



**PODER EXECUTIVO
MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA**



**PROPOSTA DE UMA METODOLOGIA PARA EXPANSÃO DE ESCOLAS PÚBLICAS
NO ESTADO DO AMAZONAS**

CLAHILDEK MATOS XAVIER

MANAUS

2020

CLAHILDEK MATOS XAVIER

PROPOSTA DE UMA METODOLOGIA PARA EXPANSÃO DE ESCOLAS PÚBLICAS
NO ESTADO DO AMAZONAS

Tese apresentada ao Programa de Pós-Graduação em Informática do Instituto de Computação da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Doutor em Informática.

Orientador: Prof. Dr. Cícero Ferreira Fernandes Costa Filho
Co-Orientador(a): Prof^ª. Dr^ª. Marly Guimarães Fernandes Costa

MANAUS

2020

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

X3m Xavier, Clahildek Matos
Proposta de uma Metodologia para Expansão de Escolas
Públicas no Estado do Amazonas / Clahildek Matos Xavier . 2020
98 f.: il. color; 31 cm.

Orientador: Cícero Ferreira Fernandes Costa Filho
Coorientadora: Marly Guimarães Fernandes Costa
Tese (Doutorado em Informática) - Universidade Federal do
Amazonas.

1. modelagem p-mediana. 2. localização de escolas de ensino
superior. 3. localização de objetivo único. 4. algoritmo genético. I.
Costa Filho, Cícero Ferreira Fernandes. II. Universidade Federal do
Amazonas III. Título



PODER EXECUTIVO
MINISTÉRIO DA EDUCAÇÃO
INSTITUTO DE COMPUTAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA



UFAM

FOLHA DE APROVAÇÃO

"PROPOSTA DE UMA METODOLOGIA PARA EXPANSÃO DE
ESCOLAS PÚBLICAS NO ESTADO DO AMAZONAS"

CLAHILDEK MATOS XAVIER

Tese de Doutorado defendida e aprovada pela banca examinadora constituída pelos Professores:

Cícero F. F. Costa Filho
Prof. Cícero Ferreira Fernandes Costa Filho - PRESIDENTE

Fabiola Guerra Nakamura
Profa. Fabiola Guerra Nakamura - MEMBRO EXTERNO

Helder Cruz da Silva
Prof. Helder Cruz da Silva - MEMBRO EXTERNO

José Raimundo Gomes Pereira
Prof. José Raimundo Gomes Pereira - MEMBRO EXTERNO

Luís Fernando Bueno
Prof. Luís Fernando Bueno - MEMBRO EXTERNO

Manaus, 24 de Março de 2020

Aos meus pais, Francisco Deolindo Xavier Filho e Juracy Matos Bandeira, pelo apoio e compreensão durante toda essa trajetória. Aos meus amigos e colegas de Mestrado/Doutorado, e a todos que contribuíram para a realização deste trabalho.

AGRADECIMENTOS

A Deus, toda a glória e toda adoração.

A Nossa Senhora, Maria Santíssima, abençoada, pelas preces e orações.

Ao meu estimado orientador Prof. Dr. Cicero Ferreira Fernandes Costa Filho agradeço pelo comprometimento, dedicação, empenho, conselhos e orientações de fundamental importância durante todo o andamento do doutorado. À minha estimada co-orientadora Prof^a. Dr^a. Marly Guimarães Fernandes Costa pelo apoio, pelo conhecimento, pela revisão dos trabalhos, dicas, ensinamentos e acompanhamento. Deixo o meu profundo agradecimento a ambos, por ensinarem o valor transformador e engrandecedor da educação.

Aos meus amigos de Mestrado/Doutorado pelos grupos de estudo formados aos sábados, domingos e feriados. Em especial, ao Leonardo Fontes, Urique Hoffmann, Diego Barros, Rayol Neto, Bernardo Gatto, Haline Oliveira, Carlos Ramos, Janaína Nascimento, Delano Campos. Agradeço pela importante companhia no estudo das disciplinas.

Aos professores do IComp, em especial ao Altigran Soares e Eduardo Nakamura, pelos ensinamentos. Ao Coordenador do PPGI, Eduardo Feitosa, e à Secretaria do IComp, pela célere prestação em atender aos alunos.

A toda minha família pelo importante apoio, especialmente aos meus pais, Francisco e Juracy e aos meus sobrinhos Emanuely, Adonias e Loreny.

Às minhas amigas Fernanda Caetano e Renata Mota pelo incentivo, motivação, apoio, opiniões e companhia durante essa jornada. Aos amigos Joel Parente, Sebastião Júnior, Rodrigo Leme, Ricardo Gorayeb, Ailton Cortez e Francisco Enedelson pela compreensão nesse período. Deixo aqui os meus sinceros agradecimentos.

Ao CENSIPAM pelos conhecimentos profissionais necessários para alcançar o doutorado.

À UFAM, ao PPGI e ao PPGEE/CETELI pela oportunidade e estrutura disponibilizada.

E a todos que contribuíram de alguma forma para a realização deste trabalho.

RESUMO

Este estudo propõe determinar localizações ótimas para a expansão de um sistema de ensino superior usando critérios populacionais e sociais. Com esse objetivo, este trabalho avalia modelos de localização de objetivo único na determinação da distribuição ótima de facilidades de ensino superior no estado do Amazonas, Brasil. Três opções de otimização são avaliadas e disponibilizadas aos tomadores de decisão: 1) priorizar cidades com baixo valor de Índice de Desenvolvimento Humano (IDH); 2) priorizar cidades com maior população; 3) favorecer ambos os critérios. Além disso, a localização deve igualar a distribuição dos alunos entre as regiões do Amazonas. Com esse objetivo, foram avaliados três modelos discretos de localização: modelo p-centro, modelo p-mediana e o modelo p-dispersão. Os modelos de localização foram projetados usando a metaheurística algoritmo genético. Uma implementação avançada do algoritmo genético que otimiza a solução e tempo de convergência foi usada. As expansões propostas aqui apresentam valores médios mais baixos de Índice de Desenvolvimento Humano em comparação com 0,619 da distribuição existente dos *campi*. Os melhores resultados foram obtidos com o modelo p-mediana.

Palavras-chave: modelagem p-mediana; localização de escolas de ensino superior; localização de objetivo único; algoritmo genético.

ABSTRACT

This study proposes determining optimal locations for expanding a higher education system by using populational and social criteria. With this aim, this work evaluates single objective location models in determining the optimal distribution of higher education facilities in Amazonas State, Brazil. Three optimization options are evaluated and made available to decision makers: 1) prioritize cities with a lower United Nation Human Development Index; 2) prioritize cities with a higher population; 3) favor both criteria. Also, the location must equalize student distribution between the regions of Amazonas State. With this aim, three discrete location models were evaluated: p-center model, p-median model, and p-dispersion model. The location models were designed using a Genetic Algorithm metaheuristic. A state-of-art implementation of the Genetic Algorithm that optimizes the solution and converge time was used. The expansions proposed here present lower mean values of United Nation Human Development Index compared with 0.619 from the existing distribution of campuses. The best results were obtained with the p-median model.

Keywords: p-median modeling; locating higher education schools; single objective location; genetic algorithm.

LISTA DE FIGURAS

Figura 1.1 – Estado do Amazonas com 62 cidades, rios (azul) e estradas (vermelho). <i>Shapefiles</i> – arquivos com dados geoespaciais – obtidos através da ANA (2016), DNIT (2016), IBGE (2016a), IBGE (2016b) e FORESTGIS (2020), respectivamente, para hidrovias, rodovias, mapa do Estado do Amazonas, mapa do Brasil e mapa da América do Sul.	16
Figura 1.2 – Mapa do estado do Amazonas mostrando mesorregiões do estado e cidades com pelo menos uma unidade de ensino superior. <i>Shapefiles</i> – arquivos com dados geoespaciais – obtidos através do IBGE (2016a), IBGE (2016b), IBGE (2019) e FORESTGIS (2020), respectivamente, para o mapa do Estado do Amazonas, mapa do Brasil, mapa das mesorregiões do Estado do Amazonas e mapa da América do Sul.	17
Figura 2.1 – Exemplo de população de indivíduos.	33
Figura 2.2 – Probabilidades de seleção da solução.	35
Figura 2.3 – Uma solução domina a população com seu valor de aptidão.	36
Figura 2.4 – Solução baseada na probabilidade de seleção através do <i>rank</i>	37
Figura 2.5 – Exemplo de um cruzamento 1-ponto.	38
Figura 2.6 – Exemplo de um cruzamento 2-pontos.	39
Figura 2.7 – Exemplo de cruzamento uniforme.	39
Figura 2.8 – Exemplo de operador de mutação.	40
Figura 4.1 – Municípios de Fonte Boa e Uarini no rio Solimões, Amazonas-Brasil.	54
Figura 4.2 – Distância fluvial entre as cidades de Fonte Boa e Juruá.	55
Figura 4.3 – Exemplo de um cromossomo codificado com 13 facilidades. O azul indica os genes fixos. Os outros genes (em verde) são aqueles que podem ser substituídos.	67
Figura 4.4 – Exemplo de operador de cruzamento de um ponto. Quadrados de cor azul indicam os genes fixos. Os outros genes (em verde) são aqueles que podem ser substituídos.	69
Figura 4.5 – Exemplo de operação de mutação de um gene. Quadrados de cor azul indicam os genes fixos. Os outros genes (em verde) são aqueles que podem ser substituídos.	69
Figura 5.1 – Resultados obtidos usando o modelo PMM para expandir de 10 para 11 a 16 cidades com IES: (a) métrica 1; (b) métrica 2, e; (c) métrica 3. Todas as distâncias estão em quilômetros.	72
Figura 5.2 – Resultados obtidos da métrica 2 para expandir de 10 para 11 a 16 cidades com IES utilizando os modelos: (a) PMM; (b) PCM, e; (c) PDM. Todas as distâncias estão em quilômetros.	74
Figura 5.3 – Resultados obtidos usando o modelo PDM para expandir de 10 para 11 a 16 cidades com IES: (a) métrica 1; (b) métrica 2, e; (c) métrica 3. Todas as distâncias estão em quilômetros.	76
Figura 5. 4 – Valores para a métrica 6 obtidos para os modelos PMM, PCM e PDM, com a expansão de 11 a 16 cidades com IES.	81

Figura 5. 5 – Valores da métrica 7 obtidos para os modelos PMM, PCM e PDM, com expansão de 11 para 16 cidades com IES.	82
Figura 5. 6 – Resultados obtidos usando o modelo PMM combinado com a estratégia C para a expansão de 10 a 16 cidades com IES. <i>Shapefiles</i> – arquivos com dados geoespaciais – obtidos através do IBGE (2016a), IBGE (2016b), IBGE (2019) e FORESTGIS (2020), respectivamente, para o mapa do Estado do Amazonas, mapa do Brasil, mapa das mesorregiões do Estado do Amazonas e mapa da América do Sul.	84

LISTA DE TABELAS

Tabela 1.1 – Comparando dados populacionais e educacionais das mesorregiões do Amazonas, Brasil,	19
Tabela 1.2 – Número médio de AEM e IDH médio das cidades com IES e todas as cidades do estado do Amazonas, (UFAM, 2016; UEA, 2012).....	19
Tabela 2.1 – Exemplo de seleção <i>rank</i> linear para um problema de maximização.....	37
Tabela 3.1 – Resumo dos artigos selecionados.	44
Tabela 4.1 – Exemplo da matriz de distância fluvial (km) com 16 municípios do estado do Amazonas utilizados no estudo.	57
Tabela 4.2 – AEM e IDH dos 62 municípios do Amazonas utilizados neste estudo. IBGE (2018, 2010).	58
Tabela 5.1 – Valores da métrica 4 para as três estratégias A, B e C, para expansão de 10 a 16 cidades com IES.	77
Tabela 5. 2 – Valores da métrica 5 com as estratégias A, B e C, para expansão de 10 a 16 cidades com IES.	79
Tabela 5. 3 – Tempos de simulação obtidos com a configuração de parâmetros ($\alpha = 1/3$, $\beta = 1/3$, $\gamma = 1/3$) para a expansão de 10 para 11, 12, 13, 14, 15 e 16 cidades com IES.	86

LISTA DE ABREVIATURAS

AEM	Alunos no Ensino Médio
ANA	Agência Nacional de Águas
ANTAQ	Agência Nacional de Transportes Aquaviários
DNIT	Departamento Nacional de Infraestrutura de Transportes
GAMS	<i>General Algebraic Modeling System</i>
GA	<i>Genetic Algorithm</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
IES	Instituição de Ensino Superior
IDH	Índice de Desenvolvimento Humano
PMM	Modelo p-mediana
PCM	Modelo p-centro
PDM	Modelo p-dispersão
SEMESP	Sindicato das Mantenedoras de Ensino Superior de São Paulo
TS	<i>Tabu Search</i>
UFAM	Universidade Federal do Amazonas
UEA	Universidade Estadual do Amazonas

SUMÁRIO

CAPÍTULO 1	15
1. INTRODUÇÃO.....	15
1.1 CONTEXTUALIZAÇÃO DO PROBLEMA.....	18
1.2 OBJETIVO GERAL.....	22
1.3 OBJETIVOS ESPECÍFICOS	22
1.4 DESENVOLVIMENTO DO TRABALHO	23
CAPÍTULO 2	24
2. FUNDAMENTOS TEÓRICOS	24
2.1 OTIMIZAÇÃO MONO-OBJETIVO E MULTIOBJETIVO	24
2.2 MODELO DE LOCALIZAÇÃO DISCRETO.....	26
2.2.1 Modelo p-mediana.....	27
2.2.2 Modelo p-centro	28
2.2.3 Modelo p-dispersão	30
2.3 ALGORITMO GENÉTICO	31
2.3.1 Algoritmo Genético	32
2.3.2 População	32
2.3.3 Inicialização.....	33
2.3.4 Avaliação.....	34
2.3.5 Seleção.....	34
2.3.5.1 Seleção através da Roleta	34
2.3.5.2 Seleção através do <i>Rank</i> Linear.....	36
2.3.6 Recombinação	38
2.3.7 Mutação	40
2.3.8 Codificação dos cromossomos ou indivíduos	40
2.4 CONSIDERAÇÕES	41
CAPÍTULO 3	41
3. REVISÃO BIBLIOGRÁFICA.....	42
3.1 TRABALHOS RELACIONADOS	43
3.2 CONSIDERAÇÕES E CONTRIBUIÇÃO DESTE TRABALHO	52
CAPÍTULO 4	53
4. METODOLOGIA.....	53

4.1	MATERIAIS	53
4.2	MÉTODOS.....	58
4.2.1	Normalização dos parâmetros	58
4.2.2	Modelagem da Locação.....	59
4.2.3	Métricas para Avaliação dos Resultados.....	63
4.2.4	Modelagem de Algoritmos Genéticos	64
4.3	CONSIDERAÇÕES	70
CAPÍTULO 5		71
5.	RESULTADOS E ANÁLISE DOS RESULTADOS.....	71
5.1	RESULTADOS DAS MÉTRICAS DO GRUPO 1.....	71
5.2	DISCUSSÃO DOS RESULTADOS PARA AS MÉTRICAS DO GRUPO 1.....	77
5.2.1	Avaliação se as soluções obtidas pela combinação de modelos e estratégias de localização citadas na Seção 1 fornecem uma melhor distribuição das IES	77
5.3	RESULTADOS DA MÉTRICA DO GRUPO 2.....	79
5.4	DISCUSSÃO DOS RESULTADOS PARA AS MÉTRICAS DO GRUPO 2.....	80
5.4.1	Avaliação se as soluções obtidas combinando modelos de localização e estratégias citadas na Seção 1 fornecem melhor equalização da taxa de AEM de todas as mesorregiões	80
5.5	RESULTADOS DAS MÉTRICAS DO GRUPO 3.....	81
5.6	DISCUSSÃO DOS RESULTADOS DAS MÉTRICAS DO GRUPO 3	82
5.6.1	Avaliação se as soluções obtidas combinando modelos de localização e algumas estratégias citadas na Seção 1 favorecem cidades com menor IDH ou cidades com maior número de AEM	82
5.7	RESULTADOS OBTIDOS PELO ALGORITMO GENÉTICO.....	85
5.7.1	Comentários sobre os Resultados do Algoritmo Genético.....	86
5.8	CONSIDERAÇÕES FINAIS	87
CAPÍTULO 6		88
6.	CONCLUSÕES E TRABALHOS FUTUROS	88
6.1	CONCLUSÕES.....	88
6.2	TRABALHOS FUTUROS.....	91
REFERÊNCIAS BIBLIOGRÁFICAS		92
APÊNDICE A		98
LISTA DE EVENTOS E PUBLICAÇÕES		98

1. INTRODUÇÃO

De acordo com o relatório do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP, 2018), a porcentagem da população no Brasil com acesso à educação superior é de apenas 19,6% na faixa etária entre 25 e 34 anos e de 13,7% na faixa entre 55 e 64 anos. Portanto, há muito espaço para o crescimento do ensino superior no Brasil. No Brasil, obter um diploma de graduação pode garantir uma renda 80% maior que a de um trabalhador médio não graduado (OECD, 2018).

