sid.inpe.br/mtc-m19/2013/09.06.14.48-TDI

CONTRIBUIÇÕES PARA A SOLUÇÃO DE DUAS APLICAÇÕES FORMULADAS COMO PROBLEMAS DE LOCALIZAÇÃO/AGRUPAMENTO

Silvio Luiz Castro Silva

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada, orientada pelo Dr. Luiz Antonio Nogueira Lorena, aprovada em 28 de agosto de 2013.

URL do documento original:
<http://urlib.net/8JMKD3MGP7W/3EPS72B>

INPE
São José dos Campos
2013

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE
Gabinete do Diretor (GB)
Serviço de Informação e Documentação (SID)
Caixa Postal 515 - CEP 12.245-970
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/6921
Fax: (012) 3208-6919
E-mail: pubtc@sid.inpe.br

CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELECTUAL DO INPE (RE/DIR-204):
Presidente:
Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)
Membros:
Dr. Antonio Fernando Bertachini de Almeida Prado - Coordenação Engenharia e Tecnologia Espacial (ETE)
Drª Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)
Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)
Dr. Germano de Souza Kienbaum - Centro de Tecnologias Especiais (CTE)
Dr. Manoel Alonso Gan - Centro de Previsão de Tempo e Estudos Climáticos (CPT)
Drª Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação
Dr. Plínio Carlos Alvalá - Centro de Ciência do Sistema Terrestre (CST)
BIBLIOTECA DIGITAL:
Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)
REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:
Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)
Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)
EDITORAÇÃO ELETRÔNICA:
Maria Tereza Smith de Brito - Serviço de Informação e Documentação (SID)
Luciana Manacero - Serviço de Informação e Documentação (SID)

sid.inpe.br/mtc-m19/2013/09.06.14.48-TDI

CONTRIBUIÇÕES PARA A SOLUÇÃO DE DUAS APLICAÇÕES FORMULADAS COMO PROBLEMAS DE LOCALIZAÇÃO/AGRUPAMENTO

Silvio Luiz Castro Silva

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada, orientada pelo Dr. Luiz Antonio Nogueira Lorena, aprovada em 28 de agosto de 2013.

URL do documento original:
<http://urlib.net/8JMKD3MGP7W/3EPS72B>

INPE
São José dos Campos
2013

Dados Internacionais de Catalogação na Publicação (CIP)

Silva, Silvio Luiz Castro.
Si38c Contribuições para a solução de duas aplicações formuladas como problemas de localização/agrupamento / Silvio Luiz Castro Silva. – São José dos Campos : INPE, 2013.
xx + 103 p. ; (sid.inpe.br/mtc-m19/2013/09.06.14.48-TDI)

Dissertação (Mestrado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2013.
Orientador : Dr. Luiz Antônio Nogueira Lorena.

1. localização. 2. agrupamento. 3. otimização. I.Título.

CDU 519.8

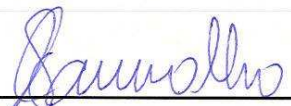
Esta obra foi licenciada sob uma Licença Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada.

This work is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License.

Aprovado (a) pela Banca Examinadora em cumprimento ao requisito exigido para obtenção do Título de ***Mestre*** em

Computação Aplicada

Dr. Solon Venâncio de Carvalho



Presidente / INPE / SJCampos - SP

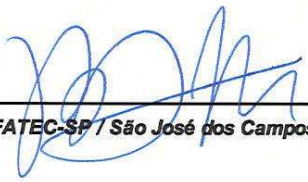
Dr. Luiz Antonio Nogueira Lorena

Orientador(a) / INPE / São José dos Campos - SP

Dr. Fabiano Luis de Sousa

Membro da Banca / INPE / SJCampos - SP

Dr. Reinaldo Gen Ichiro Arakaki



Convidado(a) / FATEC-SP / São José dos Campos - SP

Este trabalho foi aprovado por:

( ) maioria simples

(X) unanimidade

Aluno (a): ***Silvio Luiz Castro Silva***

São José dos Campos, 28 de Agosto de 2013

AGRADECIMENTOS

Agradeço, primeiramente, a Deus por ter me guiado por este caminho.

Ao meu orientador o Dr. Luiz Antônio Nogueira Lorena, pela dedicação, conhecimento e paciência; orientando-me durante o desenvolvimento desta dissertação.

Agradeço também a minha esposa Fabiana Aparecida dos Santos Silva pela paciência e apoio durante meus momentos de ausência e agradeço minhas filhas Gabriela e Julia por iluminarem o meu caminho e a minha vida.

Ao Instituto Nacional de Pesquisas Espaciais (INPE) pela oportunidade, e aos professores da CAP pelos ensinamentos transmitidos.

Ao Instituto de Fomento e Coordenação Industrial (IFI) pela liberação e apoio para cursar o Mestrado.

RESUMO

Esse trabalho apresenta contribuições para a solução de duas aplicações para problemas de localização/agrupamento. A primeira aplicação modela a realocação de laboratórios de metrologia do Comando da Aeronáutica, órgão pertencente ao Ministério da Defesa e que possui um Sistema de Metrologia Aeroespacial. Essa rede de laboratórios foi criada inicialmente sem a devida preocupação logística, localizando os laboratórios de acordo com os projetos aeronáuticos existentes. Como a demanda e os custos para a manutenção do sistema foram crescendo, foi observada a necessidade de reorganizar de forma otimizada a rede de laboratórios, melhorando assim, o atendimento das demandas de transportes de equipamentos e técnicos. Para que essa reestruturação possa ser realizada de forma otimizada, esse trabalho utiliza um modelo matemático que se baseia no Problema de Localização de Concentradores para modelar o problema capacitado em 3 (três) níveis hierárquicos. Para essa otimização foram utilizadas abordagens genéticas pela complexidade computacional do problema. A segunda aplicação apresenta o agrupamento otimizado de estudantes de acordo com seus Estilos de Aprendizagem. Esse agrupamento tem o objetivo de concentrar em cada grupo um conjunto de estudantes similares e que os aprendizes de cada grupo sejam diferentes dos elementos dos outros grupos, de acordo com as características medidas. Com esse agrupamento o profissional de educação (presencial ou à distância) poderá personalizar o material didático proposto para cada um dos grupos identificados, maximizando assim, a assimilação do conteúdo e, consequentemente, diminuindo o número de não aprendizes. Para a classificação dos alunos foi utilizado a ferramenta psicológica *Index of Learning Styles* para identificação dos estilos de aprendizagem de cada estudante. Essa identificação parte da ideia de que os indivíduos têm diferentes maneiras de "perceber" e de "processar a informação" o que caracteriza formas particulares de aprender. Cada indivíduo percebe os estímulos do ambiente de uma maneira particular e consequentemente como a percepção é a primitiva da aprendizagem cada pessoa aprende de uma forma particular. Foram utilizadas três técnicas de otimização para o agrupamento de aprendizes: otimização com o modelo para o Problema de P-Medianas, otimização com Algoritmo Genético e otimização com o método heurístico de agrupamento *K-means*.

CONTRIBUTIONS TO THE SOLUTION OF TWO APPLICATIONS FORMULATED AS PROBLEMS OF LOCATION / CLUSTERING

ABSTRACT

This work presents contributions to the solution of two applications to location problems / clustering. The first application modeling relocation of the metrology laboratories of the Aeronautical Command, agency belonging to the Ministry of Defence and that has a Metrology System Aerospace. This laboratory network was initially created without due concern logistics, locating laboratories in accordance with the existing aeronautical projects. As the demand and costs for system maintenance were growing, noted the need to reorganize optimally lab network, thus improving the care of the demands of transport of equipment and technicians. For this restructuring can be performed optimally, this work uses a mathematical model that is based on the Location Problem Concentrators able to model the capacity problem in 3 (three) hierarchical levels. For this optimization we used genetic approaches for the computational complexity of the problem. The second application presents the optimal grouping of students according to their learning styles. This grouping aims to focus on each group a set of similar students and that the learners in each group are different from the elements of the other groups, according to the characteristics measured. With this grouping professional education (classroom or distance) can customize the courseware offered for each of the identified groups, thus maximizing, the assimilation of content and, consequently, decreasing the number of non-apprentices. For the classification of students was used a psychological tool called Index of Learning Styles for identifying learning styles of each student. This identification from the idea that individuals have different ways of "perceive" and "process information" that characterizes particular forms of learning. Each individual perceives environmental stimuli in a particular way and consequently how perception is the primitive of learning each person learns in a particular way. Used three optimization techniques for grouping learners: the optimization model for the P-Median Problem, optimization with Genetic Algorithm optimization and heuristic method with K-means clustering

LISTA DE FIGURAS

Figura 1-1 - Modelo de rede *star-star* com três níveis hierárquicos 2
Figura 2-1- (a) Conexão entre origem/destino. (b) Rede eixo-raio. 13
Figura 2-2 - Rede eixo-raio de alocação múltipla 14
Figura 2-3 - Rede eixo-raio de alocação com conexão direta entre destino e origem 15
Figura 2-4 - Modelo de rede *star-star* com três níveis hierárquicos 15
Figura 2-5 – Diagrama do algoritmo subida da encosta. 21
Figura 2-6 – Funcionamento básico de um AG. 26
Figura 2-7 - Exemplo de aplicação do operador de cruzamento de um ponto 28
Figura 2-8 - Exemplo de aplicação do operador de cruzamento de múltiplos pontos 29
Figura 2-9 - Exemplo de aplicação do operador de cruzamento uniforme 29
Figura 2-10 - Exemplo de aplicação do operador de mutação 30
Figura 3-1 - Concepção matricial do SISMETRA 35
Figura 3-2 - Cadeia de rastreabilidade do SISMETRA 36
Figura 3-3 – Organograma do SISMETRA. 38
Figura 3-4 – Exemplo de rastreabilidade para calibração de paquímetro 39
Figura 3-5 – Estrutura de múltiplas demandas do SISMETRA 40
Figura 3-6 – Representação do Indivíduo 45
Figura 3-7 – Exemplo do operador de cruzamento de dois pontos 46
Figura 3-8 – Exemplo do operador de mutação M1 48
Figura 3-9 – Exemplo do operador de mutação M2 48
Figura 3-10 – Exemplo do operador de mutação M3 49
Figura 3-11 – Exemplo do operador de mutação M4 49
Figura 3-12 – Exemplo do operador de mutação M5 50
Figura 3-13 – Exemplo do operador de mutação M6 50
Figura 3-14 – Sequência de aplicação dos mecanismos de mutação 51
Figura 3-15 – Análise da convergência das 50 execuções dos AEs - Tempo 61
Figura 3-16 – Análise da convergência dos AEs – NFob 62
Figura 3-17 – Análise da diversidade por *Hamming* da população do AGS para 37 localidades 64
Figura 3-18 – Análise da diversidade por *Hamming* da população do AGM para 37 localidades 65
Figura 3-19 – Organograma otimizado do SISMETRA 66
Figura 4-1 - Relação entre aprendizagem, pensamento e o processo perceptivo. 68
Figura 4-2: Representação do cromossomo para codificação *Group-Numbers*. 75
Figura 4-3: Exemplo de aplicação do operador de cruzamento de um ponto. 77
Figura 4-4: Exemplo de aplicação do operador de mutação 77
Figura 4-5: Gráfico da análise de convergência. 81

LISTA DE TABELAS

Tabela 3-1: Matriz de múltiplas demandas 44
Tabela 3-2 - Resultados do CPLEX 10.1 para otimização do modelo MCpCLP–S..... 55
Tabela 3-3 – Resultados das 50 execuções do AGS para otimização do modelo MCpCLP–S com 37 localidades e capacidades D=210 – E=210 – F=210..... 57
Tabela 3-4 – Resultados das 50 execuções do AGM para otimização do modelo MCpCLP–S com 37 localizadas e capacidades D=210 – E=210 – F=210. 58
Tabela 3-5 – Resultados das 10 execuções do AGM para otimização do modelo MCpCLP –S com 37 localizadas e capacidades D=74 – E=66 – F=75. 60
Tabela 4-1: Exemplo de respostas do questionário (ILS)..... 73
Tabela 4-2: Escalas de estilos de aprendizagem..... 73
Tabela 4-3: Execuções do AG para agrupamento de aprendizes 80
Tabela 4-4: Agrupamento realizado pelo AG 83
Tabela 4-5: Agrupamento realizado pelo *K-means* 86
Tabela 4-6: Agrupamento realizado pelo *P*-Medianas 89

LISTA DE SIGLAS E ABREVIATURAS

AG	Algoritmo Genético
BIPM	*Bureau Internacional Des Poids Et Measures*
CE	Computação Evolutiva
CMA	Divisão de Confiabilidade Metrológica Aeroespacial
COMAER	Comando da Aeronáutica
DCTA	Departamento de Ciência e Tecnologia Aeroespacial
FOB	Valor da Função Objetivo
IA	Inteligência Artificial
INPE	Instituto Nacional de Pesquisas Espaciais
INM	Institutos Nacionais de Metrologia
INMETRO	Instituto Nacional de Metrologia, Normalização e Qualidade Industrial
LCC	Laboratório Central de Calibração
LRC	Laboratório Regional de Calibração
LSC	Laboratório Setorial de Calibração
LNM	Laboratório Nacional de Metrologia
MCCLP	Multi-Capacitated Concentrator Location Problem
MCpCLP-S	Multi-Capacitated p-Concentrator Location Problem in a Star–Star Network
MD	Ministério da Defesa
NFOB	Número de Chamadas da Função Objetivo
OC	Otimização Combinatória
PI	Programação Inteira
PIB	Programação Inteira Binária
PIM	Programação Inteira Mista
PL	Programação Linear
PLC	Problema de Localização de Concentradores
PNL	Programação Não-Linear
PO	Pesquisa Operacional

PP	Problema de Particionamento
RBC	Rede Brasileira de Calibração
SI	Sistema Internacional de Medidas
SISMETRA	Sistema de Metrologia Aeroespacial
UCLP	Uncapacitated Concentrator Location Problem
UFLP	Uncapacitated Facility Location Problem
UHLP-S	Uncapacitated Hub Location Problem in a Star–Star Network
USApHMP	Uncapacitated Single Allocation p-Hub Median Problem

LISTA DE SÍMBOLOS

λ	Custos de coleta
α	Custo de transferência
δ	Custo de distribuição
B_j	Custo de ligar o concentrador j a unidade central C_0.
C_j	Conjunto de Concentradores
C_0	Unidade Central
c_{ij}	Custo de ligar um terminal T_i ao concentrador C_j.
$d(l_i, l_j)$	Distância entre a Localidade i e a Localidade j.
d_{ij}	Distância entre os nós i e j
k_j	Capacidade máxima, aconselhável, de terminais que cada concentrador C_j pode gerenciar.
k_{js}	Múltiplas capacidades máximas de serviços que cada concentrador C_j pode atender, com $s \in \{1,2,3\}$ representando as subárea de cada concentrador (laboratório): Dimensional, Elétrica e Física.
P	Número de concentradores a serem localizados
Q_{is}	Demanda de calibração do laboratório i da área de grandeza s.
T_i	Conjunto de Terminais

Tc	Taxa de Cruzamento
Tm	Taxa de Mutação
V	Conjunto de nós da rede
w_{ij}	Quantidade de fluxo transferido entre os nós *i e j*
x_{ik}	Variável de decisão, onde $x_{ik} = 1$ se o nó i está alocado ao concentrador k ou $x_{ik} = 0$ caso contrário. Deve-se observar que quando a variável $x_{kk} = 1$ então o nó k é um concentrador ou se $x_{kk} = 0$ caso contrário.
x_{ij}	Variável binária que assume o valor 1 se o terminal T_i está ligado ao concentrador C_j e 0, caso contrário.
y_j	Variável binária que assume o valor 1 se o nó i é escolhido para concentrador e 0, caso contrário.

SUMÁRIO

Pág.

1. INTRODUÇÃO 1

2. TÉCNICAS DE OTIMIZAÇÃO PARA PROBLEMAS DE LOCALIZAÇÃO E AGRUPAMENTO 7

2.1. Programação Linear 8

2.1.1. Forma Padrão do Modelo de Programação Linear 9

2.2. Programação Linear Inteira 10

2.3. Programação Não Linear 11

2.4. Modelagens para Problemas de Localização 12

2.4.1. Problema de Localização de Concentradores 12

2.4.2. Problema de Localização de *p*-Medianas 18

2.5. Algoritmos de Busca Local 20

2.5.1. Busca de Subida da Encosta 21

2.6. *K-Means* 22

2.7. Algoritmos Genéticos 22

2.7.1. Principais Definições 25

2.7.2. Funcionamento Básico do AG 25

2.7.3. Função de Aptidão 27

2.7.4. Seleção dos Indivíduos 27

2.7.5. Cruzamento 28

2.7.6. Mutação 30

2.7.7. Parâmetros Genéticos 30

2.8. Algoritmo Genético Memético 32

3. PROBLEMA DE LOCALIZAÇÃO / ALOCAÇÃO DOS LABORATÓRIOS DE METROLOGIA DO COMANDO DA AERONÁUTICA 33

3.1. Sistema de Metrologia Aeroespacial (SISMETRA) 34

3.2. Localização-Alocação Atual dos Laboratórios do SISMETRA 37

3.3. Determinação da Matriz de Custos 43

3.4. Determinação da Matriz de Capacidades e Matriz de Demandas 44

3.5. Algoritmo Genético Simples Proposto para o Problema 45

3.5.1. Representação dos Indivíduos 45

3.5.2. População Inicial 46

3.5.3. Operador de Cruzamento 46

3.5.4. Operador de Mutação 47

3.5.5. Intervalo de Geração 51

3.5.6. Função Aptidão da Abordagem Genética proposta 51

3.6. Algoritmo Genético Memético - AGM 53

3.6.1. Busca Local Utilizada 53
3.7. Resultados Computacionais 55
3.7.1. Análise da Convergência 61
3.7.2. Análise da Diversidade da População 63
3.7.3. Nova Localização-Alocação dos Laboratórios do SISMETRA 65
4. PROBLEMA DE AGRUPAMENTO DE APRENDIZES 67
4.1. Problema de Agrupamento 67
4.2. Aprendizagem 67
4.2.1. Estilos de Aprendizagem 68
4.2.2. Codificação 74
4.2.3. População Inicial 75
4.2.4. Função de Aptidão 76
4.2.5. Seleção 76
4.2.6. Cruzamento 76
4.2.7. Mutação 77
4.2.8. Substituição da População 78
4.2.9. Critério de Parada 78
4.2.10. Preparação dos Dados 78
4.2.11. Resultados Computacionais 79
4.2.12. Método de Avaliação do Agrupamento 82
4.3. Abordagem *K-Means* 85
4.3.1. Resultados Computacionais 86
4.4. Abordagem P-Medianas 88
4.4.1. Resultados Computacionais 89
5. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS 93
REFERÊNCIAS BIBLIOGRÁFICAS 97

1. INTRODUÇÃO

O objetivo desta dissertação é apresentar as técnicas que contribuíram para a otimização de dois problemas reais, onde o primeiro problema consiste na realocação dos laboratórios de metrologia do Comando da Aeronáutica e o segundo problema consiste no agrupamento de aprendizes com base em seus Estilos de Aprendizagem.

Pode ser observado nas mais diversas áreas de atuação, uma grande demanda de soluções computacionais para problemas reais, onde essas soluções poderão servir de referência para as decisões estratégicas gerencias. Um importante conjunto de problemas a serem otimizados é composto pelos problemas de localização de facilidades, onde estes problemas abordam o desafio de localizar facilidades que terão que atender a um determinado conjunto de demandas alocadas.

Dentre os problemas de localização de facilidades destacam-se os Problemas de Localização de Concentradores (PLC). O PLC tem como objetivo determinar uma rede, *hub-and-spoke network*, de custo mínimo onde serão localizados nós concentradores e alocados os nós de demanda restantes, de forma que o fluxo de demanda chegue ao seu destino por meio de ligações entre os concentradores.

Pode-se também modelar o PLC como uma rede *star-star network*, Figura 1-1. Nessa estrutura em estrela existe uma Unidade Central (triangulo), onde os nós Concentradores localizados (retângulos) estão diretamente ligados e que posteriormente terão os nós de terminais restantes (círculos) alocados aos concentradores. Nesse tipo de rede o processo de otimização consiste em localizar as facilidades concentradoras minimizando os custos de localizar os concentradores e alocar os terminais a esses concentradores (KLINCEWICZ, 1998; YAMAN, 2008).

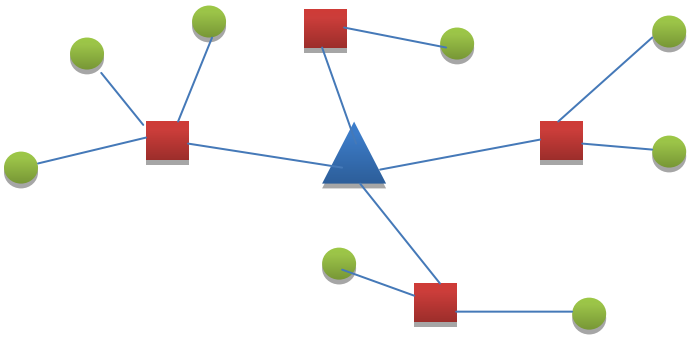


Figura 1-1 - Modelo de rede *star-star* com três níveis hierárquicos

Esse modelo segundo Sá (2011) e O' Kelly e Miller (1994) é utilizado com propriedade como suporte estratégico para implementações de atendimento de demandas de transporte público, pois como as conexões de origem/destino são inúmeras, acabam sendo inviável financeiramente a criação todas essas vias de conexão, sendo dessa forma, uma solução promissora a concentração do fluxo em pontos estratégicos (concentradores), criando mecanismos para o transporte em massa entre os concentradores até os pontos de destino.

Nesse trabalho foi utilizado o PLC para modelar um dos problemas enfrentados. Esse primeiro problema real consiste na realocação hierárquica e com múltiplas capacidades dos laboratórios de metrologia pertencentes ao Sistema de Metrologia Aeroespacial (SISMETRA), do Comando da Aeronáutica (COMAER); visando a minimização dos custos operacionais para a manutenção do sistema. Esses laboratórios estão organizados em uma estrutura composta por 3 (três) níveis hierárquicos de metrologia: Laboratório Central de Calibração (LCC), Laboratórios Regionais de Calibração (LRC) e Laboratórios Setoriais de Calibração (LSC) (BRASIL. MINISTÉRIO DA DEFESA. COMANDO DA AERONÁUTICA, 2009).

Para otimizar esse modelo foi utilizado um Algoritmo Genético Simples (AG) e um Algoritmo Genético Memético (AGM), onde foram realizados testes computacionais para duas matrizes de capacidades distintas, onde a primeira matriz de capacidades extrapola as capacidades aumentando o espaço de soluções factíveis, e a segunda matriz restringe as capacidades diminuindo o

espaço de soluções factíveis, aumentando assim, a dificuldade de resolução computacional.

Outro conjunto de problemas de grande importância e abordado nesse trabalho são os problemas de agrupamento. Esse tipo de problemas consiste em dado um conjunto de *n* elementos amostrais, o agrupamento consiste em identificar grupos com características similares, dividindo a amostra em *k* grupos, prezando pela similaridade dos elementos presentes no mesmo grupo ao invés dos os elementos dos outros grupos.

Nesse trabalho abordam-se técnicas de otimização para o problema de agrupamento de aprendizes. Essa análise tem como objetivo dividir os aprendizes da amostra, ou população, em grupos de forma que os aprendizes pertencentes a um mesmo grupo sejam similares entre si, com respeito às características que neles foram medidas, e que sejam diferentes dos elementos em grupos diferentes em relação a estas mesmas características. (MINGOTI, 2005).

Para medir a similaridade entre os aprendizes foi utilizado o instrumento denominado *Index of Learning Styles* (ILS) na sua versão reduzida, (FELDER; SOLOMAN, 2010), onde esse instrumento psicológico classifica, quantifica, as preferências de aprendizagem dos aprendizes em quatro dimensões de Estilos de Aprendizagem (EA): *Ativo/Reflexivo, Sensorial/Intuitivo, Visual/Verbal e Sequencial/Global.* O ILS é um instrumento encontrado em trabalhos de investigação dos estilos de aprendizagem de estudantes universitários (LOPES, 2002), para identificar os estilos de aprendizagem de alunos de engenharia (CURY, 2000), em problemas de agrupamento de aprendizes (ZAKRZEWSKA, 2009) e na busca de maior eficiência nas ações de capacitação das áreas de treinamento e desenvolvimento de instituições (MACHADO, et al., 2010). As dimensões utilizadas pelo instrumento ILS, são baseadas no modelo de estilos de aprendizagem de Felder e Silverman (FELDER; SILVERMAN, 1988).

Para a otimização do agrupamento dos aprendizes são utilizadas três técnicas de otimização, um Algoritmos Genéticos Simples, a heurística *K-means,* e o modelo para o Problema de P-Medianas (PPM). Essas técnicas podem ser encontradas em trabalhos para a otimização de problemas de programação de operações em um ambiente de produção *Flow Shop*, tendo como objetivo minimizar o estoque em processamento (RIBEIRO, et al., 2007); problemas de expansão de sistemas de transmissão (MIASAKI; ROMERO, 2007); problemas químicos (FILHO; POPPI, 1999), Algoritmo Evolutivo Híbrido para a Formação de Células de Manufatura em Sistemas de Produção (TRINDADE; OCHI, 2006), Algoritmos Evolutivos para o Problema de *Clusterização* de Grafos Orientados: Desenvolvimento e Análise Experimental (DIAS, 2004), *Constructive Genetic Algorithm for Clustering Problems* (LORENA; FURTADO, 2001); Integração de um Modelo de P-Medianas a Sistemas de Informação Geográficas (LORENA, et al., 1999), Aplicação do Problema das P-Medianas para Determinar a Localização de Unidades de Saúde 24 Horas (ROSÁRIO, et al., 2001); Metaheurística Híbrida com Busca por Agrupamentos Aplicada a Problemas de Otimização Combinatória (CHAVES, 2009), *Clustering Students to Help Evaluate Learning* (MERCERON; YACEF, 2005). Encontram-se também na literatura trabalhos que apresentam outras técnicas para o agrupamento de aprendizes segundo EA, tais como: *K-means* e *Self-organizing Maps* (PIMENTEL, et al., 2003), Agentes Inteligentes (FELIX; TEDESCO, 2006).

A escolha de methauristicas para a otimização dos modelos abordados nesse trabalho, tem como justificativa a complexidade dos problemas, sendo o modelo para o PLC e o modelo para PPM pertencentes à classe dos problemas NP-Difícil (LORENA; SENNE, 2000; ARAKAKI; LORENA, 2006; ALMEIDA, 2009), onde essa classe de problemas dispende de grande custo computacional para obtenção de soluções ótimas, entretanto, metaheurísticas comumente apresentam soluções de boa qualidade para o problema, como por exemplo, o AG para a configuração de uma rede *hub-and-spoke* para uma

empresa de caminhões no Brasil (CUNHA; SILVA, 2007) e abordagens neurais versus tradicional para a localização das instalações concentradoras (SMITH, et al., 1996).

Essa dissertação está dividida em 5 (cinco) capítulos.

O primeiro capítulo é responsável pela apresentação introdução da dissertação. O segundo capítulo descreve técnicas de otimização para problemas de localização e agrupamento. Já no terceiro capitulo é apresentado o problema de localização / alocação dos laboratórios de metrologia do Comando da Aeronáutica. O quarto capítulo descreve o problema de agrupamento de aprendizes. O quinto capítulo apresenta as considerações finais e trabalhos futuros da dissertação.

2. TÉCNICAS DE OTIMIZAÇÃO PARA PROBLEMAS DE LOCALIZAÇÃO E AGRUPAMENTO

A Pesquisa Operacional (PO) teve origem no desenvolvimento de atividades científicas para fins militares na Segunda Guerra Mundial. Segundo Hillier e Lieberman (2010) nessa época existia um problema operacional muito comum que era alocar recursos de forma eficiente para as diversas operações militares e atividades internas a cada operação. Dessa forma, cientistas britânicos e norte-americanos foram responsáveis pela obtenção de soluções para problemas táticos e estratégicos, proporcionando eficiência nas operações militares que se tornaram um diferencial nas vitórias dos céus da Grã-Bretanha, na Batalha do Atlântico Norte e na Campanha Britânica no Pacífico. Esse sucesso militar da PO despertaram muitas outras aplicações em problemas complexos industriais (HILLIER; LIEBERMAN, 2010).

Outro fato importante para a disseminação da PO foi o desenvolvimento de técnicas para a solução dos problemas, o mais relevante desenvolvimento foi a criação do Método Simplex para solução de problemas como programação linear, programação dinâmica e teoria de filas; desenvolvido por George Dantzig em 1974 e ganhado maturidade antes do final dos anos 1950 (HILLIER; LIEBERMAN, 2010).

O segundo fator foi de importância para o crescimento da PO foi a revolução computacional, pois o volume de processamento crescia proporcionalmente com a complexidade dos modelos, dessa forma os computadores eletrônicos tiveram e têm um papel fundamental como ferramenta de suporte a milhões de cálculos matemáticos para a democratização da PO (HILLIER; LIEBERMAN, 2010).

