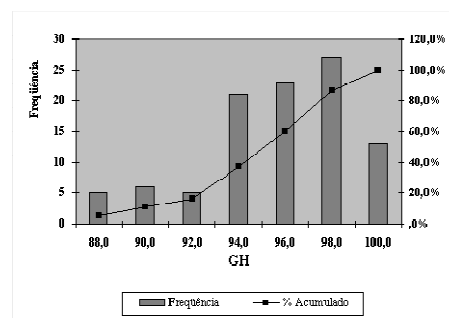


Figura 4.2b – Distribuição de GH
Lista Variável

O experimento com TS foi rodado em um computador com processador Intel, Pentium 4, com *clock* de 2,8 GHz e memória RAM de 256 MB. Com esta configuração o tempo médio de processamento foi de 4 segundos, com lista tabu de tamanho fixo, e 11 segundos, com lista tabu de tamanho variável. Quanto aos clusters gerados na solução “ótima”, estes são apresentados na figura 4.3. A metaheurística conseguiu uma evolução no GH de +5,3%, passando de 93,1% (solução inicial) para 98,4%. Pode-se considerar assim, que o TS mostrou-se bastante adequado atingindo um nível de desempenho dentro de padrões satisfatórios.

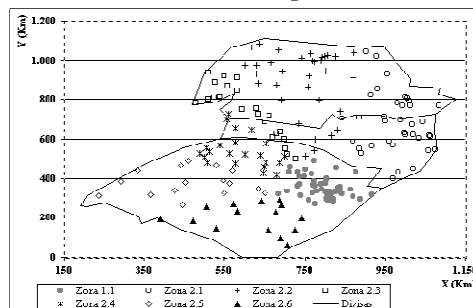


Figura 4.3 – Clusterização – Solução Final

5. CONCLUSÕES E RECOMENDAÇÕES

Esta pesquisa procurou dar uma contribuição ao problema de clusterização de pontos de demanda em uma operação logística procurando contemplar na solução do problema certos aspectos nem sempre abordados na literatura.

Sobre os seus resultados, inicialmente pode-se dizer que a primeira parte da metodologia aqui proposta, cumpriu com seus objetivos, uma vez que ficou delineado um procedimento claro para quantificar e considerar nível de serviço e restrições operacionais na solução de um problema de clusterização. Tendo ainda, possibilitado a integração destes aspectos a uma metaheurística. Chegou-se assim, a uma proposta concreta de incorporação destes importantes aspectos ao processo de clusterização. Estes são pontos que podem ser de alguma valia no desenvolvimento de pesquisas futuras na área.

Quanto ao uso da metaheurística *Tabu Search* pode-se dizer que se mostrou eficaz na solução do problema, uma vez que gerou uma solução “ótima” para uma aplicação prática que é muito próxima do máximo valor que poderia ter sido obtido para a função objetivo (GH de 98,4% vs. máximo de 100%).

Sobre as táticas e estratégias empregadas no TS têm-se algumas conclusões importantes. No princípio, trabalhou-se com uma lista tabu que mantinha todas as informações das soluções. Esta estratégia, confirmando comentários de outros autores, não se mostrou eficaz, pois não conduziu o processo à solução “ótima”. Melhor do que manter toda a solução, é utilizar atributos que caracterizem esta solução. Com isto, consegue-se uma maior diversificação da busca por alternativas de configurações para o sistema, explorando-se de forma mais abrangente o espaço de soluções. Outro aspecto a lembrar é sobre o tamanho da lista tabu. Nesta pesquisa foi testada uma lista de tamanho fixo e também uma de tamanho variável. As duas estratégias levaram à solução “ótima”, porém, a lista de tamanho variável teve uma melhor performance em termos de exploração do espaço de soluções, o que sugere que, eventualmente, em problemas mais complexos esta seja uma alternativa melhor para se alcançar a solução “ótima”. Por outro lado, a lista de tamanho fixo tem um tempo de processamento menor, o que eventualmente, pode não ser relevante, se os tempos forem da ordem de segundos.

Como recomendações para futuros trabalhos, um aspecto que mereceria destaque seria a ampliação do conceito de nível de serviço, possivelmente considerando outros indicadores além do raio máximo, bem como, de outras restrições operacionais. Por outro lado, o uso de outras metaheurísticas para a solução do problema seria também importante. Para a logística do país é fundamental que se desenvolva plena capacitação no rol completo de metaheurísticas apresentadas na literatura, para solucionar este problema de “clusterização”, que tem muitas aplicações não só em logística, mas também em diversas outras áreas.

REFERÊNCIAS BIBLIOGRÁFICAS

- Ballou, R. H. (2006). *Gerenciamento da Cadeia de Suprimentos/Logística Empresarial*. Bookman Cia. Editora. Porto Alegre, RS. 616 p.
- _____. (1991b). “Measuring Transport Costing Error in Customer Aggregation for Facility Location “; *Transportation Journal*, Spring, p.49-59.
- CEL/COPPEAD (2001). “Armazenagem em Grandes Empresas Brasileiras“ Centro de Estudos Logísticos - Universidade Federal do Rio de Janeiro.