Na região Norte do Brasil, no estado do Amazonas, a maior área da federação com 1.559.146,876 km², esse problema é agravado. Apenas 6,2% da população com 25 ou mais anos de idade têm ensino superior completo no estado do Amazonas (IBGE, 2018);

Conforme é mostrado na Figura 1.1, o estado do Amazonas compreende 62 municípios distribuídos em quatro mesorregiões: Central, Norte, Sul e Sudoeste. Essas mesorregiões podem ser vistas na Figura 1.2.

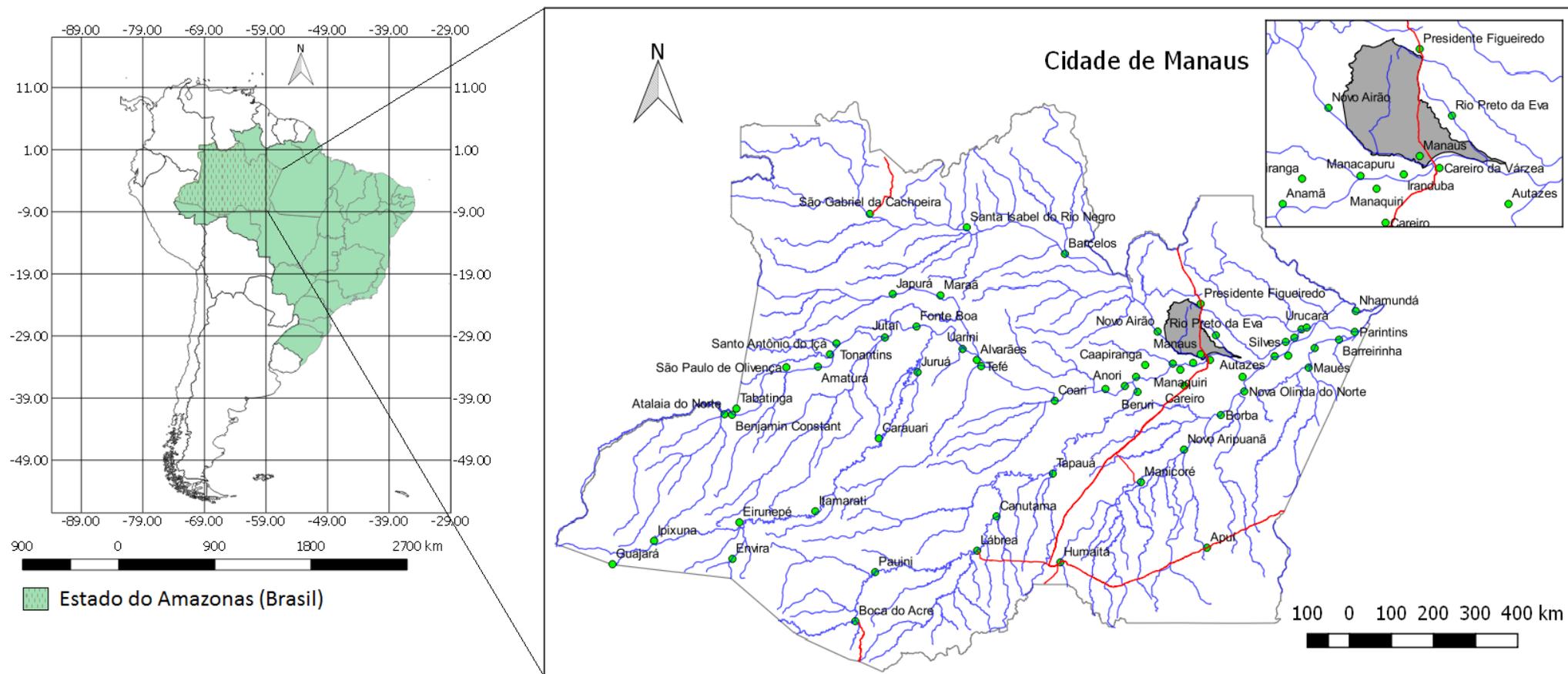


Figura 1.1 – Estado do Amazonas com 62 cidades, rios (azul) e estradas (vermelho). *Shapefiles* – arquivos com dados geoespaciais – obtidos através da ANA (2016), DNIT (2016), IBGE (2016a), IBGE (2016b) e FORESTGIS (2020), respectivamente, para hidroviás, rodovias, mapa do Estado do Amazonas, mapa do Brasil e mapa da América do Sul.

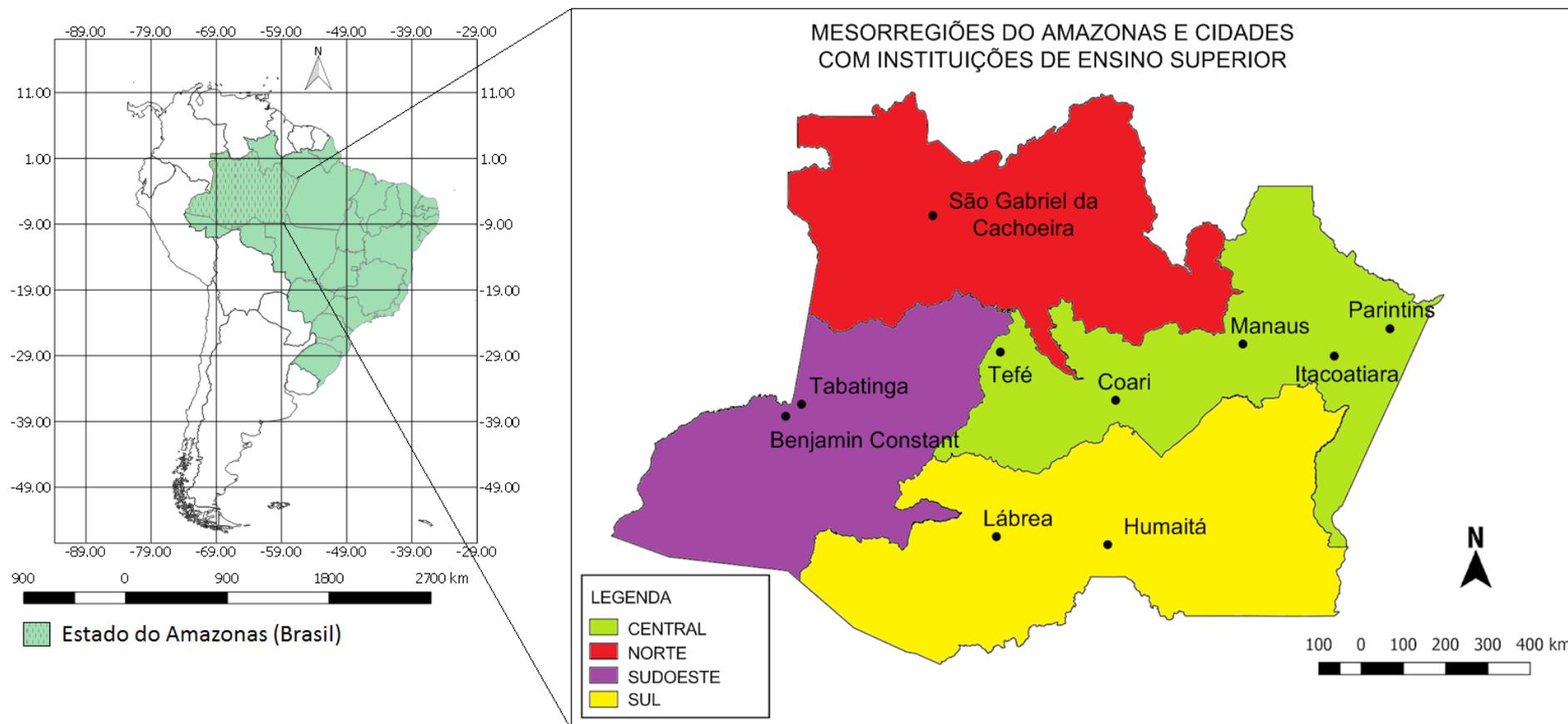


Figura 1.2 – Mapa do estado do Amazonas mostrando mesorregiões do estado e cidades com pelo menos uma unidade de ensino superior. *Shapefiles* – arquivos com dados geoespaciais – obtidos através do IBGE (2016a), IBGE (2016b), IBGE (2019) e FORESTGIS (2020), respectivamente, para o mapa do Estado do Amazonas, mapa do Brasil, mapa das mesorregiões do Estado do Amazonas e mapa da América do Sul.

No estado do Amazonas, as distâncias entre as localidades podem chegar a mais de 1.500 km, tais como os trechos Manaus–Atalaia do Norte (1.623 km), Manaus–Benjamin Constant (1.575 km), Manaus–Tabatinga (1.573 km), entre outros (ANTAQ, 2013). O acesso a essas cidades geralmente é realizado através da navegação por vias navegáveis interiores (linhas de cor azul na Figura 1.1). A rede fluvial (linhas azuis na Figura 1.1) do estado do Amazonas é a mais importante. Nas margens dos rios está localizada a maioria das cidades do estado do Amazonas. A rede ferroviária não existe e a rede rodoviária (linhas vermelhas na Figura 1.1) é muito limitada e pobremente mantida.

Metade dos alunos elegíveis (Alunos do Ensino Médio – AEM), em torno de 105.000 alunos para ingressar em uma Instituição de Ensino Superior (IES), está na capital do estado do Amazonas, Manaus, e a outra metade está distribuída em outras 61 cidades do interior do estado do Amazonas (IBGE, 2018). O transporte dos estudantes de uma cidade para a cidade mais próxima com uma IES é demorado e geralmente é feito pelo rio, que nem sempre é a forma de transporte mais modernizada.

Um aspecto importante a ser considerado são os longos tempos de viagens envolvidos no deslocamento dos alunos do interior para uma cidade com IES mais próxima. Uma alternativa para reduzir essa dificuldade de deslocamento é a localização de novas unidades de IES no interior.

Neste trabalho, estamos interessados na modelagem e solução de um problema específico de locação de unidades de IES em um conjunto de localidades. Esse conjunto de localidades são as cidades do estado do Amazonas. Especificamente, a localização de unidades de ensino em um conjunto de localidades é um problema discreto.

1.1 CONTEXTUALIZAÇÃO DO PROBLEMA

O estado do Amazonas possui dois sistemas de IES públicos e dezoito IES privadas. Um dos sistemas públicos está vinculado ao governo federal e o outro ao governo estadual. Como mostra a Figura 1.2, as instituições de ensino superior estão localizadas em apenas dez cidades das quatro mesorregiões. Nas mesorregiões Central, Norte, Sudoeste e Sul, existem 5, 1, 2 e 2 cidades com IES, respectivamente, denominadas na Figura 1.2.

Para cada mesorregião, a Tabela 1.1 mostra os seguintes campos:

- a) A quantidade de cidades;
- b) O número de alunos no ensino médio (AEM);

- c) A taxa de AEM, definida como o número de AEM em cada mesorregião dividido pelo número total de AEM no estado do Amazonas;
- d) O número de estudantes do ensino superior, AES;
- e) A taxa de AES, definida como o número de AES em cada mesorregião dividido pelo número total de AES no estado do Amazonas;
- f) O número de instituições de ensino superior (IES), e;
- g) A taxa de IES, definida como o número de IES em cada mesorregião dividido pelo número total de IES no estado do Amazonas.

Tabela 1.1 – Comparando dados populacionais e educacionais das mesorregiões do Amazonas, Brasil, (Semesp, 2018).

Mesorregião	Cidades	AEM	Taxa de AEM	AES	Taxa de AES **	Número de IES	Taxa de IES ***
Central	30	165525	0,783	133.645	0,949	20	0,800
Norte	6	6109	0,0289	623	0,004	1	0,040
Sudoeste	16	25651	0,121	4.044	0,029	2	0,080
Sul	10	13993	0,0662	2.596	0,018	2	0,080

*AEM/Total de AEM (211.278) **AES/Total de AES (140.908) ***IESs/Total de IESs (25)

As IES privadas estão localizadas apenas na capital. Na Tabela 1.2, mostramos uma comparação entre o valor médio do IDH e o número médio de AEM das 10 cidades com IES e de todas as 62 cidades. O índice IDH leva em consideração a expectativa de vida ao nascer, a média de anos de escolaridade, os anos esperados de escolaridade e a renda nacional bruta per capita (UNDP, 2018).

Tabela 1.2 – Número médio de AEM e IDH médio das cidades com IES e todas as cidades do estado do Amazonas, (UFAM, 2016; UEA, 2012).

Cidades	Número médio de AEM	IDH médio
10 cidades atuais com IES	14.170,40	0,619
Todas as cidades do Estado do Amazonas	3.407,70	0,565

Das Tabelas 1.1 e 1.2, podem ser tiradas as seguintes conclusões:

1. No estado do Amazonas, Brasil, os alunos do ensino superior (AES) estão concentrados na mesorregião Central, a mais populosa e onde Manaus, capital do estado, está localizada. No entanto, a taxa de AES dessa mesorregião em comparação com a mesma taxa de outras mesorregiões é maior que a taxa de AEM dessa mesorregião em comparação com a mesma taxa de outras mesorregiões, sugerindo que mais AEM dessa mesorregião têm acesso ao ensino superior;

2. O IDH médio das cidades com IES é maior que o IDH médio de todas as cidades, sugerindo que a atual distribuição das IESs favorece as cidades com um IDH mais alto.

Para equalizar a taxa de AES e a taxa de AEM de todas as mesorregiões, dentro do escopo deste trabalho, há duas opções:

1. Aumentar o número de estudantes nas instituições de ensino superior localizadas nas mesorregiões com menor taxa de estudantes, ou;
2. Aumentar o número de cidades com instituições de ensino superior nessas mesorregiões.

A primeira opção é uma decisão administrativa. A segunda opção pode ser alcançada expandindo o número de IES, buscando uma melhor distribuição delas. Isso pode ser feito usando modelos de localização discretos.

Especificamente, tomadores de decisão devem selecionar locais de um conjunto finito de locais candidatos (como em problemas de localização discretos) ou podem estabelecer facilidades em uma rede (por exemplo, rede de estradas). Além disso, os modelos de localização discretos não fazem suposições particulares sobre a demanda e a localização das facilidades. Simplesmente, são fornecidos os locais ou coordenadas dos nós de demanda e os locais candidatos. E as distâncias entre a demanda e os locais candidatos não precisam aderir a nenhuma fórmula específica (Daskin, 2013).

Neste trabalho, pretendemos avaliar soluções obtidas, para o problema de locação de expansão de IES no estado do Amazonas, com os seguintes modelos:

- 1) PMM (Modelo P-Mediana): Um que minimiza a distância média entre um nó de demanda (cidade sem IES) e um nó de facilidade mais próximo (cidade com IES);
- 2) PCM (Modelo P-Centro): Um que minimiza a distância máxima entre nós de demanda e um nó de facilidade mais próximo;
- 3) PDM (Modelo P-Dispersão): Um que maximiza a distância mínima entre pares de facilidades;

O modelo PCM (Tutunchi e Fathi, 2019) é um tipo de modelo baseado em cobertura. Esse tipo de modelo que minimiza a distância máxima entre nós de demanda e uma facilidade pressupõe que exista alguma distância ou tempo crítico de cobertura dentro do qual as demandas precisam ser atendidas. Tais modelos são normalmente usados no projeto de serviços de emergência, como o estabelecimento de unidades para combate a incêndios ou unidades de saúde para atendimento de emergência. O modelo PMM (Daskin e Maass, 2015)

é um tipo de modelo baseado em mediana que minimiza a distância média entre nós de demanda e facilidades atribuídas. Tais modelos são normalmente usados em contextos de planejamento de distribuição nos quais é essencial minimizar o custo total de transporte. O modelo PDM (Tutunchi e Fathi, 2019) maximiza a distância mínima entre pares de facilidades.

Esses três modelos são os modelos de localização discretos mais utilizados na literatura. Conforme declarado no último parágrafo, cada um otimiza um determinado critério. A escolha desses modelos visa avaliar qual deles alcança melhor desempenho em termos dos parâmetros avaliados neste trabalho: taxa de IES e taxa de AEM.