A representação dos modelos de otimização expressos matematicamente são modelos, frequentemente, estudados pela teoria da PO, sendo estruturados de forma lógica e tendo como objetivo a clareza e eficiência de funcionamento para as técnicas de solução empregadas. As técnicas utilizadas para a

solução de formulações matemáticas de otimização podem ser agrupadas segundo Goldbarg e Luna (2005), em:

Programação Linear (PL): modelagem em que as variáveis são continuas, e as restrições e a função objetivo apresentam comportamento linear.

Programação Linear Inteira (PI): quando as variáveis não podem assumir valores contínuos. Essa característica normalmente implica em complexidade computacional.

Programação Não-Linear (PNL): modelagem em que é observado algum tipo de não linearidade seja na função objetivo ou nas restrições.

2.1. Programação Linear

Os modelos de Programação Linear (PL) são modelos de otimização são funções lineares de variáveis continuas $\boldsymbol{x}$ e que possuem algoritmos eficientes para solução, robustez para cálculos e de implementação fácil. Esses modelos possuem segundo Goldbarg, Luna (2005) e Hiller, Lieberman (2010, pp. 36-42) as seguintes características:

Proporcionalidade: a quantidade de recurso consumido e o custo de uma dada atividade devem ser proporcionais ao nível dessa atividade na solução final do problema.

Não Negatividade: a representação das atividades deve ser sempre realizada no nível não negativo.

Aditividade: toda função é a soma das parcelas das respectivas atividades.

Separabilidade: pode-se identificar de forma separada o custo ou consumo de recurso de operação de cada atividade.

Divisibilidade: as variáveis de decisão podem assumir quaisquer valores, respeitando as restrições funcionais e de não negatividade.

Certeza: o valor de cada parâmetro é uma constante conhecida.

Segundo Goldbarg e Luna (2005, p. 28) os passos para eficiente elaboração de modelos de programação linear consiste na:

- Definição das atividades que compõe o problema enfrentado, normalmente, associada a uma unidade de medida.
- Definição dos recursos utilizados durante as atividades.
- Determinação dos coeficientes de relacionamento entre as atividades e os recursos, onde cada coeficiente representa os recursos necessários por unidade de atividade.
- Determinação das condições externas, limites de disponibilidade dos recursos utilizados pelas atividades.
- Formalização do modelo através de uma equação matemática que modela as relações das quantidades, não negativas, de cada atividade com as condições externas existentes.

2.1.1. Forma Padrão do Modelo de Programação Linear

Segundo Hiller e Lieberman (2010) o modelo matemático padrão para Programação Linear é apresentado nas formulações (2-1) a (2-5).

Maximizar $$Z = c_1x_1 + c_2x_2 + \cdots + c_nx_n \quad (2\text{-}1)$$

Sujeito às restrições:
$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1 \quad (2\text{-}2)$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leq b_2 \quad (2\text{-}3)$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq b_m \quad (2\text{-}4)$$
$$x_1 \geq 0, \quad x_2 \geq 0, \quad \ldots \quad x_n \geq 0 \quad (2\text{-}5)$$

onde:

Z: representa a medida de desempenho global Z
m: representa os recursos
n: representa as atividades
x_j: representa o nível da atividade j

A função Z a ser maximizada ou minimizada é chamada de **função objetivo**, as restrições compostas das variáveis de decisão são as **restrições funcionais** e as restrições restantes representam as restrições de **não negatividade.**

O modelo de PL possui todos os conceitos presentes nos demais modelos, sendo muito empregado e utilizado para estudos acadêmicos, pois permite a compreensão dos demais modelos de otimização (GOLDBARG; LUNA, 2005, p. 25).

2.2. Programação Linear Inteira

Os modelos de Programação Linear (PL) são eficientes e muito utilizados em aplicações em que a variável de decisão $\boldsymbol{x} \in \mathbf{R}$ assume valores contínuos. Entretanto, um grande número de aplicações tem como característica a localização de facilidades, sendo necessário para essa formulação que a variável de decisão $\boldsymbol{x}$ assuma somente valores inteiros $\boldsymbol{x} \in \mathbf{N}$. Essa característica da variável $\boldsymbol{x}$ é a única diferenciação aplicada ao modelo PL, tornando essa nova forma de formulação em um modelo Programação Inteira (PI).

Segundo Hillier e Lieberman (2010, p. 463), outra variável da PL é a formulação em que todas as variáveis de decisão representam decisões “sim ou não”. Essa é formulação em que as variáveis de decisão são conhecidas como *variáveis binárias* (valores 0 ou 1), proporcionando estrutura para a construção de modelos de Programação Inteira Binária (PIB). Tal variável binária é apresentada na Equação (2-6).

$$x_j = \begin{cases} 1 & se\ a\ decisão\ j\ for\ \ sim \\ 0 & se\ a\ decisão\ j\ for\ não \end{cases} \quad (2\text{-}6)$$

Existem situações em que é necessário permitir que apenas algumas variáveis de decisão assumam valores inteiros, esses modelos se enquadram na classificação de modelos de Programação Inteira Mista (PIM) (HILLIER; LIEBERMAN, 2010, p. 462). Um exemplo de utilização de um modelo de PIM pode ser encontrado em Espejo e Galvão (2002), nesse artigo são aplicados técnicas de relaxação *lagrangeana* e *surrogate* em problemas de programação inteira.

2.3. Programação Não Linear

Muitos problemas reais normalmente necessitam de representações não-lineares em sua modelagem. Essa não linearidade em PO é modelada a partir modelos de Programação Não-Linear (PNL).

Hillier e Lieberman (2010) definem que o problema de PNL consiste em encontrar $x = (x_1, x_2, \dots, x_n)$ de modo a:

Maximizar $f(x)$, (2-7)

Sujeito a: $g_i(x) \leq b_i$, para $i = 1,2, \dots, m,$ (2-8)

$$x \geq 0 \quad (2\text{-}9)$$

Em que $f(x)$ e os $g_i(x)$ são funções dadas das n variáveis de decisão.

Segundo Hillier e Lieberman (2010) existem vários métodos para a resolução de problemas de PNL, a escolha do melhor método para cada formulação depende muito das características das funções $f(x)$ e $g_i(x)$, pois em PNL mesmo os problemas pequenos, muitas vezes, são de difícil solução.

2.4. Modelagens para Problemas de Localização

O problema de localização abordado nesse trabalho pode ser modelado através de modelos matemáticos clássicos de localização-alocação. Esse modelo tem o objetivo de localizar facilidades concentradoras que serão alocadas as demandas, maximizando o atendimento ou minimizando os custos operacionais. Segundo Yaman (2008) e Arakaki (2002), muitos desses problemas de localização-alocação pertencem à classe dos problemas NP-Difícil.

2.4.1. Problema de Localização de Concentradores

Existem formulações de problemas de localização-alocação que são modelados como Problemas de Localização de Concentradores (PLC). O modelo para o PLC idealiza uma estrutura de rede tipo eixo-raio (*hub-and-spoke network*): Figura 2-1, onde são localizados concentradores para aglomerar o fluxo de demanda existente, sendo criadas conexões de maior fluxo entre todos os concentradores, para que dessa forma, o deslocamento da demanda seja realizado através do concentrador em que o nó de demanda está alocado, até o concentrador em que o nó de destino está alocado, que por sua vez, distribuirá a demanda para o destino pertinente. A localização dos concentradores pode ser realizada nos próprios nós de demanda ou em nós novos na rede, sendo essa premissa considerada no modelo. Esse tipo de rede pode ser encontrado nos trabalhos de (CAMPBELL; 1994; O' KELLY; MILLER, 1994; O'KELLY, et al., 1996; Sá, 2011).

Esse modelo segundo Sá (2011), O' Kelly e Miller (1994) é utilizado com muita propriedade como suporte estratégico para implementações de atendimento de demandas de transporte público, pois como as conexões de origem/destino são inúmeras, acabam sendo inviável financeiramente a criação todas essas vias de conexão, sendo dessa forma, uma solução promissora a concentração do fluxo em pontos estratégicos (concentradores), criando mecanismos para o transporte em massa entre os concentradores até os pontos de destino.

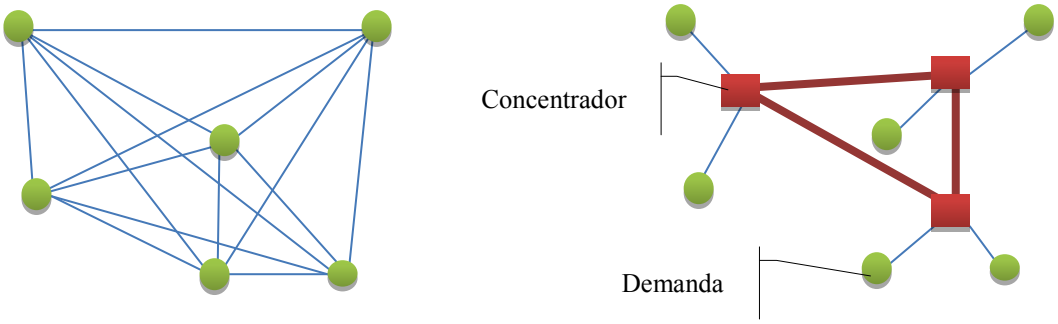


Figura 2-1- (a) Conexão entre origem/destino. (b) Rede eixo-raio.

Fonte: adaptada de Sá (2011, p. 4).

A primeira publicação realizada para apresentar o problema de localização de concentradores foi realizada por Goldman (1969), mas O'Kelly (1987) publica a primeira formulação quadrática para o problema de localização não capacitada de *p*-concentradores, (USApHMP - *Uncapacitated Single Allocation p-Hub Median Problem)*, formulado da seguinte forma:

$$f(x) = Min \sum_i \sum_j w_{ij} \sum_k \lambda d_{ik} x_{ik} + \sum_i \sum_j w_{ji} \sum_l \delta d_{jl} x_{jl} + \sum_i \sum_k x_{ik} \sum_j \sum_l \alpha x_{jl} d_{kl} w_{ij} + \sum_j x_{jj} f_j \quad (2\text{-}10)$$

Sujeito a:

$$\sum_k x_{kk} = p \quad (2\text{-}11)$$

$$\sum_k x_{ik} = 1, \quad \forall\, i,k \in V \quad (2\text{-}12)$$

$$x_{kk} - x_{ik} \geq 0, \qquad \forall\, i,k \in V \quad (2\text{-}13)$$

$$x_{ik} \in \{0,1\}, \qquad \forall\, i,k \in V \quad (2\text{-}14)$$

onde:

- V é o conjunto de nós da rede;
- p é o número de concentradores a serem localizados;
- d_{ij} é a distância entre os nós i e j;
- f_j é custo de se localizar um concentrado na localidade j;
- w_{ij} é a quantidade de fluxo transferido entre os nós *i e j;*

- λ, α, δ são, respectivamente, os custos de coleta, transferência e distribuição;
- x_{ik} é uma variável de decisão, onde $x_{ik} = 1$ se o nó i está alocado ao concentrador k ou $x_{ik} = 0$ caso contrário. Deve-se observar que quando a variável $x_{kk} = 1$ então o nó k é um concentrador ou se $x_{kk} = 0$ caso contrário.

A função-objetivo (2-10), a ser minimizada, representa o custo total da soma dos custos de coleta, distribuição e transferência de uma rede eixo-raio não capacitada com *p*-concentradores, a restrição (2-11) fixa o número de concentradores igual a *p*, as restrições (2-12) garantem que cada nó de demanda será alocado a um único concentrador, as restrições (2-13) asseguram que as alocações serão feitas apenas para nós que são concentradores, e as restrições (2-14) correspondem às condições de integralidade das variáveis de decisão.

A estrutura de redes eixo-raio prioriza a inclusão de concentradores que concentram os fluxos recebidos das origens de demanda, para posteriormente, direcionar o tráfego aglomerado, até o seu destino, através das vias de maior volume. Existem derivações da rede eixo-raio, onde é permitido que as origens e os destinos sejam conectados a mais de um concentrador (HALL, 1989; O'KELLY; LAO, 1991; O'KELLY, et al., 1996; CAMPBELL, 1991). Essas redes eixo-raio de múltipla alocação podem ser representadas de acordo com a Figura 2-2.

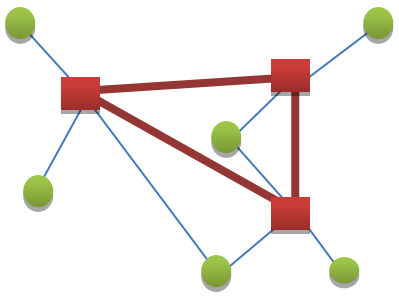


Figura 2-2 - Rede eixo-raio de alocação múltipla

Outra concepção de rede eixo-raios é a rede que permite que origem e destino tenham comunicação direta, não sendo necessário que o fluxo passe por um concentrador, Figura 2-3. Esse tipo de rede pode ser encontrado no trabalho de Aykin (1991) e O'Kelly, Miller (1994).

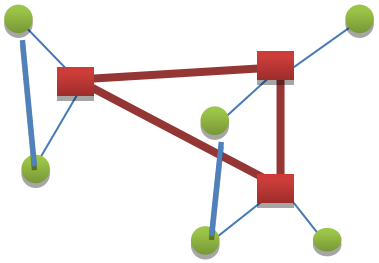


Figura 2-3 - Rede eixo-raio de alocação com conexão direta entre destino e origem

Segundo Sá (2011, p. 4) os modelos clássicos para o PLC consideram que a rede tipo eixo-raio possui interconexão entre todos os concentradores. Na prática para soluções de problemas reais se torna inviável, muitas vezes, implementar essa estrutura, pois existem muitas limitações técnicas e grande impacto financeiro.

Existe outro tipo de rede conhecido com *star-star network*, Figura 2-4, nessa estrutura em estrela existe uma Unidade Central, onde os nós Concentradores localizados estão diretamente ligados. Os nós restantes são alocados aos concentradores. Nesse tipo de rede o processo de otimização consiste em localizar as facilidades concentradoras minimizando os custos de localizar os concentradores e alocar os terminais a esses concentradores (KLINCEWICZ, 1998; YAMAN, 2008).

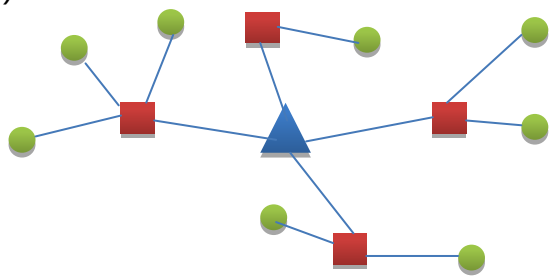


Figura 2-4 - Modelo de rede *star-star* com três níveis hierárquicos

Nesse tipo de rede *star-star* os nós de demanda direcionam o fluxo através de um nó concentrador para a unidade central. Esse tipo de rede é um caso especial do problema de *p*-medianas onde é acrescentado um custo de ligar os concentradores, localizados, a unidade central. Esse tipo de problema é chamado na literatura de Problema de Localização de Concentradores Não Capacitado em uma Rede Estrela-Estrela (UHLP-S - *Uncapacitated Hub Location Problem in a Star–Star Network*). O UHLP-S também é modelado como um Problema de Localização de Facilidades (UFLP – *Uncapacitated Facility Location Problem*). Segundo (LABBÉ; YAMAN, 2008) a modelagem UFLP aplicada a uma rede *star-star* é introduzida por Hakimi (1964) quando ele aplica os modelos 1-mediana e o *p*-mediana para a localização de centros de troca em redes de comunicação.

Uma forma para modelar o problema é através do modelo para o Problema de Localização de Concentradores Não Capacitado (UCLP - *Uncapacitated Concentrator Location Problem*) (LABBÉ; YAMAN, 2008). Essa formulação pode ser representada da seguinte forma: I é um conjunto de nós (localidades) e quando a variável de decisão $x_{ij} = 1$ representa que o nó $i \in I$ está alocado ao nó $j \in I$ e quando $x_{jj} = 1$ representa que o nó j é um concentrador (*hub*). Existe também um custo B_j aplicado à escolha do concentrador j e o custo c_{ij} de alocar o nó i ao concentrador j.

Minimizar: $$z = \sum_{i \in I} \sum_{j \in I} c_{ij} x_{ij} + \sum_{j \in I} B_j x_{jj} \tag{2-15}$$

Sujeito a: $$\sum_{k} x_{ij} = 1, \quad \forall\, i,j \in I \tag{2-16}$$

$$x_{ij} \leq x_{jj}, \qquad \forall\, i,j \in I \tag{2-17}$$

$$x_{ij} \in \{0,1\}, \qquad \forall\, i,j \in I \tag{2-18}$$

Nesse modelo os nós i são ligados diretamente aos concentradores j e os concentradores j ligados diretamente à unidade central C_0. A restrição (2-16) e

(2-17) garante que cada terminal i será ligado somente a um concentrador j. A restrição (2-18) garante a integralidade das variáveis de decisão.

Para problemas capacitados existe a modelagem para o Problema de Localização de Concentradores Capacitados em uma Rede *Star-Star.* Esse modelo é composto por um conjunto de *Terminais*, $T_i, i \in N, N = \{1, \dots, n\}$, normalmente idênticos, dispersos geograficamente e de localização conhecida; um conjunto de *Concentradores*, $C_j, j \in M, M = \{0, \dots, m\}$, unidades de porte e capacidade superior aos terminais, podendo ou não estarem localizadas juntos aos terminais, e cujo número e localização, normalmente, é desconhecido; e uma *Unidade Central,* C_0, unidade de grande porte quando comparado com os demais elementos e de localização normalmente conhecida. A formulação para o Problema de Localização de Concentradores Capacitados em uma Rede *Star-Star* (CHLP-S - *Capacitated Hub Location Problem in a Star–Star Network*) é apresentada na Equação (2-19) (GEN, et al., 1996):

Minimizar:
$$z = \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_{ij} + \sum_{j=1}^{m} d_j y_j \qquad (2\text{-}19)$$

Sujeito a:
$$\sum_{j=1}^{m} x_{ij} = 1, \qquad i = \{1, \dots, n\} \qquad (2\text{-}20)$$

$$\sum_{i=1}^{n} x_{ij} \leq k_j y_j, \qquad j = \{1, \dots, m\} \qquad (2\text{-}21)$$

$$x_{ij}, y_j \in \{0,1\}, \qquad \forall\, i, j \qquad (2\text{-}22)$$

onde:

- y_j é uma variável binária que assume o valor 1 se o nó j é escolhido para concentrador e 0, caso contrário.
- x_{ij} é uma variável binária que assume o valor 1 se o terminal T_i está ligado ao concentrado C_j e 0, caso contrário.
- d_j representa o custo de implementação de uma unidade concentradora no ponto j.

- c_{ij} representa o custo de ligar um terminal T_i ao concentrador C_j.
- k_j representa a capacidade máxima, aconselhável, de terminais que cada concentrador C_j pode gerenciar.

Nesse modelo os terminais i são ligados diretamente aos concentradores j e os concentradores j ligados diretamente à unidade central C_0. A restrição (2-20) garante que cada terminal i será ligado somente a um concentrador j. A restrição (2-21) restringe o número máximo, aconselhável, de terminais que cada concentrador C_j pode gerenciar. A restrição (2-22) garante a integralidade das variáveis de decisão.

A formulação procura otimizar uma solução que minimize os custos de implantação dos concentradores e da alocação dos terminais, permitindo responder as seguintes perguntas (GOLDBARG; LUNA, 2005):

1. Quantos concentradores são necessários?
2. Onde eles deverão estar localizados?
3. De que forma os concentradores estarão ligados à central?
4. Quais terminais serão designados a quais concentradores?
5. De que maneira os terminais estarão conectados aos concentradores a eles designados?

As duas formulações apresentadas UCLP e UHLP-S são classificados como modelos NP-Difícil em Labbé e Yaman (2008).

2.4.2. Problema de Localização de *p*-Medianas

Na literatura existem vários problemas de otimização clássicos, sendo o problema de localização de facilidades (*p*-medianas) um desses problemas que são bastante estudados e utilizados como base para a tomada de grandes decisões de gerenciamento tanto no setor público como no setor privado. Segundo Lorena (1999) o problema de localização de facilidades no setor público tenta por um lado maximizar a satisfação da sociedade e por outro

minimizar o custo para o atendimento das necessidades básicas, como: localização de escolas, postos de saúde, corpo de bombeiros, ambulâncias, viaturas de polícia, pontos de ônibus, entre outras. Já no setor privado vemos algumas aplicações, como: localização comercial e industrial, telecomunicações (localização de torres de transmissão), logística, etc.

Segundo Arakaki (2002) a modelagem do problema de *p*-medianas tem como objetivo localizar *p* facilidades (recursos), de forma a minimizar o custo total da soma das distâncias de cada vértice à sua mediana mais próxima.

Esse tipo de problema foi modelado primeiramente por Hakimi (1964), onde esse modelo contemplava a localização de apenas uma mediana. Já em 1965 o próprio Hakimi (1965) formaliza um modelo para a localização de múltiplas medianas.

Como modelo clássico de localização na literatura, comumente, são encontrados trabalhos abordando a utilização do modelo de *p*-medianas em situação práticas, onde soluções otimizadas fornecem suporte para decisões estratégicas. Um desses exemplos na literatura é abordado em (PIZZOLATO; SILVA, 1997) onde é analisado o problema para localização de escolas públicas numa região pobre do Rio de Janeiro e em (DEUTSCH, et al., 1998) na formação de células de manufatura.

Esse tipo de problema é classificado como NP-Difícil em (GAREY; JOHNSON, 1979), e como isso, muitos trabalhos sugerem a utilização de heurísticas e metaheurísticas para a resolução desses problemas, como a Substituição de Vértices em (TEITZ; BART, 1968), Algoritmo Genético (LORENA; FURTADO, 2001), Busca Local em (LORENA, SENNE, 2000) e a Busca Tabu em (ROLLAND, et al., 1996).

O problema de *p*-medianas considerado neste trabalho pode ser formulado como o seguinte problema de programação inteira binária:

$$Pmed(x) = Min \sum_{i=1}^{n}\sum_{j=1}^{n} D_{ij}x_{ij} \quad (2\text{-}23)$$

Sujeito a: $$\sum_{i=1}^{n} x_{ij} = 1; j \in N \quad (2\text{-}24)$$

$$\sum_{i=1}^{n} x_{ii} = p \quad (2\text{-}25)$$

$$x_{ij} \leq x_{ii};\ \forall\, i,j \in N \quad (2\text{-}26)$$

$$x_{ij} \in \{0,1\};\ \forall\, i,j \in N \quad (2\text{-}27)$$

onde:

- D_{ij} é uma matriz simétrica de distâncias, com $D_{ii} = 0, \forall\, i$;
- x_{ij} é a matriz de alocações, onde $x_{ij} = 1$ se a o nó *j* está alocado a mediana *i* e $x_{ij} = 0$ caso contrário; $x_{ii} = 1$ se o nó *i* é uma mediana e $x_{ii} = 0$ caso contrário;
- *p* é o n° de medianas (facilidades a serem localizadas);
- *n* é o número total de nós da rede, com *N = {1,2,3,...,n}.*

Na formulação, (2-23), procura-se minimizar a função objetivo, que depende de um número finito de variáveis de entrada x_{ij} e que serão restringidas de 4 maneiras: (2-24) e (2-25) garantem que cada nó *j* é alocado a somente um nó *i*, que é uma mediana. Já a restrição (2-26) determina o número exato de medianas a serem localizadas, e (2-27) corresponde às condições de integralidade.

2.5. Algoritmos de Busca Local

Os algoritmos de busca local trabalham com a geração de um conjunto de soluções vizinhas, tendo como base uma solução corrente. Dada uma solução corrente a busca local consiste em implementar um algoritmo que gere pequenas alterações na solução corrente, gerando soluções vizinhas. Essas soluções vizinhas serão avaliadas e se existir uma solução vizinha melhor que

a solução corrente tal solução se torna a nova solução corrente e o algoritmo continua. Caso contrário, a solução corrente é mantida.

2.5.1. Busca de Subida da Encosta

É um algoritmo que se move no sentido crescente (encosta acima) da função objetivo. Nesse algoritmo a cada iteração é criado um conjunto de soluções vizinhas e caso a melhor entre elas seja melhor que a solução corrente a solução corrente é substituída. O algoritmo encerra no momento que encontrar uma solução corrente melhor que cada um dos seus estados vizinhos (RUSSEL, 2004).

Segundo Russel (2004) a subida da encosta também é conhecida como Busca Gulosa Local, pois seleciona o primeiro estado vizinho encontrado sem analisar a direção da busca. Russel (2004) também afirma que mesmo gulosa a Busca de Subida da Encosta frequentemente funciona muito bem, melhorando com grande rapidez uma solução ruim. O diagrama do algoritmo de busca de Subida da Encosta é apresentado na Figura 2-5.

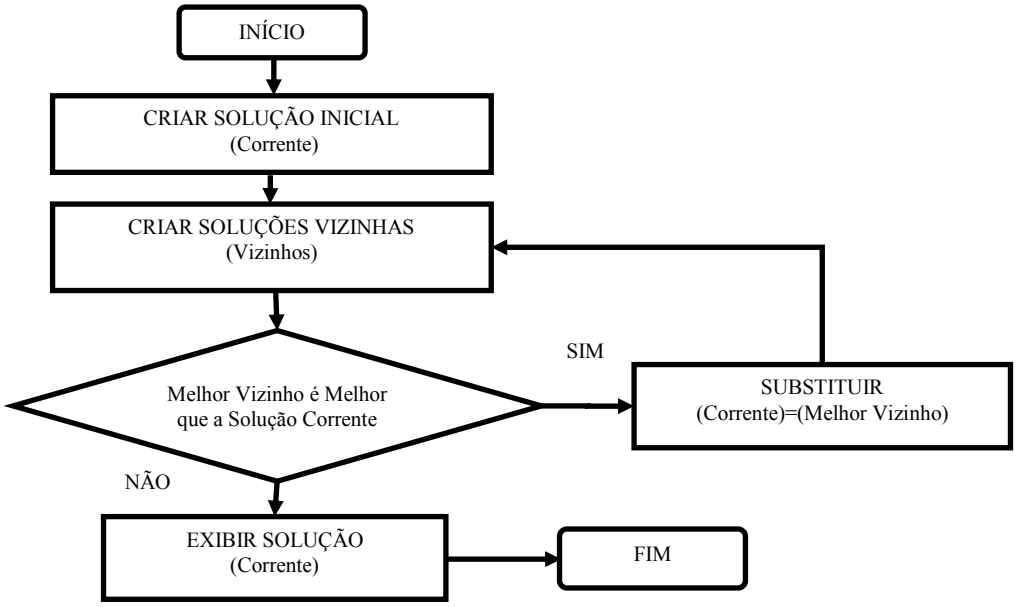


Figura 2-5 – Diagrama do algoritmo subida da encosta.
Fonte: adaptado de Russel (2004).

2.6. *K-Means*

O método *K-means* é um método heurístico que organiza os elementos de uma amostra em *k* grupos, através da continua geração de centroides, com base nas médias dos grupos gerados a cada iteração; os centros desses grupos são comparados, continuamente, com os elementos da amostra, sendo, os elementos, agrupados ao grupo mais próximo. Dessa forma, são gerados novos grupos, com novos centros, até que os grupos não se alterem mais, minimizando assim, a soma das distâncias totais entre os dados de cada grupo e seu respectivo centro. (JAIN, et al., 1999).

2.7. Algoritmos Genéticos

Segundo Krasnogor e Smith (2005), Algoritmos Evolutivos são um classe de ferramentas de otimização que tem origem nos princípios da evolução das espécies por seleção natural proposto por Darwin (1859).

Essas ferramentas codificam as soluções para um dado problema em forma de cromossomos, onde um conjunto de cromossomos representa uma população de soluções candidatas que irão evoluir através dos mecanismos de seleção, cruzamento e mutação.

Algoritmos Genéticos (AGs) fazem parte do conjunto de algoritmos computacionais pertencentes à Computação Evolutiva (CE) na área de Inteligência Artificial (IA). Esses algoritmos são baseados nos mecanismos de seleção natural e da genética introduzidos por John Holland (1975) e popularizados por David Goldberg (1989). Normalmente são usados para resolver problemas que possuem espaço de soluções complexas e que não possuam um algoritmo de solução em tempo polinomial. Essa metaheurística utiliza de mecanismos para encontrar regiões de soluções viáveis para procurar o ponto com “alta aptidão” que melhor representa uma solução para o problema.

Essa classe de algoritmos estocásticos Darwinistas são normalmente aplicados a problemas que não possuem métodos exatos de solução. Esses problemas são conhecidos como problemas NP-Difícil e os AGs frequentemente encontram soluções, aproximadamente, ótimas. (MICHALEWICZ, 1996, p. 13).

Os AGs possuem reconhecimento científico por servirem de ferramenta robusta para a eficiente otimização de problemas complexos. Algumas das implementações de sucesso dos AGs são: otimização do plano de tratamento em radioterapia conformal 3D (GOLDBARG, et al., 2009), configuração de redes de distribuição (BENTO; KAGAN, 2008), sequenciamento de projetos com recursos limitados (MENDES; GONÇALVES, 2003), no agrupamento de aprendizes de acordo com seus estilos de aprendizagem (SILVA; LORENA, 2011), problemas de agrupamento de grafos orientados (DIAS, 2004); problemas de localização de *p*-medianas (LORENA; FURTADO, 2001); problemas de programação de operações em um ambiente de produção *Flow Shop*, tendo como objetivo minimizar o estoque em processamento (RIBEIRO, et al., 2007); problemas de expansão de sistemas de transmissão (MIASAKI; ROMERO, 2007); problemas químicos (FILHO; POPPI, 1999), etc.