- Colorni, A., Dorigo, M., Maffioli, F., Maniezzo, V., Righini, G., Trubian, M. (1997). "Heuristics from Nature for Hard Combinatorial Optimization Problems"; *International Transactions in Operational Research*, No.3.1 ,p. 1 - 38.
- Crainic, T.G. e Laporte, G. (1997). "Planning Models for Freight Transportation". *European Journal of Operational Research* 97, p. 409-438.
- D'Amico, S. J.; Shouu-Jiun, W.; Batta, R.; Rump, C. M. (2002). "A Simulated Annealing Approach To Police District Design". *Computers and Operations Research*. Vol 29, pp. 667-684.
- Daskin, M. S. (1995). "*Network and Discrete Optimization*". John Wiley & Sons Inc. N.Y.
- Eaton, J.L. (1983). Children's Temperament: Cluster Analysis of Parent Questionnaires. PhD Thesis. University of Florida. Gainesville, FL.
- Francis, R.L.; Rushton, G.; Rayco, M.B. (1999). "A Synthesis Of Aggregation Methods For Multi-Facility Location Problems: strategies for containing error," *Geographical Analysis*. 67-87,31.
- Furuta, T.; Suzuki A.; Inakawa, K. (2005). The kth Nearest Network Voronoi Diagram and its Application to Districting Problem of Ambulance Systems. *Technical Report 0501*. Nanzan University.
- Galvão, L.C., Novaes, A. G., Cursi, J. E. S., Souza, J. C. (2006). A Multiplicatively-Weighted Voronoi Diagram Approach to Logistics Districting. *Computers and Operations Research*. Volume 33 , Issue 1 . pp: 93 – 114.
- Glover, F. (1986). Future Paths for Integer Methods for Integer Programming and Links to Artificial Intelligence. *Computers and Operations Research*. n. 13, p. 533-549.
- _____. (1990). Tabu Search: a Tutorial. *Interfaces*. July-August, 1990, pp. 74-94.
- _____. (1993). A User's Guide to Tabu Search. *Annals of Operations Research*, 41, pp. 3-28.
- Matteucci, M. (2006). Tutorial in Clustering. *Politecnico di Milano. Department of Eletronics and Information*. Milano, Italy. Disponível em 15/05/2006 no endereço: http://www.elet.polimi.it /upload /matteucc/Clustering/tutorial_html/
- Muyldermans, L.; Cattrysse, D.; Van Oudheusden, D.; Lotan, T. (2002). "Districting for salt spread operations. *European Journal of Operations Research*. Vol. 139, p. 521-532.
- Novaes, A.G. (1978). "*Métodos de Otimização*". Editora Edgard Blücher., São Paulo, SP. 463 p.
- _____. (1989). "Sistemas logísticos". Editora Edgard Blücher, São Paulo. 372 p.
- Novaes, A.G.; Gracioli, O. D. (1999). "Designing Multi-Vehicle Delivery Tours in a Grid-Cell Format". *European Journal of Operations Research*. Vol. 119, p. 613-614.
- Novaes, A.G.; Cursi, J. E. S.; Gracioli, O. D. (2000). "A Continuous Approach to the Designing of Physical Distribution Systems". *Computers and Operations Research*. Vol. 27, pp. 877-893.
- Ochi, L. S., Dias, C. R., Soares, S. S. F. (2005). Clusterização em Mineração de Dados. Documento Técnico do Programa de Pós-Graduação em Computação. Instituto de Computação da Universidade Federal Fluminense. Niterói, RJ.
- Oliveira, A.L. R. (2005). Otimização de Recebimento e Distribuição em Unidades Armazenadoras de Soja. Dissertação de Mestrado. Setores de Ciência Exatas e Tecnologia da *Universidade Federal do Paraná*. Curitiba, PR.
- SAS Institute (1982). SAS User's Guide: Statistics. *SAS Institute Inc*. Cary, NC. 584 p.
- Thorndike, R. M. (1978). *Correlational Procedures for Research*. Gardner Press. New York, NY. 340 p.
- Vallim F^o , A. R. A. (2004) Localização de Centros de Distribuição de Carga: Contribuições à Modelagem Matemática. Tese de Doutorado, *Escola Politécnica da Universidade de São Paulo*. São Paulo, 286p..
- Vallim F^o , A. R. A., Gualda, N. D. F. (2003). "Heurística para Seleção de 'Melhores' Candidatos para Localização de Centros de Distribuição". *XVII ANPET - Congresso de Pesquisa e Ensino em Transportes*, 2003, Rio de Janeiro. v.2. p.632-643.
- _____. (2004). "Heuristics, Metaheuristics and Optimization Models for the Distribution Centers Location Problem". *XII Panamerican Conference on Traffic and Transportation Engineering*, Albany, NY.
- Zhou, G.; Hokey, M., Gen M. (2002). "The Balanced Allocation of Customers to Multiple Distribution Centers in the Supply Chain Network: a Genetic Algorithm Approach". *Computers and Industrial Engineering*. V. 43, pp. 251-261.

Arnaldo Rabello de Aguiar Vallim Filho

Rua da Consolação 896 - Prédio João Calvino – sala 317

São Paulo – SP – 01302-907 - e-mail: aavallim@mackenzie.br