Para fornecer ao tomador de decisão várias opções, pretendemos também modelar o problema de expansão de maneira a fornecer as seguintes estratégias:

- a) Estratégia A: Uma expansão que favorece cidades com menor IDH;
- b) Estratégia B: Uma expansão que favorece as cidades mais populosas;
- c) Estratégia C: Uma expansão que favorece igualmente ambos os critérios (IDH e cidades mais populosas).

Especificamente, neste trabalho, as expansões consideradas serão para as instituições de ensino superior públicas, uma vez que os objetivos, restrições e variáveis podem ser diferentes se considerarmos o tipo da IES. E as estratégias mencionadas acima são mais adequadas para instituições públicas do que para as instituições privadas.

Os problemas de localização das facilidades (FLP) são geralmente empregados para resolver problemas públicos, comerciais, industriais e militares. Nesses problemas, os pontos de demanda de serviço devem ser atendidos por um número limitado de facilidades. Esses problemas possuem natureza estratégica, pois envolvem grandes somas de recursos de capital, com impactos econômicos, sociais e ambientais de longo prazo. São difíceis de solucionar, pelos menos de forma otimizada, uma vez que possuem alta complexidade computacional. São aplicações específicas, pois seus objetivos, restrições e variáveis são determinados pelo problema particular de locação em estudo. Segundo Drezner e Hamacher (2002), não existe um modelo geral que possa ser aplicado a todos os problemas de localização.

Localização ou locação de facilidades é a determinação do melhor local geográfico para estabelecer escolas, fábricas, hospitais, entre outros, com base em critérios predeterminados. Esses critérios podem ser, por exemplo, a proximidade com os clientes, fornecedores, redes de transporte, custo da mão de obra, custo de transporte, entre outros, e que tenham impacto sobre os objetivos que norteiam o processo de localização (Reid e Sanders, 2011).

De acordo com Fotheringham et al. (1995), modelos para locação-alocação consistem de um conjunto de técnicas para análise de localização. Eles definiram alocação como sendo o relacionamento ou atribuição da demanda (população) a uma facilidade estabelecida.

A teoria da complexidade computacional classifica o PMM como um problema NP-difícil não polinomial – problema NP-difícil (Kariv e Hakimi, 1979a). Métodos meta-heurísticos como Intercâmbio Guloso (GI), Vizinhança (N) e Intercâmbio (Whitaker, 1983), Relaxamento Semi-Lagrangeano (Beltran et al., 2004), Recozimento Simulado (SA) (Chiyoshi e Galvão, 2000), Busca Tabu (TS) (Salhi, 2002) e Algoritmo Genético (GA) (Alp et al., 2003; Satoglu et al., 2016) são geralmente usados para resolver problemas NP-difícil, para os quais um método de solução ideal não é conhecido ou não existe. Nas comparações feitas em (Salhi, 2002) e (Drezner et al., 2015), a heurística GA destaca-se como a melhor em termos de tempo e precisão da solução.

A implementação do GA, proposta por Silva et al. (2019), quando usada para resolver problemas de PMM da biblioteca de Pesquisa Operacional (OR) (Beasley et al., 1990), superou as implementações propostas em (Alp et al., 2003) e (Satoglu et al., 2016), em termos de tempo e precisão. Neste trabalho, adaptamos essa implementação do GA para resolver os três tipos de problemas de localização mencionados anteriormente.

1.2 OBJETIVO GERAL

Neste trabalho, estamos interessados em alcançar o seguinte objetivo principal:

- Elaborar uma metodologia para o problema de expansão de unidades de ensino superior levando em consideração critérios populacional e social, atendendo a um maior número de cidades.

1.3 OBJETIVOS ESPECÍFICOS

1. Avaliar modelos de locação de objetivo único tais como o p-mediana, o p-centro e o p-dispersão;
2. Propor métricas para avaliação das soluções de locação obtidas;
3. Elaborar uma modelagem do Algoritmo Genético eficiente para os modelos discretos usados neste trabalho;

1.4 DESENVOLVIMENTO DO TRABALHO

Este trabalho está organizado nos seguintes capítulos:

- i. Introdução (Capítulo I);
- ii. Fundamentos Teóricos (Capítulo II);
- iii. Revisão Bibliográfica (Capítulo III)
- iv. Metodologia (Capítulo IV);
- v. Resultados e Análise dos Resultados (Capítulo V);
- vi. Conclusão e Trabalhos Futuros (Capítulo VI);
- vii. Referências Bibliográficas.

No Capítulo II são abordados tópicos tais como: Modelagem mono-objetivo e multiobjetivo; Modelos de localização discreto p-mediana, p-cento e p-dispersão, e; Algoritmo Genético. Esses tópicos são necessários para descrever as técnicas utilizadas e os modelos elaborados neste trabalho.

No Capítulo III são mostrados: Resumo dos trabalhos relacionados na literatura no que diz respeito à locação de escolas; Visão Geral desses trabalhos com os modelos utilizados, técnicas de otimização e o que cada um desses trabalhos aperfeiçoa.

O Capítulo IV mostra os Materiais e Métodos elaborados neste trabalho. Na seção de Materiais é apresentado como foram obtidos os dados utilizados neste trabalho, tais como: as distâncias entre as cidades, as populações e os valores de IDH dos municípios. Na seção de Métodos são apresentadas as técnicas desenvolvidas/aplicadas nesta pesquisa, tais como: os três tipos de normalização dos parâmetros usados; a modelagem para o problema de expansão das cidades com IES; as métricas para avaliação das soluções de locação, e; a modelagem do Algoritmo Genético para os problemas de localização discretos.

O Capítulo V mostra os resultados para os três grupos de métricas obtidos para cada modelo de locação e configuração de parâmetros. São discutidos, nesse capítulo, se a combinação de modelos de locação com as estratégias de expansão obtém melhores valores de taxas para cada mesorregião do que a solução existente.

No Capítulo VI são apresentadas as principais conclusões deste trabalho, os objetivos alcançados, bem como sugestões para continuação e aprimoramentos.

2. FUNDAMENTOS TEÓRICOS

O Capítulo I delineou a principal contribuição desse trabalho, que versa sobre a otimização da alocação de unidades de ensino para a expansão do sistema superior no estado do Amazonas: a proposição de uma metodologia para expansão de unidades de ensino utilizando os modelos p-mediana, p-centro e p-dispersão. Além disso, todas as soluções de locação foram obtidas por meio do Algoritmo Genético.

Neste capítulo introduzimos conceitos teóricos relacionados aos tópicos acima citados. Os modelos de localização discretos p-mediana, p-centro e p-dispersão são apresentados com o intuito de mostrar suas funções objetivo, variáveis e restrições. Esses modelos são adaptados, posteriormente, no capítulo de Metodologia, para a modelagem do problema específico desta tese.

Embora seja recomendada a leitura deste capítulo, fica a critério do leitor prosseguir para os próximos capítulos e recorrer a este capítulo quando assim o desejar.

Estruturalmente, este Capítulo consiste de três seções. Na primeira, apresentamos os conceitos relacionados com a otimização mono-objetivo e multiobjetivo. Na segunda, abordamos o métodos de localização discretos p-mediana, p-centro e p-dispersão. Por último, apresentamos conceitos relacionados com o Algoritmo Genético.

2.1 OTIMIZAÇÃO MONO-OBJETIVO E MULTI-OBJETIVO

A otimização mono-objetivo tenta obter a melhor solução que usualmente é um mínimo ou máximo global, dependendo se o problema de otimização é de minimização ou maximização. No caso de múltiplos objetivos, pode não existir uma solução que seja a melhor (máximo ou mínimo global) com respeito aos objetivos. Tipicamente, nos problemas de otimização multiobjetivo, existe um conjunto de soluções que são superiores ao restante das soluções no espaço de busca quando todos os objetivos são considerados, mas são inferiores a outras soluções no espaço em um ou mais objetivos (Srinivas e Deb, 1994).

Em problemas de otimização multiobjetivo com objetivos conflitantes, no qual o ganho em um objetivo sacrifica os demais, não há solução simples e existe um número de

soluções que são ótimas. Nenhuma solução desse conjunto de soluções ótimo pode ser considerada a melhor entre elas, sem qualquer informação adicional. Nesse caso, várias soluções ótimas são importantes por causa de seus objetivos conflitantes. Essa é a principal diferença entre um problema de otimização multiobjetivo e mono-objetivo, cuja solução ótima é a única importante (Deb, 2001).

Matematicamente, o problema de otimização multiobjetivo é dado como segue (Chong e Zak, 2013):

$$\text{minimizar } \mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_\ell(x_1, x_2, \dots, x_n) \end{bmatrix}$$

Sujeito a $\mathbf{x} \in \Omega$,

Onde:

$\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^\ell$ e $\Omega \subset \mathbb{R}^n$. O conjunto de restrições Ω pode ter a forma]

ℓ é o número de funções objetivo;

f_1, \dots, f_ℓ são a ℓ funções objetivo;

$\Omega = \{\mathbf{x}: \mathbf{h}(\mathbf{x}) = \mathbf{0}, \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$, em que $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g: \mathbb{R}^n \rightarrow \mathbb{R}^p$, $m \leq n$.

Em geral, podemos ter três diferentes tipos de problemas de função multiobjetivo:

- Minimiza todas as funções objetivo;
- Maximiza todas as funções objetivo;
- Minimiza algumas e maximiza outras.

Usualmente, alguns desses problemas podem ser convertidos na forma de um problema de otimização equivalente. Segundo Chong e Zak (2013), um problema de otimização multiobjetivo pode ser convertido em um problema mono-objetivo por meio das seguintes técnicas:

- Método da soma ponderada. Nesse método, são dadas importâncias relativas, pesos, aos componentes individuais no vetor objetivo. Esses pesos são os coeficientes da combinação linear (ou seja, componentes \mathbf{c}):

$$f = \mathbf{c}^T \mathbf{f}(\mathbf{x})$$

- Método *minimax*. Nesse método, o problema de minimização multiobjetivo é convertido em um de minimização do máximo dos componentes,

$$f(\mathbf{x}) = \max\{f_1(x), \dots, f_\ell(x)\}$$

- Método p -norm. Nesse método, o problema multiobjetivo é convertido para um problema mono-objetivo, assumindo que os componentes do vetor objetivo são não negativo;

$$f(\mathbf{x}) = \|\mathbf{f}(\mathbf{x})\|_p$$

- Método de minimizar um dos componentes do vetor objetivo sujeito às restrições sobre os outros componentes. Por exemplo, dado o vetor de funções objetivo $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), \dots, f_\ell(\mathbf{x})]^T$, temos:

Minimizar:

$$f_1(\mathbf{x})$$

Sujeito a:

$$\begin{aligned} f_2(\mathbf{x}) &\leq b_2 \\ &\vdots \\ f_\ell(\mathbf{x}) &\leq b_\ell \end{aligned}$$

Onde b_2, \dots, b_ℓ são constantes que refletem satisfatoriamente valores, quando puderem ser determinados, para as funções objetivo f_2, \dots, f_ℓ , respectivamente.

Neste trabalho, utilizamos o método de conversão soma ponderada, com matriz de pesos $\mathbf{c} = [\alpha \ \beta \ \gamma]^T$, de um problema de otimização multiobjetivo para um problema mono-objetivo.

2.2 MODELO DE LOCALIZAÇÃO DISCRETO

Nos modelos de localização discretos, as facilidades podem ser estabelecidas em pontos discretos, bem como os nós de demanda tem lugar em locais discretos. As distâncias podem ser obtidas de forma mais precisa, incluindo os parâmetros necessários para uma descrição adequada da realidade. De fato, o conjunto completo de pares de locais nó de demanda-facilidade candidata é inteiramente conhecido, sendo a principal dificuldade prática, nesses modelos, os dados a serem coletados (Plastria, 2001).

Neste trabalho, utilizamos modelos discretos que levam em conta otimizar as distâncias nó de demanda-facilidade ou facilidade-facilidade. Entre esses modelos estão o p -mediana, o p -centro e o p -dispersão. O objetivo do modelo p -mediana é minimizar a distância total ponderada pela demanda para estabelecer um número predeterminado de facilidades (Turkoglu e Genevois, 2019). O modelo p -mediana também é conhecido como problema

minisum (Daskin, 2013). O problema p-centro busca minimizar a distância máxima entre o ponto de demanda e sua facilidade mais próxima, localizando um número predefinido de facilidades, é conhecido como problema *minimax* (Daskin, 2013). O p-dispersão (*maxmin*) busca estabelecer certo número de facilidades o mais longe possível umas das outras, ou seja, maximiza a distância mínima entre quaisquer pares de facilidades. Esse modelo é usado para localizar facilidades militares, dificultando a ação de inimigos, franquias de lojas, evitando o competição desnecessária, entre outros (Daskin, 2013).

2.2.1 Modelo p-mediana

Segundo Daskin e Maas (2015), a modelagem p-mediana pressupõe o estabelecimento de p facilidades e a minimização da distância total média entre os pontos de demanda e a facilidade mais próxima.

Para o modelo p-centro é definida a seguinte notação:

- d_{ij} é a distância entre o ponto de demanda i e o local candidato j ;

Os parâmetros utilizados nas expressões de (2.1) à (2.5) são:

- n é o número total de locais candidatos para as facilidades;
- a_i é a população da i -ésima comunidade, ou seja, o ponto de demanda i ;
- p é o número de facilidades a ser estabelecido.

A variável de decisão otimizada pelo método é definida como:

- $x_{ij} = \begin{cases} 1, & \text{se a zona de demanda } i \text{ for atribuída à facilidade } j \\ 0, & \text{caso contrário} \end{cases}$

A formulação mais conhecida para problema p-mediana foi proposta por ReVelle e Swain (1970) e é definida como:

Minimizar

$$\sum_{i=1}^n \sum_{j=1}^n a_i d_{ij} x_{ij} \quad (2.1)$$

Sujeito a:

$$\sum_{i=1}^n x_{ii} = p \quad (2.2)$$

$$\sum_{j=1}^n x_{ij} = 1, \quad i = 1, 2, \dots, n \quad (2.3)$$

$$x_{jj} \geq x_{ij}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, n, \quad i \neq j \quad (2.4)$$

$$x_{ij} \in \{0, 1\}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, n \quad (2.5)$$

A função objetivo (2.1) minimiza a distância total ponderada (ponderação feita pela população) entre os pontos de demanda e a facilidade. A restrição (2.2) diz que exatamente p facilidades devem ser alocadas. A restrição (2.3) requer que cada ponto de demanda i seja atribuído a apenas uma única facilidade no local j . A restrição (2.4) determina que o ponto de demanda seja atribuído a um ponto com facilidade alocada. A restrição (2.5) evidencia a natureza binária das variáveis de decisão.

Com relação à complexidade computacional, Kariv e Hakimi (1979a) mostraram que o problema p -mediana é NP-difícil. Eles propuseram um algoritmo com tempo de execução $O(n^2p^2)$, para $n > p > 1$, com o objetivo de encontrar p medianas em uma rede com estrutura em árvore.

2.2.2 Modelo p -centro

Segundo Daskin (2013), o modelo p -centro pressupõe o estabelecimento de p facilidades de forma a minimizar a distância máxima entre o local de demanda e a facilidade mais próxima do nó de demanda.

Para o modelo p -centro é definida a seguinte notação:

- W é a distância máxima entre o nó de demanda e a facilidade mais próxima;

Os parâmetros utilizados nas expressões de (2.7) à (2.12) são:

- I é o conjunto de nós de demanda;
- J é o conjunto de localidades candidatas;
- d_{ij} é a distância do nó de $i \in I$ para a localidade candidata $j \in J$ para a facilidade;
- h_i é a demanda no nó $i \in I$;
- p é o número de facilidades para estabelecer;

As variáveis de decisão:

$$X_j = \begin{cases} 1, & \text{se estabelecemos uma facilidade no local candidato } j \in J \\ 0, & \text{caso contrário} \end{cases}$$

- Y_{ij} é uma parte da demanda no nó $i \in I$ atendida pela facilidade no nó $j \in J$;

A formulação matemática modelo p-centro foi proposta por Hakimi (1964, 1965), definida como:

Minimizar

$$W \quad (2.6)$$

Sujeito a:

$$\sum_{j \in J} Y_{ij} = 1 \quad (2.7)$$

$$\sum_{j \in J} X_j = p \quad (2.8)$$

$$Y_{ij} \leq X_j \quad \forall i \in I; j \in J \quad (2.9)$$

$$W \geq \sum_{j \in J} d_{ij} Y_{ij} \quad \forall i \in I \quad (2.10)$$

$$X_j \in \{0, 1\} \quad \forall j \in J \quad (2.11)$$

$$Y_{ij} \geq 0 \quad \forall i \in I; j \in J \quad (2.12)$$

A função objetivo (2.6) minimiza a distância máxima entre um nó de demanda e a facilidade mais próxima ao nó de demanda. As restrições (2.7) afirmam que, para todos os nós $i \in I$, toda a demanda $i \in I$ deve ser atribuída para uma facilidade $j \in J$. A restrição (2.8) estipula que p facilidades estejam estabelecidas. As restrições (2.9) afirmam que demandas no nó $i \in I$ só podem ser atribuídas a uma facilidade localizada no nó $j \in J$. As restrições (2.10) declaram que a distância máxima (W) entre um nó de demanda e a facilidade mais próxima ao nó deve ser maior que a distância entre qualquer nó de demanda $i \in I$ e a facilidade $j \in J$, à qual está atribuído. As restrições (2.11) e (2.12) são as restrições de integralidade e não negatividade das variáveis, respectivamente (Daskin, 2013).