A publicação do naturalista Charles Darwin: “A Origem das Espécies” em 1859 e que inspirou, no futuro, Holland para a criação do AG descreve a teoria da evolução biológica, concluindo que as espécies que melhor conseguem se adaptar ao ambiente terão mais chances de sobreviverem e gerar descendentes mais adaptados (DARWIN, 1859).

O conceito de evolução natural das espécies aplicado à otimização utiliza os conceitos de sobrevivência dos indivíduos mais aptos. Cada cromossomo (possível solução para o problema) é composto por um conjunto de *genes* que quando decodificados representam o *fenótipo*, o fenótipo atribuído a uma “função aptidão” determina a adaptabilidade de um indivíduo, capacidade do *genótipo* em representar uma solução para o problema proposto e quando

maior sua adaptabilidades maiores serão suas chances de gerarem descendentes para a próxima geração (MICHALEWICZ, 1996).

Os AGs durante sua execução mantêm um conjunto de soluções possíveis (uma população de cromossomos), esse conjunto de soluções é responsável pela busca paralela no espaço de solução (*Exploration)*, sendo um diferencial dos AGs, pois alguns outros algoritmos realizam a busca em um único ponto do espaço de busca que é processado a cada iteração (MICHALEWICZ, 1996, p. 16).

No AG proposto por Holland (1975) os indivíduos eram representados por um vetor de valores binários (0-1). A representação por um vetor de valores binários é a representação mais tradicional para o genótipo e segundo Carvalho, Braga, Ludemir (2003), pode não ser a mais adequada para representar uma possível solução para todos os tipos de problemas propostos.

“O princípio básico do funcionamento dos Algoritmos Genéticos é que um critério de seleção vai fazer com que, depois de muitas gerações, o conjunto inicial de indivíduos gere indivíduos mais aptos” (CARVALHO, et al., 2003).

O processo de seleção de um indivíduo é realizado através de uma função que irá avaliar a qualidade do seu “código genético” e, dessa forma, mensurar qual é a sua chance de gerar novos descendentes. Esta função também conhecida como função de aptidão, é uma codificação da função-objetivo do problema e mensura a adaptabilidade do indivíduo, isto é, a qualidade do cromossomo para representar uma ótima, ou quase ótima para o problema. Quando é conhecido a solução ótima do problema podemos observação claramente a adaptabilidade do indivíduo, senão só é possível verificar a adaptabilidade relativa a adaptabilidade média da população.

2.7.1. Principais Definições

Segundo Michalewicz (1996), as principais definições de vocabulário relacionadas aos algoritmos genéticos são emprestadas da teoria genética:

Cromossomo ou Genótipo: vetor que representa as variáveis do problema. Cada cromossomo representa uma solução possível para o problema proposto;

Gene: é a unidade básica do cromossomo. Na codificação binária um conjunto de genes, de acordo com a precisão, representa uma variável do problema. Na codificação real cada gene representa uma variável do problema;

Alelos: são os valores atribuídos a cada gene.

População: conjunto de cromossomos ou soluções;

Fenótipo: cromossomo decodificado. É a solução candidata para o problema;

Geração: corresponde a cada iteração do algoritmo genético, que executa os operadores genéticos para gerar uma nova população;

Operações Genéticas: operações que o algoritmo genético realiza sobre cada um dos cromossomos para gerar as novas populações;

Espaço de busca: é o conjunto de soluções possíveis para o problema;

Função objetivo ou de aptidão: é a forma de se avaliar a qualidade de um fenótipo (solução candidata), refletindo, através de um valor, sua adaptabilidade a um determinado ambiente (problema).

2.7.2. Funcionamento Básico do AG

A estrutura básica de funcionamento, Figura 2-6, de um AG inicia-se com a codificação do espaço de soluções do problema, através de um vetor binário ou de números reais, essa representação é chamada de *Cromossomo* ou *Genótipo*. Posteriormente criamos uma *População Inicial* composta por um

conjunto de cromossomos preenchidos aleatoriamente ou por heurísticas. Durante o processo evolutivo, essa população é avaliada e cada cromossomo recebe uma nota (*fitness*) que representa a adaptabilidade do cromossomo ao problema. Após esse processo são aplicados os operadores genéticos: *Seleção, Cruzamento e Mutação* para a construção das novas gerações mais adaptadas. Esse processo é realizado iterativamente até que o critério de parada seja atingido (MICHALEWICZ, 1996).

O cruzamento e a mutação são dois mecanismos que garantem exploração de novos pontos do espaço de busca (*exploration*), enquanto a seleção dirige a busca para os melhores pontos do espaço de soluções (*exploitation*), e a pressão de seleção é caracterizada pela razão entre a aptidão máxima da população e a aptidão média, sendo que essa pressão de seleção influencia diretamente na quantidade de *Exploitation* e *Exploration* do algoritmo.

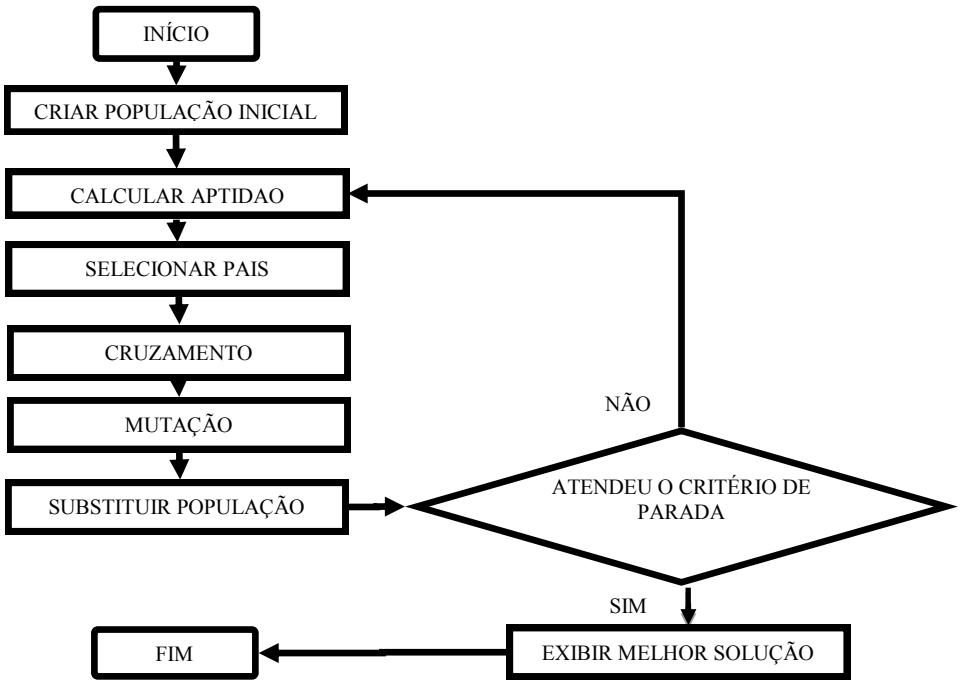


Figura 2-6 – Funcionamento básico de um AG.

2.7.3. Função de Aptidão

A função de Aptidão é a forma como o AG mede o quão adaptado o indivíduo está no ambiente em relação aos outros indivíduos da população. Essa função normalmente corresponde à função-objetivo do problema, fornecendo um valor que permite avaliar a qualidade, adaptabilidade, da solução representada pelo indivíduo.

Uma característica importante da estrutura do AG que possibilita o seu emprego em inúmeras aplicações, sendo a busca pela solução ótima baseada na avaliação de aptidão individual de cada cromossomo, permitindo a modelagem dessa mesma função para cada tipo de problema enfrentado.

2.7.4. Seleção dos Indivíduos

Nesse procedimento o AG executa, com o auxilio da função de aptidão, a seleção dos indivíduos que melhor representam a solução para o problema proposto. A seleção dos indivíduos é um dos mais importantes procedimentos de um AG, pois é o procedimento responsável por selecionar os indivíduos que serão utilizados para a reprodução e que passarão suas características de sobrevivência às gerações futuras.

2.7.4.1. Método da Seleção por Roleta

A seleção pelo Método da Roleta consiste em representar cada indivíduo da população em uma fatia da roleta e a área cedida para cada indivíduo corresponde, proporcionalmente, ao seu valor de aptidão; dessa forma, os indivíduos com maiores aptidões terão mais chances de serem escolhidos para formar a população intermediária que sofrerá as operações genéticas criando uma nova geração. A execução da seleção pelo método da roleta é realizada quantas vezes forem necessárias para formar a população intermediária com os indivíduos escolhidos.

2.7.4.2. Método da Seleção por Torneio

Outro método utilizado para a seleção dos indivíduos é o Método do Torneio que se baseia na formação de subpopulações com *m* indivíduos da população atual que posteriormente serão selecionadas de acordo com a sua aptidão. A formação das subpopulações é aleatória e o cromossomo com maior aptidão dentre estes *m* cromossomos é selecionado. Esse processo é realizado até que a população intermediária fique completa.

2.7.5. Cruzamento

É o operador responsável pela recombinação de características dos pais para formarem um novo indivíduo, garantindo que as novas gerações possuirão as características adaptativas dos pais. Possui grande importância para os Algoritmos Genéticos, pois o operador de cruzamento é quem possui uma taxa de probabilidade (Tc) predominante no algoritmo. De acordo com Carvalho, Braga e Ludemir, (2003), a taxa de cruzamento (Tc) tem usualmente uma de probabilidade entre 60% < Tc < 99%.

As formas mais usuais dos operadores de cruzamento serão apresentadas abaixo:

2.7.5.1. Cruzamento de um Ponto

O ponto de cruzamento é obtido de forma aleatória e os valores após o ponto de cruzamento são trocados entre os indivíduos, Figura 2-7:

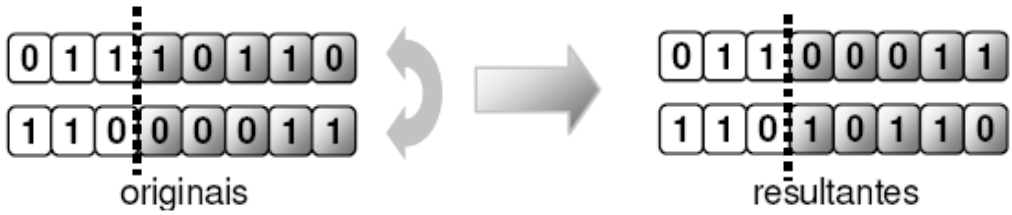


Figura 2-7 - Exemplo de aplicação do operador de cruzamento de um ponto
Fonte: (DIAS, 2004).

2.7.5.2. Cruzamento de Múltiplos Pontos

Os dois pontos são obtidos de forma aleatória e os valores dos genes localizados entre os pontos são trocados entre os indivíduos, Figura 2-8:

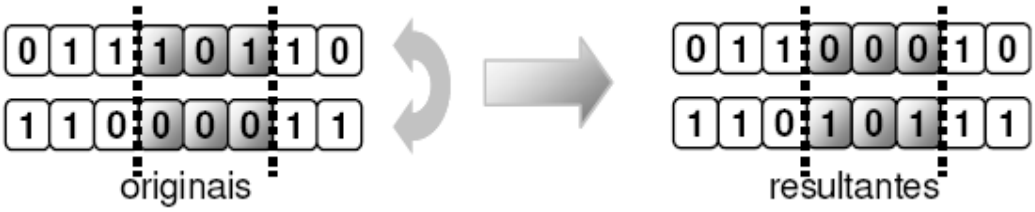


Figura 2-8 - Exemplo de aplicação do operador de cruzamento de múltiplos pontos
Fonte: (DIAS, 2004).

2.7.5.3. Cruzamento uniforme

Uma máscara de dígitos binários é obtida de forma aleatória, onde o dígito 1 (um) indica que o valor na respectiva posição dos indivíduos deverá ser trocado e o dígito 0 (zero) indica que os valores originais serão conservados nas respectivas posições, Figura 2-9:

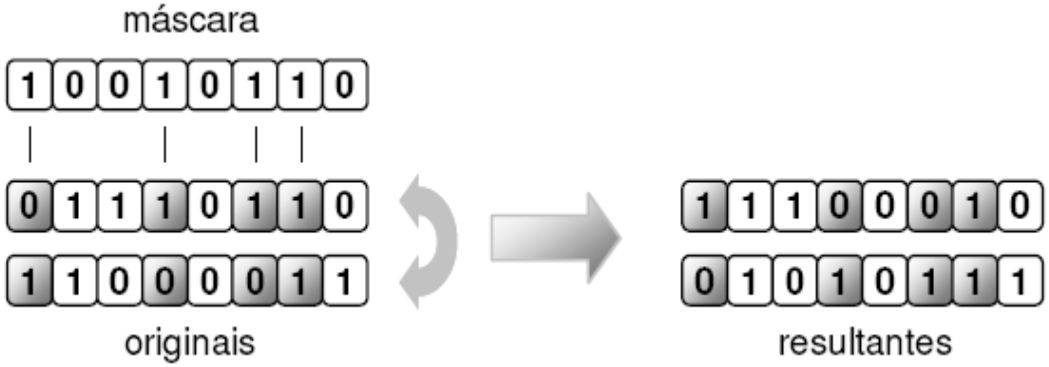


Figura 2-9 - Exemplo de aplicação do operador de cruzamento uniforme
Fonte: (DIAS, 2004).

2.7.6. Mutação

Diferentemente do operador de cruzamento, o operador mutação realiza trocas aleatórias de alguns genes dos indivíduos, com o intuito de pesquisar novas áreas do espaço de busca, a partir de indivíduos selecionados. Com a realização da mutação é acrescentada à população uma diversidade genética o que contorna o problema de mínimos locais, não permitindo que a população fique estagnada em uma região de ótimo local. A Figura 2-10 exemplifica o operador de mutação.

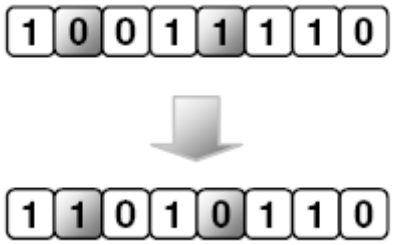


Figura 2-10 - Exemplo de aplicação do operador de mutação
Fonte: (DIAS, 2004).

O operador de mutação possui uma taxa de mutação (Tm), aplicada a cada gene, menor que a taxa de cruzamento (Tc) e essa taxa de probabilidade segundo Carvalho, Braga e Ludemir, (2003), deve ser de 0,1% < Tm < 10%.

2.7.7. Parâmetros Genéticos

A eficiência de um AG está diretamente relacionada com as escolhas de certos atributos como: *tamanho da população, taxa de cruzamento, taxa de mutação, intervalo de geração e critério de parada*; caso esses parâmetros sejam escolhidos de forma equivocada o algoritmo genético pode ser ineficiente e não conseguir convergir para uma solução ótima global.

Tamanho da População: Caso o tamanho da população seja muito pequeno o espaço de busca do Algoritmo Genético se tornar muito restrito, não permitindo uma cobertura adequada do espaço de busca do problema. Já com uma

população grande a representatividade do espaço de busca se torna mais evidente prevenindo que o AG tenha convergência prematura, mas para trabalhar com populações muito grandes é necessário um poder computacional maior ou um tempo maior de trabalho.

Taxa de Cruzamento: O operador de cruzamento é o operador predominante na estrutura do AG e quanto maior é sua taxa de cruzamento mais rapidamente serão inserida novas estruturas promissoras, aumentando assim a velocidade de evolução. Se a taxa for muito alta, estruturas com boas aptidões poderão ser perdidas e se a taxa for muito baixa o algoritmo poderá estagnar (CARVALHO, et al., 2003).

Taxa de Mutação: Com uma taxa de mutação muito alta a busca se torna aleatória, mas com uma taxa adequada previne-se o problema de mínimos locais, isto é, proporciona mecanismos para que a busca não fique parada em sub-regiões vasculhando ao máximo o espaço de soluções.

Intervalo de Geração: Controla a porcentagem, (λ/μ) "*gap* geracional", de μ pais da população atual que serão substituídos por λ filhos da população temporária, gerando assim, uma nova população. Com um intervalo geracional λ (=μ), indivíduos com alta aptidão poderão ser substituídos por indivíduos com baixa aptidão mais facilmente, desestruturando a busca. Já com um intervalo muito pequeno λ (< μ), o algoritmo pode se tornar lento.

Critério de Parada: É o responsável pela parada do algoritmo. A escolha de um Critério de Parada errado pode tornar o AG ineficiente, proporcionando uma parada prematura. Existem alguns critérios muito utilizados para terminar a execução de um AG, tais como: após um número *n* de iterações, quando a aptidão dos indivíduos não melhorar, quando as aptidões se tornarem parecidas ou quando for encontrada a solução ótima para o problema.

2.8. Algoritmo Genético Memético

Em algumas situações em que o espaço de busca do problema é muito complexo observou-se um aumento de eficiência do Algoritmos Evolutivo (AE) quando é acrescenta, durante os ciclos do algoritmo, uma ou mais buscas locais, com a finalidade de melhorar a eficiência de *exploitation* do AE. Essa combinação de técnicas de busca é inspirada em sistemas naturais que combinam adaptação evolucionária de uma população com uma aprendizagem individual durante o tempo de vida de seus membros. Os Algoritmos Genéticos Meméticos (AGM) tiveram inspiração nos conceitos de um *meme* de Richard Dawkins (1976), onde os *memes,* nos AE, são estratégias para realizar refinamento local das soluções da população corrente (KRASNOGOR; SMITH, 2005).

Segundo Krasnogor e Smith (2005) os AGM têm sido utilizados com grande sucesso em várias áreas do conhecimento, como: pesquisa operacional, aprendizagem de máquinas; na otimização de modelos de economia, ecologia e sistemas sociais.

Nesse capítulo foi apresentada a metaheurística destinada a resoluções de problemas de forma alternativa. Essa metaheurística utiliza estruturas baseadas na teoria da evolução natural, sendo uma forma elegante de se realizar buscas paralelas em diferentes áreas do espaço de soluções, sendo muito eficiente para busca de soluções ótimas, ou aproximadamente ótimas, em uma grande variedade de problemas.

No próximo capítulo serão apresentadas as abordagem utilizadas para o agrupamento de laboratórios de metrologia do COMAER.

3. PROBLEMA DE LOCALIZAÇÃO / ALOCAÇÃO DOS LABORATÓRIOS DE METROLOGIA DO COMANDO DA AERONÁUTICA

Metrologia é a ciência da medição e suas aplicações e tem como objetivo primordial fornecer confiabilidade ás medições. Essa confiabilidade é conseguida através de uma cadeia ininterrupta de medições, sendo assim, rastreadas a padrões internacionais de medição. A metrologia é dividida em metrologia científica responsável pelo desenvolvimento e manutenção dos padrões de referência utilizados no território nacional. Existe também a metrologia industrial que é responsável pela adequação dos instrumentos e ensaios utilizados na indústria. Em outra vertente temos a metrologia legal que é responsável pelo controle metrológico das áreas da metrologia que possuem exigências legais compulsórias para as unidades de medição, para os métodos de medição e para os instrumentos de medição (RECHE; ASSIS, 2010).

Para conseguir padrões confiáveis os laboratórios realizam a *Calibração* de seus padrões, essa calibração é uma operação que estabelece uma relação entre os valores e as incertezas dos Padrões de Referência com as indicações correspondentes do *Mensurando (Equipamento/Instrumento de medição)* (INMETRO, 2009).

Essas calibrações são realizadas ininterruptamente entre os níveis metrológicos, definindo assim, a *Rastreabilidade Metrológica*, que é a propriedade de um resultado de medição pela qual tal resultado pode ser relacionado a uma referência através de uma cadeia ininterrupta e documentada de calibrações, cada uma contribuindo para a incerteza de medição (INMETRO, 2009).

Segundo o Vocabulário Internacional de Metrologia a *Incerteza de Medição* é definida como*:* Cálculo estatístico que radiografa a qualidade e a exatidão de um processo de medição, representado por um parâmetro não negativo que caracteriza a dispersão dos valores atribuídos a um mensurando. A incerteza de medição é avaliada através de duas abordagens, a primeira estimada as

componentes por uma avaliação do TIPO A da incerteza de medição, a partir dos desvios padrão das medições realizadas no mensurando. A segunda abordagem de avaliação é denominada avaliação do TIPO B da incerteza de medição que é estimada a partir de funções densidade de probabilidade baseadas na experiência ou em outras informações (INMETRO, 2009).

3.1. Sistema de Metrologia Aeroespacial (SISMETRA)

Ao buscar uma estrutura organizacional para a construção da confiabilidade metrológica no Comando da Aeronáutica (COMAER) foi criado em 07 de dezembro de 1988 (Portaria N° 858/GM3) o Sistema de Metrologia Aeroespacial, o qual tem por finalidade o planejamento, o controle, a coordenação, a normalização e o aprimoramento das atividades relacionadas com a metrologia no âmbito do Comando da Aeronáutica. Essa estrutura organizacional é apresentada na Figura 3-1.

Essa *Confiabilidade Metrológica* é um atributo caracterizado por um nível desejável de eficiência metrológica, obtido por meio da aplicação de técnicas organizacionais e operacionais apropriadas, permitindo assim, que os laboratórios pertencentes ao SISMETRA possam executar de medições confiáveis, adequadas aos fins a que se destinam, bem como a emissão de relatórios, certificados e laudos confiáveis. (BRASIL. MINISTÉRIO DA DEFESA. COMANDO DA AERONÁUTICA, 2008).

O SISMETRA é constituído de organizações militares dispostos em uma estrutura matricial atuando, tanto verticalmente quanto horizontalmente nas atividades metrológicas de interesse do COMAER, sendo composto no seu Nível Estratégico pela Coordenadoria Geral de Metrologia, no Nível Tático pelo Órgão Central do SISMETRA e, no Nível Operacional pelos Elos do SISMETRA (BRASIL. MINISTÉRIO DA DEFESA. COMANDO DA AERONÁUTICA, 2008).

O Nível Estratégico é representado pela Coordenadoria de Metrologia do SISMETRA a qual é um conselho tem a responsabilidade de assessorar o Comandante do Departamento de Ciência e Tecnologia Aeroespacial (DCTA) na elaboração de diretrizes metrológicas, visando operacionalidade para o sistema. Essa coordenadoria é assessorada por especialistas nacionais e/ou internacionais, podendo essa assessoria participar de atividades específicas, relacionadas à Coordenadoria de Metrologia do SISMETRA.

O Nível Tático é representado pela Divisão de Confiabilidade Metrológica Aeroespacial (CMA) pertencente ao Instituto de Fomento e Coordenação Industrial (IFI), localizado em São José dos Campos dentro do DCTA, a qual tem a responsabilidade de planejar, controlar, normalizar e aprimorar o SISMETRA, garantindo recursos humanos e o gerenciamento das informações metrológicas.

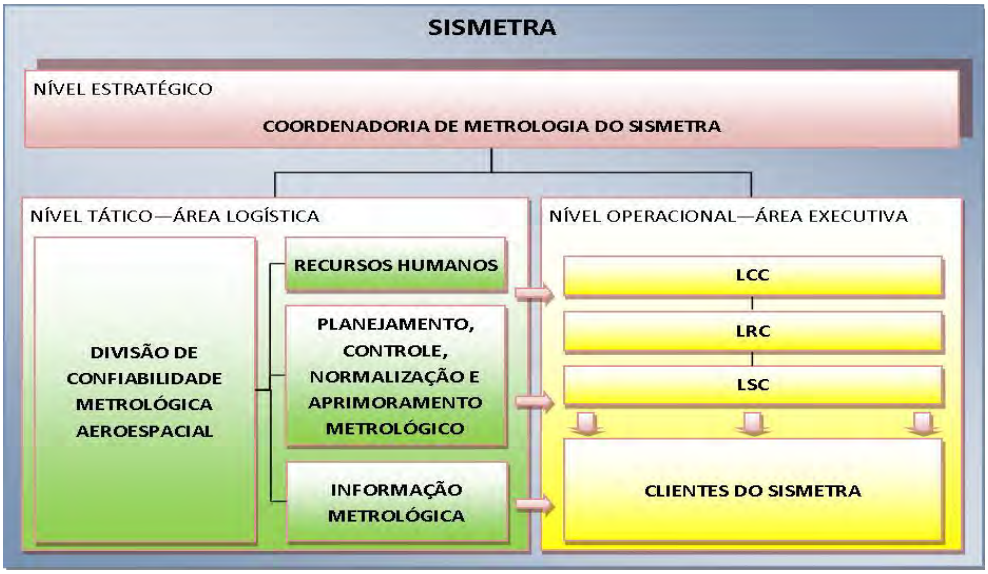


Figura 3-1 - Concepção matricial do SISMETRA

Fonte: Adaptada de BRASIL. MINISTÉRIO DA DEFESA. COMANDO DA AERONÁUTICA (2008).

O Nível Operacional é formado pelo Laboratório Central de Calibração (LCC), conjuntos de laboratórios pertencentes ao DCTA, que é o laboratório de mais alto nível metrológico dentro da estrutura do SISMETRA, estando no topo da

cadeia de rastreabilidade metrológica do COMAER, garantindo que os padrões metrológicos do sistema sejam rastreáveis a padrões nacionais e internacionais de medição. No nível nacional a rastreabilidades é garantida diretamente pelo Laboratório Nacional de Metrologia (LNM) do Instituto Nacional de Metrologia, Normalização e Qualidade Industrial (INMETRO) ou por laboratórios da Rede Brasileira de Calibração (RBC). Quando os padrões do SISMETRA não puderem ser calibrados por laboratórios nacionais de metrologia, os mesmos poderão ser enviados para calibração em laboratórios internacionais de metrologia.

Ainda no Nível Operacional temos os Laboratórios Regionais de Calibração (LRC), laboratórios subordinados ao LCC, tendo seus padrões calibrados pelo LCC. Os LRC são laboratórios alocados em Organizações Militares (OM) no território nacional, sendo esses laboratórios autorizados pelo Órgão Central para fornecer rastreabilidade aos padrões dos Laboratórios Setoriais de Calibração (LSC). Os LSC são laboratórios dispostos no território nacional e que tem a reponsabilidade de calibrar instrumentos de medição, equipamento, sistemas de medição para fins de medição e controle dos processos tecnológicos de interesse do COMAER.

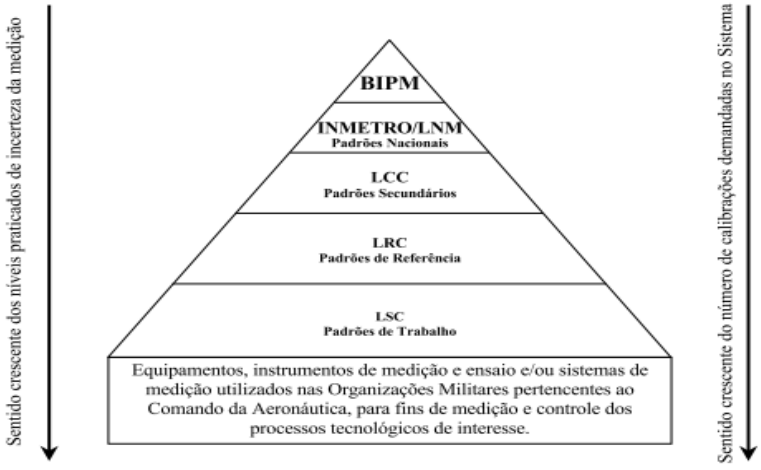


Figura 3-2 - Cadeia de rastreabilidade do SISMETRA

Fonte: (BRASIL. MINISTÉRIO DA DEFESA. COMANDO DA AERONÁUTICA, 2008)

A Figura 3-2 ilustra os níveis hierárquicos que proporcionam rastreabilidade dos padrões metrológicos do SISMETRA a padrões nacionais e internacionais. Nessa estrutura o INMETRO, através de seu LNM, envia seus padrões ao *Bureau Internacional Des Poids Et Measures* (BIPM), na França, ou realiza a comparação de seus padrão com padrões de outros países mediantes as comparações chave coordenados pelo próprio BIPM, ou por outros Institutos Nacionais de Metrologia (INM) de reconhecimento internacional que dão suporte ao Sistema Internacional de Unidades (SI).

3.2. Localização-Alocação Atual dos Laboratórios do SISMETRA

A estrutura hierárquica do SISMETRA é utilizada para o aprimoramento dos recursos metrológicos humanos e instrumentais, onde os laboratórios referência fornecem suporte para a formação de recursos humanos, realizando visitas técnicas para unificação do conhecimento técnico necessário, visando à realização confiável das atividades de medição e calibração demandadas pelas atividades de manutenção aeronáutica.

Essa estrutura em árvore com níveis metrológicos hierárquicos possibilita um fluxo de demanda em que os laboratórios de nível inferior (demanda) são atendidos pelo seu laboratório (facilidade concentradora) do nível superior, concentrando assim, as demandas de rastreabilidade e fornecendo estrutura para o entendimento de todo o sistema.

Os laboratórios do SISMETRA estão estruturados de acordo com o seguinte organograma, apresentado na Figura 3-3.

Nesse organograma pode-se visualizar a hierarquia dos níveis metrológicos onde o DCTA representa o LCC (raiz de arvore), sendo o laboratório ligado diretamente ao LNM e que dissemina a rastreabilidade nacional ao SISMETRA. Temos também os laboratórios concentradores de demanda, os LRC (nós intermediários), que são os laboratórios que ampliam a abrangência da rastreabilidade atendendo as demandas dos LSC (nós folhas) que são os

laboratórios responsáveis pelo controle da instrumentação diretamente ligada a manutenção aeronáutica.

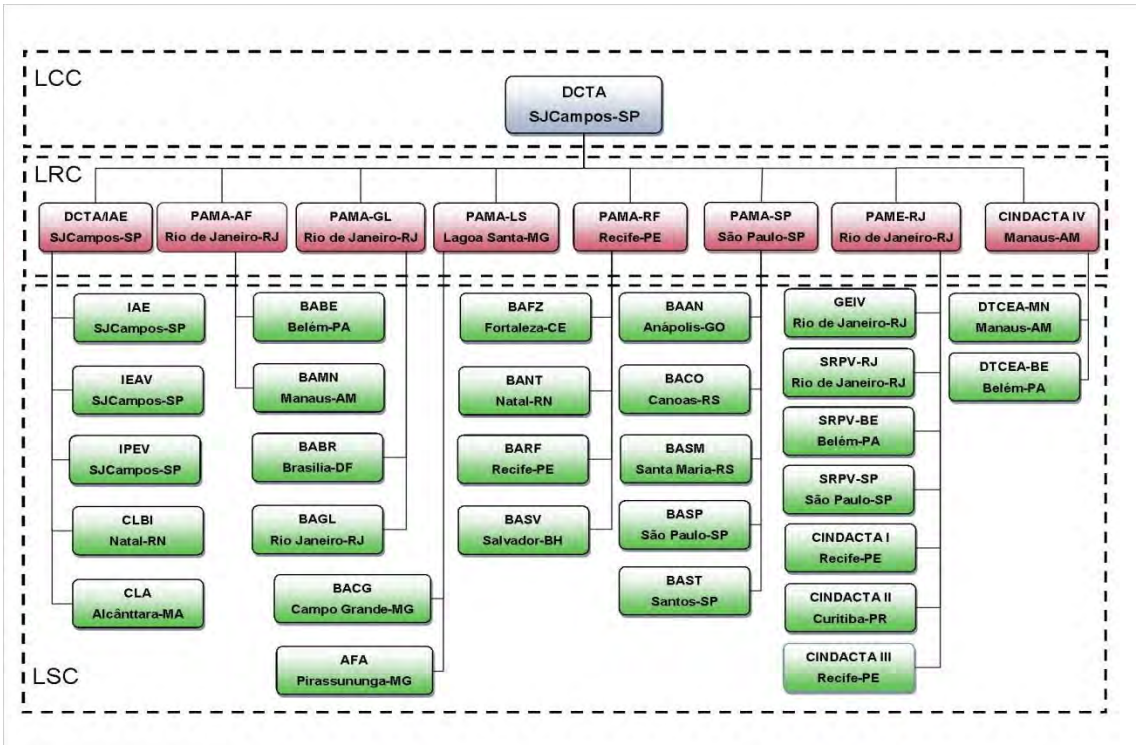


Figura 3-3 – Organograma do SISMETRA.

Os laboratórios do SISMETRA pertencem a uma estrutura hierárquica composta por 3 (três) níveis, sendo que o LCC está no topo da pirâmide os LRCs estão na parte intermediária da cadeia e os LSC estão na base da pirâmide. Consequentemente, essa estrutura permite que todos os laboratórios sejam rastreados ao Sistema Internacional de Unidades (SI) (INMETRO, 2009).

Os laboratórios do COMAER atendem as necessidades de calibração dos laboratórios de nível inferior, clientes, do sistema em três áreas de grandeza: Dimensional (D), Elétrica (E), e Física (F), onde os clientes possuem demandas distintas para cada uma das áreas. Essa particularidade consiste das características tanto de padrões metrológicos como de conhecimentos técnicos diferentes.

Como cada laboratório possui setores com responsabilidades independentes, possuindo capacidades de atendimento distintas de suas respectivas demandas. Temos então, um atendimento em que cada setor dos laboratórios subordinados exerce uma carga de demanda particular para cada setor metrológico do nível acima.

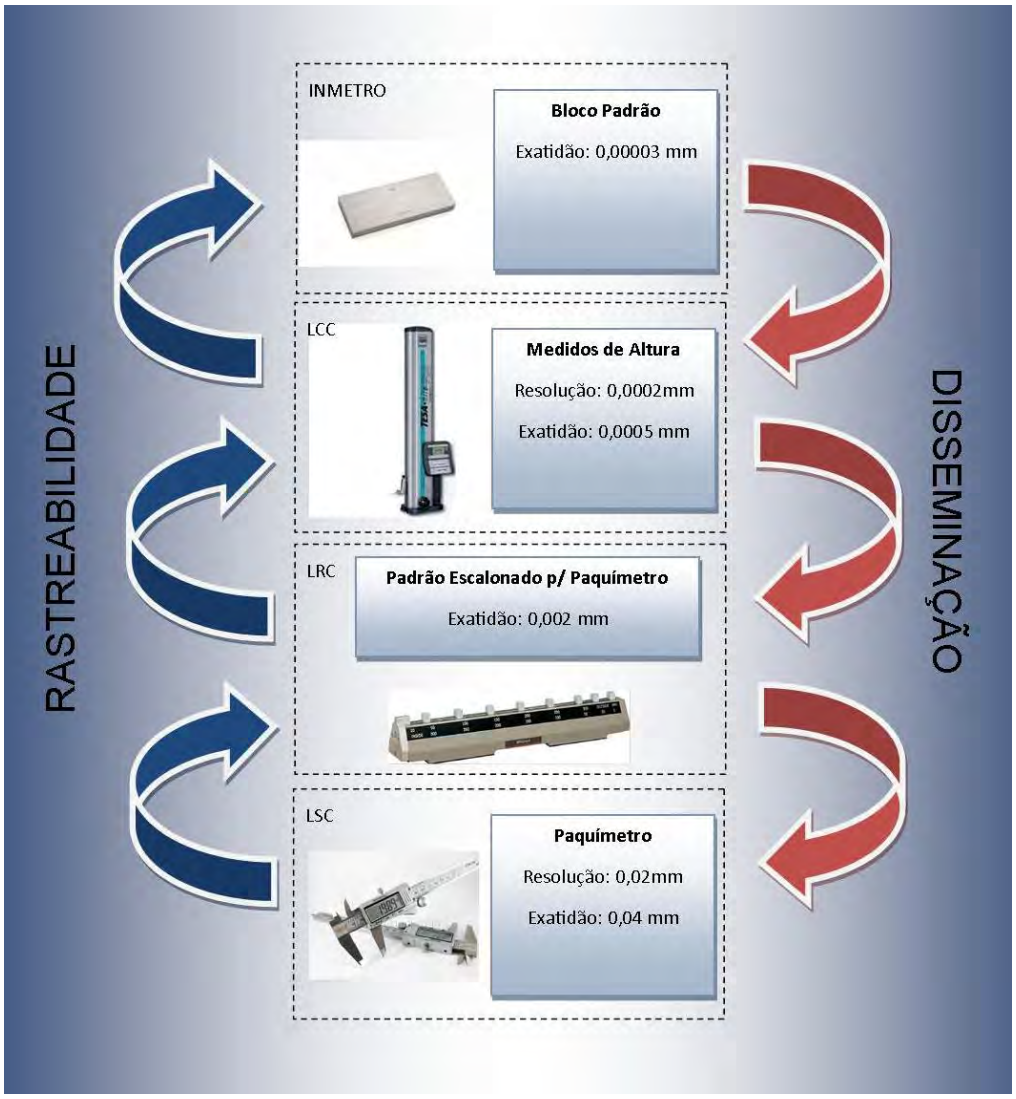


Figura 3-4 – Exemplo de rastreabilidade para calibração de paquímetro

A Figura 3-4, exemplifica a cadeia de rastreabilidade para calibração de paquímetro (instrumento da área dimensional para medição de comprimentos até 150 mm com resolução de 0,02mm). Nessa exemplificação pode-se observar que as sucessivas calibrações iniciam-se com o órgão de referência

nacional INMETRO, que de acordo com suas atribuições dissemina a rastreabilidade aos padrões primários do SISMETRA, através da calibração dos padrões do LCC. No exemplo o INMETRO calibra o Bloco Padrão do LCC que internamente calibra o seu padrão de transferência (Medidor de Altura), depois o LCC recebe e calibra o Padrão Escalonado para Paquímetro do LRC (concentrador) que posteriormente realizará a calibração dos Paquímetros dos Laboratórios Setoriais de Calibração – LSC alocados a ele e que serão utilizados nos serviços de manutenção aeronáutica do COMAER. Dessa forma, o SISMETRA consegue garantir confiabilidade metrológica das suas atividades de manutenção.

Essa múltipla capacidade dos laboratórios, apresentada na Figura 3-5, busca atender as múltiplas demandas dos laboratórios alocados a ele, e que durante as janelas de calibração (período agendado para o envio e calibrações dos instrumentos do cliente) enviam de forma paralela os instrumentos a cada área de calibração do seu laboratório de referência.

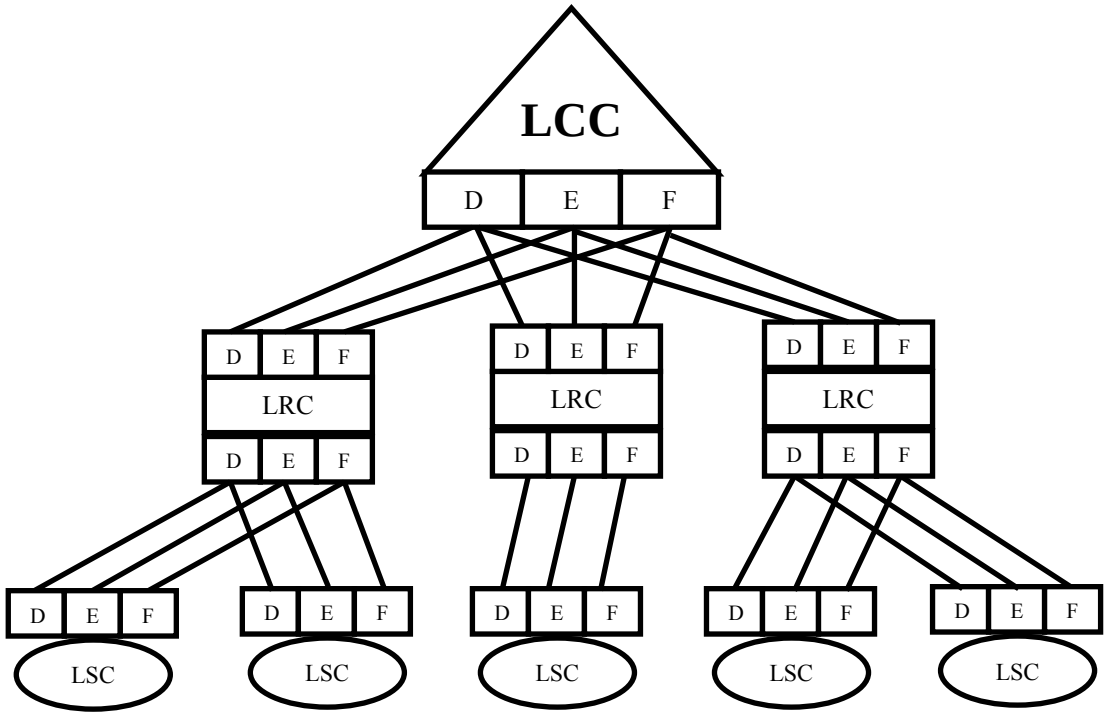


Figura 3-5 – Estrutura de múltiplas demandas do SISMETRA

Para que seja possível representar efetivamente a estrutura e funcionalidade do SISMETRA, Figura 3-5, foram realizadas algumas mudanças na formulação apresentada na Equação (2-19), transformando o modelo CFLP-S em um modelo para o Problema de Localização de Concentradores com Múltiplas Capacidades em uma Rede Estrela-Estrela (MCCLP-S – *Multi-Capacitated Concentrator Location Problem in a Star–Star Network*):

Minimizar $$z(x,y) = \sum_{s\in S}\sum_{i=1}^{n}\sum_{j=1}^{m} Q_{is}C_{ij}x_{ij} + \sum_{j=1}^{m} D_j y_j \quad (3\text{-}1)$$

Sujeito a: $$\sum_{j=1}^{m} x_{ij} = 1, \qquad i = \{1,\dots,n\} \quad (3\text{-}2)$$

$$\sum_{i=1}^{n} Q_{is}x_{ij} \le K_{js}y_j, \qquad j = \{1,\dots,m\}\ e\ s = \{1,2,3\} \quad (3\text{-}3)$$

$$x_{ij}, y_j \in \{0,1\}, \qquad \forall\, i,j \quad (3\text{-}4)$$

onde:

- *Terminais*, $T_i, i \in N, N = \{1,\dots,n\}$, Laboratório Setoriais de Calibração LSC com estrutura metrológica normalmente idênticas, dispersos geograficamente e de localização conhecida.
- *Concentradores*, $C_j, j \in M, M = \{1,\dots,m\}$, Laboratórios Regionais de Calibração com capacidade metrológica superior aos LRC, podendo ou não estarem localizadas juntos aos terminais, e cujo número é conhecido e a localização é desconhecida.
- *Unidade Central,* C_0, Laboratório Central de Calibração LCC, laboratório de grande porte quando comparado com os demais laboratórios e de localização conhecida.
- *Áreas de Grandeza*, $s \in S, S = \{1 = Dimensional, 2 = Elétrica, 3 = Física\}$, representa as três áreas de atendimento de calibração.
- y_j é uma variável binária que assume o valor 1 se a localidade j é escolhida para LRC e 0, caso contrário.

- x_{ij} é uma variável binária que assume o valor 1 se o LSC T_i está ligado ao LRC C_j e 0, caso contrário.
- D_j representa a distância do LRC na localidade j ao LCC.
- C_{ij} representa a distância do LSC T_i ao LRC C_j.
- Q_{is} representa as múltiplas demandas de calibração do laboratório i das áreas de grandeza s.
- K_{js} representa as múltiplas capacidades máximas de serviços que cada setor do LRC C_j pode atender, com $s \in \{1,2,3\}$ representando as subárea de cada laboratório: Dimensional, Elétrica e Física.

No modelo modificado $n = m$ e os LSC i são ligados diretamente aos LRC j, e os LRC j ligados diretamente ao LCC C_0. A restrição (3-2) garante que cada LSC i será ligado a um LRC j. Já a restrição (3-3) foi alterada e agora gerencia a capacidade de atendimento de cada setor $s \in \{1,2,3\}$ do LRC j da demanda Q_{is} de cada setor do LSC i, para que a mesma não seja excedida. A restrição (3-4) ainda garante a integralidade das variáveis de decisão do modelo.

Pode-se transformar o modelo para o MCCLP-S em um modelo para o Problema de Localização de p-Concentradores com Múltiplas Capacidades em uma Rede Estrela-Estrela (MCpCLP-S *Multi-Capacitated p-Concentrator Location Problem in a Star–Star Network*), acrescentando a restrição abaixo:

$$\sum_{i=1}^{m} y_j = p \qquad (3\text{-}5)$$

Com a restrição (3-5) pode-se definir o número *p* de concentradores que serão localizados pelo modelo matemático. Sendo que no modelo utilizou-se ***p***=8, pois a estrutura atual do SISMETRA possui 8 Laboratórios Regionais de Calibrações e esses laboratórios representados os concentradores através das variáveis de decisão y_j do modelo.

3.3. Determinação da Matriz de Custos

Para determinar um parâmetro de similaridade entre as localidades militares, utilizadas na realocação dos laboratórios de metrologia do Sistema de Metrologia Aeroespacial (SISMETRA), foi levado em considerado que a maioria das atividades logísticas para a manutenção do sistema, são realizadas através das aeronaves militares. Dessa forma, como as organizações militares estão dispostas no território nacional, pode-se mensurar a distância mais representativa entre as cidades que hospedam esses laboratórios, através da Distância Geodésica (menor distância entre dois pontos sobre uma elipsoide). Essa distância é medida sobre o circulo formado quando se realiza um corte bem no centro do globo terrestre, gerando um circulo máximo o “great circle” (exemplo: Equador e Meridiano de *Greenwich*). Essa medição do comprimento do semicírculo gerado entre as duas localidades define a distância geodésica, podendo ser determinada em radianos pela equação (3-6) (LOVE, et al., 1988, p. 38).

$$d(l_i, l_j) = cos^{-1}\left[\cos\ l_{i1}\ \cos\ l_{j1} \cos\ (l_{i2} -\ l_{j2}) + \sin l_{i1} \sin l_{j1}\right] \quad (3\text{-}6)$$

onde:

l_i representa, respectivamente, as coordenadas de l_{i1} =latitude e l_{i2} =longitude do laboratório i.

l_j representa, respectivamente, as coordenadas de l_{j1} =latitude e l_{j2} =longitude do laboratório j.

Através da equação (3-6) pode-se determinar a matriz de distâncias, C_{ij}, que será utilizada nos processos de otimização da solução para a realocação hierárquica e de múltiplas capacidades dos laboratórios do SISMETRA, onde as coordenadas de Latitude e Longitude foram coletadas do site http:// https://maps.google.com.br/ tendo como referência os endereços das unidades militares do COMAER que hospedam laboratórios de metrologia. Multiplicando

as distâncias calculadas em radianos pelo raio terrestre em quilometro temos a matriz de distâncias final em quilômetros.

3.4. Determinação da Matriz de Capacidades e Matriz de Demandas

A matriz de múltiplas demandas, Q_{is}, foi determinada aleatoriamente, utilizando para isso, a função de números aleatórios do Excel 2010.

Tabela 3-1: Matriz de múltiplas demandas

N°	LOCALIDADE	DEMANDAS			N°	LOCALIDADE	DEMANDAS		
		D	E	F			D	E	F
1	LRC - DCTA/IAE	17	14	19	**20**	LSC – BAFZ	23	20	4
2	LRC - PAMA-AF	11	12	18	**21**	LSC-BANT	23	17	23
3	LRC - PAMA-GL	6	7	20	**22**	LSC-BARF	24	5	19
4	LRC - PAMA-LS	16	4	11	**23**	LSC-BASV	23	9	16
5	LRC - PAMA-RF	22	20	20	**24**	LSC-BAAN	9	6	4
6	LRC - PAMA-SP	11	16	24	**25**	LSC-BACO	5	6	7
7	LRC - PAME-RJ	21	13	14	**26**	LSC-BASM	12	5	8
8	LRC - CINDACTA IV	24	25	14	**27**	LSC-BASP	19	16	24
9	LSC - IAE	5	21	4	**28**	LSC-BAST	22	4	9
10	LSC - IEAV	12	17	10	**29**	LSC – GEIV	4	7	9
11	LSC - IPEV	10	4	21	**30**	LSC - SRPV-RJ	15	15	19
12	LSC - CLBI	6	18	10	**31**	LSC - SRPV-BE	16	24	18
13	LSC - CLA	14	4	9	**32**	LSC - SRPV-SP	15	19	25
14	LSC-BABE	21	24	21	**33**	LSC - CINDCTA I	8	19	10
15	LSC-BAMN	18	7	10	**34**	LSC - CINDCTA II	22	9	5
16	LSC-BABR	18	10	15	**35**	LSC - CINDCTA III	15	21	25
17	LSC-BAGL	10	23	25	**36**	LSC - DTCEA-MN	8	11	17
18	LSC-BACG	14	9	9	**37**	LSC - DTCEA-BE	8	11	17
19	LSC-AFA	14	9	9					

Essa matriz de demandas, Q_{is}, foi utilizada para a determinação das matrizes de múltiplas capacidades, K_{is}, utilizadas pelas abordagens de otimização. Para determinar essas matrizes foram utilizadas duas abordagens. A primeira abordagem extrapola o valor de capacidade de cada área de grandeza dos laboratórios, esse valor foi determinado somando-se todas as demandas de todos os laboratórios e seus respectivos setores e dividindo esse valor pelo

número de concentradores, ***p***=8. Com isso amplia-se a região de soluções viáveis, tendo com isso, a distâncias entre as localidades um peso maior na otimização.

A segunda abordagem determina as capacidades somando as demandas de todos os laboratórios, separadamente, para cada área de grandeza e dividindo esses 3 (três) somatórios pelo número de concentradores, *p*=8. Dessa forma, cada área de grandeza obtém um valor capacidade de atendimento, sendo essa capacidade atribuída a todos os laboratórios. Dessa forma minimiza-se a região de soluções viáveis, gerando mais dificuldade para a otimização.

3.5. Algoritmo Genético Simples Proposto para o Problema

3.5.1. Representação dos Indivíduos

Buscando a simplicidade e eficiência para representação dos indivíduos esse trabalho utiliza a representação proposta por Topcuoglu, Corut, Ermis, & Ylmaz (2005) e utilizada por Almeida (2009, p. 44). Essa representação é formada por dois vetores que, respectivamente, representam os concentradores localizados e as alocações selecionadas. É utilizada uma representação binária para localizar os concentradores, onde em, S_{1j}, cada *j-ésima* posição do *hubArray* quando igual a 1 (um) representa que foi localizado um concentrador na *j-ésima* localidade e 0 (zero) caso contrário. Na representação das alocações utiliza-se uma representação inteira para S_{2i}, onde na *i-ésima* posição do *AssignArray* armazena-se a *j-ésima* posição do concentrador ao qual a *i-ésima* localidade está alocado.

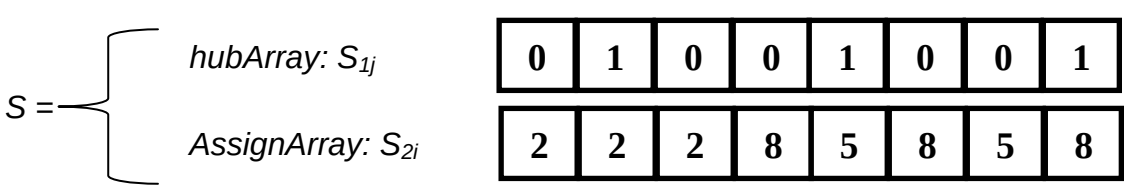


Figura 3-6 – Representação do Indivíduo

A Figura 3-6 demonstra uma alocação dos nós 1 e 3 ao concentrador 2; do nó 7 ao concentrador 5 e dos nós 4 e 6 ao concentrador 8. Os nós onde são localizados os concentradores são alocados a própria localidade no vetor *AssignArray*. Essa representação proporciona eficiente movimentação sobre o espaço de busca tornando fácil o cálculo da função objetivo.

3.5.2. População Inicial

A População Inicial de tamanho *Np* é composta por um conjunto de cromossomos $S=\{S_1,..., S_n\}$, onde inicialmente é preenchido, de forma aleatória, o vetor *hubArray* com um conjunto de concentradores = *p* e depois cada nó Terminal T_i é alocado, aleatoriamente, a um dos C_j concentradores localizados anteriormente

3.5.3. Operador de Cruzamento

O operador de cruzamento utilizado na abordagem proposta é o operador de dois pontos, onde os vetores *hubArray e AssignArray* são empilhados e dois pontos de cruzamento aleatórios cortam os dois cromossomos e os genes entre os dois pontos são trocados, gerando dois novos filhos.

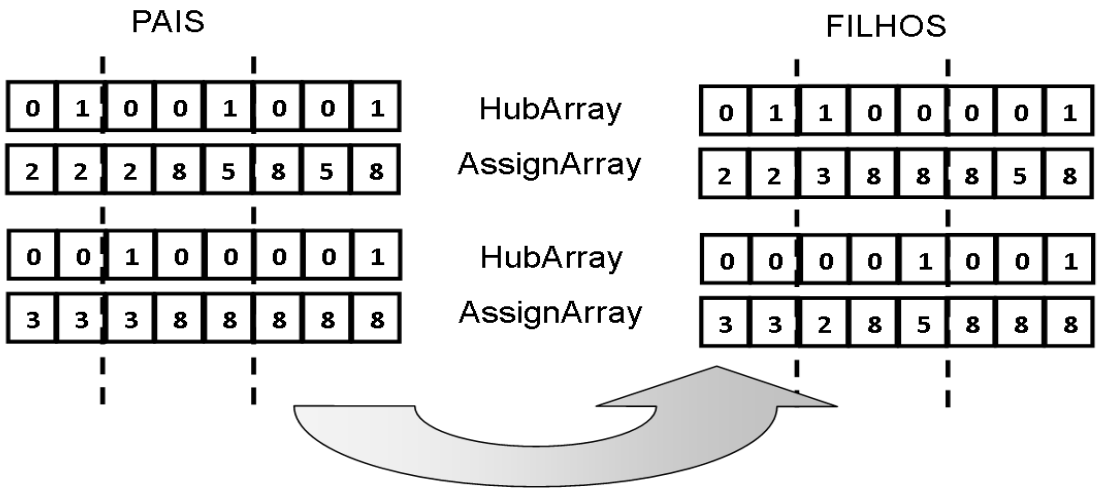


Figura 3-7 – Exemplo do operador de cruzamento de dois pontos

Com a operação de cruzamento alguns filhos podem não representar uma solução válida para o problema, sendo necessário realizar um ajuste no cromossomo. Esse ajuste pode ser realizado em duas situações:

1. Quando o cruzamento gerar uma solução em que uma alocação não tenha o seu respectivo concentrador localização.

    - **Ajuste:** Verificar, no sentido crescente, qual *j-ésimo* concentrador não foi localizado no vetor *hubArray* e realizar e alteração para 1 (um) da *j-ésima* posição.

2. Quando uma solução tiver um concentrador localizado sem nenhuma alocação referenciada no vetor *AssignArray.* Esses erros de representação são corrigidos através dos seguintes procedimentos:

    - **Ajuste:** Verificar, no sentido crescente, qual *j-ésimo* concentrador não possui alocações referenciadas e atribuir ao *j-ésimo* gene do vetor *hubArray* o valor 0 (zero).

Abordagens de ajustes semelhantes são apresentadas em Topcuoglu, Corut, Ermis e Ylmaz (2005).

3.5.4. Operador de Mutação

Pelo tipo de problema e pela forma de representação utilizada existem alguns movimentos com a representação dos concentradores e com representação dos terminais que podem ser implementados, aplicando diversidade à população, permitindo que ela tenha recursos para fugir dos ótimos locais e encontrar regiões promissoras para o refinamento da solução. Nesse trabalho foram aplicados 6 (seis) tipos de movimentos permitidos pela representação (*M1, M2, M3, M4, M5, M6*), onde os operadores de mutação *M3* e *M4* se baseiam nos procedimentos de mutação utilizados em Topcuoglu, et al. (2005).

Os operadores de mutação utilizados possuem suas respectivas taxa de probabilidades de mutação (*Tm1, Tm2, Tm3, Tm4, Tm5, Tm6*).

- **M1:** Nesse procedimento é selecionado um concentrador C_j com probabilidade *Tm1* e o movimento de mutação consiste em trocar a localização do concentrador para um nó terminal pertencente ao grupo do concentrador C_j, e após a troca alocar o antigo concentrador e nós nos pertencentes ao grupo ao novo concentrador. Esse processo de seleção de concentradores é realizado no sentido crescente do índice j. O Operador de Mutação M1 é apresentado na Figura 3-8.

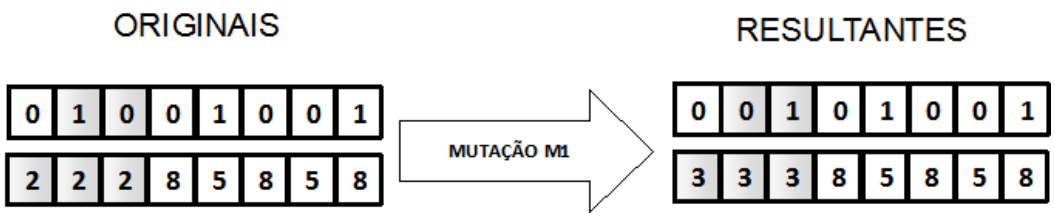


Figura 3-8 – Exemplo do operador de mutação M1

- **M2:** Nessa operação realiza-se a permutação entre dois concentradores selecionados, atribuindo à localidade de um para o outro e vice e versa. Durante o processo o primeiro concentrador é selecionado com probabilidade *Tm2* no sentido crescente do índice j e o segundo concentrador é escolhido de forma aleatória. O Operador de Mutação M2 é apresentado na Figura 3-9.

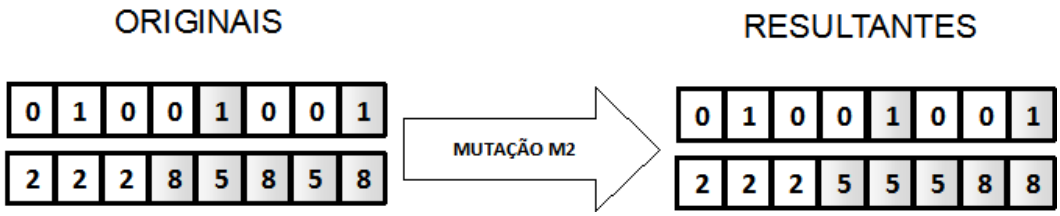


Figura 3-9 – Exemplo do operador de mutação M2

- **M3:** Nesse procedimento são selecionados nós terminais T_i e o movimento de mutação consiste em realocar esse nó para outro concentrador caso esse nó seja o único alocado ao concentrador a operação de mutação não é realizada. A taxa de mutação *Tm3* é aplicada a cada gene do vetor *assignArray*. Esse processo de seleção de terminais é realizado no sentido crescente do índice i. O Operador de Mutação M3 é apresentado na Figura 3-10.