Em alguns casos, queremos considerar a distância ponderada pela demanda h_i . Nesse sentido, a restrição (2.10) pode ser substituída por:

$$W \geq \sum_{j \in J} h_i d_{ij} X_j \quad \forall i \in I$$

Kariv e Hakimi (1979b) mostraram que o problema de encontrar p-centro vértice ou p-centro absoluto de uma rede – com vértice ponderado ou não ponderado – é NP-difícil, para $1 < p < n$, mesmo no caso em que a rede tenha uma estrutura simples. Por exemplo, um gráfico planar de grau máximo de vértice igual a 3. Eles obtiveram um algoritmo $O(n^2 \lg n)$

para encontrar p-centro (vértice ou absoluto) para qualquer $1 < p < n$. Em que n é o número de vértices.

No problema p-centro vértice, as facilidades podem ser estabelecidas somente nos nós da rede. Nos problemas p-centro absoluto, as facilidades podem ser estabelecidas nos nós e nas interligações desses nós.

2.2.3 Modelo p-dispersão

O modelo p-dispersão visa o estabelecimento de p facilidades de forma a maximizar a distância mínima de separação entre qualquer par de facilidades. Segundo Erkut (1990), o problema p-dispersão é NP-Completo. A formulação desse modelo foi proposta por Kuby (1987).

Para o modelo p-dispersão é definida a seguinte notação:

- D é a distância mínima de separação entre quaisquer pares de facilidades;

Os parâmetros utilizados nas expressões de (2.14) à (2.16) são:

- J é o conjunto de localidades;
- d_{ij} é a distância entre as localidades candidatas $i \in J$ e $j \in J$;
- M é um número muito grande, tal que $M \geq \max_{i,j \in J} \{d_{ij}\}$;
- p é o número de facilidades para estabelecer;

As variáveis de decisão são:

$$X_j = \begin{cases} 1, & \text{se nós localizamos uma facilidade no local candidato } j \in J \\ 0, & \text{caso contrário;} \end{cases}$$

A formulação matemática para esse modelo é dada a seguir (Kuby, 1987):

Maximizar

$$D \tag{2.13}$$

Sujeito a:

$$\sum_{i \in J} X_j = p \tag{2.14}$$

$$D \leq d_{ij} + (M - d_{ij})(1 - X_i) + (M - d_{ij})(1 - X_j), \quad \forall j \in J; i \in J, i < j \tag{2.15}$$

$$X_j \in \{0,1\} \forall j \in J \tag{2.16}$$

A função objetivo (2.13) maximiza D , a distância mínima entre as facilidades. A restrição (2.14) estipula que estabelecemos exatamente p facilidades. As restrições (2.15) definem a distância mínima entre facilidades em termos das localizações das facilidades selecionadas. A restrição (2.15) pode ser reorganizada para obter

$$D + (M - d_{ij})X_i + (M - d_{ij})X_j \leq 2M - d_{ij}$$

Caso $X_i = X_j = 1$, então a restrição (2.15) reduz-se a $D \leq d_{ij}$. Em outras palavras, se estabelecemos em ambos locais candidatos $i \in J$ e $j \in J$, então a distância máxima entre qualquer par de facilidades não pode exceder a distância entre os locais $i \in J$ e $j \in J$. Caso apenas um dos dois locais seja escolhido (ou seja, $X_i = 0, X_j = 1$ ou $X_i = 1, X_j = 0$), então a restrição (2.15) torna-se $D \leq M$. Assim, queremos que M seja suficientemente grande de forma que a restrição não seja obrigatória. Consequentemente, define-se $M = \max_{i,j \in J} \{d_{ij}\}$. Finalmente, se $X_i = 0$ e $X_j = 0$, a restrição (2.15) torna-se $D \leq 2M - d_{ij}$. Assim, não é desejável que essa restrição seja obrigatória e definindo $M = \max_{i,j \in J} \{d_{ij}\}$ irá garantir que o valor de D não seja restrito. As restrições (2.16) são simples restrições de integridade (Daskin, 2013).

O problema p-dispersão é NP-Difícil (Paschos, 2013). Ravi et al. (1991) obtiveram um algoritmo $O(n^2)$ para solucionar o problema de dispersão unidimensional.

2.3 ALGORITMO GENÉTICO

No problema tratado neste trabalho, a solução combinatória para localização de unidades de ensino superior nas cidades do estado do Amazonas é altamente complexa em termos de tempo e recursos computacionais e inviáveis para ser solucionada por programação inteira com *branch and bound*, ou seja, por um método exato. Isso significa que precisamos empregar métodos heurísticos, tais como o Algoritmo Genético (Xavier et al., 2016).

Heurísticas são métodos que não garantem que a solução ótima seja encontrada, mas encontram soluções aproximadas. Entre esses métodos, destaca-se o Algoritmo Genético.

2.3.1 Algoritmo Genético

Inicialmente, Algoritmo Genético (GA) foi introduzido por John Henry Holland em 1975 na obra intitulada “Adaptação em Sistemas Artificiais e Naturais”. É um método baseado no princípio da genética e seleção natural vindos da teoria da evolução.

Segundo Rao (2009), as principais características dos Algoritmos Genéticos são: a) utilizam uma população de indivíduos ou cromossomos; b) utilizam somente valores da função objetivo; c) os indivíduos podem ser representados com codificação binária, decimal, caracteres, etc. O que torna esse método de busca ideal para ser aplicado em problemas de programação inteira ou discreto; d) o valor da função objetivo de um indivíduo representa a aptidão na genética natural; e) em cada nova geração, um novo conjunto de indivíduos é produzido, utilizando a seleção aleatória de pais e o cruzamento entre os mesmos para gerar filhos; f) possuem elementos comuns, tais como população de indivíduos, seleção de acordo com a aptidão, cruzamento entre indivíduos selecionados para gerar novos indivíduos e mutação aleatória de novos indivíduos.

Segundo Zapfel et al. (2010), os componentes básicos em GA são: população, inicialização, avaliação, seleção, recombinação e mutação. A seguir cada um desses termos é descrito de forma mais detalhada.

2.3.2 População

A população em algoritmos genéticos é um conjunto de soluções. Uma solução dessa população é denominada de indivíduo ou cromossomo. Os cromossomos são compostos por genes e os valores de um gene são denominados alelos. Um gene é uma propriedade de uma solução, que pode assumir valores de um domínio predefinido. A posição de um gene é chamada de *locus*. Caso um gene tome certo valor, o mesmo é denominado alelo. A Figura 2.1 ilustra uma população de indivíduos (soluções).

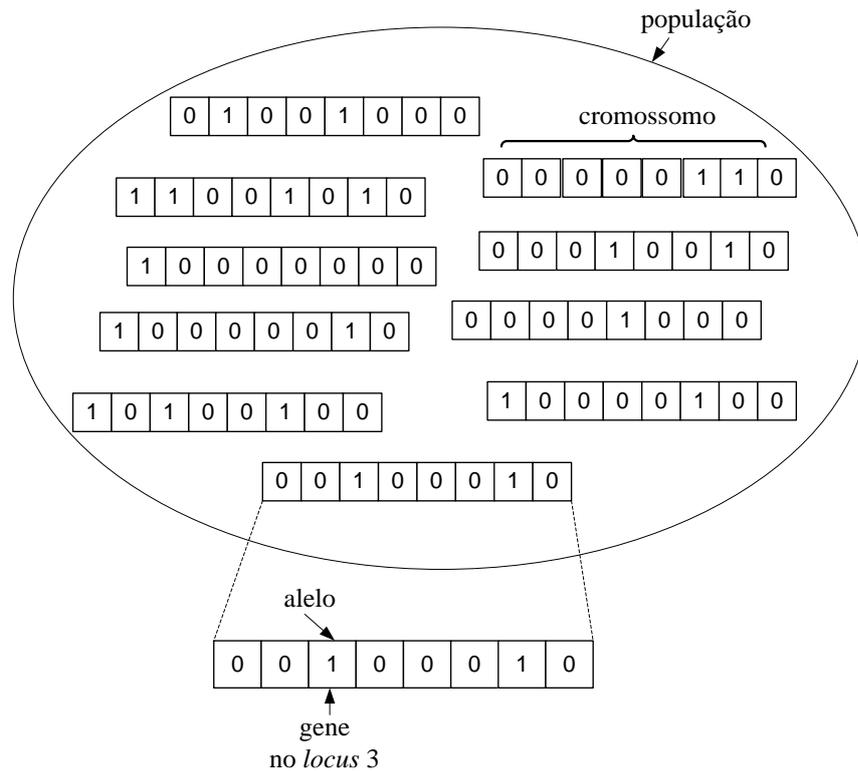


Figura 2.1 – Exemplo de população de indivíduos.
Fonte: Adaptado de Zapfel et al. (2010)

É importante que no início do processo de busca, a população contenha a maior diversidade possível. À medida que a busca progride, a qualidade das soluções na população torna-se melhor até que uma solução aproximada seja encontrada e a população converge para essa solução.

Uma escolha importante é o tamanho da população, uma vez que, caso a mesma seja muito pequena, há risco de convergência prematura. E para uma população com grande número de indivíduos, há um aumento considerável do esforço computacional (Sastry et al., 2005). Não há um número ideal para o tamanho da população. Geralmente, inicia-se com uma população pequena, digamos de tamanho 50 e que vai sendo aumentada, de acordo com a natureza do problema em questão.

2.3.3 Inicialização

Nessa etapa são gerados os cromossomos da população, para a primeira geração. Os cromossomos gerados possuem a mesma codificação, ou seja, decimal, binária, caracteres, etc. Outro objetivo dessa etapa, é produzir indivíduos diferentes entre si, com distintas aptidões. A propriedade de amostragem aleatória é bastante utilizada nesse método de

inicialização, mas nem sempre é aplicado a problemas altamente restritivos, ou seja, problemas em que nem toda representação da solução é uma solução viável. Esses aspectos devem ser levados em consideração quando for projetado o método de inicialização (Zapfel et al., 2010).

2.3.4 Avaliação

Para cada solução candidata é necessário verificar a qualidade da mesma. Essa verificação é feita através da função de avaliação conhecida como aptidão em GA. De forma geral, essa função pode corresponder à função objetivo. O valor de aptidão é necessário para seleção e substituição, pois as soluções com melhor aptidão são preferidas na etapa de seleção. Uma vez que a aptidão das soluções conduz o processo de busca, ela fornece informação sobre a qualidade comparativa da solução, (Zapfel et al., 2010).

2.3.5 Seleção

É o operador que escolhe quais soluções são usadas para recombinação. Tipicamente, essa decisão é baseada na aptidão da solução. Frequentemente, o valor de aptidão reflete a qualidade da solução e é usado para seleção, comparando-o aos dos outros indivíduos na população. A melhor solução é, em termos da otimização, a que possui o melhor valor de aptidão e, provavelmente, será escolhida para recombinação.

No processo de seleção, normalmente, os melhores indivíduos que direcionam a busca são escolhidos para formar o *mating pool* (agrupamento de acasalamento). No entanto, diferentes métodos de seleção foram desenvolvidos nos últimos anos, tais como Seleção através da Roleta, Rank Linear, Torneio, etc. A seguir são mostrados em mais detalhes os dois primeiros, (Zapfel et al., 2010).

2.3.5.1 Seleção através da Roleta

É um método de seleção simples. Nele, as soluções são selecionadas de acordo com sua aptidão proporcional. Por exemplo, vamos considerar um problema de maximização com as seguintes soluções e seus valores de aptidão: $f(s_1) = 15$; $f(s_2) = 10$; $f(s_3) = 5$, $f(s_4) = 30$, $f(s_5) = 18$, $f(s_6) = 22$.

Fazendo uma comparação entre s_3 e s_4 , percebemos que a aptidão do primeiro é seis vezes menor do que a do segundo. Podemos afirmar que a probabilidade de selecionar s_3 é menor do que a de selecionar s_4 . A equação (2.17) define o cálculo da probabilidade para uma solução s_i :

$$p(s_i) = \frac{f(s_i)}{\sum_{k=1}^n f(s_k)} \quad (2.17)$$

Em que:

- $p(s_i)$ é a probabilidade de seleção para a solução s_i ;
- $f(s_i)$ é o valor de aptidão para a solução s_i ;
- $\sum_{k=1}^n f(s_k)$ é o somatório dos valores de aptidão de cada solução s_i ;
- n é o número de soluções considerado.

Calculando as probabilidades para s_1, s_2, s_3, s_4, s_5 e s_6 , temos: $p(s_1) = 0,1$; $p(s_2) = 0,10$; $p(s_3) = 0,05$; $p(s_4) = 0,30$; $p(s_5) = 0,18$; $p(s_6) = 0,22$. A soma dessas probabilidades é igual a um e as mesmas são ilustradas na Figura 2.2.

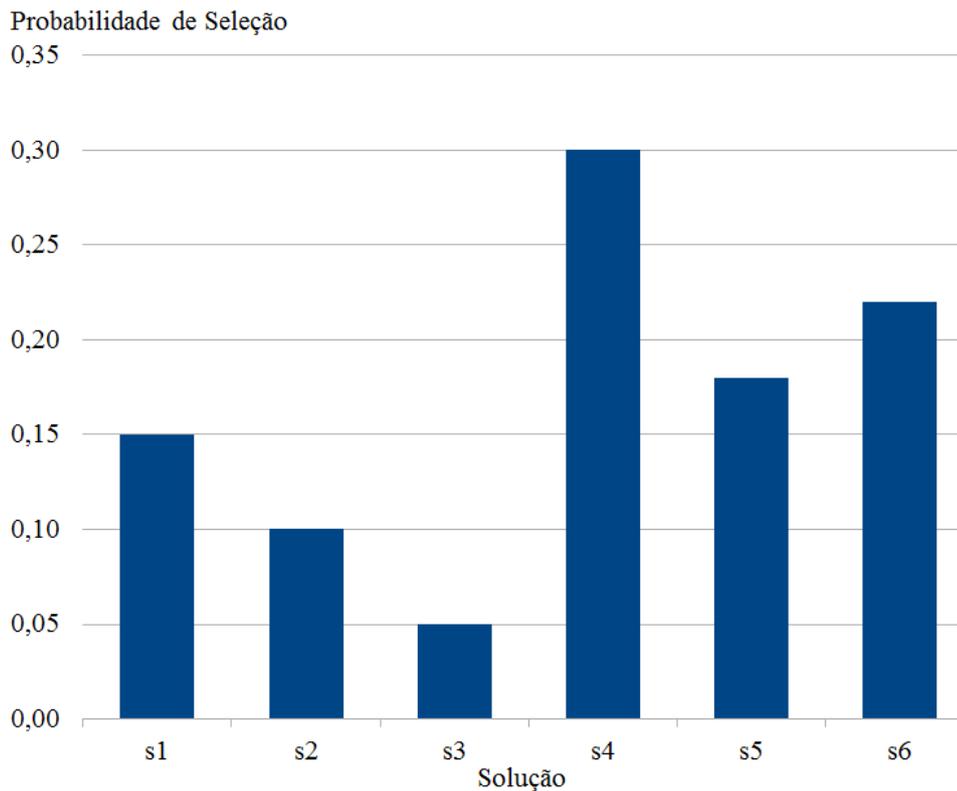


Figura 2.2 – Probabilidades de seleção da solução.
Fonte: Adaptado de Zapfel et al. (2010).

Esse esquema de seleção possui algumas falhas, que mostraremos a seguir. Considere as seguintes seis soluções e os valores de aptidão delas: $f(s_1) = 5$; $f(s_2) = 15$; $f(s_3) = 110$; $f(s_4) = 10$; $f(s_5) = 12$; $f(s_6) = 20$.