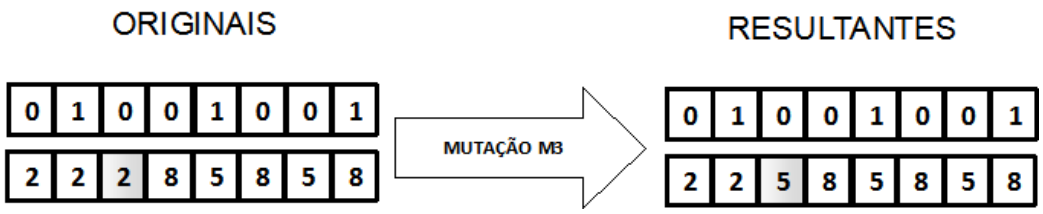


Figura 3-10 – Exemplo do operador de mutação M3

- **M4:** Nessa operação realiza-se a permutação entre dois nós terminais, atribuindo o concentrador de um para o outro e vice e versa, caso esses nós não sejam os únicos nós alocados aos concentradores. A taxa de mutação *Tm4* é aplicada a cada gene do vetor *assignArray*. O processo de seleção do primeiro terminal é realizado no sentido crescente do índice i e o segundo terminal e escolhido de forma aleatória. O Operador de Mutação M4 é apresentado na Figura 3-11.

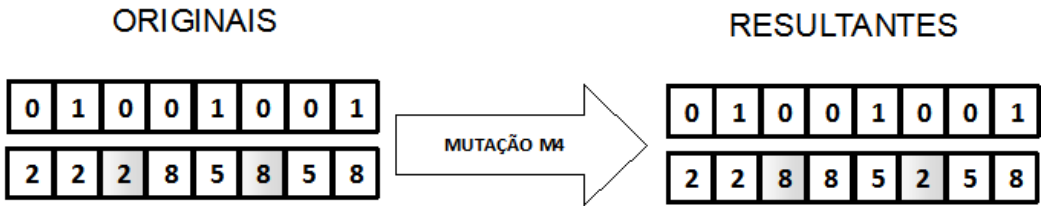


Figura 3-11 – Exemplo do operador de mutação M4

- **M5:** Nessa operação realiza-se a fusão entre dois grupos. São escolhidos dois concentradores aleatoriamente com probabilidade *Tm5* e as alocações do segundo concentrador e o segundo concentrador são alocados ao primeiro concentrador. O Operador de Mutação M5 é apresentado na Figura 3-12.

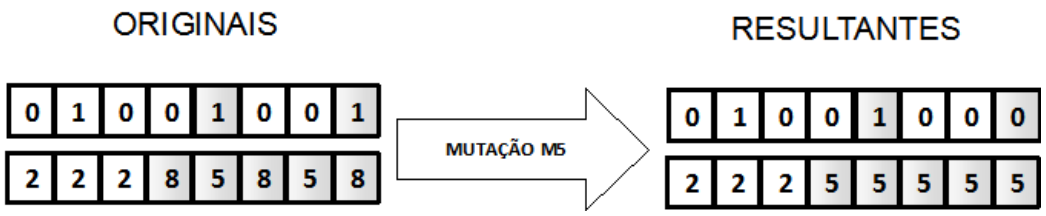


Figura 3-12 – Exemplo do operador de mutação M5

- **M6:** Nessa operação um grupo é dividido em dois. É escolhido aleatoriamente com probabilidade *Tm6* um terminal T_i e esse terminal é alocado ao concentrador C_j, onde esse terminal escolhido é transformando em um novo concentrador e aleatoriamente alguns terminais alocados ao concentrador C_j são alocados ao novo concentrador C_i. O Operador de Mutação M6 é apresentado na Figura 3-13.

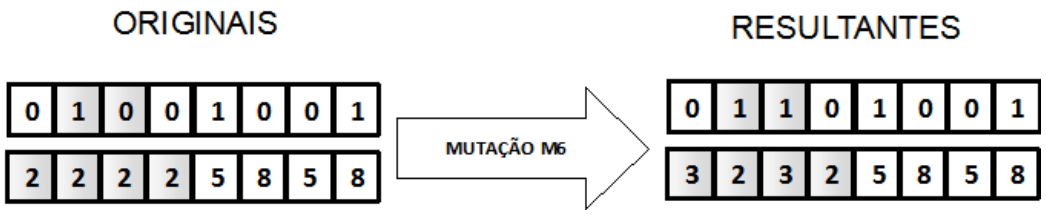


Figura 3-13 – Exemplo do operador de mutação M6

As abordagens de mutação M1, M2, M3, M4, M5 e M6 são aplicadas sequencialmente, Figura 3-14, na população corrente da iteração.

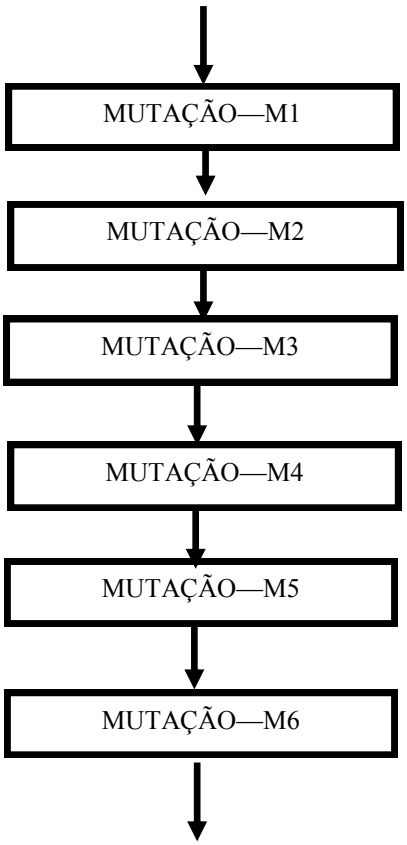


Figura 3-14 – Sequência de aplicação dos mecanismos de mutação

3.5.5. Intervalo de Geração

O intervalo de geração utilizado e o Intervalo Geracional λ (=μ) (μ=n° de pais, λ=n° de filhos) acrescido de Elitismo, onde o melhor indivíduo da população corrente é mantido na próxima geração, substituindo os *k* piores indivíduos dessa nova população.

3.5.6. Função Aptidão da Abordagem Genética proposta

A função de Aptidão é a forma como o AG mede a capacidade do indivíduo de se adaptar ao ambiente, isto é, o quanto o cromossomo consegue representar uma solução para o problema proposto. Essa função normalmente corresponde à função-objetivo do problema, fornecendo um valor que permite avaliar a qualidade, adaptabilidade, da solução representada pelo indivíduo (DIAS, 2004).

A função utilizada na abordagem utiliza a função objetivo do modelo (3-1) acrescida de penalidades para o gerenciamento das restrições de capacidade e do número de concentradores.

$$Aptidao = \sum_{s\in S}\sum_{i=1}^{n}\sum_{j=1}^{m} Q_{is}C_{ij} + \sum_{j=1}^{m} D_j + \theta(S) + \rho(S);\ \forall\, S_{1j} = 1\ e\ S_{2i} = j\ e\ i \neq j \qquad (3\text{-}7)$$

Os modelos para otimização de problemas do mundo real, frequentemente, são compostos de restrições, para as variáveis de decisão, que precisam ser gerenciadas pelo algoritmo de otimização. O modelo, (3-1), proposto para localizar e alocar os laboratórios do SISMETRA possui 4 (quatro) restrições que garantem:

1. Restrição (3-2) garante que cada LSC i será ligado a um LRC j. Essa restrição é gerenciada pela própria representação do cromossomo, pois cada *i-ésimo* gene do vetor *AssignArray* terá apenas um fenótipo representando o seu *j-ésimo* concentrador.

2. Restrição (3-3) gerencia a capacidade de atendimento de cada setor $s \in \{1,2,3\}$ do LRC j da demanda Q_{is} de cada setor do LSC i, para que a mesma não seja excedida. Essa restrição é gerenciada com a penalização da função de aptidão, onde:

$$\theta(S) = Tpenal * \sum_{s\in S}\sum_{j=1}^{m} Máx\left(0; \sum_{i=1}^{n} Q_{is} - K_{js}\right);\ \forall\, S_{1j} = 1,\ S_{2i} = j\ e\ i \neq j \qquad (3\text{-}8)$$

3. Restrição (3-5) garante o numero de concentradores que serão localizados. Essa restrição é gerenciada com a penalização da função de aptidão, onde:

$$\rho(S) = abs\left(\sum_{i=1}^{m} S_{1j} - p\right) * Tpenal \qquad (3\text{-}9)$$

4. A Restrição (3-4) garante que as variáveis de decisão são valores binários é gerenciada, intrinsicamente, pela representação proposta para o cromossomo utilizado.

A constante *Tpenal* é um fator de penalidade utilizado para ajustar o peso da penalidade na função objetivo.

3.6. Algoritmo Genético Memético - AGM

Para transforma o AGS em um Algoritmo Genético Mémetico (AGM) foi acrescentada uma busca local após a substituição da população com elitismo, com o intuito de potencializar o mecanismo de refinamento de soluções promissoras do AGS.

Com a busca local implementada pode-se realizar buscas locais em ambos os vetores *hubArray* e *assignArray*. Fornecendo assim, robustez para o processo de *exploitation* do algoritmo genético.

3.6.1. Busca Local Utilizada

Para melhorar a eficiência do AGS foi incorporada algumas buscas locais durante as iterações do algoritmo. As buscas escolhidas se baseiam no algoritmo de subida da encosta. Nesse algoritmo a cada iteração é criado um conjunto de soluções vizinhas conseguidas pela função *N(s)* que mapeia toda solução $s \in S=\{Espaço\ de\ Busca\ do\ Problema\}$ para um subconjunto $N(s) \subseteq S$, onde um elemento de *N(s)* é denominado vizinho de *s*. Para conseguir um vizinho $s' \in N(s)$ realiza-se uma operação denominada movimento. Na busca local implementada são utilizados 6 operadores de movimentos, onde é realizada uma busca independente para cada operador de movimento. Após a criação de cada vizinhança é selecionado o melhor vizinho *s'* e caso ele seja melhor que a solução corrente *s* a solução corrente é substituída por *s'*. Os operadores de movimento são descritos abaixo:

- **Movimento 1:** O primeiro operador de movimento consiste em escolher o *i-ésimo* terminal no sentido crescente do índice *i* e trocar a alocação desse terminal para os outros concentradores localizados. A vizinhança criada por esse movimento é composta por todas as alocações possíveis para os *i-ésimos* terminais.

- **Movimento 2:** O segundo movimento consiste em fundir dois grupos. São escolhidos dois concentradores e as alocações do segundo concentrador e o segundo concentrador são alocados ao primeiro concentrador. A vizinhança criada por esse movimento é composta por todos os agrupamentos possíveis de dois concentradores da solução corrente.

- **Movimento 3:** Esse movimento consiste em escolher um *j-ésimo* concentrador localizado e trocar a localização dele para a localização dos *i-ésimos* terminais. Após a troca o concentrador antigo e os terminais alocados a ele, são alocados ao novo *i-ésimo* concentrador. A vizinhança formada por esse operador é composta por todas as trocas possíveis trocas entres os *j-ésimos* concentradores e os *i-ésimos* terminais.

- **Movimento 4:** Esse operador de movimento realiza a permutação entre dois terminais, alocando um ao concentrador do outro e você e versa. A vizinhança criada com esse operador e composta por todas as combinações possíveis entre dois terminais da solução corrente.

- **Movimento 5:** Esse operador de movimento realiza a permutação entre dois concentradores, alocando os terminais de um ao outro concentrador e você e versa. A vizinhança criada com esse operador é composta por todas as permutações possíveis entre dois concentradores da solução corrente.

3.7. Resultados Computacionais

Para a otimização do problema proposto foi utilizado o software CPLEX na versão 10.1 (*Software* para otimização de modelos de programação matemática), um Algoritmo Genético Simples e um Algoritmo Genético Memético. As execuções foram realizadas para a instância do problema com 37 localidades. Essa instância com 37 localidades, ilustrada na Figura 3-3, utiliza as localidades que compõem, atualmente, o SISMETRA. Também é apresentado o Melhoramento Logístico (ML) conseguido comparando a atual distribuição dos laboratórios do SISMETRA com a realocação proposta pela otimização.

São utilizadas as siglas: **FOB** para apresentar o valor da função objetivo da abordagem, **NFO** para apresentar o número de chamadas da função objetivo durante a execução da abordagem, $\mathbf{FOB_{médio}}$ para apresentar o valor médio dos valores de função objetivo das execuções realizadas, $\mathbf{NFO_{médio}}$ para apresentar o número médio de chamadas da função objetivo para as execuções realizadas, $\mathbf{FOB_{desvPad}}$ para apresentar o desvio padrão de FOB das execuções realizadas, $\mathbf{NFO_{desvPad}}$ para apresentar o desvio padrão de NFOB das execuções realizadas.

As execuções computacionais foram realizada em um computador *Atom* 1,66 GHz, 2GB de memória *ram*, com *Windows 7 Starter* 32 *bits*.

Para essa proposta são utilizados valores de demanda e capacidade definidos de acordo com o item 3.4. Já a determinação das matrizes de distâncias foi realizada através da equação (3-6).

Tabela 3-2 - Resultados do CPLEX 10.1 para otimização do modelo MCpCLP–S

Capacidades D – E – F	N=M	P	Fob	Tempo (seg)
210 – 210 - 210	37	8	130722	0,59
74 – 66 - 75	37	8	361514	34666

A execução do CPLEX, apresentada pela Tabela 3-2, foi realizada na configuração *default* do *software*, conseguindo encontrar a solução ótima com um tempo de execução de 0,59 segundos para a instância de 37 localidades e com uma matriz de capacidades extrapolada (D=210, E=210, F=210). Já para a matriz de capacidade apertada (D=74, E=66, F=75) o CPLEX não conseguiu encontrar a solução ótima, estourando a memória aos 34666 (9,6 horas) segundos e encontrado uma solução factível com **FOB**=361514. Essa dificuldade encontrada durante a otimização com o CPLEX foi proporcionada pela diminuição do espaço de soluções factíveis acrescido de muitos mínimos locais.

As soluções encontradas pelo *software* CPLEX foram utilizadas como referência para comparação com as abordagens evolutivas propostas AGS e AGM. Todas as propostas de otimização utilizam o modelo MCpCLP-S apresentado na formulação (3-1).

Na Tabela 3-3, são apresentadas as 50 execuções do AGS para um instancia de 37 localidades para a otimização do modelo MCpCLP –S com a matriz de capacidades extrapolada (D=210, E=210, F=210). Essas execuções produziram os seguintes valores médios e seus respectivos desvios padrão: $\mathbf{Fob_{médio}}$=132471 e $\mathbf{Fob_{desvPad}}$=2357; $\mathbf{NFob_{médio}}$=782586 e $\mathbf{NFob_{desvPad}}$=542315; $\mathbf{Tempo_{médio}}$=43 segundos e $\mathbf{Tempo_{desvPad}}$=30 segundos. Com esses resultados constata-se que o AGS apresenta tempo computacional maior que o tempo gasto pela otimização com o CPLEX, e mesmo assim, não conseguiu encontrar a solução ótima em 50% das execuções (25 de 50 execuções).

Tabela 3-3 – Resultados das 50 execuções do AGS para otimização do modelo MCpCLP–S com 37 localidades e capacidades D=210 – E=210 – F=210.

Exec.	NFob	Fob	Tempo (seg)	Exec.	NFob	Fob	Tempo (seg)
1	1158336	131545	64,44	**26**	1075056	136003	58,50
2	1249056	131545	71,03	**27**	1037406	131545	64,37
3	1011252	136003	60,07	**28**	320370	**130722**	17,45
4	1011558	136003	57,67	**29**	180546	**130722**	10,45
5	31938	**130722**	1,71	**30**	1608528	136003	91,62
6	299004	**130722**	16,95	**31**	124764	**130722**	6,61
7	1503912	136003	89,47	**32**	43224	**130722**	2,17
8	1421412	136003	78,99	**33**	162048	**130722**	9,15
9	229776	**130722**	13,04	**34**	1668198	**130722**	94,96
10	846840	**130722**	46,07	**35**	56718	**130722**	3,06
11	856818	**130722**	49,24	**36**	1191144	136003	68,31
12	1040970	131545	57,20	**37**	354600	**130722**	19,35
13	252606	**130722**	14,21	**38**	398580	**130722**	20,28
14	1170996	131545	65,38	**39**	411570	**130722**	20,63
15	1124310	136003	56,56	**40**	1733916	131545	88,54
16	12054	**130722**	0,70	**41**	1120584	131545	76,15
17	1092624	131545	58,20	**42**	94596	**130722**	4,89
18	1090194	131545	67,31	**43**	1208418	136003	66,76
19	1234518	136003	67,93	**44**	1678116	136003	92,70
20	1376214	136003	78,85	**45**	210984	**130722**	12,32
21	420390	**130722**	23,21	**46**	1173054	131545	59,37
22	232554	**130722**	13,36	**47**	1070664	136003	54,22
23	1387452	136003	73,36	**48**	1194078	136003	57,89
24	247260	**130722**	12,52	**49**	108606	**130722**	5,43
25	396984	**130722**	20,47	**50**	204486	**130722**	9,98

	$NFOB_{médio}$	**Desvio Padrão NFOB**	**$FOB_{médio}$**	**Desvio Padrão FOB**	**$Tempo_{Médio}$**
MÉDIA	782586	542315	132471	2357	43

As execuções do AGS foram realizadas utilizando o modelo MCpHLP–S com ***p***=*8* e ***n***=***m***=37 localidades. Foi utilizado o método de seleção pelo método da roleta, o operador de cruzamento de dois pontos e os operadores de mutação *M1, M2, M3, M4*, onde, através de testes empíricos, foram padronizados os seguintes parâmetro genéticos: ***Tc***=0,8, ***Tm1***=0,1, ***Tm2***=0,001, ***Tm3***=0,005, ***Tm4***=0,005, ***Tpenal***=10000 e critério de parada igual a 1 milhão de iterações

ou se o algoritmo encontrar a solução ótima (solução conseguida com a otimização do CPLEX). Também utilizou um intervalo de geração geracional *λ (=μ)* (*μ*=n° de pais, *λ*=n° de filhos) acrescido de Elitismo, onde o melhor indivíduo da população corrente é mantido na próxima geração, substituindo os *k=2* piores indivíduos dessa nova população. Também foram padronizados o tamanho da população, ***Np***=6.

Tabela 3-4 – Resultados das 50 execuções do AGM para otimização do modelo MCpCLP–S com 37 localizadas e capacidades D=210 – E=210 – F=210.

Exec.	NFob	Fob	Tempo (seg)	Divers.	Exec.	NFob	Fob	Tempo (seg)	Divers.
1	5700	**130722**	0,73	11	**26**	7206	**130722**	0,29	22
2	3132	**130722**	0,40	9	**27**	5718	**130722**	0,27	19
3	6762	**130722**	0,33	5	**28**	5730	**130722**	0,30	10
4	5232	**130722**	0,22	7	**29**	3894	**130722**	0,17	22
5	3234	**130722**	0,22	22	**30**	7248	**130722**	0,29	8
6	6654	**130722**	0,51	11	**31**	15582	**130722**	0,63	12
7	5196	**130722**	0,35	7	**32**	3852	**130722**	0,16	18
8	3234	**130722**	0,20	13	**33**	4830	**130722**	0,21	4
9	9150	**130722**	0,39	9	**34**	17496	**130722**	0,78	5
10	1782	**130722**	0,08	16	**35**	4530	**130722**	0,18	26
11	4980	**130722**	0,21	21	**36**	3210	**130722**	0,14	24
12	3210	**130722**	0,14	6	**37**	3234	**130722**	0,14	7
13	3696	**130722**	0,15	17	**38**	10044	**130722**	0,43	7
14	2634	**130722**	0,09	11	**39**	8256	**130722**	0,32	16
15	11736	**130722**	0,72	17	**40**	10392	**130722**	0,41	18
16	8466	**130722**	0,41	9	**41**	4182	**130722**	0,19	10
17	22374	**130722**	1,14	17	**42**	1506	**130722**	0,09	33
18	11148	**130722**	0,46	34	**43**	24042	**130722**	1,01	14
19	8856	**130722**	0,46	12	**44**	15000	**130722**	0,57	8
20	9180	**130722**	0,42	9	**45**	3354	**130722**	0,13	18
21	17970	**130722**	0,87	13	**46**	5694	**130722**	0,28	10
22	9948	**130722**	0,42	20	**47**	3072	**130722**	0,13	8
23	2118	**130722**	0,08	13	**48**	12870	**130722**	0,51	21
24	3648	**130722**	0,18	29	**49**	4548	**130722**	0,19	17
25	4950	**130722**	0,20	6	**50**	8010	**130722**	0,35	8

	$NFOB_{médio}$	**Desvio Padrão NFOB**	**$FOB_{médio}$**	**Desvio Padrão FOB**	**$Tempo_{Médio}$**
MÉDIA	7369	5176	130722	0	0,35

O AGM conseguiu encontra a melhor solução para o problema em todas as execuções da Tabela 3-4, apresentando um tempo de execução médio de **Tempo$_{médio}$**=0,35 segundos e desvio padrão de **Tempo$_{desvPad}$**=0,24 segundos, as execuções também apresentaram um valor médio de chamadas da função objetivo **NFob$_{médio}$**=7369 com desvio padrão de **NFob$_{desvPad}$**=5176. Com essas execuções é observado que o AGM apresentou menor tempo médio de execução do que o utilizado pelo CPLEX para encontrar a solução ótima para o problema. Nessas execuções foram utilizados os operadores genéticos: seleção pelo método da roleta, o operador de cruzamento de dois pontos com taxa de cruzamento ***Tc***=0,7; as abordagens de mutação *M1, M2, M3, M4* com suas respectivas taxas de mutação ***Tm1***=0,1, ***Tm2***=0,01, ***Tm3***=0,005, ***Tm4***=0,005, ***Tpenal***=10000; também utilizou um intervalo de geração geracional *λ (=μ)* (*μ*=n° de pais, *λ*=n° de filhos) acrescido de Elitismo, onde o melhor indivíduo da população corrente é mantido na próxima geração, substituindo os ***k***=*2* piores indivíduos dessa nova população. Também foram padronizados o tamanho da população ***Np***=6, o número de concentradores ***p***=8 e o critério de parada utilizado é o critério que encerra o AGM quando o algoritmo encontrar a solução ótima para o problema (solução encontrada pelo CPLEX).

Na Tabela 3-5 são apresentados os resultados da otimização para a instância com a matriz de capacidades restringida (D=74 – E=66 – F=75). Nesse teste computacional comparou-se o resultado conseguido pelo CPLEX com a otimização realizada pelo AGM.

Tabela 3-5 – Resultados das 10 execuções do AGM para otimização do modelo MCpCLP –S com 37 localizadas e capacidades D=74 – E=66 – F=75.

TEMPO (seg)	NFOB	FOB
21112,00	5,13E+08	361033,00
6725,00	1,98E+08	361033,00
827,00	3,89E+07	361033,00
4398,00	2,19E+08	361033,00
3375,00	1,72E+08	361033,00
2835,00	1,43E+08	361033,00
1031,00	5,23E+07	361033,00
21490,00	1,09E+09	361033,00

$TEMPO_{Médio}$	$NFOB_{Médio}$	$FOB_{Médio}$
7724,13	3,03E+08	361033,00

$TEMPO_{Desvpad}$	$NFOB_{Desvpad}$	$FOB_{Desvpad}$
8585,56	3,43E+08	0,00

O AGM em todas as 10 execuções da Tabela 3-5 para o problema com capacidades (D=74, E=66, F=75) obteve os seguintes resultados, $\mathbf{Tempo_{médio}}$=7724,13 segundos (2,14 horas) e desvio padrão de $\mathbf{Tempo_{desvPad}}$=8585,56 segundos, as execuções também apresentaram um valor médio de chamadas da função objetivo $\mathbf{NFob_{médio}}$=3,03E+08 com desvio padrão de $\mathbf{NFob_{desvPad}}$=3,43E+08. Com essas execuções é observado que o AGM apresentou menor tempo médio de execução do que o utilizado pelo CPLEX, e o AGM também encontrou soluções melhores do que a do CPLEX. Nessas execuções foram utilizados os operadores genéticos: seleção pelo método da roleta, o operador de cruzamento de dois pontos com taxa de cruzamento ***Tc***=0,3; as abordagens de mutação *M1, M2, M3, M4, M5, M6* com suas respectivas taxas de mutação ***Tm1***=0,2, ***Tm2***=0,002, ***Tm3***=0,1, ***Tm4***=0,002, ***Tm5***=0,03, ***Tm6***=0,03, ***Tpenal***=10000; também utilizou um intervalo de geração geracional $\lambda\ (=\mu)$ (μ=n° de pais, λ=n° de filhos) acrescido

de Elitismo, onde o melhor indivíduo da população corrente é mantido na próxima geração, substituindo os ***k***=*1* piores indivíduos dessa nova população. Também foram padronizados o tamanho da população ***Np***=100, o número de concentradores ***p***=8 e o critério de parada utilizado é o critério que encerra o algoritmo após 20000 iterações sem melhora da melhor solução corrente.

3.7.1. Análise da Convergência

Para avaliar a capacidade de convergência dos algoritmos evolutivos propostos, foi criada uma ilustração gráfica, Figura 3-15, contendo o tempo de execução utilizado para cada uma das 50 execuções das Tabelas Tabela 3-3 e Tabela 3-4. A convergência média do AGS acontece por volta de **Tempo$_{médio}$**=43 segundos com **NFob$_{médio}$**=782586 chamadas da função objetivo. Já no AGM a convergência acontece por volta de **Tempo$_{médio}$**=0,35 segundos com **NFob$_{médio}$**=7370 chamadas da função objetivo. Essa diferença de desempenho acontece por que, o AGS não possui um mecanismo robusto para refinar regiões promissoras do espaço de soluções, ficando estagnado em regiões de mínimo local. Isso demostra que o mecanismo de busca local acrescentado no AGM potencializou o mecanismo de refinamento do AE, tornando-o mais eficiente.

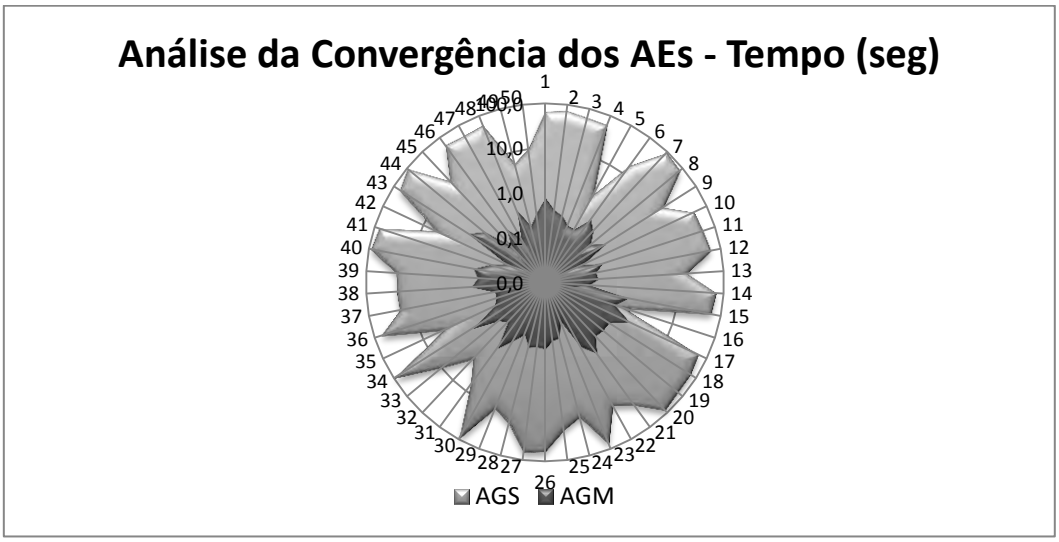


Figura 3-15 – Análise da convergência das 50 execuções dos AEs - Tempo

Na apresentação da convergência realizada pela Figura 3-15 pode-se observar que o AGM possui um tempo de convergência médio, $melhoraTempo = \frac{TempoAGM_{Médio} - TempoAGS_{Médio}}{TempoAGS_{Médio}} x100 = \frac{0,35-43}{43} = -99\%$, menor que a do AGS.

Na ilustração, Figura 3-16, é apresentada a análise de convergência utilizando o parâmetro *NFob*, número de chamadas da função objetivo, e a melhor solução corrente, Fob_{melhor}. Pode-se observar que a convergência do AGM é bem mais acentuada do que a do AGS. A analise foi realizada utilizando uma das 50 execuções realizadas para cada AE.

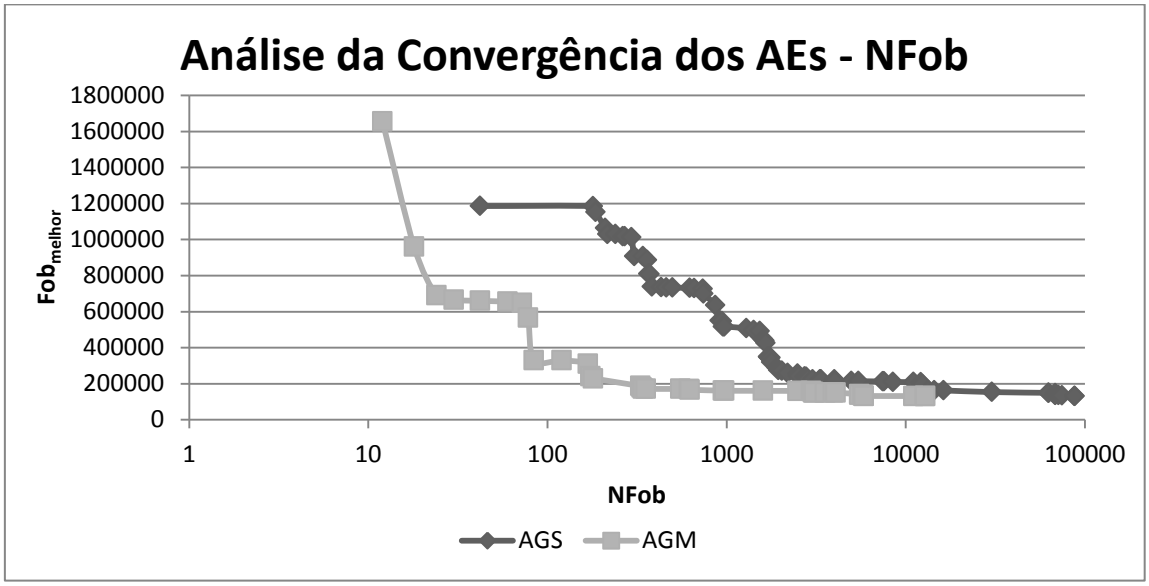


Figura 3-16 – Análise da convergência dos AEs – NFob

Nessa análise a convergência do AGS acontece com **NFob**=88068 e a convergência do AGM acontece com **NFob**=12864. A diferença entre os pontos de convergência é de **NFob**=75204 chamadas da função objetivo, $MelhoraNFOB = \frac{12864-88068}{88068} x100 = -85\%$, essa diferença reforça a eficiência do AGM em refinar pontos de alta aptidão, convergindo mais rapidamente para o ótimo global, pois mesmo com uma execução do AGM que necessitou de mais chamadas da função objetivo para convergir do que o valor médio apresentado na Tabela 3-4, foi conseguido um melhor desempenho que a

execução do AGS que possui um melhor desempenho que o resultado médio da Tabela 3-3.

3.7.2. Análise da Diversidade da População

Para um bom desempenho de um AE para problemas de otimização é imprescindível que o mesmo realize uma boa exploração do espaço de busca, procurando pontos de alta aptidão que representem regiões promissoras. Por que essas regiões promissoras sofreram um processo de refinamento até encontrar a solução ótima para o problema. Segundo Topcuoglu, et al. (2005), um AG com uma população com baixa diversidade pode proporcionar uma convergência prematura do algoritmo, dificultando que a abordagem evolutiva consiga escapar dos mínimos ou ótimos locais.

Nesse trabalho utiliza-se para medir a diversidade da população o conceito apresentado no trabalho de Topcuoglu, et al. (2005), onde é utilizada a distância de *hamming* (número de posições nas quais, duas *strings* de mesmo tamanho, se diferem) como métrica para avaliar a diversidade da população. Essa métrica é aplicada ao vetor de alocações, *assignArray*, onde é avalida a quantidade média de alterações necessária para transformar o vetor de alocações de cada um dos S_n indivíduos da população corrente *S,* no vetor de alocações da melhor solução S_{melhor}.

As diversidades das populações dos AEs geradas durante as otimizações: Tabela 3-3 e Tabela 3-4, são apresentadas pelas ilustrações gráficas: Figura 3-17 e Figura 3-18, onde é possível visualizar a diversidade populacional de duas execuções ao longo das chamadas da função objetivo.

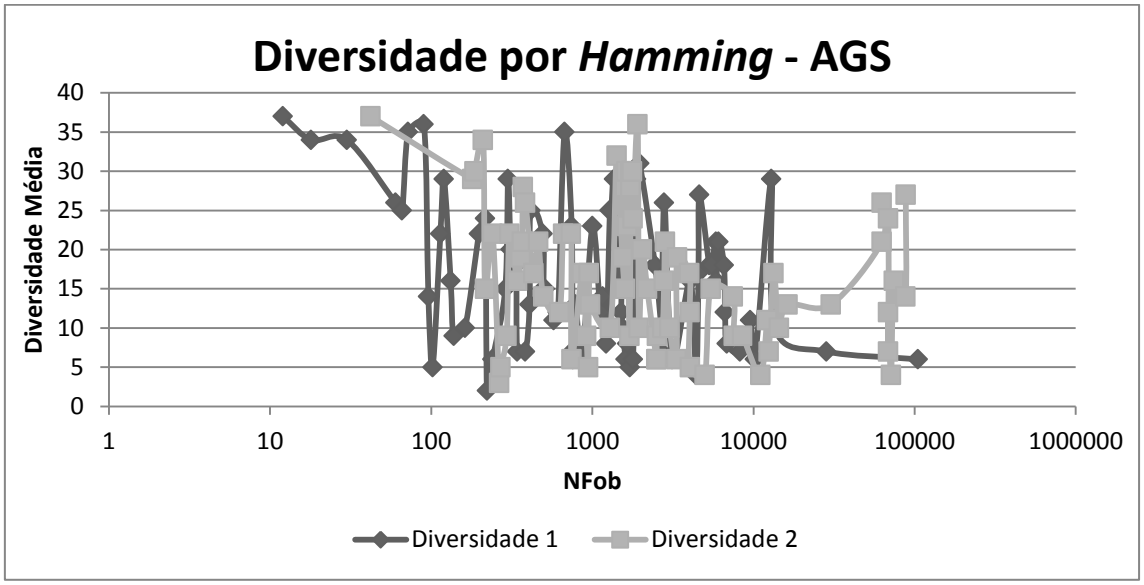


Figura 3-17 – Análise da diversidade por *Hamming* da população do AGS para 37 localidades

A analise de diversidade do AGS foi realizada com duas execuções, Figura 3-17, onde a primeira execução (Diversidade 1) não encontrou a solução ótima, **Fob**=136003 e **NFob**=105114, tendo um índice de diversidade inicial para a população de 37 alterações. Já a segunda execução (Diversidade 2) analisada para o AGS encontrou a melhor solução com **Fob**=130722 e com **NFob**=88068 e a população inicial também possui uma diversidade de 37 alterações. Analisando essas duas execuções visualiza-se que as duas possuem constantes oscilações de diversidade caracterizando que o AGS possui mecanismos que convergem à população e quando a população converge para um mínimo local o AGS também possui recursos para, novamente, diversificar a população, fugindo dos mínimos locais.

Essa capacidade de convergência pode ser medida pela Amplitude Total (R=Máximo Valor - Mínimo Valor) das entradas de diversidade, onde a primeira execução possui um valor de amplitude igual a **R**= 37-2 = 35 e para a segunda execução tem-se **R**= 37-3 = 34.

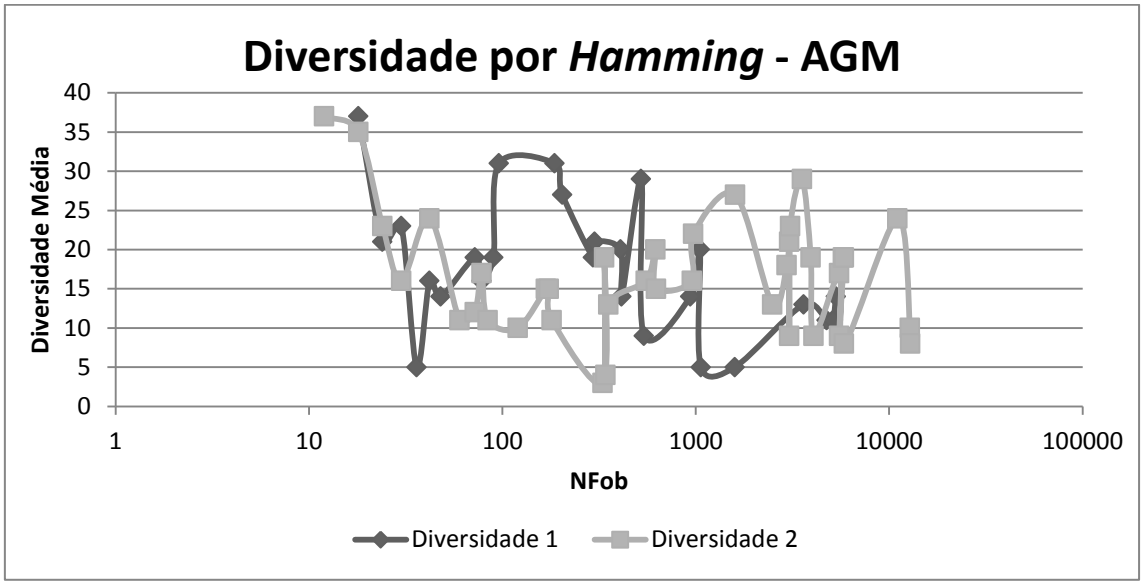


Figura 3-18 – Análise da diversidade por *Hamming* da população do AGM para 37 localidades

Já na análise do AGM pode-se perceber que a população inicial, das duas execuções analisadas, também começam com uma diversidade de 37 alterações e ao longo das execuções vão oscilando, fugindo dos mínimos locais até encontrar a solução ótima. O poder de convergência do AGM é confirmado por **R**=32 e **R**=34 respectivamente para cada execução ilustrada na Figura 3-18.

Com essa análise de diversidade observa-se que o mecanismo de exploração utilizado pelos AEs propostos, funciona bem durante toda a busca, e que o diferencial para garantir o melhor desempenho do AGM foi à inclusão de uma busca local (busca de subida da encosta), o que forneceu ao AGM um mecanismo de refinamento, de regiões promissoras, mais robusto do que a do AGS.

3.7.3. Nova Localização-Alocação dos Laboratórios do SISMETRA

A otimização das localizações-alocações dos laboratórios de metrologia do SISMETRA foi realizada utilizando o conjunto atual de localidades, apresentada na Figura 3-3, sendo composta de 38 localidades militares, distribuídas em 1

(um) LCC, 8 (oito) LRC e 29 (vinte e nove) LSC. Essa otimização gerou uma nova proposta de distribuição dos laboratórios, apresentada na Figura 3-19. Essa nova disposição dos laboratórios caso seja aceita para implementação no sistema pode sofrer algumas alterações para se enquadrar nas determinações hierárquicas das organizações militares do COMAER.

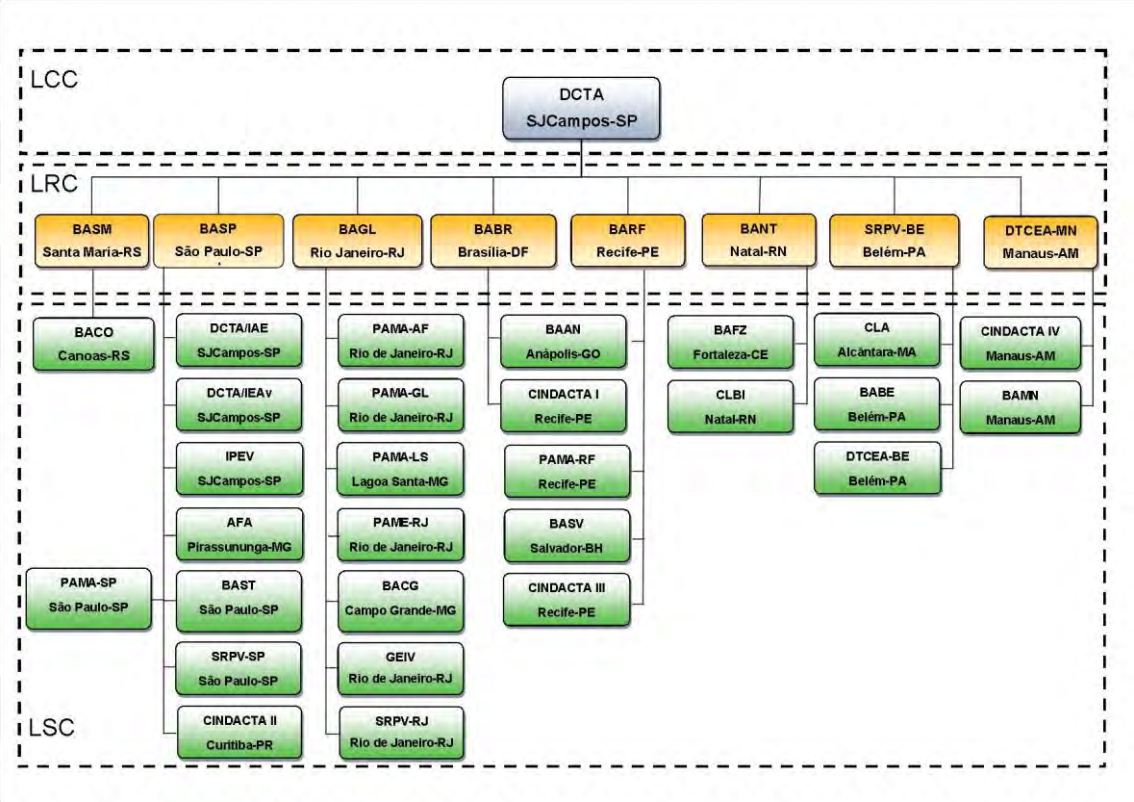


Figura 3-19 – Organograma otimizado do SISMETRA

No próximo capítulo serão apresentadas as abordagens utilizadas para a otimização do problema de agrupamento de aprendizes.

4. PROBLEMA DE AGRUPAMENTO DE APRENDIZES

A análise de agrupamentos, também conhecida como análise de conglomerados, classificação ou agrupamento, tem como objetivo dividir os elementos da amostra, ou população, em grupos de forma que os elementos pertencentes a um mesmo grupo sejam similares entre si, com respeito às variáveis características que neles foram medidas, e que sejam diferentes dos elementos em grupos diferentes em relação a estas mesmas características. (MINGOTI, 2005).

4.1. Problema de Agrupamento

Com um conjunto de *n* elementos amostrais (aprendizes) com p-variáveis (variáveis de classificação de aprendizagem) em cada um deles $A_{ij} = \{a_{ij},..., a_{np}\}$, o agrupamento consiste em identificar grupos com características similares divididos em *k* grupos, $C = \{C_1, C_2, ..., C_k\}$, proporcionando que os elementos de cada grupo C_i tenham a maior similaridade entre os elementos do mesmo grupo do que com os elementos de qualquer um dos outros grupos do conjunto *C*.

Segundo Dias (2004), o conjunto *C* é considerado um agrupamento com *k* grupos, caso as seguintes condições sejam satisfeitas: $\sum_{i=1}^{k} C_i = A, C_i \neq \emptyset, para\ 1 \leq i \leq k, C_i \cap C_j = \emptyset, para\ i \leq i, j \leq k\ e\ i \neq j$.

Segundo Mingoti (2005) existem algumas formas para se mensurar a similaridade e dissimilaridade dos elementos de uma amostra, as mais comuns para variáveis quantitativas são: *Distância Euclidiana, Distância Generalizada ou Ponderada e Distância de Minkowsky*.

4.2. Aprendizagem

Segundo Forgus (1971), a relação entre aprendizagem e pensamento no processo de percepção consiste em que os estímulos possuem a informação que é extraída pelo organismo sob a forma de aprendizagem, modificando o organismo de modo que a próxima percepção do mesmo estímulo seja

diferente. O processo de pensamento (resultante de aprendizagem prévia) também modifica o organismo porque ocorre uma nova aprendizagem. (FORGUS, 1971).

Apresentação da relação entre aprendizagem, pensamento e o processo perceptivo será apresentada na Figura 4-1.

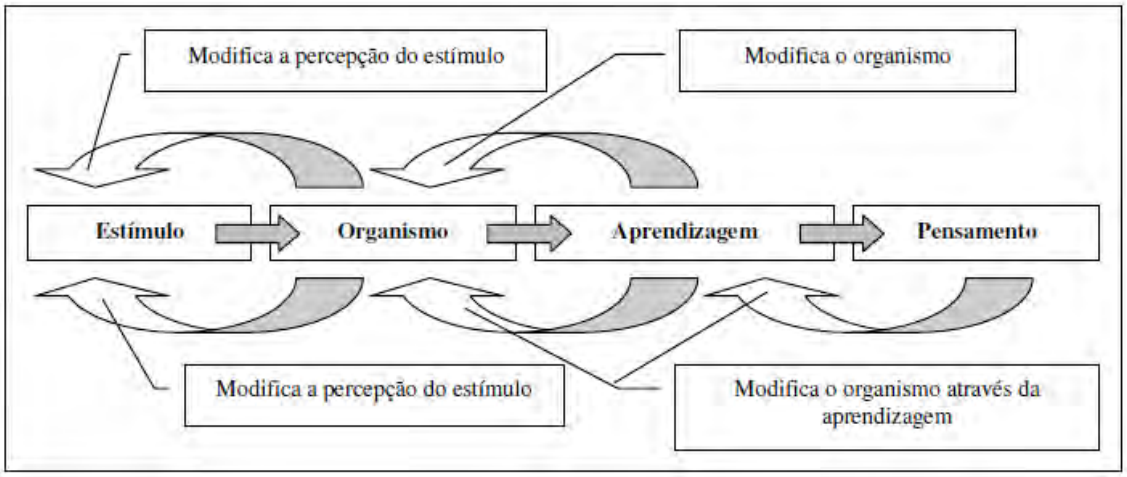


Figura 4-1 - Relação entre aprendizagem, pensamento e o processo perceptivo. Fonte: (FORGUS, 1971).

4.2.1. Estilos de Aprendizagem

Os estilos de aprendizagem podem ser entendidos como o comportamento do aprendiz durante o processo de aprendizado. O conhecimento prévio do estilo de aprendizagem pode, então, maximizar a assimilação do conhecimento, encurtando os laços do material didático apresentado com o aprendiz.

Existem instrumentos que permitem o levantamento dos estilos de aprendizagem dos alunos, indicando os prováveis pontos fortes e fracos de sua vida acadêmica. Esse levantamento não reflete a sua adequação ou inadequação para uma determinada disciplina, curso ou profissão, mas ajuda a melhorar a aprendizagem na medida em que o professor propõe não só atividades que vão ao encontro do estilo preferencial de suas turmas, como

também a escolha de métodos instrutivos que desafiem outros estilos, a fim de estimular e fortalecer os estilos menos desenvolvidos. (LOPES, 2002).

4.2.1.1. Modelo de Estilos de Aprendizagem de Felder e Silverman

Agora será apresentado o modelo desenvolvido por Richard M. Felder e Linda K. Silverman que é o modelo de classificação dos estilos de aprendizagem utilizado na implementação do trabalho proposto.

Felder e Silverman (1988) afirmam que os estudantes têm diferentes "estilos de aprendizagem". Eles definem os "estilos de aprendizagem" como sendo características fortes e preferenciais na maneira como o estudante processa a informação.

Características dos aprendizes de acordo com os estilos de aprendizagem, de Felder-Silverman:

Aprendizes Ativos:

- Frase característica: "Vamos experimentar e ver como funciona".
- Tendem a reter e compreender melhor a informação fazendo algo ativo, discutindo, aplicando ou explicando para os outros.
- Trabalham melhor em grupos, são hábeis administradores e coordenadores de projetos de trabalho.

Aprendizes Reflexivos:

- Frase característica: "Vamos primeiro meditar sobre o assunto".
- Processam melhor a informação de forma reflexiva, ou seja, pensar sobre algo primeiro.
- Tem mais eficiência trabalhando sozinhos em projetos e pesquisas.

Aprendizes Sensoriais:

- São práticos e cuidadosos, gostam de aprender através de fatos e observações e a resolver problemas usando procedimentos bem estabelecidos, sem complicações e surpresas.
- Procuram cursos que tenham conexão com o mundo real e tendem a ser pacientes com detalhes.
- Sentem-se menos confortáveis com símbolos e estes devem ser traduzidos em imagens mentais concretas, a fim de compreendê-los.
- São atentos, detalhistas, pacientes e metódicos e podem ser bons experimentalistas.

Aprendizes Intuitivos:

- Tendem a ser mais imaginativos e inovadores.
- Preferem conceitos e interpretações e se sentem mais confortáveis em lidar com símbolos e abstrações.
- Adoram novidades e se aborrecem com rotina.
- Realizam trabalhos mais rapidamente.

Aprendizes Visuais:

- Aprendem melhor quando a informação é apresentada visualmente através de fotos, diagramas, fluxogramas, filmes, gráficos e demonstrações.
- Relembram melhor o que viram.

Aprendizes Verbais:

- Aprendem melhor quando a informação é apresentada verbalmente.
- Recordam com mais facilidade aquilo que escutam ou leem.

Aprendizes Sequenciais:

- Tendem a aprender de forma linear, seguindo uma progressão lógica de pequenas etapas, para encontrar soluções.
- São muitas vezes bons analistas e hábeis para resolver problemas.

Aprendizes Globais:

- Aprendem quase que aleatoriamente em saltos grandes, holisticamente, sem ver as conexões para, então, repentinamente, “compreender” tudo.
- Podem ser hábeis em resolver rapidamente problemas complexos ou unir coisas, mas têm dificuldade de explicar como fizeram.
- São muitas vezes bons sintetizadores.

Aprendizes Indutivos:

- Organizam a informação partindo do particular para o geral, com teorias ou ideias unificadoras, resultantes do estudo de situações individuais.

Aprendizes Dedutivos:

- Organizam a informação onde as soluções e aplicações particulares são consequências de uma ideia geral.
- Estilo utilizado principalmente em assuntos técnicos e de organização do conhecimento adquirido.

4.2.1.2. Instrumento para a Medição de Estilos de Aprendizagem (ILS)

O instrumento para a identificação dos estilos de aprendizagem utilizado neste trabalho é a versão reduzida do *Index of Learning Styles (ILS)*, desenvolvido por Richard M. Felder e Bárbara A. Soloman (FELDER; SOLOMAN, 2010), com o objetivo de investigar e identificar as preferências de aprendizagem em quatro dimensões de estilos de aprendizagem: *Ativo/Reflexivo, Sensorial/Intuitivo, Visual/Verbal e Sequencial/Global*. (FELDER; SILVERMAN, 1988).

O instrumento é composto por quarenta e quatro questões e duas alternativas de respostas para cada questão, sendo onze questões para cada uma das quatro dimensões dos estilos de aprendizagem de Felder e Silverman.

Das duas alternativas de resposta (“a” ou “b”) será forçado que para cada questão a pessoa que estiver respondendo o questionário tenha apenas a opção de escolher uma das duas alternativas, mesmo se as duas se aplicassem igualmente, a escolha deverá ser feita àquela que acontece mais frequente.

Realizado o preenchimento de todo o questionário, Tabela 4-1, começará a contagem das respostas, somando quantas vezes a alternativa “a” ou “b” foi escolhida para cada uma das dimensões correspondentes. Após a contagem das alternativas para cada dimensão realizamos a subtração, em cada dimensão, entre o acumulado das alternativas “a” e “b”, sempre subtraindo o menor do maior, conservando a letra do maior valor, para a posterior comparação com a Escala de Estilos de Aprendizagem, Tabela 4-2.

Tabela 4-1: Exemplo de respostas do questionário (ILS)

ATI/REF			SEN/INT			VIS/VER			SEG/GLO		
Questão	A	B	Questão	A	B	Questão	A	B	Questão	A	B
1	X		2	X		3		X	4	X	
5	X		6		X	7		X	8	X	
9	X		10		X	11		X	12		X
13	X		14	X		15		X	16		X
17		X	18		X	19		X	20	X	
21		X	22	X		23		X	24	X	
25		X	26	X		27		X	28	X	
29		X	30	X		31		X	32		X
33	X		34	X		35		x	36		X
37	X		38	X		39	X		40		X
41	X		42	X		43	X		44	X	
TOTAL (SOMA DE CADA COLUNA)											
ATI/REF			SEM/INT			VIS/VER			SEG/GLO		
A	B		A	B		A	B		A	B	
7	4		8	3		2	9		6	5	
(Maior – Menor) + Letra do Maior											
3A			5B			7B			1A		

Tabela 4-2: Escalas de estilos de aprendizagem

ATIVO						REFLEXIVO					
11A	9A	7A	5A	3A	1A	1B	3B	5B	7B	9B	11B
SENSORIAL						INTUITIVO					
11A	9A	7A	5A	3A	1A	1B	3B	5B	7B	9B	11B
VISUAL						VERBAL					
11A	9A	7A	5A	3A	1A	1B	3B	5B	7B	9B	11B
SEGUÊNCIAL						GLOBAL					
11A	9A	7A	5A	3A	1A	1B	3B	5B	7B	9B	11B

Observação:

- Escore com valor de 1 ou 3: você tem leve preferência entre ambas dimensões da escala.

- Escore com valor de 5 ou 7: você tem uma preferência moderada por uma das dimensões da escala e aprenderá mais facilmente se o ambiente de ensino favorecer esta dimensão.

- Escore com valor de 9 ou 11: você tem uma forte preferência por uma das dimensões da escala. Você pode ter dificuldades de aprendizagem em um ambiente que não favoreça essa preferência.

O instrumento da pesquisa é a primeira versão da tradução para o português, realizada por Marcius F. Giorgetti e Nídia Pavan Kuri, (s/d), da escola de Engenharia de São Carlos – USP.

O grande desafio para a otimização de rotinas de agrupamento está diretamente ligada ao tamanho do espaço de soluções, decorrentes das diferentes formas de agrupamento dos *n* elementos da amostra em *k* grupos. Existem métodos de otimização que encontram boas soluções para esses problemas, os mais utilizados são os métodos baseados em mecanismos de evolução de populações, sendo os mais utilizados deles os Algoritmos Genéticos proposto por Holland (1975).

4.2.2. Codificação

Cada cromossomo da população deve representar uma solução para o agrupamento dos aprendizes, sendo a representação escolhida para esse trabalho a codificação *group-number* (KIM, et al., 2011).

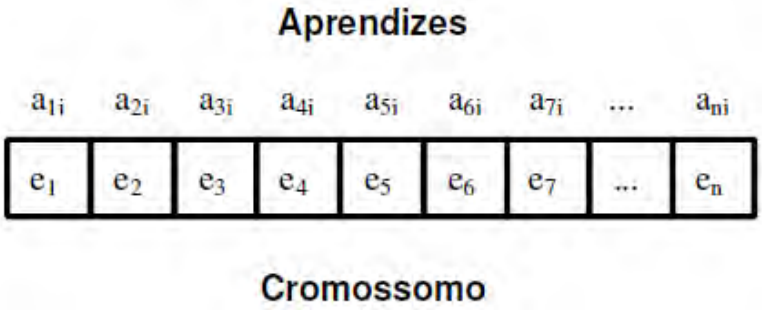


Figura 4-2: Representação do cromossomo para codificação *Group-Numbers*.

onde:

- *N={1, 2,...,n}*, e n representa o número de aprendizes.
- *P={1, 2,...,p},* e *p=8* representa os 8 (oito) Estilos de Aprendizagem.
- $i \in N$ e representa os índices dos aprendizes.
- $j \in P$ e representa os índices das variáveis de classificação de aprendizagem.
- $e_i \in \{1, 2,...,8\}$ e representa o grupo do aprendiz da *i-ésima* posição de acordo com os 8 (oito) Estilos de Aprendizagem de Felder e Silverman (1988).
- $A=\{a_{ij},...,a_{np}\}$ e representa o conjunto de aprendizes e suas variáveis de classificação de aprendizagem (KIM, et al., 2011).

4.2.3. População Inicial

O AG desenvolvido cria aleatoriamente a *PI* com um número pré-definido de soluções. Após testes computacionais definiu-se o melhor tamanho para a *PI: Np = 20* cromossomos, onde os genes $e_i \in \{1,2,...,8\}$ são escolhidos aletoriamente criando a *PI.* A criação dos cromossomos iniciais torna-se simples de implementação através da linguagem de programação JAVA na versão 5.0, pela eficiência da classe *Random* para criação de números pseudo-aleatórios, através de um algoritmo gerador de congruência linear. (ORACLE, 2011).

4.2.4. Função de Aptidão

A função de aptidão desenvolvida é representada pelo somatório das distâncias euclidianas das variáveis de classificação de estilos de aprendizagem dos aprendizes pertencentes ao mesmo grupo. A função proposta procura minimizar as distâncias entre os aprendizes do mesmo grupo, procurando grupos que tenham um conjunto de aprendizes similares, tendo como referência a distância euclidiana de suas variáveis de aprendizagem.

$$a_{iv} = \{ATI_{i1}, REF_{i2}, INT_{i3}, SEN_{i4}, VIS_{i5}, VER_{i6}, SEG_{i7}, GLO_{i8}\} \quad (4\text{-}1)$$

$$d_{lk}(a_l, a_k) = \left[\sum_{v=1}^{8} (a_{lv} - a_{kv})^2\right]^{1/2} \quad (4\text{-}2)$$

$$\text{Aptidão} = \sum_{i=1}^{n-1} \sum_{s=i+1}^{n} d(a_i, a_s)\,;\ \forall\, e_i = e_s \quad (4\text{-}3)$$

Utilizamos o índice $s \in N$ para possibilitar a comparação entre os aprendizes agrupados no mesmo grupo.