A Figura 2.3, com as probabilidades dessas soluções, ilustra um problema proeminente no uso da seleção através da roleta. Nesse método, soluções com um grande valor de aptidão dominam a população, uma vez que sua probabilidade é muito alta e a chance de selecionar outros indivíduos é mínima. Isso pode facilmente levar a convergência prematura. Repare que a aptidão da solução s_3 é muito maior do que as das demais soluções.

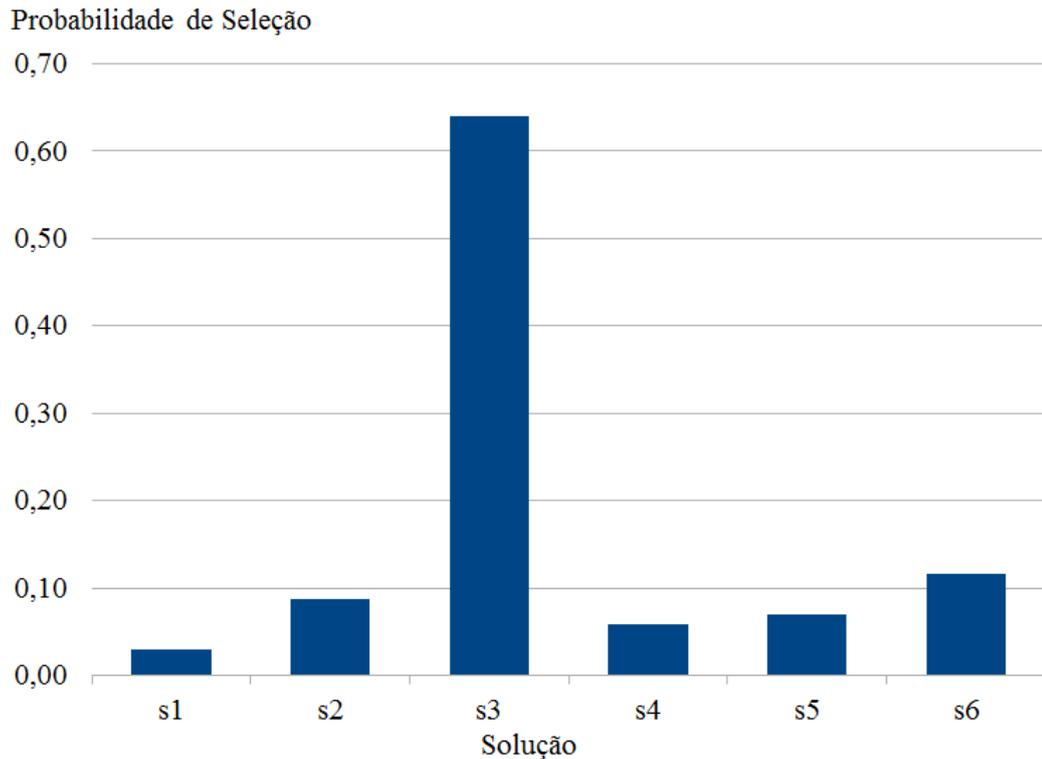


Figura 2.3 – Uma solução domina a população com seu valor de aptidão.
Fonte: Adaptado de Zapfel et al. (2010).

Na próxima subseção, veremos uma das formas de prevenir a convergência prematura ocasionada pelo uso da função de aptidão no cálculo da probabilidade de seleção.

2.3.5.2 Seleção através do *Rank* Linear

Para sobrepor os problemas de seleção proporcional, o método de seleção através do *rank* linear pode ser usado. Esse esquema de seleção usa o *rank* da solução comparada a outras soluções em vez de usar o valor de aptidão diretamente, (Zapfel et al., 2010).

Vamos utilizar o exemplo da Figura 2.3, da seleção através da roleta. O primeiro passo é ordenar os valores de aptidão de forma crescente (para um problema de maximização), o que leva a: $(s_1, s_4, s_5, s_2, s_6$ e $s_3)$. Aqui, usamos o *rank* da solução em vez do valor de aptidão. Por exemplo, a solução s_1 tem *rank* 1 e a solução s_3 tem *rank* 6. Em seguida, calculamos a probabilidade de seleção da mesma forma usada para os valores de aptidão, equação (2.18):

$$p(s_i) = \frac{\text{rank}(s_i)}{\sum_{k=1}^n \text{rank}(s_k)} \quad (2.18)$$

A Tabela 2.1 mostra as soluções, aptidão, *rank* e probabilidades de seleção das soluções s_1, s_2, s_3, s_4, s_5 e s_6 .

Tabela 2.1 – Exemplo de seleção *rank* linear para um problema de maximização.

s_i	$f(s_i)$	$\text{rank}(s_i)$	$p(s_i)$
s_1	5	1	0,05
s_4	10	2	0,10
s_5	12	3	0,14
s_2	15	4	0,19
s_6	20	5	0,24
s_3	110	6	0,29
Total	172	21	1

Um gráfico de probabilidade de seleção da Tabela 2.1 é mostrado na Figura 2.4.

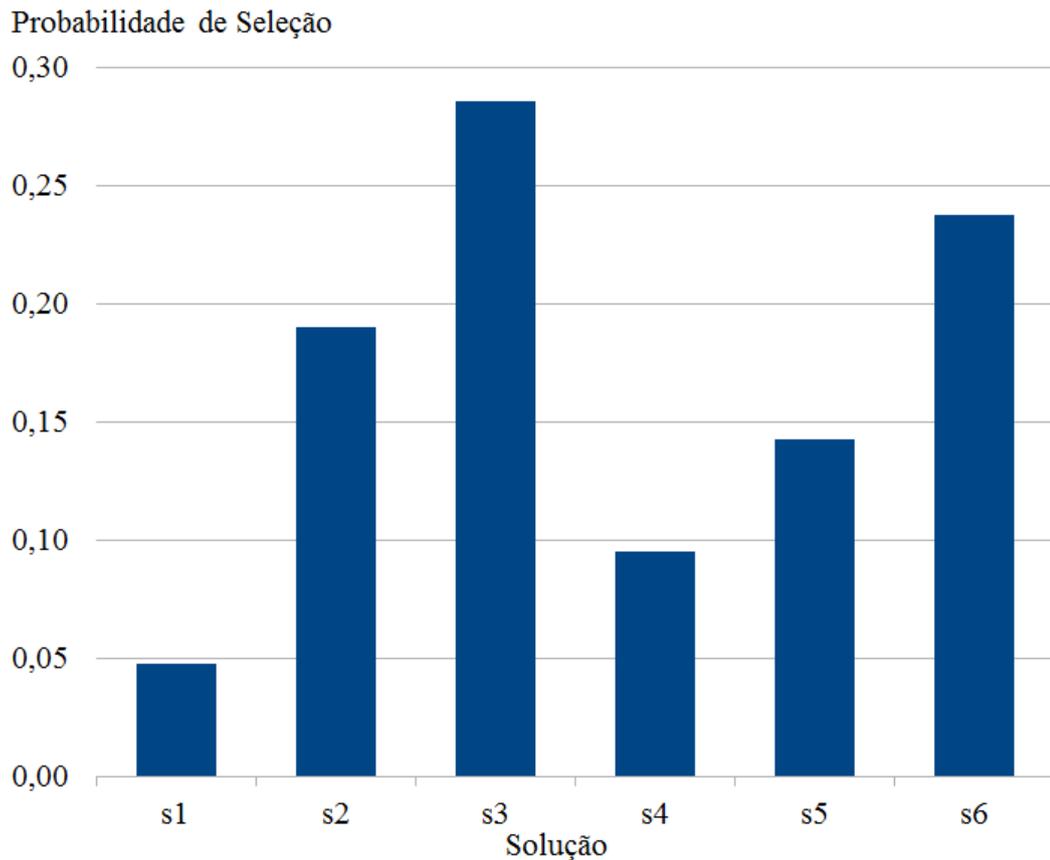


Figura 2.4 – Solução baseada na probabilidade de seleção através do *rank*.
Fonte: Adaptado de Zapfel et al. (2010).

Como pode ser visto na Figura 2.4, o esquema de seleção através do *rank* faz com que soluções dominantes tenham sua influência reduzida. Veja que a probabilidade de s_3 foi reduzida de 0,64, na Figura 2.3, para 0,29, na Figura 2.4.

2.3.6 Recombinação

É o processo de combinar genes (propriedades) de soluções selecionadas para obter uma nova solução. Soluções que são recombinadas são chamadas de pais e as soluções resultantes são denominadas de filhos. O operador usado no processo de recombinação é conhecido como cruzamento. Esse operador escolhe aleatoriamente um ou mais pontos de corte e troca as subsequências antes e depois desse ponto entre dois cromossomos pais para gerar dois novos filhos. É esperada que a qualidade da nova solução fosse melhor do que a da progenitora. Podemos dizer que o cruzamento é um operador m -ponto, (Zapfel et al., 2010).

Considere que os indivíduos 10101010 e 00001011 cortados a partir do terceiro *locus*, de cada um, formem dois novos indivíduos 10001011 e 00101010, conforme mostrado na Figura 2.5.

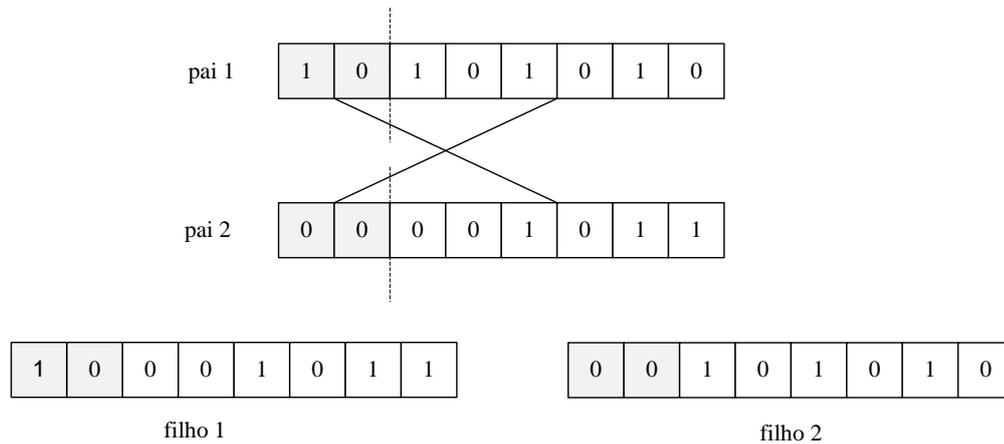


Figura 2.5 – Exemplo de um cruzamento 1-ponto.
Fonte: Adaptado de Zapfel et al. (2010).

O operador de cruzamento imita a recombinação biológica entre dois organismos haploides simples. A Figura 2.6 ilustra o operador de cruzamento 2-pontos, em que os pontos de corte foram aplicados nos *locus* 3 e 6. Nesse caso, alternam-se entre os dois pais as subsequências para formar dois filhos.

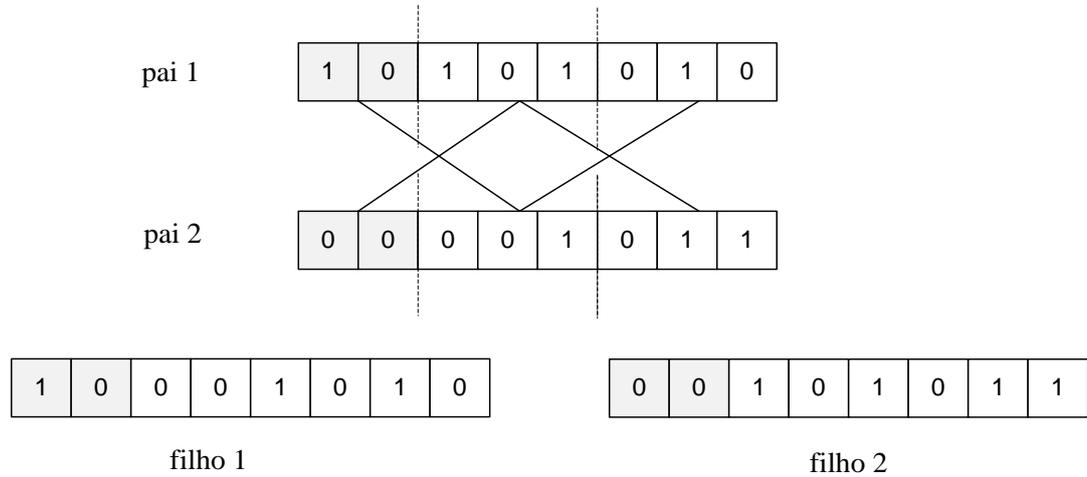


Figura 2.6 – Exemplo de um cruzamento 2-pontos.
Fonte: Adaptado de Zapfel et al. (2010).

Outra forma é o cruzamento uniforme que seleciona ao acaso genes das soluções pais. Para isso, usa-se uma máscara de cruzamento que indica a propriedade ou gene selecionado de cada pai. A Figura 2.7 exemplifica esse tipo de cruzamento. Uma entrada 0 na máscara de cruzamento significa que a propriedade correspondente ao pai 1 é usada. Por outro lado, uma entrada 1 indica que a propriedade do pai 2 é utilizada.

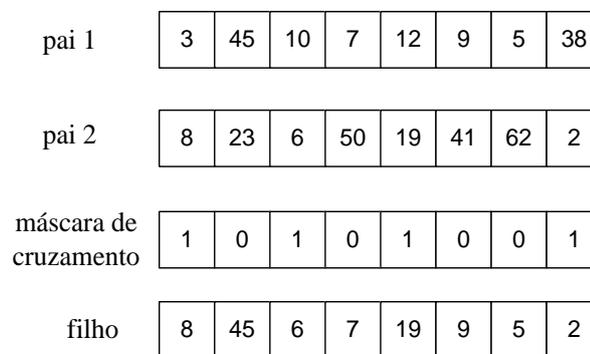


Figura 2.7 – Exemplo de cruzamento uniforme.
Fonte: Adaptado de Zapfel et al. (2010).

Conforme vimos, a recombinação tipicamente utiliza genes de soluções existentes e não produz soluções com novas informações. Essa perda de informação pode levar a convergência prematura. Esse é o motivo pelo qual existe o operador de mutação.

2.3.7 Mutação

A mutação é operador secundário que provoca uma pequena perturbação aleatória na solução candidata, com baixa probabilidade de ocorrência. Sastry et al. (2005) propuseram uma taxa de mutação experimental de 0,05. O objetivo do operador de mutação é recuperar informação que foi perdida durante a busca, ou seja, deve impedir a convergência prematura, Zapfel et al. (2010).

O operador de mutação troca de forma aleatória algum alelo no cromossomo. Por exemplo, o cromossomo 10000010 deve sofrer mutação em sua segunda posição para obter o cromossomo 11000010, conforme mostrado na Figura 2.8 A mutação pode ocorrer em qualquer posição com uma probabilidade muito pequena.

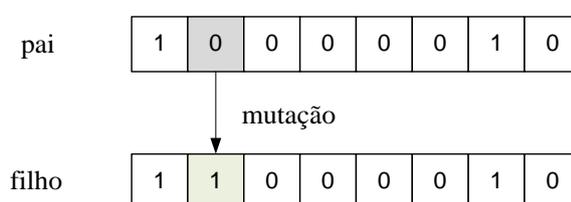


Figura 2.8 – Exemplo de operador de mutação.
Fonte: Adaptado de Mitchel (1999).

2.3.8 Codificação dos cromossomos ou indivíduos

De acordo com Mitchel (1999), a forma na qual a solução candidata é codificada constitui-se um fator central no sucesso do Algoritmo Genético. A maioria dos cromossomos, em GA, são codificados como *strings* binárias, contudo, a codificação com múltiplos caracteres e valores decimais também é utilizada.

No caso específico para o nosso problema, optamos por modelar cada cromossomo com valores (alelos) decimais, em que cada *locus* representa um identificador vinculado a uma cidade.

2.4 CONSIDERAÇÕES

Neste Capítulo foram mostrados os fundamentos teóricos das técnicas utilizadas neste trabalho. Em pontos específicos foram feitos alguns comentários sobre como determinado conceito, técnica ou método se encaixa no escopo deste trabalho.

Mostramos ainda o método metaheurístico Algoritmo Genético, bem como algumas particularidades dos mesmos em relação ao problema a ser tratado, qual seja, o conjunto completo de cidades.

Ressaltamos que, para um melhor entendimento deste tópico, foram apresentados alguns exemplos tais como nas diversas etapas do método Algoritmo Genético. Evidentemente, o aprofundamento dos tópicos abordados pode ser realizado a partir da consulta às referências indicadas neste capítulo.

No próximo capítulo, faremos uma revisão dos principais trabalhos da literatura referentes à localização de escolas nos últimos dez anos, no que tange aos seus propósitos, modelos e métodos de resolução com único objetivo e múltiplos objetivos.

3. REVISÃO BIBLIOGRÁFICA

Neste capítulo serão mostrados alguns trabalhos na literatura relacionados com a localização de escolas ou que utilizaram múltiplos modelos discretos para locação de facilidades. A intenção do capítulo é mostrar ao leitor abordagens utilizadas pelos trabalhos publicados na literatura na área de locação/alocação de unidades de ensino. Nesse contexto, para cada trabalho revisado, destacam-se alguns pontos importantes, tais como os objetivos a serem otimizados e os modelos/técnicas empregados na otimização. Após a apresentação dessa revisão, ficará clara a diferença entre o trabalho ora proposto e os trabalhos anteriormente publicados na literatura sob o mesmo tema. De antemão, a principal diferença e contribuição advém do fato de que o nosso trabalho é o único a propor a modelagem de um problema de locação levando em consideração critérios populacional e social para expansão de unidades de ensino superior.

Estruturalmente, esse capítulo foi dividido nas seguintes seções: a) Trabalhos relacionados, onde é apresentado um resumo das principais características dos trabalhos selecionados; b) Considerações e Contribuição.

Na seção Trabalhos relacionados introduz-se o que cada artigo revisado se propõe a fazer, incluindo: contexto e os objetivos; a técnica de otimização empregada para solução do problema e resultados alcançados. Na seção Considerações e Contribuição, são elencadas as conclusões das características em comum sobre os trabalhos revisados e no que o trabalho proposto nesta tese se diferencia.

3.1 TRABALHOS RELACIONADOS

A Tabela 3.1 mostra um resumo das principais características dos trabalhos publicados na literatura sobre localização/alocação de escolas analisados nesta tese, tais como: aplicação, que se refere ao contexto de aplicação do trabalho (região, município, etc.); modelo, que diz respeito à modelagem utilizada para representação matemática do problema; técnica de otimização, que diz respeito à técnica empregada para obtenção da solução (métodos heurísticos, programação linear, etc.); propósito, que diz respeito à tomada de decisão sobre a localização, e; objetivos a serem otimizados;

Tabela 3.1 – Resumo dos artigos selecionados.

Referência	Aplicação	Modelos e características	Técnica de Solução	Decisão/ Propósito	Objetivo único/ Multiobjetivo
Teixeira et al. (2007)	Rede de escolas secundárias em Coimbra – Portugal, com 11 escolas secundárias e 43 centros populacionais.	Uma variante do modelo p-mediana, abertura/fechamento de escolas	Otimizador Comercial	Localização no centro/ Objetiva maximizar a acessibilidade de estudantes à escola, com restrições às capacidades de ocupação mínima e máxima das escolas.	Abordagem com objetivo único: minimizar o deslocamento aluno-escola, com a população como peso.
Teixeira e Antunes (2008)	Rede de escolas primárias da cidade de Coimbra (Portugal), com 25 escolas do primeiro 14 escolas do segundo ciclo e 68 centros populacionais.	Uma variante do modelo p-mediana, modelo capacitado	Otimizador Comercial	Localização no centro/ Objetiva maximizar a acessibilidade de estudantes à escola, com restrições às capacidades de ocupação mínima e máxima das escolas e níveis da demanda.	Abordagem com objetivo único: minimizar o deslocamento aluno-escola, com a demanda como peso.
Gac et al. (2009)	Rede com escolas privadas, subsidiadas e públicas secundárias, na cidade de Santiago, Chile, com 6 grandes zonas populacionais.	Modelo linear determinístico	Otimizador Comercial	Localização no centro e atribuição de alunos/ Objetiva a determinação de alocação ótima de diferentes tipos de escolas às zonas de maneira a maximizar os benefícios sociais do sistema de educação.	Abordagem multiobjetivo: maximiza a utilidade do indivíduo e lucro escolar.
Muller et al. (2009)	Redes de escolas secundárias na cidade de Dresden, Alemanha, com 26 escolas e 64 centros populacionais.	Modelo de alocação probabilístico, modelo capacitado, abertura/fechamento de escolas.	Otimizador Comercial	Localização de centros e atribuição de alunos/ Esse artigo propõe uma abordagem para planejamento e localização de escola multiperíodo em áreas urbanas.	Abordagem multiobjetivo: minimiza o custo de uma rede escolar e maximiza a utilidade do aluno.

continua

Tabela 3.1 – Resumo dos artigos selecionados.

continuação

Referência	Aplicação	Modelos e características	Técnica de Solução	Decisão/ Propósito	Objetivo único/ Multiobjetivo
Ndiaye et al. (2012)	Rede de escolas secundárias no subúrbio de Dakar, Senegal, com 15 centros populacionais.	Modelo p-mediana	Otimizador Comercial	Localização no centro/ Este artigo apresenta a localização ótima de escolas.	Abordagem com objetivo único: minimiza o deslocamento aluno-escola, com o centro populacional como peso.
Bruno et al. (2014)	Rede de escolas de ensino pré-primário, primário e secundário, na cidade de Nápoles, Itália, com 100.000 habitantes e 29 escolas.	Modelo p-mediana, clusterização, agrupamento de escolas.	Otimizador Comercial	Clusterização de centros/ Nesse estudo é apresentado um modelo de localização que objetiva descrever e solucionar o problema de agrupamento de escolas definido no caso Italiano.	Abordagem com objetivo único: minimiza o deslocamento aluno-escola, com a população como peso.
Menezes e Pizzolato (2014)	Rede de escolas na Área 5 da região de Guaratiba, Rio de Janeiro, Brasil, com 5.778 alunos e 4 escolas.	Modelo p-median capacitado e problema de localização de cobertura máxima	Otimizador Comercial	Localização de centros/ O objetivo desse artigo é contribuir para o planejamento da rede escolar pública na região, considerando o forte crescimento populacional para 2015 e 2020.	Abordagem com objetivo único: minimiza o deslocamento aluno-escola, com o centro populacional como peso.
Amaya et al. (2015)	Redes de escolas primárias e secundárias em dois municípios no Chile, com 10.320 e 19.946 alunos e 15 e 24 escolas, respectivamente.	Modelo de otimização inteiro, níveis de turmas a serem oferecidas nas escolas, decisão sobre fechamento de escolas.	Otimizador Comercial	Localização de centros e atribuição de alunos/ Esse artigo apresenta um modelo de decisão econômica para planejamento em educação pública, considerando a satisfação total da demanda por educação em um município e o custo envolvido no sistema.	Abordagem multiobjetivo. A função objetivo essencialmente procura minimizar os custos total e variável das escolas.

continua

Tabela 3.1 – Resumo dos artigos selecionados.

continuação

Referência	Aplicação	Modelos e características	Técnica de Solução	Decisão/ Propósito	Objetivo único/ Multiobjetivo
Castillo-López e López-Ospina (2015)	Rede de escolas secundárias em áreas rurais do Chile, com 45 centros populacionais e 34 escolas.	Modelo usando uma abordagem microeconômica.	Otimizador Comercial e Heurística Busca Tabu	Localização de centros e atribuição de alunos/ Este artigo propõe uma localização e modificação do modelo de capacidade escolar para reduzir as desigualdades na educação rural.	Abordagem multiobjetivo que procura uma configuração escolar que minimiza o tempo de deslocamento médio aluno-escola, maximiza o número médio de alunos por escola e o número de escolas com turmas multiníveis.
Chen et al. (2018)	Esses autores aplicaram seu modelo no sistema escolar de Charlotte-Mecklenburg, na Carolina do Norte – Estados Unidos), considerando que 20 escolas poderiam ter sido abertas em um conjunto viável de 25 locais com 37.851 alunos.	Modelo capacitado <i>min-max</i>	Para locação, Busca Tabu. Para alocação, Algoritmo Guloso.	Modelo de localização-alocação de escolas levando em conta restrições de capacidade mínima e máxima.	Abordagem multiobjetivo que procura uma configuração escolar para minimizar a impedância agregada de deslocamento (tempo ou distância) aluno-escola, para maximizar o número de alunos enviados para a escola mais próxima e para minimizar o número de alunos enviados para escolas tão distantes de suas casas.
Karatas et al. (2017)	Geraram demanda e locais candidatos de facilidade dentro de uma região de 100 x 100 km quadrados. Fixaram os parâmetros básicos como $m = 20$, $n = 200$ e $p = 10$.	Modelos de locação p-mediana, cobertura máxima (MCLP) e p-centro	Otimizador Comercial	Modelos de alocação que atenda a múltiplos critérios.	Abordagem multicritério. O objetivo é medir e examinar a eficiência das soluções em termos de distâncias médias entre demanda-facilidade, número de demandas atendidas e distância máxima demanda-facilidade.

continua

Tabela 3.1 – Resumo dos artigos selecionados.

continuação

Referência	Aplicação	Modelos e características	Técnica de Solução	Decisão/ Propósito	Objetivo único/ Multiobjetivo
Lotfian e Najafi (2018)	Rede da mina subterrânea de Tabas, província de Khorasan do Sul, Iran, com 46 locais candidatos.	Modelos de locação p-mediana, cobertura máxima e p-centro.	Otimizador Comercial. GAMS [®] .	Localizar facilidades de emergência dentro de uma mina subterrânea com uma rede complexa usando o conceito de grafo e modelos discretos de locação, minimizando o tempo de atendimento aos feridos.	Abordagem multicritério, levando em consideração os seguintes objetivos: minimizar a distância entre estação de emergência-mineradores; minimização do número facilidades; maximizar a cobertura de todos os pontos de demanda.

Vários autores usaram a formulação do PMM na otimização de problemas de localização de escolas com objetivo único (Teixeira et al., 2007; Ndiaye et al., 2012; Menezes e Pizzolato, 2014; Teixeira e Antunes, 2008; Bruno et al., 2014).

Em Teixeira et al. (2007), uma variante do PMM foi empregada para determinar localidades para estabelecer p escolas em Coimbra, Portugal. O critério de otimização maximiza a acessibilidade dos alunos às escolas, com restrições à ocupação de capacidade máxima e mínima. O problema de otimização considerou 11 escolas secundárias existentes, 43 centros populacionais e a distância total entre alunos e a escola de 21.097 km. Foi utilizado um otimizador XPRESS-MP e um valor otimizado de 10.861 km foi alcançado para a soma das distâncias, com 14 escolas.

Teixeira e Antunes (2008) propuseram um modelo hierárquico capacitado (HCM) para localizar p escolas de forma a minimizar o deslocamento total aluno-escola na cidade de Coimbra (Portugal), melhorando a taxa de ocupação total das escolas. Esses autores definiram o HCM como um modelo no qual uma escola de nível s pode atender aos alunos de nível 1 a s , ou seja, uma rede educacional composta por escolas de jardim de infância, primárias e secundárias. Para aplicação do modelo HCM, eles consideraram 25 escolas EB1 (primeiro ciclo) e 14 escolas EB12 (primeiro e segundo ciclo) e 68 centros populacionais. Esses autores implementaram o modelo HCM com o pacote XPRESS-MP e utilizaram, como solucionador, o otimizador XPRESS versão 15.30. Os resultados obtidos por eles, após a solução do HCM foram: a) Cenário 1 – aumento da taxa de ocupação para 74%, fechamento de 3 escolas EB12 (primeiro e segundo ciclo), e total de distância aluno-escola para 12.527 km; b) Cenário 2 – taxa de ocupação de 88%, fechamento de 12 escolas EB1 e 1 Escola EB12, e aumento de 15% no deslocamento total aluno-escola para 14.391 km, em relação ao cenário 1; c) Cenário 3 – taxa de ocupação de 85%, localização de uma nova escola EB12 e diminuição do deslocamento total aluno-escola para 13.183 km, 5% maior do que o cenário 1.

Ndiaye et al. (2012) resolveram um problema de otimização com objetivo único. O critério de otimização foi a distância média percorrida pelos alunos de casa para a escola. A aplicação foi em uma rede de escolas suburbanas secundárias de Dakar, Senegal, com 15 centros. Os autores consideraram 8 locais candidatos ao estabelecimento de escolas, incluindo 5 locais com escolas existentes, de modo que a distância média entre estudantes e escolas foi limitada a 2.000 km. Os autores usaram um solucionador IBM-CPLEX para realizar os experimentos. Eles alcançaram os melhores resultados para $p = 5$. Quatro dos cinco sites candidatos já tinham escolas e apenas um site candidato não. A distância média de viagem entre o aluno e a escola era de 1.620 km.

Bruno et al. (2014) propuseram um modelo p -mediana para identificar o melhor local para centros de agrupamento e determinar a alocação de escolas em cada grupo, com a finalidade de minimizar o deslocamento aluno-escola, na cidade de Nápoles, Itália. A abordagem deles é baseada em grupos compreendidos por 29 escolas de ensino pré-escolar, primárias e secundárias. O cenário base era constituído por 11 grupos de escolas. O caso de estudo utilizado por Bruno et al. (2014) consistiu de 29 escolas com diferentes níveis e um total de 9.077 alunos (distribuídos entre escolas com mínimo de 40 e máximo de 798 estudantes) e 11 grupos. Essa organização, inicial, não satisfaz os requisitos governamentais que estabelecem um número mínimo de alunos por grupo e que contenha escolas de cada nível no mesmo. O modelo deles foi otimamente solucionado usando o CPLEX 12.2. Dessa forma, a menor distância média que eles obtiveram foi de 0,6 km entre a escola e o centro do grupo, para $p = 7$ grupos, com distância máxima de 2,50 km. O número de estudantes, nos sete grupos, variou de 1.040 a 1.466. Assim, eles diminuíram o número de grupos de 11 para 7.

Menezes e Pizzolato (2014) propuseram a localização de novas escolas municipais na área de Guaratiba - região de avaliação 5, Rio de Janeiro, Brasil, baseando-se em modelos p -mediana capacitado e de cobertura máxima. No primeiro modelo, eles procuraram minimizar o deslocamento total dos alunos até a escola mais próxima. Cada escola tinha uma capacidade máxima de 1.300 alunos e o número total de novas escolas era 15. No segundo modelo, eles procuraram maximizar a cobertura da população dentro de uma distância máxima predefinida de 1.500m. Ambas as abordagens foram realizadas usando a ferramenta *Advanced Interactive Multidimensional Modeling System*, baseada em um solucionador CPLEX.

Outros estudos conduzidos por Gac et al. (2009), Muller et al. (2009), Amaya et al. (2015) e Castillo-López (2015) resolvem problemas de localização escolar de múltiplos objetivos.

Gac et al. (2009) propuseram um modelo para maximizar os benefícios sociais do aluno e maximizar os ganhos econômicos das instituições de ensino na cidade de Santiago, Chile. Eles consideraram 18 escolas secundárias, sendo 7 privadas, 9 subsidiadas (mistas) e 2 públicas e 6 zonas populacionais classificadas por critério socioeconômico, variando de alta para baixa renda. Os autores incorporaram em uma única expressão esse dois termos considerados, em um modelo linear determinístico. Para o primeiro termo, eles levam em conta: a escolha do aluno; o fato de que todos os alunos devam ser atendidos; o nível de renda do aluno. Para o segundo termo, eles levam em conta: a capacidade da escola, nível de preço e o lucro econômico da escola. Os autores implementarem o modelo proposto na ferramenta

GAMS, solucionando a otimização com o CPLEX 7.0. De forma geral, os autores minimizaram o gasto do governo em 10%, reduzindo de 2 para 1 o número de escolas públicas, realocando as 17 escolas restantes em 6 zonas populacionais.

Muller et al. (2009) propuseram um modelo de localização-alocação de escolas com o objetivo de minimizar os custos totais fixo e variável e maximizar a utilidade dos alunos em relação às escolas. A área de estudo compreendeu a cidade de Dresden, Alemanha, composta por 64 zonas populacionais e 26 escolas secundárias. Os autores utilizaram um modelo de alocação probabilístico e modelaram a maximização de utilidade do aluno como um problema não linear que leva em conta a minimização da capacidade de expansão da escola e da rejeição da mesma pelo aluno. Os autores consideram, devido ao alto número de combinações, 11 escolas fixas, em vez das 26 escolas, e 25 períodos (de 1995 a 2020), totalizando 819.200 combinações de escolas abertas e fechadas por períodos. Após a otimização do modelo de localização realizada com o GAMS/CPLEX 22.2, eles obtiveram: a) Para Cenário I (sem fixar variáveis), o fechamento de 9 escolas, localização de 2 novas escolas e uma redução do custo total de € 1.787.000 por ano; b) Para o Cenário II (fixando o período de 1995), fechamento de 8 escolas, localização de 3 novas escolas e uma redução do custo total de € 1.731.000 por ano. O trabalho desses autores obteve uma redução entre 8% e 10% do custo total em relação à solução política adotada ou planejada.