4.2.5. Seleção

O método de seleção escolhido para a implementação foi o Método da Roleta. A execução da seleção pelo método da roleta é realizada quantas vezes forem necessárias para formar a população intermediária com os indivíduos escolhidos.

4.2.6. Cruzamento

Consiste na combinação de segmentos de 2 (dois) cromossomos selecionados (Pais) para gerar dois novos cromossomos (Filhos). O método de cruzamento escolhido para a implementação foi o método de cruzamento de um ponto, o ponto de cruzamento é obtido de forma aleatória, utilizando a classe *Random*

da linguagem JAVA (ORACLE, 2011), e os genes após o ponto de cruzamento são trocados entre os indivíduos, Figura 4-3.

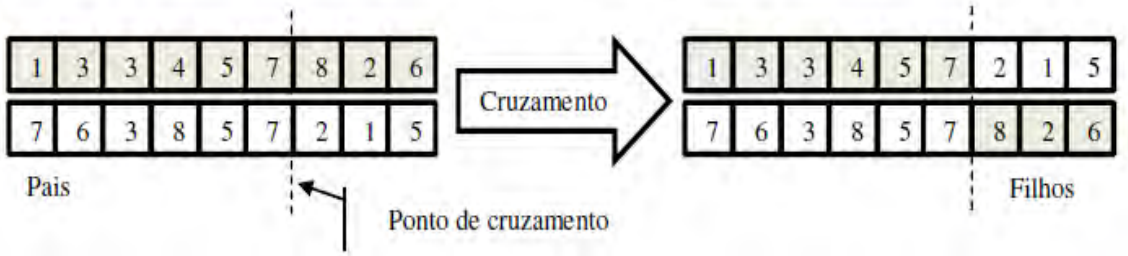


Figura 4-3: Exemplo de aplicação do operador de cruzamento de um ponto.

A taxa de cruzamento utilizada no AG desenvolvido é de 0,8 da população.

4.2.7. Mutação

Na mutação alteram-se os valores dos genes (alelos) de um cromossomo, após o operador cruzamento, inserindo diversidade genética em uma população, e com isso, realizando buscas em diferentes áreas do espaço de solução, não permitindo que a população fique estagnada em uma região de ótimo local,

Figura 4-4

.

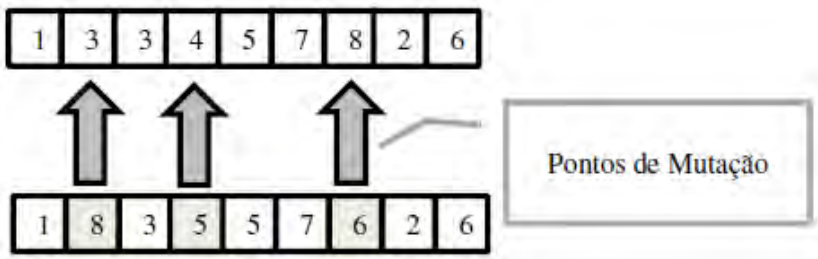


Figura 4-4: Exemplo de aplicação do operador de mutação

A taxa de mutação utilizada na implementação é **Tm**=0,005, sendo aplicada sobre todos os genes e_i do cromossomo. Os alelos dos genes e_i selecionados são alterados aleatoriamente com $e_i \in \{1, 2,...,8\}$.

4.2.8. Substituição da População

O intervalo de geração utilizado nesse artigo é o Intervalo Geracional λ (=μ) acrescido de Elitismo, onde o melhor indivíduo de todas as anteriores gerações substitui o pior indivíduo dessa nova população.

4.2.9. Critério de Parada

Existem alguns critérios muito utilizados para terminar o critério de parada de um algoritmo genético, tais como: após um número de iterações, com base na convergência, isto é, quando após um dado numero de gerações não for encontrado um indivíduo com melhor aptidão que a do melhor indivíduo das gerações anteriores, quando as aptidões se tornarem parecidas ou quando for encontrada a solução ótima para o problema.

O critério de parada utilizado nesse trabalho é o critério que leva em consideração a convergência da população. Nesse critério a execução do AG é interrompida quando ocorrer 100000 gerações sequenciais sem melhora da melhor solução encontrada nas gerações passadas.

4.2.10. Preparação dos Dados

Utilizou-se uma amostra de 54 alunos universitários regularmente matriculados no curso de Ciência da Computação no Centro Universitário Salesiano de São Paulo – Unidade de Lorena. A pesquisa foi realizada com todas as quatro séries do curso no ano letivo de 2008.

Tendo como população para o experimento os alunos do 1º, 2º, 3º, 4º séries do curso de Ciência da Computação e considerando o número médio de 30 alunos por série, teremos uma população de 120 alunos, e o tamanho da amostra para

estimar a proporção de uma população finita para esse experimento é definida pela fórmula. (MARTINS, 2008):

Dados: N=120; Z=1,96 para 95%; p'=0,50; q'=(p'-1)=0,50; d=10%= 0,10.

$$\text{Amostra} = \frac{Z^2 . p' . q' . N}{d^2(N-1) + Z^2 . p' . q'} \quad (4\text{-}4)$$

onde:

- N = tamanho da população (número de aprendizes)
- Z = abscissa da distribuição normal padrão, fixado um nível de (1-α)% de confiança do intervalo de confiança para a média. (Z=1,96 para 95%).
- p' = estimativa da verdadeira proporção de um dos níveis da variável escolhida. (Admite-se p=0,50, caso não haja estimativa prévia).
- q' = 1-p'.
- d = erro amostral expresso em decimais. Máxima diferença que o investigador admite suportar entre a verdadeira proporção *(p)* e a proporção do evento *(p')* a ser calculada com base na amostra.

Aplicando a fórmula acima obtemos o número mínimo de 54 alunos para representar a população para um erro amostral de 10%. Dessa forma, como o experimento está sendo realizado com uma amostra de 54 alunos, validamos a representatividade da amostra.

4.2.11. Resultados Computacionais

O AG foi implementado em linguagem de programação JAVA na versão 5.0, foi criado um ambiente web para que os aprendizes pudessem responder o questionário *ILS – Index of Learning Styles* (FELDER; SOLOMAN, 2010) on-line, a máquina responsável pela execução da aplicação possui a seguinte

configuração: processador Intel Atom, 1,66 GHz; 2 Gb de memória *ram* e sistema operacional Windows 7 Starter. Os testes realizados são apresentados na Tabela 4-3, de acordo com as siglas **nFOB** – número de execuções da função objetivo, **FOB** – valor da função objetivo para a solução encontrada.

Tabela 4-3: Execuções do AG para agrupamento de aprendizes

Execuções	nFOB	FOB	Execuções	nFOB	FOB
1	142836	755,7087	**26**	104304	759,6803
2	137070	781,3475	**27**	241974	753,1103
3	118236	776,4259	**28**	166164	763,5906
4	104910	778,5111	**29**	112548	757,4209
5	102564	761,6021	**30**	101640	753,5603
6	181404	747,2395	**31**	197310	769,5435
7	155112	771,4897	**32**	137406	767,4463
8	179424	791,6596	**33**	228606	754,1723
9	139116	807,4047	**34**	90342	768,9969
10	92916	766,5646	**35**	116580	769,8406
11	173646	783,1562	**36**	87210	795,308
12	119010	747,2395	**37**	199644	762,3616
13	261708	754,0238	**38**	191034	771,2457
14	148092	758,932	**39**	92184	801,7921
15	222546	753,1368	**40**	132054	761,6831
16	189642	741,5339	**41**	119310	751,9338
17	98820	774,6985	**42**	109488	761,1869
18	191304	765,437	**43**	129360	760,0371
19	151362	775,3405	**44**	109614	775,8424
20	249006	753,4571	**45**	144522	769,4648
21	155202	751,9619	**46**	145872	744,091
22	281760	794,5603	**47**	168918	754,738
23	139854	775,3062	**48**	92268	765,3691
24	147732	764,2382	**49**	133716	762,3476
25	111312	753,8145	**50**	159462	752,1037
	$nFOB_{médio}$	**Desvio Padrão nFOB**		**$FOB_{médio}$**	**Desvio Padrão FOB**
MÉDIA	150122,00	48024,9814		765,8331	14,6748

Todos os 50 testes foram realizados com os seguintes parâmetros genéticos Taxa de Cruzamento –**Tc**=0,8, Taxa de Mutação – **Tm**=0,005, Critério de

Parada = 100000 gerações sequenciais sem melhora da melhor solução e Intervalo de Geração Geracional λ(=μ) com de Elitismo do melhor indivíduo de todas as anteriores gerações. Esses parâmetros genéticos foram determinados empiricamente, onde combinações aleatórias eram utilizadas até que se encontrasse a melhor, após encontrar essa combinação foram executados os 50 testes e uma dessas execuções da Tabela 4-3 foi analisada para validar o agrupamento. A análise dos testes realizados encontrou um valor médio para o número de execuções da função objetivo – $\mathbf{nFOB_{médio}}$=150122 com um desvio padrão $\mathbf{S_{nFOB}}$=48025,9814 e um valor médio para as 50 melhores soluções encontradas – $\mathbf{FOB_{médio}}$=765,8331, com desvio padrão – $\mathbf{S_{FOB}}$=14,6748.

Verificando o Coeficiente de Variação da amostra FOB, $\mathbf{CV} = \frac{S}{\bar{x}}.100\% = 1{,}9\% < 15\%$, constatou-se que a amostra consiste de um conjunto de dados homogêneos (MARTINS, 2008).

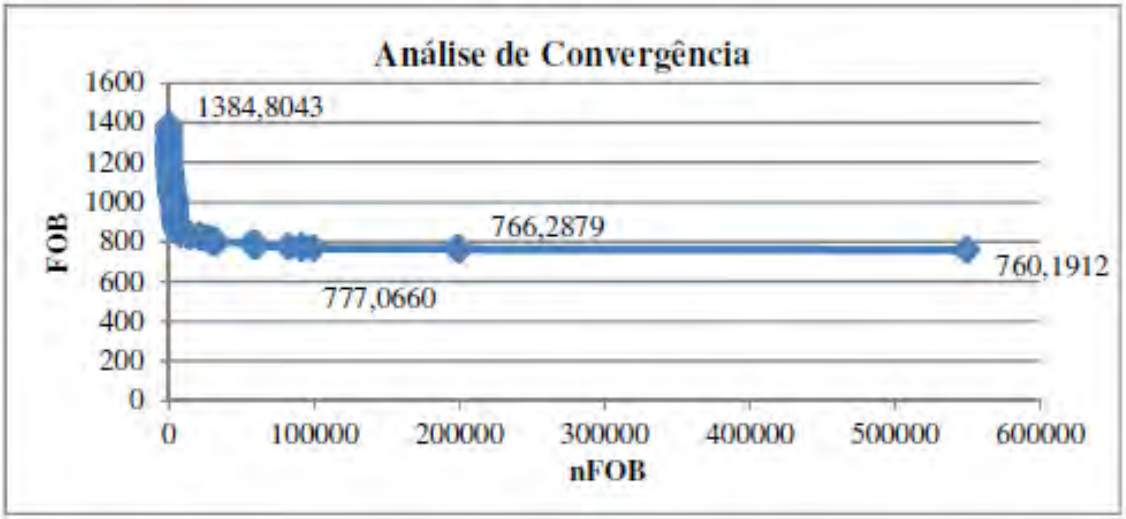


Figura 4-5: Gráfico da análise de convergência.

Pode-se verificar, para uma execução aleatória com melhor **FOB**=760,1912, na Figura 4-5, que a convergência do AG se estabiliza, aproximadamente, a partir de **nFOB**=200000 chamadas da função objetivo, tendo **FOB**=766,2879, pois a

próxima melhor solução encontrada pelo AG, nessa execução, precisou de mais **nFOB**=350901 chamadas da função objetivo. Comparando o melhor resultado da

Figura 4-5, com o resultado médio da Tabela 4-3, verificou-se que o valor **FOB**=760,1912 esta a -0,3844 desvios padrão do valor médio encontrado, sendo dessa forma, um valor pertencente à população.

A Tabela 4-4 mostra os grupos formados pelo AG durante a execução 42 (um) com **FOB**=761,1869 da Tabela 4-3, também são apresentadas as variáveis de classificação de estilos de aprendizagem de cada aprendiz: *ATI: Ativo, REF: Reflexivo, SEN: Sensitivo, INTU: Intuitivo, VIS: Visual, VER: Verbal, SEQ: Sequencial, GLO: Global.*

4.2.12. Método de Avaliação do Agrupamento

Para a validação do agrupamento foi utilizada uma técnica estatística conhecida como: contagem-z. Essa técnica tem a finalidade de avaliar o quanto um elemento está afastado da média dos elementos de seu grupo. Segundo Anderson et al. (2008), usando a média e o desvio padrão, é possível determinar a posição relativa de qualquer observação, pois cada observação além de estar relacionada à média e ao desvio padrão da amostra, também estará associada a um outro valor que se chama contagem-z.

$$Z_{e_i} = \text{MÀX}\left|\frac{a_{sj} - \bar{a}_{e_ij}}{S_{e_ij}}\right|;\ \forall\, e_i = e_s;\ i, s \in \{1, \dots, n\}\ e\ j \in p \qquad (4\text{-}5)$$

onde:

- S_{e_ij} - representa o vetor dos desvios padrão amostral das variáveis de classificação de aprendizagem *j* dos aprendizes do grupo representado pelo valor de e_i.

- $\bar{a}_{e_ij}$ - representa o vetor das médias das variáveis de classificação de aprendizagem j dos aprendizes do grupo representado pelo valor de e_i.

Na análise da contagem-z ou também conhecida como valor padronizado podemos interpretar Z_{e_i} como o valor de desvios padrão que o aprendiz, mais distante, do grupo e_i, está afastado da média $\bar{a}_{e_ij}$ desse mesmo grupo. (ANDERSON, et al., 2008).

Foi realizado o teste de *Shapiro-Wilk* para cada conjunto de variáveis de aprendizagem (ATI: Ativo, REF: Reflexivo, SEN: Sensitivo, INTU: Intuitivo, VIS: Visual, VER: Verbal, SEQ: Sequencial, GLO: Global) dos 54 alunos, onde se determinou a estatística de teste média $\mathbf{w_{média}}$=0,9622 e a probabilidade ***p-valor***=0,1016 com um nível de significância de **α**=0,05, validando assim, a normalidade do conjunto de variáveis de aprendizagem.

Anderson, et al. (2008), também afirmam que para a detecção de pontos fora da curva podemos utilizar valores padronizados (contagem-z), pois em relação a dados com uma distribuição em forma de sino, quase todos os valores estão contidos em três desvios padrão da média. Sendo consideradas observações fora da curva àquelas que apresentarem valores z menores que -3 e valores z maiores que +3.

Tabela 4-4: Agrupamento realizado pelo AG

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
1	Aprendiz 4	7	4	8	3	8	3	4	7
	Aprendiz 7	9	2	8	3	7	4	4	7
	Aprendiz 15	8	3	8	3	7	4	6	5
	Aprendiz 23	6	5	10	1	9	2	6	5
	Aprendiz 28	11	0	9	2	8	3	5	6
	Aprendiz 31	9	2	10	1	8	3	6	5
	Aprendiz 47	11	0	7	4	7	4	5	6
MÉDIA		8,7	2,3	8,6	2,4	7,7	3,3	5,1	5,9
DESVPAD		1,89	1,89	1,13	1,13	0,76	0,76	0,9	0,9
Z_1		1,70							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
2	Aprendiz 1	1	10	2	9	8	3	6	5
	Aprendiz 17	2	9	6	5	6	5	10	1

	Aprendiz 27	0	11	5	6	9	2	4	7
	Aprendiz 37	4	7	4	7	6	5	9	2
	Aprendiz 39	3	8	8	3	8	3	7	4
	Aprendiz 53	4	7	8	3	10	1	8	3
MÉDIA		2,3	8,7	5,5	5,5	7,8	3,2	7,3	3,7
DESVPAD		1,63	1,63	2,35	2,35	1,6	1,6	2,16	2,16
Z_2		1,54							

Tabela 4-4: Agrupamento realizado pelo AG (Continuação)

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
3	Aprendiz 10	7	4	9	2	6	5	6	5
	Aprendiz 20	8	3	8	3	5	6	9	2
	Aprendiz 24	6	5	9	2	4	7	5	6
	Aprendiz 32	6	5	8	3	4	7	3	8
	Aprendiz 41	7	4	8	3	5	6	4	7
	Aprendiz 45	7	4	8	3	6	5	7	4
MÉDIA		6,8	4,2	8,3	2,7	5	6	5,7	5,3
DESVPAD		0,75	0,75	0,52	0,52	0,89	0,89	2,2	2,2
Z_3		1,55							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
4	Aprendiz 2	2	7	4	10	1	9	2	8
	Aprendiz 8	8	7	4	9	2	11	0	8
	Aprendiz 18	18	9	2	8	3	9	2	8
	Aprendiz 26	26	9	2	9	2	9	2	8
	Aprendiz 34	34	8	3	11	0	10	1	8
	Aprendiz 38	38	6	5	8	3	10	1	9
	Aprendiz 46	46	10	1	10	1	11	0	11
MÉDIA		8	3	9,3	1,7	9,9	1,1	8,6	2,4
DESVPAD		1,41	1,41	1,11	1,11	0,9	0,9	1,1	1,1
Z_4		2,14							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
5	Aprendiz 11	6	5	6	5	9	2	6	5
	Aprendiz 16	7	4	6	5	6	5	5	6
	Aprendiz 22	6	5	6	5	8	3	7	4
	Aprendiz 33	6	5	4	7	6	5	5	6
	Aprendiz 40	7	4	5	6	8	3	4	7
	Aprendiz 49	6	5	4	7	6	5	4	7
	Aprendiz 52	6	5	6	5	6	5	6	5
MÉDIA		6,3	4,7	5,3	5,7	7	4	5,3	5,7
DESVPAD		0,49	0,49	0,95	0,95	1,29	1,29	1,1	1,1
Z_5		1,55							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
6	Aprendiz 5	9	2	7	4	11	0	7	4
	Aprendiz 14	10	1	6	5	10	1	8	3
	Aprendiz 19	6	5	5	6	8	3	9	2
	Aprendiz 30	9	2	7	4	8	3	8	3

	Aprendiz 44	8	3	5	6	9	2	6	5
	Aprendiz 51	8	3	5	6	7	4	9	2
	Aprendiz 54	8	3	5	6	10	1	8	3
MÉDIA		8,3	2,7	5,7	5,3	9	2	7,9	3,1
DESVPAD		1,25	1,25	0,95	0,95	1,41	1,41	1,1	1,1
Z_6		1,82							

Tabela 4-4: Agrupamento realizado pelo AG (Continuação)

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
7	Aprendiz 6	9	2	6	5	5	6	5	6
	Aprendiz 9	9	2	5	6	5	6	6	5
	Aprendiz 25	8	3	6	5	5	6	7	4
	Aprendiz 29	11	0	6	5	8	3	7	4
	Aprendiz 36	9	2	5	6	6	5	6	5
	Aprendiz 43	10	1	6	5	8	3	6	5
	Aprendiz 48	10	1	4	7	6	5	9	2
MÉDIA		9,4	1,6	5,4	5,6	6,1	4,9	6,6	4,4
DESVPAD		0,98	0,98	0,79	0,79	1,35	1,35	1,3	1,3
Z_7		1,91							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
8	Aprendiz 3	5	6	6	5	4	7	4	7
	Aprendiz 12	6	5	2	9	7	4	1	10
	Aprendiz 13	4	7	6	5	4	7	5	6
	Aprendiz 21	3	8	6	5	2	9	5	6
	Aprendiz 35	4	7	6	5	6	5	4	7
	Aprendiz 42	4	7	6	5	4	7	1	10
	Aprendiz 50	4	7	5	6	7	4	3	8
MÉDIA		4,3	6,7	5,3	5,7	4,9	6,1	3,3	7,7
DESVPAD		0,95	0,95	1,5	1,5	1,86	1,86	1,7	1,7
Z_8		2,20							

O agrupamento apresentado na Tabela 4-4 apresentou um agrupamento, no grupo 8, com o máximo valor de contagem-z $\mathbf{Z_8}$ = 2,20, representando assim, um agrupamento sem pontos fora da curva (*outliers*), comprovando estatisticamente a homogeneidade dos grupos formados pelo AG.

4.3. Abordagem *K-Means*

O agrupamento foi realizado utilizando-se o software *Minitab* na sua versão 15.1.0.0. Foi utilizado o método *K-means* que é um método heurístico que

organiza os elementos de uma amostra em *k* grupos, através da continua geração de *k* grupos, com base nas médias dos grupos gerados a cada iteração; o centro desses grupos são comparados, continuamente, com os elementos da amostra, sendo, os elementos, agrupados ao grupo mais próximo, dessa forma, são gerados novos grupos, com novos centros, até que os grupos não se alterem mais.

4.3.1. Resultados Computacionais

Tabela 4-5: Agrupamento realizado pelo *K-means*

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
1	Aprendiz 3	7	4	6	5	8	3	4	7
	Aprendiz 12	6	5	7	4	6	5	4	7
	Aprendiz 13	4	7	6	5	7	4	3	8
	Aprendiz 28	6	5	7	4	6	5	5	6
	Aprendiz 49	6	5	9	2	7	4	1	10
MÉDIA		6	5	7	4	7	4	3	8
DESVPAD		1	1	1	1	1	1	2	2
Z_1		1,64							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
2	Aprendiz 6	10	1	5	6	8	3	6	5
	Aprendiz 10	11	0	4	7	7	4	5	6
	Aprendiz 11	10	1	7	4	6	5	9	2
	Aprendiz 23	11	0	2	9	8	3	5	6
	Aprendiz 24	11	0	5	6	8	3	7	4
	Aprendiz 25	9	2	4	7	8	3	8	3
	Aprendiz 31	9	2	6	5	6	5	6	5
	Aprendiz 46	9	2	6	5	5	6	6	5
	Aprendiz 51	10	1	5	6	10	1	8	3
MÉDIA		10	1	5	6	7	4	7	4
DESVPAD		1	1	1	1	2	2	1	1
Z_2		1,99							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
3	Aprendiz 4	7	4	3	8	5	6	4	7
	Aprendiz 19	6	5	2	9	4	7	5	6
	Aprendiz 27	6	5	3	8	4	7	3	8
	Aprendiz 41	7	4	3	8	8	3	4	7
	Aprendiz 43	9	2	5	6	5	6	5	6
	Aprendiz 44	9	2	3	8	7	4	4	7
	Aprendiz 53	7	4	5	6	6	5	5	6
MÉDIA		7	4	3	8	6	5	4	7
DESVPAD		1	1	1	1	2	2	1	1
Z_3		1,70							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO

4	Aprendiz 7	8	3	6	5	9	2	6	5
	Aprendiz 14	8	3	6	5	7	4	9	2
	Aprendiz 17	8	3	6	5	10	1	8	3
	Aprendiz 34	6	5	6	5	8	3	9	2
	Aprendiz 37	6	5	5	6	8	3	7	4
	Aprendiz 42	9	2	4	7	11	0	7	4
	Aprendiz 48	6	5	5	6	9	2	6	5
MÉDIA		7	4	5	6	9	2	7	4
DESVPAD		1	1	1	1	1	1	1	1
Z_4		1,82							

Tabela 4-5: Agrupamento realizado pelo K-means (Continuação)

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
5	Aprendiz 2	3	8	3	8	8	3	7	4
	Aprendiz 22	0	11	6	5	9	2	4	7
	Aprendiz 32	4	7	7	4	6	5	9	2
	Aprendiz 38	1	10	9	2	8	3	6	5
	Aprendiz 54	2	9	5	6	6	5	10	1
MÉDIA		2	9	6	5	7	4	7	4
DESVPAD		2	2	2	2	1	1	2	2
Z_5		1,34							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
6	Aprendiz 1	6	5	3	8	10	1	9	2
	Aprendiz 9	10	1	1	10	11	0	11	0
	Aprendiz 16	4	7	3	8	10	1	8	3
	Aprendiz 18	6	5	1	10	9	2	6	5
	Aprendiz 21	9	2	2	9	9	2	8	3
	Aprendiz 29	8	3	0	11	10	1	8	3
	Aprendiz 33	9	2	3	8	9	2	8	3
	Aprendiz 39	7	4	1	10	9	2	8	3
	Aprendiz 45	7	4	2	9	11	0	8	3
MÉDIA		7	4	2	9	10	1	8	3
DESVPAD		2	2	1	1	1	1	1	1
Z_6		2,13							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
7	Aprendiz 5	4	7	5	6	4	7	1	10
	Aprendiz 30	4	7	5	6	6	5	4	7
	Aprendiz 36	3	8	5	6	2	9	5	6
	Aprendiz 40	5	6	5	6	4	7	4	7
	Aprendiz 50	4	7	5	6	4	7	5	6
MÉDIA		4	7	5	6	4	7	4	7
DESVPAD		1	1	0	0	1	1	2	2
Z_7		1,70							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
8	Aprendiz 8	7	4	3	8	6	5	7	4
	Aprendiz 15	6	5	5	6	6	5	6	5
	Aprendiz 20	8	3	5	6	5	6	7	4
	Aprendiz 26	9	2	1	10	8	3	6	5
	Aprendiz 35	8	3	3	8	5	6	9	2

	Aprendiz 47	7	4	2	9	6	5	6	5
	Aprendiz 52	8	3	3	8	7	4	6	5
MÉDIA		8	3	3	8	6	5	7	4
DESVPAD		1	1	1	1	1	1	1	1
Z_8		2,05							

O agrupamento realizado pela heurística *K-means*, Tabela 4-5, apresentou um valor máximo de contagem-z no grupo 6 igual a $\mathbf{Z_6}$ = 2,13, representando assim, um agrupamento sem pontos fora da curva (*outliers*). Também se utilizou a equação (4-6) para avaliar a qualidade do agrupamento, onde se obteve um valor da função objetivo de **FOB**=812,01.

4.4. Abordagem P-Medianas

O agrupamento foi realizado utilizando o modelo de localização de facilidades (p-medianas), padronizando o *p=8* para o número de medianas, pois o instrumento utilizado para a classificação dos estilos de aprendizagem (*ILS – Index of Learning Styles,* desenvolvido por Richard M. Felder e Bárbara Saloman) identifica as preferências de aprendizagem em quatro dimensões de estilos de aprendizagem: *Ativo/Reflexivo, Sensorial/Intuitivo, Visual/Verbal e Sequencial/Global.* Essas dimensões são baseadas no modelo de estilos de aprendizagem de (FELDER; SILVERMAN, 1988).

A matriz de distância utilizada como entrada para a formulação do problema de *p*-medianas foi gerada utilizando a distância euclidiana como medida de similaridade, com a formulação (4-7):

$$X_{iv} = \{ATI_{i1}, REF_{i2}, INT_{i3}, SEN_{i4}, VIS_{i5}, VER_{i6}, SEG_{i7}, GLO_{i8}\} \quad (4\text{-}6)$$

$$D_{lk}(X_l, X_k) = \left[\sum_{v=1}^{8} (X_{lv} - X_{kv})^2\right]^{1/2} \; ; \; \forall\, l, k \; \in N \quad (4\text{-}7)$$

4.4.1. Resultados Computacionais

O otimização do modelo de *p*-medianas (2-23) foi implementado com o *software* CPLEX 10.1. Os resultados dessa otimização são apresentados na Tabela 4-6.

Tabela 4-6: Agrupamento realizado pelo *P*-Medianas

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
1	Aprendiz 3	7	4	6	5	8	3	4	7
	Aprendiz 12	6	5	7	4	6	5	4	7
	Aprendiz 13	4	7	6	5	7	4	3	8
	Aprendiz 28	6	5	7	4	6	5	5	6
	Aprendiz 49	6	5	9	2	7	4	1	10
	Aprendiz 53	7	4	5	6	6	5	5	6
MÉDIA		6,00	5,00	6,67	4,33	6,67	4,33	3,67	7,33
DESVPAD		1,10	1,10	1,37	1,37	0,82	0,82	1,51	1,51
Z_1		1,83							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
2	Aprendiz 6	10	1	5	6	8	3	6	5
	Aprendiz 7	8	3	6	5	9	2	6	5
	Aprendiz 10	11	0	4	7	7	4	5	6
	Aprendiz 11	10	1	7	4	6	5	9	2
	Aprendiz 20	8	3	5	6	5	6	7	4
	Aprendiz 24	11	0	5	6	8	3	7	4
	Aprendiz 31	9	2	6	5	6	5	6	5
	Aprendiz 43	9	2	5	6	5	6	5	6
	Aprendiz 46	9	2	6	5	5	6	6	5
MÉDIA		9,44	1,56	5,44	5,56	6,56	4,44	6,33	4,67
DESVPAD		1,13	1,13	0,88	0,88	1,51	1,51	1,22	1,22
Z_2		2,18							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
3	Aprendiz 9	10	1	1	10	11	0	11	0
	Aprendiz 21	9	2	2	9	9	2	8	3
	Aprendiz 23	11	0	2	9	8	3	5	6
	Aprendiz 25	9	2	4	7	8	3	8	3
	Aprendiz 33	9	2	3	8	9	2	8	3
	Aprendiz 42	9	2	4	7	11	0	7	4
	Aprendiz 51	10	1	5	6	10	1	8	3
MÉDIA		8,55	1,42	2,85	7,22	8,45	1,58	7,15	3,02
DESVPAD		2,99	0,73	1,38	2,56	3,00	1,18	2,59	1,68
Z_3		2,40							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
4	Aprendiz 2	3	8	3	8	8	3	7	4
	Aprendiz 14	8	3	6	5	7	4	9	2
	Aprendiz 15	6	5	5	6	6	5	6	5
	Aprendiz 17	8	3	6	5	10	1	8	3
	Aprendiz 32	4	7	7	4	6	5	9	2
	Aprendiz 34	6	5	6	5	8	3	9	2
	Aprendiz 37	6	5	5	6	8	3	7	4
	Aprendiz 48	6	5	5	6	9	2	6	5
	Aprendiz 54	2	9	5	6	6	5	10	1
MÉDIA		5,44	5,56	5,33	5,67	7,56	3,44	7,89	3,11
DESVPAD		2,07	2,07	1,12	1,12	1,42	1,42	1,45	1,45
Z_4		2,09							

Tabela 4-6: Agrupamento realizado pelo *P*-Medianas (continuação)

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
5	Aprendiz 22	0	11	6	5	9	2	4	7
	Aprendiz 38	1	10	9	2	8	3	6	5
MÉDIA		0,50	10,50	7,50	3,50	8,50	2,50	5,00	6,00
DESVPAD		0,71	0,71	2,12	2,12	0,71	0,71	1,41	1,41
Z_5		0,71							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
6	Aprendiz 1	6	5	3	8	10	1	9	2
	Aprendiz 16	4	7	3	8	10	1	8	3
	Aprendiz 18	6	5	1	10	9	2	6	5
	Aprendiz 29	8	3	0	11	10	1	8	3
	Aprendiz 39	7	4	1	10	9	2	8	3
	Aprendiz 45	7	4	2	9	11	0	8	3
MÉDIA		6,33	4,67	1,67	9,33	9,83	1,17	7,83	3,17
DESVPAD		1,37	1,37	1,21	1,21	0,75	0,75	0,98	0,98
Z_6		1,86							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
7	7	4	3	8	5	6	4	7	7
	4	7	5	6	4	7	1	10	4
	6	5	3	8	4	7	3	8	6
	4	7	5	6	6	5	4	7	4
	3	8	5	6	2	9	5	6	3
	5	6	5	6	4	7	4	7	5
	4	7	5	6	4	7	5	6	4
MÉDIA		4,71	6,29	4,43	6,57	4,14	6,86	3,71	7,29
DESVPAD		1,38	1,38	0,98	0,98	1,21	1,21	1,38	1,38
Z_7		1,97							

GRUPO	APRENDIZ	ATI	REF	SEN.	INTU	VIS	VER	SEG	GLO
8	Aprendiz 8	7	4	3	8	6	5	7	4
	Aprendiz 19	6	5	2	9	4	7	5	6
	Aprendiz 26	9	2	1	10	8	3	6	5

	Aprendiz 35	8	3	3	8	5	6	9	2
	Aprendiz 41	7	4	3	8	8	3	4	7
	Aprendiz 44	9	2	3	8	7	4	4	7
	Aprendiz 47	7	4	2	9	6	5	6	5
	Aprendiz 52	8	3	3	8	7	4	6	5
MÉDIA		7,63	3,38	2,50	8,50	6,38	4,63	5,88	5,13
DESVPAD		1,06	1,06	0,76	0,76	1,41	1,41	1,64	1,64
Z_8		1,98							

O agrupamento realizado pelo modelo de *p*-medianas encontrou a solução ótima para o modelo (2-23) com um valor de função objetivo igual a **FOB**=162,00 com um tempo de execução de 0,22 segundos. O agrupamento também apresentou um valor máximo de contagem-z no grupo 3 igual a $\mathbf{Z_3}$ = 2,40, representando assim, um agrupamento sem pontos fora da curva (*outliers*). Utilizou-se a equação (4-3) para avaliar a qualidade do agrupamento, onde se obtém um valor de função objetivo de **FOB**=856,91.

Como o objetivo da otimização é gerar grupos onde a soma das distâncias entre os aprendizes de cada grupo seja minimizada, foi escolhido o teste de hipótese unicaudal à esquerda para analisar as soluções. Dessa forma, foram estabelecidas as seguintes hipóteses:

A hipótese de que o resultado médio do AGS é melhor que o resultado da heurística *K-means,* onde:

$H_0: AGS_{FOB} \geq K-means_{FOB}$ A média das soluções do AGS é melhor ou igual a solução do *K-means.*

$H_a: AGS_{FOB} < K-means_{FOB}$ A média das soluções do AGS é menor que a solução do *K-means.*

Como foram realizadas 50 execuções do AGS pode-se calcular a estatística do teste da seguinte forma, $z = \frac{AGS_{FOB}-K-means_{FOB}}{AGS_{DesvPad}/\sqrt{n}} = \frac{765{,}8331-812{,}01}{14{,}6748/\sqrt{50}} = -22{,}25$. Para o valor $z = -22{,}25$, temos $p-valor = 5{,}64x10^{-110}$, e como o valor de $p-valor$ é extremamente menor que $\alpha = 0{,}05$, deve-se rejeitar H_0. Comprovando

estatisticamente que a solução encontrada pelo AGS é melhor que a solução encontrada pela heurística *K-means*, quando comparadas pela equação (4-6).

Automaticamente, como o resultado do agrupamento realizado pelo método das P-Medianas obteve um resultado maior que o resultado conseguido pela heurística *K-means.* Pode-se afirmar estatisticamente que o agrupamento realizado pelo AGS é melhor que os agrupamentos realizados pela heurística *K-means* e pelo método das P-Medianas, quando comparadas pela equação (4-6).

5. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

A otimização das localizações/alocações dos laboratórios de metrologia do SISMETRA foi realizada utilizando o conjunto atual de localidades, apresentada na Figura 3-3, sendo composta de 38 localidades militares, distribuídas em 1 (um) LCC, 8 (oito) LRC e 29 (vinte e nove) LSC. Essa otimização gerou uma nova proposta de localização/alocação apresentada na Figura 3-19.

Com os resultados apresentados para o problema de agrupamento dos laboratórios de metrologia do comando da aeronáutica, pode-se avaliar que as abordagens genéticas propostas são promissoras para a otimização do modelo MCpCLP-S apresentado na formulação (3-1). Observou-se que o AGS apresentou dificuldades de encontrar a melhor solução para o problema de realocação de laboratórios com a matriz de capacidades extrapolada, pois não possui um mecanismo de refinamento robusto (*exploitation*). Por outro lado, no AGM, para a mesma configuração do problema, essa robustez é conseguida com a inclusão de uma busca local, fornecendo assim, eficiência para o mecanismo de refinamento do algoritmo. Esse melhoramento de eficiência conseguiu tornar o AGM ((43-0,35)/43)*100=99% mais rápido que o AGS e ((0,59-0,35)/0,59*100)=40% mais rápido que a otimização realizada pelo CPLEX.

Utilizando (3-1) para medir a atual localização/alocação dos laboratórios do SISMETRA obtém-se $\mathbf{FOB_{Atual}}$=1034797, e comparando a atual configuração, Figura 3-3 , com a configuração conseguida pela otimização do modelo para o problema de realocação de laboratórios com a matriz de capacidades extrapolada, Figura 3-19, com valor $\mathbf{FOB_{Otimizado}}$=130722, obtemos um Melhoramento Logístico (ML), $\boldsymbol{ML} = \frac{PLConc_{Atual} - PLConc_{Otimizado}}{PLConc_{Atual}} * 100 = 87\%$.
Dessa forma, esse ML poderá melhorar a eficiência na realização das atividades metrológicas facilitando o transporte de materiais e funcionários para a realização das atividades que proporcionam a confiabilidade metrológica do SISMETRA.

Também foi realizada a otimização para o problema com a segunda abordagem para construção da matriz de capacidades restrita (D=74, E=66, F=75). Nessa otimização foram comparados os resultados do CPLEX e do AGM, onde o CPLEX não conseguiu encontrar a solução ótima, estourando a memória aos 34666 segundos (9,6 horas) e encontrado uma solução sub-ótima factível com **FOB**=361514. Já com o AGM conseguiu-se encontrar uma solução melhor para o problema em todas as execuções da Tabela 3-5, apresentando um tempo de execução médio de $\mathbf{Tempo_{médio}}$=7724,13 segundos (2,14 horas) e desvio padrão de $\mathbf{Tempo_{desvPad}}$=8585,56 segundos (2,38 horas), as execuções também apresentaram um valor médio de chamadas da função objetivo $\mathbf{NFob_{médio}}=3{,}03\text{x}10^{8}$ com desvio padrão de $\mathbf{NFob_{desvPad}}=3{,}43\text{x}10^{8}$. Com essas execuções é observado que o AGM apresentou menor tempo médio de execução do que o utilizado pelo CPLEX para encontrar a solução ótima para o problema.

Para o segundo problema analisado foi realizada a classificação e o agrupamento de um conjunto de alunos das quatros séries do curso de Ciência da Computação do Centro Universitário Salesiano de São Paulo – Unidade de Lorena, e constatou-se uma predominância do curso pelos estilos de aprendizagem: ATIVO, SENSORIAL, VISUAL e SEQÜENCIAL.

O primeiro método proposto para realizar o agrupamento dos alunos em 8 grupos distintos utiliza o modelo de localização de facilidades (*p*-medianas) executado através do software CPLEX 10.1 conseguiu uma solução com valor para a função objetivo igual a **FOB**=162 e utilizando a equação (4-3) para avaliar a qualidade do agrupamento obtém-se um valor de função objetivo de **FOB**=856,91.

O segundo método utilizado foi a heurística *k-means* (k-médias) implementada através do software Minitab na sua versão 15.1.0.0 e utilizando a equação (4-6) para avaliar a qualidade do agrupamento, obtém-se um valor da função objetivo de **FOB**=812,01.

O terceiro agrupamento utiliza um algoritmo genético simples como ferramenta para realizar o agrupamento dos aprendizes. Foram realizados testes para a escolha dos melhores parâmetros genéticos, sendo os escolhidos: *taxa de cruzamento* = 80%, *taxa de mutação* = 0,5%, substituição da população *geracional* com *elitismo* do melhor individuo, *tamanho da população* = 20 indivíduos, *critério de parada* = 100000 iterações sem melhora da melhor solução corrente. Após a determinação dos parâmetros genéticos foram realizadas 50 execuções, onde foram encontrados: $\mathbf{FOB_{médio}}$=765,8331, com desvio padrão – $\mathbf{FOB_{desvPad}}$ =14,6748.

Com auxilio do método estatístico contagem-z (valor padronizado, escore padronizado z) conseguiu-se avaliar a similaridade interna dos grupos formados, pois não foram encontrados pontos foras da curva (*outliers*) dentro dos limites de -3 a +3 desvios padrão da média. Dessa forma, o procedimento proposto possui eficiência na classificação dos aprendizes de acordo com o método proposto por Richard M. Felder e Linda K. Silverman (1988), e os métodos analisados para o agrupamento dos aprendizes de acordo com os seus estilos de aprendizagem apresentaram eficiência no agrupamento, se destacando o agrupamento realizado pelo Algoritmo Genético que apresentou menor valor de função objetivo, equação (4-6), e o agrupamento realizado pelo *k-means* que apresentou melhor valor $\mathbf{Z_{max}}$=2,13.

Avaliou-se também através de um teste de estatístico, Teste Z, com nível de significância de α=0,05, a afirmação de que o agrupamento de aprendizes conseguido pelo AGS, e que apresentou **FOB**=765,8331 é melhor que os agrupamento dos métodos *K-means*, com **FOB**=812,01 e P-Medianas, com **FOB**=856,91. Esses valores de função objetivo foram conseguidos com a equação (4-6).

Para trabalhos futuros seria interessante utilizar as técnicas abordadas de localização e agrupamento para instâncias maiores e diferentes dos problemas analisados, verificando assim, sua robustez e escalabilidade.

Para o problema de localização apresentado pode-se aumenta o número de localidades consideradas, verificando assim, uma maior cobertura do território nacional. Pode-se também realizar a otimização com valores reais de demanda e capacidade, alterando o número *p* de concentradores para verificar qual o número ideal de concentradores devem ser utilizados para melhor atender as demandas de calibração.

Pode-se utilizar as técnicas apresentadas nesse trabalho para a classificação e agrupamento dos aprendizes em ambientes de educação a distância, aumentando assim o tamanho do problema, podendo também variar o número de *k* grupos e realizar o agrupamento com aprendizes de cursos diferentes, verificando assim, se existe uma relação dos estilos de aprendizagem com a escolha do curso.

As otimizações realizadas também poderão ser confrontadas com outras técnicas de otimização, como: Algoritmos Genéticos Construtivos, *Simulated Annealing*. Busca Tabu, Algoritmo de Otimização Extrema Generalizada (*Generalized Extremal Optimization* - GEO), etc.

REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, W. G. *Métodos heurísticos para o problema de localização de concentradores*. 2009. 112 p. (INPE-15744-TDI/1489). Dissertação (Mestrado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2009. Disponível em: http://urlib.net/8JMKD3MGP8W/34PGRBS>. Acesso em: 10 set. 2013.

ANDERSON, D. R.; SWEENEY, D. J.; WILLIAMS, T. A. *Estatística aplicada à administração e economia.* São Paulo: CENGAGE Learning, 2008. v.2.

ARAKAKI, R. G. I. *Heurística de localizacão-alocacão para problemas de localizacão de facilidades*. 2002. 79 p. (INPE-9572-TDI/837). Tese (Doutorado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2002. Disponível em:<http://urlib.net/sid.inpe.br/jeferson/2004/03.08.16.47>. Acesso em: 16 set. 2013.

ARAKAKI, R. G. I.; Lorena, L. A. N.. Uma heurística de localização-alocação (HLA). *Produção*, v. 16, n.2, p. 319-328, maio-ago. 2006.

AYKIN, T. Networking policies for hub-and-spoke systems with application to the air transportation system. *Transportation Science*, Aug 95, v. 29, n. 3, p. 201, 1995.

BENTO, E. P.; KAGAN, N. Algoritmos genéticos e variantes na solução de problemas de configuração de redes de distribuição. *Controle & Automação*, v. 19, n. 3, p. 302-315, 2008.

BRASIL. MINISTÉRIO DA AERONÁUTICA*Portaria N° 858/GM3*. Brasília, 1988. Institui o Sistema de Metrologia Aeroespacial do Ministério da Aeronáutica.

BRASIL. MINISTÉRIO DA DEFESA. COMANDO DA AERONÁUTICA. 2008. *NSCA 9-1 Sistema de Metrologia Aeroespacial (SISMETRA).* São José dos Campos(SP): s.n.

BRASIL. MINISTÉRIO DA DEFESA. COMANDO DA AERONÁUTICA. *Estrutura Funcional do Sistema de Metrologia Aeroespacial (SISMETRA): São José dos Campos, 2006. NSCA 9-4.*

CAMPBELL, J. F. Hub location problems and the p-hub median problem. *Operations Research*, v. 44, pp. 923-935, 1991.

CAMPBELL, J. F. Integer programming formulations of discrete hub location problems. *European Journal Of Operational Research,* v. 72, pp. 387-405, 1994.

CARVALHO, A. C. P. L. F., BRAGA, A. P., LUDEMIR, T. B. Computação Evolutiva. In: REZENDE, S. O. (org.). *Sistemas inteligentes*: fundamentos e aplicações. Barueri, SP: Manole, 2003. p. 225-239.

CHAVES, A. A. *Uma meta-heurística híbrida com busca por agrupamentos aplicada a problemas de otimização combinatória*. 2009. 197 p. (INPE-15685-TDI/1459). Tese (Doutorado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2009. Disponível em: <http://urlib.net/8JMKD3MGP8W/34NDEC8>. Acesso em: 16 set. 2013.

CUNHA, C. B., SILVA, M. R. A genetic algorithm for the problem of configuring a hub-and-spoke network for a LTL trucking company in Brazil. *European Journal of Operational Research*, v. 179, pp. 747-758, 2007.

CURY, H. N. Estilos de aprendizagem de alunos de engenharia. In: CONGRESSO BRASILEIRO DE ENSINO DE ENGENHARIA, 28., 2000, Ouro Preto. *Anais...* Natal: ABENGE, 2000. p. 9. CD-Rom.

DARWIN, C. *The origin of species by means of natural selection*. 1859. Disponível em:: http://www.talkorigins.org/faqs/origin.html
Acesso em: 11 Outubro 2012.

DAWKINS, R. *The selfish gene.* New York: Oxford University Press, 1976.

DEUTSCH, S. J.; FREEMAN, S. F.; HELANDER, M. Manufacturing cell formation using an improved P-median model. *Computers and Industrial Engineering*, v. 34, n.1, p. 135–146, 1998.

DIAS, C. R. *Algoritmos evolutivos para problemas de clusterização de grafos orientados: desenvolvimento e análise experimental.* 2004. Dissertação (Mestrado em Computação Aplicada e Automação) - Universidade Federal Fluminense, Niterói, 2004. Disponível em: <http://www.bdtd.ndc.uff.br/tde_arquivos/41/TDE-2008-03-05T114032Z-1361/Publico/Carlos%20Dias-dissert.pdf >. Acesso em: 16 set. 2013.

ESPEJO, L. G. A.; GALVÃO, R. D. O uso de ralaxações langrangeana e surrogate em problemas de programação inteira. *Pesquisa Operacional*, v. 22, n.3, pp. 387-402, Junho a Dezembro 2002.

FELDER, R. M.; SILVERMAN, L. K. Learning and teaching styles. *Journal of Engineering Education*, v. 78, n.7, pp. 674–681, 1988.

FELDER, R. M.; SOLOMAN, B. A. *Index of learning styles*. Disponível em: <http://www.engr.ncsu.edu/learningstyles/ilsweb.html>. Acesso em: 20 Setembro 2010.

FELIX, Z.; TEDESCO, P*. Formação de grupos de aprendizagem em ambientes CSCL ciente de contexto.* 2006. Disponível em: < http://www.aedb.br/seget/artigos06/646_GruposSeget2006.pdf>. Acesso em: 16 set. 2013.

FILHO, P. A. C.; POPPI, R. J. Algoritmo genético em química. *Química Nova*, v.22, n.3, pp. 405-411, 1999.

FORGUS, R. H. *Percepção o processo básico do desenvolvimento cognitivo.* São Paulo: Editora da Universidade de São Paulo, 1971.

GAREY, M.; JOHNSON, D. *Computers and Intractability: a guide to the theory of NP-completeness.* São Francisco: W. H. FREEMAN, 1979. 338 p.

GEN, M.; TSUJIMURA, Y.; ISHIZAKI, S. Optimal design of a star-LAN using neural networks. *Computers & Industrial Engineering,*v.31, p. 855-859, 1996.

GOLDBARG, M. C. et al. Algoritmo evolucionário para otimização do plano de tratamento em radioterapia conformal 3D. *Pesquisa Operacional*, v.29, n.2, p. 239-267, maio a ago., 2009.

GOLDBARG, M. C.; LUNA, H. P. L. *Otimização combinatória e programação linear: modelos e algoritmos.* 2. ed. Rio de Janeiro - RJ: Elsevier, 2005.

GOLDBERG, D. E. *Genetic algorithms in search, optimization, and machine learning.* The Universiy Alabama: Addison-Wesley, 1989

GOLDBERG, D. E.; VOESSNER, S. *Optimizing global-local search hybrids*. Disponível em: http://reference.kfupm.edu.sa/content/o/p/optimizing_global_local_search_hybrids_125387.pdf. 1999. Acesso em: 10 Outubro de 2012.

GOLDMAN, A. Optimal location for centers in a network. *Transportation Science,* v.3, p. 352–360, 1969.

HAKIMI, S. Optimum location of switching centers and the absolute centers and the. *Operations Research*, v.12, n.3, p. 450–459, 1964.

HAKIMI, S. Optimum location of switching centers in a communications network and some related graph theoretic problems. *Operations Research*, v.13, p. 462-475, 1965.

HAKIMI, S. L. Optimum locations of switching centers and the absolute centers and medians of a graph. *Oper Res,* v.12, p. 450–459, 1964.

HALL, R. W. Configuration of an overnight package air network. *Transportation Research,* v.23, p. 139-149, 1989.

HILLIER, F. S.; LIEBERMAN, G. J. *Introdução à pesquisa operacional.* 8. ed. Porto Alegre - RS: AMGH, 2010.

HOLLAND, J. H. *Adaptation in natural and artificial systems*. MIT Press*,* 1975. p. 11-147.

INMETRO. *Vocabulário Internacional de Metrologia Conceitos Fundamentais e Gerais e Termos Associados (VIM 2008).* 1. ed. Rio de Janeiro(RJ): INMETRO, 2009.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. *ACM computing surveys*, v.31, n.3, p. 264-323, set., 1999.

KIM, J.; YONG-HYUK, K.; INWOOD, H.; BYUNG-RO, M. Genetic approaches for graph partitioning: a survey. In: ANNUAL CONFERENCE ON GENETIC AND EVOLUTIONARY COMPUTATION, 13., 2011, New York. *Proceedings…* New York, NY, USA: ACM, 2011. p. 8.ISBN: 978-1-4503-0557-0 doi>10.1145/2001576.2001642.

KLINCEWICZ, J. G. Hub location in backbone/tributary network design: a review. *Location Science,* v.6, p. 307-335, 1998.

KRASNOGOR, N.; SMITH, J. A Tutorial for competent memetic algorithms: model, taxonomy, and design issues. *IEEE Transactions On Evolutionary Computation,* out, v.9, n.5, p. 474-488, 2005.

LABBÉ, M.; YAMAN, H. Solving the Hub Location Problem in a Star – Star Network. *Networks, v.*51, n.1, p. 19-33, 2008.

LOPES, W. M. G. *Inventário de estilos de aprendizagem de felder-soloman: investigação de sua validade em estudantes universitários de Belo Horizonte.* 2002. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal de Santa Catarina, Centro Tecnológico, Florianópolis, 2002 Disponível em: < http://www.tede.ufsc.br/teses/PEPS3508-D.pdf >. Acesso em: 16 set. 2013.

LORENA, L. A. N.; FURTADO, J. C. Constructive Genetic Algorithm for Clustering Problems. *Evolutionary Computation,* v.9, n.3, pp. 309-327, 2001.

LORENA, L. A. N.; SENNE, E. L. F. *Local search heuristics for capacitated p-median problems.* Budapest - Hungria, s.n, 2000.

LORENA, L. A. N.; SENNE, E. L. F.; PAIVA, J. A. M.; MARCONDES. S. P. B. Integracao de um modelo de p-medianas a sistemas de informacoes geograficas. In*:* SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 31., 1999, Juiz de Fora, MG. *Anais...* Juiz de Fora, 1999. P. 635-647.

LOVE, R. F.; MORRIS, J. G.; WESOLOWSKY, G. O. *Facilities location*: models & methods. New York: Elsevier Science Publishing, 1988.

MACHADO, C. S.; PALHANO, M. D. M.; PECONICK, M. L.; ÁVILA, V. A. *Estilos de aprendizagem* – uma abordagem utilizando o ILS - index of learning styles, 2010. Disponível em: http://www.abepro.org.br/biblioteca/ENEGEP2001_TR111_0240.pdf Acesso em: 19 fev. 2010.

MARTINS, G. A. *Estatística geral e aplicada.* São Paulo: Atlas, 2008.

MENDES, J. J. d. M.; GONÇALVES, J. F. Um algoritmo genético para o problema de sequenciamento de projectos com recursos limitados. *Investigação Operacional*, v.23, p. 179-195, 2003.

MERCERON, A.; YACEF, K. Clustering students to help evaluate learning. *Technology Enhanced Learning,* v.171, pp. 31-42, 2005.

MIASAKI, C. T.; ROMERO, R. Um algoritmo genético especializado aplicado ao planejamento da expansão do sistema de transmissão com alocação de dispositivos de compensação série. *Revista Controle & Automação,* v.18, n.2, p. 210-22, 2007.

MICHALEWICZ, Z. *Genetic algorithms + data structures* = evolution programs. 3. ed. New York: Spring-Verlag, 1996.

MINGOTI, S. A. *Análise de dados através de métodos de estatística multivariada.* Belo Horizonte: Editora UFMG, 2005.

O' KELLY, M. E.; MILLER, H. J. The hub network design problem A review and synthesis. *Journal of Transport Geography,* v.2, n.1, p. 31-40, 1994.

O'KELLY, M. A quadratic integer program for the location of interacting hub facilities. *European Journal of Operational Research,* v.32, p. 393-404, 1987.

O'KELLY, M. E.; BRYAN, D.; SKORIN-KAPOV, D.; SKORIN-KAPOV, J. Hub network design with single and multiple allocation : a computational study. *Locafion Scienc,* v. 4, n.3, p. 125-138, 1996

O'KELLY, M.; LAO, Y. Mode choice in a hub-and-spoke network: a zero-one linear programming approach. *Geographical Analysis,* v.23, p. 283-297, 1991.

ORACLE. *API Specification Java 2 Platform.* 2011.. Disponível em: <http://download.oracle.com/javase/1.4.2/docs/api/java/util/Random.html>. Acesso em: 15 jun. 2011.

PIMENTEL, E. P.; FRANÇA, V. F.; OMAR, N. A identificação de grupos de aprendizes no ensino presencial utilizando técnicas de clusterização, 2003. In:

SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 14., 2003. *Anais...* Disponível em: <http://www.nce.ufrj.br/sbie2003/publicacoes/paper52.pdf>. p.10. 2003

PIZZOLATO, N. D.; SILVA, H. B. F. The location of public schools: evaluation of practical experiences. *International Transactions in Operational Research*, v.4, n.1, p. 13–22, 1997.

RECHE, M. M.; ASSIS, A. S. *Acervo digital.* Rio de Janeiro: INMETRO, p. 10, 2010.

REEVES, C. R. *Modern heuristic techniques for combinatorial problems.* Great Britain: McGraw-Hill Book Company Europe, 1995.

RIBEIRO, G. F.; NAGANO, M. S.; LORENA, L. A. N. *Metaheurística híbrida algoritmo genético-clustering search para redução de estoques em processamento em sistemas de produção flow shop permutacional.* 2007. Disponível em:<http://www.lac.inpe.br/~lorena/geraldo/ECS-flow-shop.pdf>. Acesso em: 16 set. 2013.

ROLLAND, E.; SCHILLING, D.; CURRENT, J. An efficient tabu search procedure for the p-median problem. *European Journal of Operational Research*, v.96, n.2, p. 329–342, 1996.

ROSÁRIO, R. R. L; CARNIERI, C.; STEINER, M. T. A. Aplicação do problema das p-medianas para determinar a localização de unidades de saúde 24 horas. 2002. In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO, 22., 23 a 25 de outubro de 2002. Curitiba – PR. *Anais...* Disponível em: <http://www.abepro.org.br/biblioteca/ENEGEP2002_TR61_0949.pdf>. p.8. 2002.

RUSSEL, S. J. *Inteligência artificial.* 2. ed. Rio de Janeiro: Elsevier, 2004.

SÁ, E. M. *Localização de concentradores aplicada ao transporte público.* Belo Horizonte - MG: Universidade Federal de Minas Gerais, 2011.

SILVA, S. L. C.; LORENA, L. A. N. Algoritmo genético para a clusterização de grupos de aprendizes utilizando o inventário de estilos de aprendizagem de Felder-Soloman. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 22., 2011. Aracaju-SE, 21 a 25 de novembro 2011. *Anais…* 2011. P.160-163. Disponível em: <http://www.br-ie.org/pub/index.php/sbie/article/view/1821>.

SMITH, K.; KRISHNAMOORTHY, M.; PALANISWAMI, M. Neural versus traditional approaches to the location of interacting hub facilities. *Location Science, v.*4, n.3, p. 155-171, 1996.

TEITZ, M.; BART, P. Heuristics methods for estimating the generalized vertex median of a weighted graph. *Operations Research*, v.16, v.5, p. 955–961, 1968.

TOPCUOGLU, H., CORUT, F., ERMIS, M., YLMAZ, G. Solving the uncapacitated hub location problem using genetic algorithms. *Computers and Operations Reseach,* v.32, p. 967-984, 2005.

TRINDADE, Á. R., OCHI, L. S. Um algoritmo evolutivo híbrido para a formação de células de manufatura em sistemas de produção. *Pesquisa Operacional,* maio.ago., p. 255-294, 2006.

YAMAN, H. Star p-hub median problem with modular arc capacities. *Computers & Operations Research,* v.35, p. 3009-3019, 2008.

ZAKRZEWSKA, D. Cluster analysis in personalized e-learning systems. *Intelligent Systems for Knowledge Managemente Studies in Computational Intelligence,* v.252, p. 229-250, 2009.