Amaya et al. (2015) propuseram um modelo para determinar a localização das escolas com o objetivo de minimizar os custos fixos e gerais das escolas, como as taxas por hora. Esse trabalho levou em consideração os custos de abertura e fechamento de instalações escolares existentes. O problema foi otimizado usando o CPLEX 9.1. Eles conseguiram reduzir mais de 30% dos custos totais por um período de 5 anos e economizaram mais de \$ 6.000.000 por ano.

Castillo-López e López-Ospina (2015) propuseram um modelo multiobjetivo para localizar escolas secundárias em áreas rurais no Chile, considerando 45 centros populacionais e 34 escolas existentes. A função objetivo minimiza os custos operacionais e de investimento no sistema educacional, minimiza o tempo médio de viagem do aluno à escola, maximiza o número médio de alunos matriculados por escola e minimiza o número de escolas multiníveis. Esses autores obtiveram uma redução de 4,5% no custo. GAMS/MINOS (Sistema Modular de Otimização Não-Linear *In-core*) e Busca Tabu foram utilizados para resolver o modelo. Esses autores deixaram de lado a abordagem p-mediana e adotaram o procedimento de otimizar uma função ponderada.

Chen et al. (2018) formularam um modelo para o problema de alocação da escola para minimizar a impedância agregada de deslocamento (tempo ou distância) aluno-escola, para

maximizar o número de alunos enviados para a escola mais próxima e para minimizar o número de alunos enviados para escolas tão distantes de suas casas. Esses autores aplicaram seu modelo no sistema escolar de Charlotte-Mecklenburg, na Carolina do Norte – Estados Unidos) e em seu condado, considerando que 20 escolas poderiam ter sido abertas em um conjunto viável de 25 locais com 37.851 alunos entre 14 e 17 anos. O modelo foi resolvido pelo CPLEX e pelo Sistema Interativo de Alocação de Localização Gráfica para Escolas (iGLASS), uma abordagem proposta por esses autores que integra os algoritmos Busca Tabu, Guloso e Genético. Para o iGLASS e o CPLEX, respectivamente, esses autores obtiveram: uma impedância total de 144,8 km e 138,3 km; porcentagem de alunos designados para a escola mais próxima de 80,61% e 75,45%, e; porcentagem de alunos designados para uma escola que está duas vezes mais longe que a escola mais próxima de 5,13% e 2,68%.

Os trabalhos a seguir fizeram estudos comparativos usando dois ou mais métodos para resolver problemas de localização. Karatas et al. (2017) compararam o PMM clássico, o problema de localização de cobertura máxima e o modelo PCM aplicados a um problema de localização de serviços de emergência. Esses modelos foram avaliados com relação a vários critérios de desempenho, variando $p = \{10, 15, 20\}$ e usando $n = 200$. Os autores resolvem os modelos de otimização no Sistema Geral de Modelagem Algébrica (GAMS[®]), usando o CPLEX 12.2.0.2. Os autores alcançaram os melhores resultados para o PMM. Este modelo superou outros modelos em quatro de sete critérios.

Lotfian e Najafi (2018) aplicaram os modelos de localização PMM, cobertura máxima e PCM para determinar a localização de facilidades de postos de emergência dentro de uma rede de mina subterrânea de carvão para minimizar o tempo de alcance aos feridos. Os autores aplicaram o método proposto em uma mina localizada a 85km de Tabas, na província de Khorasan do Sul – Irã, considerando 46 locais candidatos. O modelo matemático foi resolvido usando o GAMS[®] e 7 locais foram selecionados dentre 46 candidatos. Como resultado, o tempo padrão de atendimento aos feridos que ficava entre 4 a 7min diminuiu para 3,3min, pois qualquer estação pode prestar atendimento aos feridos em um raio de 400m.

3.2 CONSIDERAÇÕES E CONTRIBUIÇÃO DESTE TRABALHO

Nenhum dos trabalhos encontrados na literatura trata da localização de escolas de ensino superior, nem resolve problemas que abrangem escolas localizadas em várias cidades, levando em consideração parâmetros sociais dessas cidades, como o IDH.

A maioria dos trabalhos utiliza otimizador comercial para solucionar os modelos de locação. Apenas Chen et. al (2018) utilizam heurísticas tais como a Busca Tabu, para locação das facilidades e algoritmos guloso/genético para alocação demanda-facilidade. Nenhum dos trabalhos utilizou uma abordagem com estratégias combinando os modelos p-mediana, p-centro e p-dispersão. E não realizam a normalização dos termos da expressão a ser otimizada.

Trabalhos como o de Castillo-López e López-Ospina (2015) e Gac et al. (2009) levam em conta a escolha do aluno e abordam escolas públicas, privadas e mistas. O que não é o caso do nosso trabalho.

Amaya et al. (2015) levam em conta custos fixos e variáveis separados em uma única expressão linear, no entanto, não utilizam o p-mediana.

A partir da revisão da literatura recém-apresentada sobre problemas de locação de unidades de ensino e conforme exposto anteriormente, concluímos que:

- Nenhum dos trabalhos citados acima usa uma coleção geral de conjuntos de dados de teste para uma variedade de problemas de pesquisas operacionais para comparar seus resultados obtidos. Isso é devido à especificidade da modelagem de problema de locação de unidades de ensino em cada trabalho;
- Nenhum desses trabalhos refere-se à unidade de ensino superior nem englobam diversas locações, mas a mesma locação dentro de uma jurisdição particular;

Portanto, no que diz respeito à localização das IES, o presente estudo trata de um problema diferente em comparação com os artigos publicados na literatura mencionados anteriormente.

4. METODOLOGIA

Neste capítulo é mostrado como foram obtidos e tratados os materiais e como foram elaborados os métodos para modelagem do problema de locação de unidade de IES em um conjunto de cidades do estado do Amazonas. Os seguintes tópicos são abordados: a) Mecanismos usados para coleta dos dados utilizados neste trabalho; b) Tratamento/normalização dos dados; c) Modelagem para o problema de locação; d) Métricas para Avaliação dos Resultados; e) Modelagem do Algoritmo Genético.

Este capítulo está organizado em três seções. A seção de Materiais compreende os dados coletados e os meios usados para a coleta. A seção Métodos é composta pelas seguintes partes: a) Normalização dos Dados; b) Modelos de locação p-mediana, p-centro e p-dispersão – em atendimento ao Objetivo Específico 1; c) Métricas para avaliar as soluções de expansão obtidas – em atendimento ao Objetivo Específico 2; d) Algoritmo Genético modelado para os modelos de locação discretos utilizados – em atendimento ao Objetivo Específico 3.

Na terceira seção empreendem-se considerações dos principais tópicos abordados neste capítulo.

4.1 MATERIAIS

Os materiais utilizados neste trabalho consistem de:

- 1) Populações dos 62 municípios, e;
- 2) IDH's dos 62 municípios do estado do Amazonas.
- 3) Distâncias entre as 62 cidades do estado do Amazonas;

As populações e IDH's dos 62 municípios do estado do Amazonas foram obtidos através do Instituto Brasileiro de Geografia e Estatística (IBGE, 2018 e IBGE, 2010). A população que utilizamos neste trabalho refere-se ao número de Alunos que cursam o Ensino Médio (AEM) nesses municípios, ou seja, a população com potencial de cursar o ensino superior. É importante destacar que as distâncias consideradas entre as cidades foram as distâncias fluviais, que correspondem ao meio de transporte mais utilizado na região. Distância fluvial (DF), população de alunos do ensino médio e IDH constituem os parâmetros utilizados neste trabalho.

Algumas distâncias fluviais foram determinadas a partir de dados da Associação Nacional de Transportes Aquaviários – ANTAQ (2013). As distâncias que não constavam nesses dados foram obtidas por um dos seguintes métodos:

- 1) Pela diferença entre as distâncias de trechos conhecidos que ficam no mesmo rio. Por exemplo, na Figura 4.1, o trecho entre Fonte Boa e Uarini, ambos no rio Solimões, foi obtido pela diferença entre o trecho Fonte Boa–Manaus e o trecho Uarini–Manaus;



Figura 4.1 – Municípios de Fonte Boa e Uarini no rio Solimões, Amazonas-Brasil.

- 2) Pela utilização da API *Google Maps* para medir a rota entre duas cidades localizadas em rios diferentes ou afluentes. Por exemplo, o trecho entre Juruá, que fica às margens do rio Juruá, e Fonte Boa, no rio Solimões, na Figura 4.2.

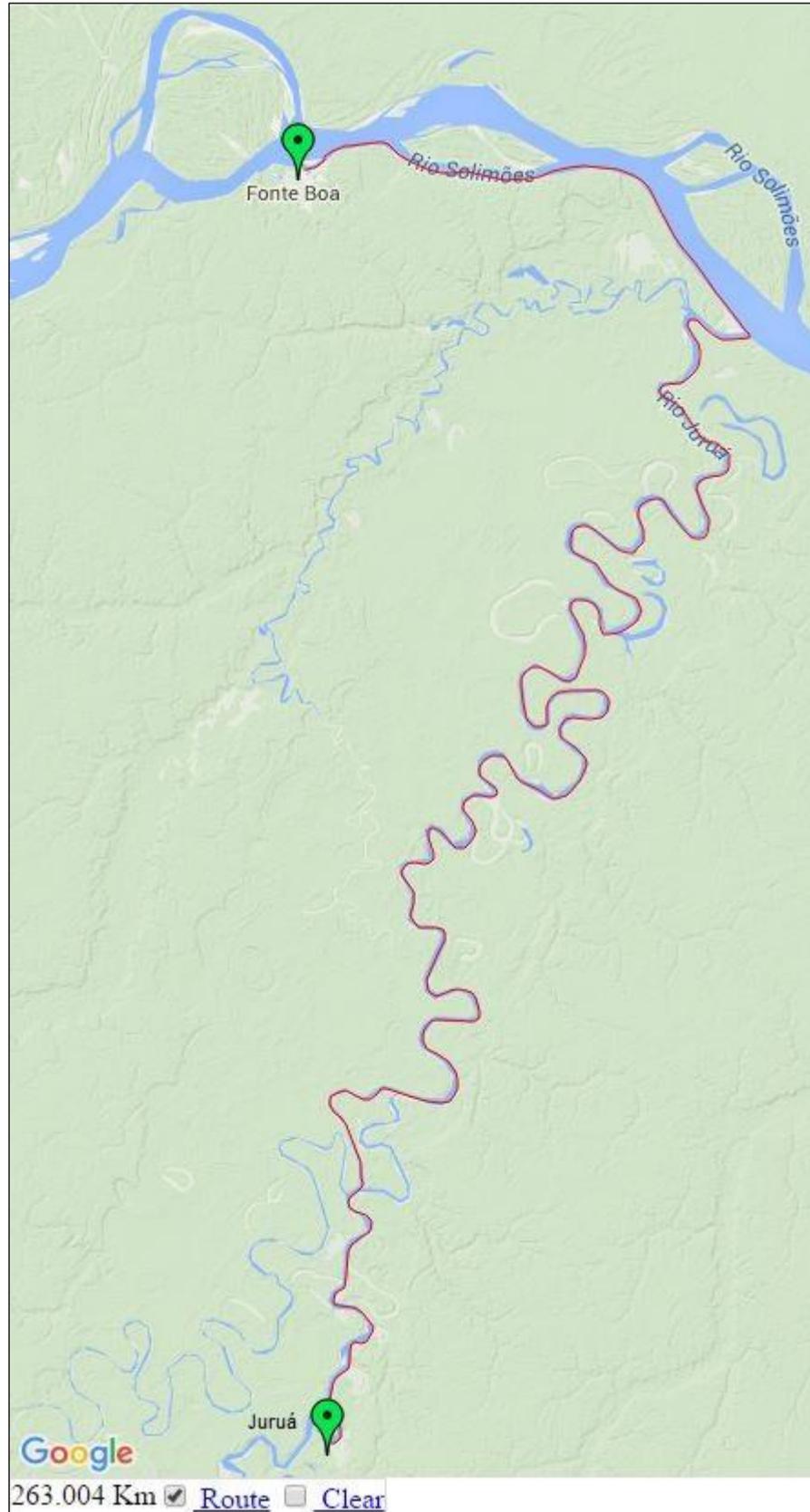


Figura 4.2 – Distância fluvial entre as cidades de Fonte Boa e Juruá.

Para se ter uma ideia da matriz de distância utilizada para as 62 cidades, a Tabela 4.1 mostra as distâncias fluviais entre um conjunto reduzido de localidades. Esses dados foram organizados em formato de planilha. A planilha contém apenas 16 cidades para efeito de demonstração. Esses municípios são listados tanto na vertical, quanto na horizontal, de tal forma que a distância entre duas cidades pode ser determinada pelo número situado no cruzamento da linha a que um deles pertence e da coluna a que o outro pertence. A diagonal principal indica a distância entre uma cidade para ela própria, e é igual a zero. Pode-se observar a simetria da planilha mostrada.

Tabela 4.1 – Exemplo da matriz de distância fluvial (km) com 16 municípios do estado do Amazonas utilizados no estudo.

	SGC	JAP	SIRN	MAR	FTB	JUT	TON	SAI	AMA	UAR	JUR	SPO	ALV	TEF	BAR	CAR
São Gabriel da Cachoeira (SGC)	0	1920	264	1797	1881	2002	2165	2196	2252	1688	1995	2346	1645	1632	547	2412
Japurá (JAP)	1920	0	1656	123	511	632	795	826	882	318	625	976	275	288	1373	1042
Santa Isabel do Rio Negro (SIRN)	264	1656	0	1533	1617	1738	1901	1932	1988	1424	1731	2082	1533	1368	283	2148
Maraã (MAR)	1797	123	1533	0	388	509	672	703	759	195	502	853	152	165	1250	919
Fonte Boa (FTB)	1881	511	1617	388	0	121	284	315	371	193	263	465	236	249	1334	680
Jutaí (JUT)	2002	632	1738	509	121	0	163	194	250	314	384	344	357	370	1455	801
Tonantins (TON)	2165	795	1901	672	284	163	0	31	87	477	547	181	520	533	1618	964
Santo Antônio do Içá (SAI)	2196	826	1932	703	315	194	31	0	56	508	578	150	551	564	1649	995
Amaturá (AMA)	2252	882	1988	759	371	250	87	56	0	564	634	94	607	620	1705	1051
Uarini (UAR)	1688	318	1424	195	193	314	477	508	564	0	307	658	43	56	1141	724
Juruá (JUR)	1995	625	1731	502	263	384	547	578	634	307	0	728	350	363	1448	417
São Paulo de Olivença (SPO)	2346	976	2082	853	465	344	181	150	94	658	728	0	701	714	1799	1145
Alvarães (ALV)	1645	275	1533	152	236	357	520	551	607	43	350	701	0	13	1098	767
Tefé (TEF)	1632	288	1368	165	249	370	533	564	620	56	363	714	13	0	1085	780
Barcelos (BAR)	547	1373	283	1250	1334	1455	1618	1649	1705	1141	1448	1799	1098	1085	0	1865
Carauari (CAR)	2412	1042	2148	919	680	801	964	995	1051	724	417	1145	767	780	1865	0

A Tabela 4.2 mostra o AEM e o IDH para o conjunto completo de cidades do estado do Amazonas considerado neste trabalho, além de exibir o identificador (ID) único que definimos para cada cidade.

Tabela 4.2 – AEM e IDH dos 62 municípios do Amazonas utilizados neste estudo. IBGE (2018, 2010).

ID	Município	AEM	IDH	ID	Município	AEM	IDH
1	São Gabriel da Cachoeira	2.707	0,609	32	Presidente Figueiredo	1.453	0,647
2	Japurá	290	0,522	33	Manaquiri	1.248	0,596
3	Santa Isabel Do Rio Negro	797	0,479	34	Lábrea	1.928	0,531
4	Maraã	676	0,498	35	Pauini	649	0,496
5	Fonte Boa	1.066	0,530	36	Irاندuba	3.080	0,613
6	Jutaí	1.089	0,516	37	Careiro	1.797	0,557
7	Tonantins	728	0,548	38	Manaus	108.182	0,737
8	Santo Antônio do Içá	1.498	0,490	39	Rio Preto da Eva	1.331	0,611
9	Amaturá	777	0,560	40	Envira	457	0,509
10	Uarini	514	0,527	41	Careiro da Várzea	1.155	0,568
11	Juruá	506	0,522	42	Manicoré	2.649	0,582
12	São Paulo de Olivença	2.278	0,521	43	Ipixuna	889	0,481
13	Alvarães	714	0,527	44	Novo Aripuanã	1.081	0,554
14	Tefé	3.621	0,639	45	Humaitá	2.486	0,605
15	Barcelos	821	0,500	46	Borba	1.862	0,560
16	Carauari	1.504	0,549	47	Boca do Acre	1.623	0,588
17	Tabatinga	3.899	0,616	48	Autazes	2.270	0,577
18	Benjamin Constant	2.183	0,574	49	Nova Olinda do Norte	1.853	0,558
19	Atalaia Do Norte	604	0,450	50	Itacoatiara	5.976	0,644
20	Coari	4.106	0,586	51	Guajará	678	0,532
21	Codajás	1.068	0,563	52	Silves	713	0,632
22	Itamarati	5.976	0,477	53	Urucurituba	811	0,588
23	Anori	675	0,561	54	Itapiranga	582	0,654
24	Novo Airão	818	0,570	55	São Sebastião Do Uatumã	633	0,577
25	Anamá	722	0,594	56	Urucará	1.018	0,620
26	Caapiranga	637	0,569	57	Boa Vista do Ramos	1.201	0,565
27	Tapauá	759	0,502	58	Maués	3.140	0,588
28	Beruri	1.413	0,506	59	Apuí	571	0,637
29	Canutama	385	0,530	60	Barreirinha	2.446	0,574
30	Eirunepá	1.519	0,563	61	Nhamundá	1.117	0,586
31	Manacapuru	5.433	0,614	62	Parintins	6.616	0,658

4.2 MÉTODOS

4.2.1 Normalização dos parâmetros

Neste trabalho, propomos que os parâmetros distância fluvial entre as cidades, população de AEM e IDH coletados de todos os 62 municípios sejam normalizados através de três diferentes formas: a) através dos valores máximo e mínimo, equação (4.1) (Xu e Wunsch, 2008); b) através da mediana, equação (4.2), e; c) através do desvio padrão, equação (4.3).

$$x_i = \frac{x_i^* - \min(x_i^*)}{\max(x_i^*) - \min(x_i^*)} \quad (4.1)$$

$$x_i = \frac{x_i^*}{X_m} \quad (4.2)$$

$$x_i = \frac{x_i^*}{s} \quad (4.3)$$

Em que:

- x_i^* representa os parâmetros não normalizados: distância, população ou IDH das cidades;
- x_i é o valor do parâmetro normalizado (Bussab e Morettin, 2009);
- X_m e s são, respectivamente, a mediana e o desvio padrão desses parâmetros (Spiegel, 2004).

A normalização apresentada nessa seção é necessária, uma vez que a diferença das faixas dinâmicas das variáveis distância fluvial, população e IDH é muito grande. Se não houvesse a normalização, numa expressão que envolvesse a soma desses parâmetros, aqueles que têm uma faixa dinâmica maior dominariam os demais, impossibilitando que, num processo de otimização do valor da expressão da soma, os termos com menor faixa dinâmica fossem considerados. Como, neste trabalho, conforme será mostrado na seção 4.2.2, propomos o uso desses parâmetros em uma única expressão, o tratamento desses dados constitui uma etapa de importância capital.

Foi demonstrado em Xavier et al. (2016) que, com relação à normalização, a mediana leva a valores mais baixos para a distância fluvial entre cidades com e sem IES e a um maior número de AEM para cidades com IES. O desvio padrão e a normalização máxima levam a valores mais baixos do IDH para cidades com IES. Assim, neste trabalho, usaremos a normalização através da mediana.

4.2.2 Modelagem da Locação

Conforme elencado anteriormente, os parâmetros utilizados na modelagem do problema de locação de objetivo único são o AEM, o IDH e a distância fluvial entre cidades sem unidades de ensino superior e cidades com unidades de ensino superior.

A escolha dos municípios a serem estabelecidas as facilidades deve atender aos seguintes critérios:

- 1) devem-se priorizar os municípios com maior população de AEM;
- 2) devem-se priorizar municípios com menor IDH;
- 3) devem-se minimizar as distâncias entre cidades com e sem unidades estabelecidas.

Neste trabalho, avaliamos soluções obtidas para a expansão do número de cidades com IES de 10 para 11 a 16 cidades. Propomos um aumento de 60% no número de cidades com IES, pois acreditamos que são soluções mais viáveis uma vez que há custos relevantes para o estabelecimento de uma IES.

As equações (4.4), (4.5) e (4.6) são submetidas às restrições (4.7), (4.8), (4.9) e (4.10). É importante ressaltar que, nas equações (4.4), (4.5) e (4.6), a população de AEM das cidades com uma facilidade é maximizada (sinal negativo), priorizando cidades com maior população, e o IDH é minimizado (sinal positivo), priorizando cidades com menor IDH. Por fim, as distâncias entre as cidades com IES são maximizadas, pois o primeiro termo também tem sinal negativo.

$$\text{Minimizar } f_1 = \alpha \frac{\sum_{i=1}^n \sum_{j=1}^n d_{ij} x_{ij}}{n-p} - \beta \frac{\sum_{j=1}^n p_j x_{jj}}{p} + \gamma \frac{\sum_{j=1}^n HDI_j x_{jj}}{p} \quad (4.4)$$

$$\text{Minimizar } f_2 = \alpha \max\{d_{ij}^w\} - \beta \frac{\sum_{j=1}^n p_j x_{jj}}{p} + \gamma \frac{\sum_{j=1}^n HDI_j x_{jj}}{p} \quad (4.5)$$

$$\text{Minimizar } f_3 = -\alpha \max\{d_{ij}^f\} - \beta \frac{\sum_{j=1}^n p_j x_{jj}}{p} + \gamma \frac{\sum_{j=1}^n HDI_j x_{jj}}{p} \quad (4.6)$$

Restrito a:

$$\sum_{i=1}^n x_{ii} = p \quad (4.7)$$

$$\sum_{j=1}^n x_{ij} = 1, \quad \forall i \in \{1, \dots, n\} \quad (4.8)$$

$$\sum_{i=1}^n x_{ij} \leq n x_{jj}, \quad \forall j \in \{1, \dots, n\} \quad (4.9)$$

$$x_{ii} \text{ and } x_{ij} \in \{0,1\}, \quad \forall i \in I, j \in J \quad (4.10)$$

Os parâmetros utilizados nas expressões de (4.7) à (4.10) são:

- d_{ij} é a distância fluvial normalizada entre as cidades i e j ;
- d_{ij}^w é a distância fluvial normalizada entre as cidades i sem uma IES e j com uma IES;
- d_{ij}^f é a distância fluvial normalizada entre as cidades i e j com uma IES;
- p_j é o número normalizado de AEM para a cidade j ;
- p é o número de facilidades a ser estabelecido;
- HDI_j é o IDH normalizado para a cidade j ;
- α, β e γ são os pesos para distância, população e IDH, respectivamente;
- I é o conjunto de locais de demanda;
- J é o conjunto de locais candidatos.

As variáveis de decisão são:

$$x_{ij} = \begin{cases} 1, & \text{se } \exists \text{ uma IES na cidade } j \text{ e se a cidade } i \text{ for mais próxima de } j \\ & \text{do que qualquer cidade com IES;} \\ 0, & \text{caso contrário.} \end{cases}$$

$$x_{jj} = \begin{cases} 1, & \text{se } \exists \text{ uma IES na cidade } j; \\ 0, & \text{caso contrário.} \end{cases}$$

Neste trabalho, temos três problemas com objetivos únicos que são otimizados separadamente pelo algoritmo genético. O primeiro problema com objetivo único otimiza um modelo PMM (equação (4.4)). O segundo problema com objetivo único otimiza um modelo PCM (equação (4.5)). O terceiro problema com objetivo único otimiza um modelo PDM (equação (4.6)). Em cada uma dessas equações, temos três parâmetros, um relacionado à distância, outro relacionado à população e o último ao IDH.

A modelagem p-mediana acima apresentada é clássica. A função objetivo da equação (4.4) minimiza a distância média entre as cidades e sua facilidade mais próxima, incluindo também os critérios AEM e IDH. A restrição (4.7) exige que exatamente p facilidades possam ser estabelecidas. A restrição (4.8) exige que cada cidade tenha uma única IES. A restrição (4.9) exige que cidades sem IES alocadas sejam associadas apenas a cidades com IES alocadas. A equação (4.10) indica a natureza binária das variáveis de decisão.

Na modelagem p-centro, a função objetivo da equação (4.5) minimiza a distância máxima entre uma cidade sem IES e sua cidade mais próxima com uma IES. As restrições (4.7), (4.8), (4.9) e (4.10) são as mesmas da modelagem p-mediana.

Na modelagem p-dispersão, a função objetivo da equação (4.6) maximiza a distância entre as duas cidades mais próximas com IES alocadas. As restrições (4.7), (4.8), (4.9) e (4.10) são as mesmas da modelagem p-mediana.

O primeiro termo da função objetivo (4.1) minimiza distância entre o nó de demanda e facilidade mais próxima é semelhante ao termo da função objetivo (2.1). O primeiro termo da função objetivo (4.2) minimiza a distância máxima entre o nó de demanda e semelhante à função objetivo (2.6). O primeiro termo da função objetivo (4.3) maximiza a distância mínima de separação entre quaisquer pares de facilidades é semelhante à função objetivo (2.13). A restrição (4.4) é similar às restrições (2.2), (2.8) e (2.14) dos modelos p-mediana, p-centro e p-dispersão, respectivamente. A restrição (4.5) é semelhante às restrições (2.3) e (2.3), respectivamente, dos modelos p-mediana e p-centro. A restrição (4.6) é similar às restrições (2.4) e (2.9). A restrição (4.7) é semelhante às restrições (2.5), (2.11) e (2.16), respectivamente, dos modelos p-mediana, p-centro e p-dispersão.

É interessante notar que os parâmetros população de AEM e IDH são conflitantes. Esse conflito surge porque cidades pequenas tem baixo valor de IDH, no caso do Estado do Amazonas. Por exemplo, a cidade de Atalaia do Norte tem uma população de AEM de 900 habitantes e um IDH de 0,450. A cidade de Itamarati tem uma população de AEM de 830 habitantes e um IDH de 0,477. Assim sendo, nas equações (4.4), (4.5) e (4.6), algumas cidades com menor população apresentam o segundo termo menor e o terceiro termo maior.

Na Introdução, foi mencionado que ofereceremos nove opções de expansão de IES nas localidades do interior do Estado do Amazonas para os tomadores de decisão. Essas opções são as seguintes:

1. PMM, PCM e PDM que favorecem cidades com menor IDH, definindo $\alpha = 0,5$, $\beta = 0$ e $\gamma = 0,5$ nas equações (4.4), (4.5) e (4.6), respectivamente;
2. PMM, PCM e PDM que favorecem igualmente a maioria das cidades populosas, definindo $\alpha = 0,5$, $\beta = 0$ e $\gamma = 0,5$ nas equações (4.4), (4.5) e (4.6), respectivamente;
3. PMM, PCM e PDM que favorecem igualmente os dois critérios (IDH e cidades mais populosas), definindo $\alpha = 0,5$, $\beta = 0$ e $\gamma = 0,5$ nas equações (4.4), (4.5) e (4.6), respectivamente.

4.2.3 Métricas para Avaliação dos Resultados

Neste estudo, avaliamos os resultados das soluções de localização usando sete tipos de métricas, divididos em três grupos.

- 1) Grupo 1: Para avaliar se as soluções obtidas pela combinação de modelos e estratégias de localização citadas na Seção 1 estão fornecendo uma melhor distribuição das IES, usaremos as seguintes métricas:
 - a) Métrica 1: distância fluvial média entre uma cidade sem IES e uma cidade com IES;
 - b) Métrica 2: distância fluvial máxima entre uma cidade sem IES e uma cidade com IES;
 - c) Métrica 3: distâncias fluviais máximas entre cidades com IES;
 - d) Métrica 4: taxa de IES.
- 2) Grupo 2: Para avaliar se as soluções obtidas pela combinação de modelos e estratégias de localização citadas na Seção 1 estão proporcionando uma melhor equalização do número de AEM de todas as mesorregiões, usaremos a seguinte métrica:
 - a) Métrica 5: taxa de AEM.
- 3) Grupo 3: Para avaliar se as soluções obtidas pela combinação de modelos de localização e algumas estratégias citadas na Seção 1 podem favorecer cidades com menor IDH ou com maior valor para AEM, usaremos as seguintes métricas:
 - a) Métrica 6: AEM médio, calculado pelo somatório dos AEM das cidades com IES dividido pela quantidade dessas cidades;
 - b) Métrica 7: IDH médio, calculado pelo somatório do IDH das cidades com IES dividido pela quantidade dessas cidades.

Com relação às métricas mencionadas anteriormente, os melhores resultados das soluções de localização devem: a) apresentar os menores valores para as métricas 1, 2, 3 e 7, e; b) maiores valores para as métrica 4, 5 e 6.

4.2.4 Modelagem de Algoritmos Genéticos

Um algoritmo genético é um algoritmo de otimização estocástica, inspirado na teoria da evolução de Charles Darwin (Holland, 1975). Desde sua proposição inicial, ele tem sido efetivamente aplicado na solução de problemas complexos, como o *Travelling Salesman Problem* (TSP) – Problema do Caixeiro Viajante – (Reinelt, 1991) e o PMM (Alp et al., 2003; Satoglu et al., 2016, Drezner et al. 2015).

O algoritmo genético começa com uma população $P(0)$ de N cromossomos, gerados aleatoriamente. O valor de aptidão de cada cromossomo é avaliado através de uma função objetivo do problema. Para o modelo proposto, conforme a seção 4.2.2, o caminho natural é a utilização de uma aplicação genérica, tal como o Algoritmo Genético para encontrar soluções aproximadas, solucionando o modelo proposto.

A implementação do Algoritmo Genético geralmente consiste em três etapas: construção do modelo cromossômico, definição da função objetivo e parametrização dos operadores genéticos. Neste trabalho, a função de aptidão é uma das equações (4.4), (4.5) ou (4.6), utilizada em problemas modelados com PMM, PCM e PDM, respectivamente. O modelo cromossômico usa os índices das facilidades: $\{1, 2, \dots, 62\}$.

Neste trabalho, usamos uma versão adaptada do algoritmo GA proposto por Silva et al. (2019). Esse algoritmo é apresentado abaixo em oito etapas do Algoritmo 1.

Algoritmo 1: Algoritmo Genético Adaptado (Silva et al., 2019).

Etapas

1 Geração aleatória da população.

2 Cálculo da aptidão da população.

Repita por x gerações

3 Selecione 2 pais através do operador Seleção da Roleta.

4 Cruzamento de um ponto, com 95% de probabilidade.

5 Mutação aleatória de um gene, com 5% de probabilidade.

6 Calcula a aptidão.

7 Substitui os pais pelos filhos com melhor aptidão.

Até que a população tenha convergido, vá para a Etapa 3.

8 Pare.

O Algoritmo 2, a seguir, ilustra como foram implementadas as funções para o algoritmo genético utilizado neste trabalho, mostrando o número de repetições no pior caso de cada comando para a definição da complexidade computacional em termos de tempo de execução.

Para isso, definimos a seguinte notação:

- c_i é o custo em tempo que cada comando i leva para executar;
- n é o número de gerações;
- $|p|$ é o tamanho da população;
- p é o tamanho do cromossomo, com $p > p^*$;
- p^* é número de genes que não entram na etapa de evolução do GA;
- l é o número de cidades no Estado do Amazonas;

Algoritmo 2: Algoritmo Genético implementado neste trabalho.

Função	Custo	Repetições
ExecutaAlgoritmoGenetico(populacao, probabilidadeCruzamento, probabilidadeMutacao)		
1 Para i de 1 até n	c_1	n
2 CalculaProbabilidadeSelecao($populacao$)	c_2	n
3 $pc \leftarrow$ CalculaProbabilidadeCruzamento()	c_3	n
4 Se($pc \leq$ probabilidadeCruzamento)	c_4	n
5 $pais \leftarrow$ SeleccionaPaisCruzamento($populacao$)	c_5	n
6 $filhos \leftarrow$ RealizaCruzamento($pais$)	c_6	n
7 $filhos \leftarrow$ SubstituiGenesRepetidos($filhos$)	c_7	$n[(p - p^*)l]$
8 CalculaAptidaoFilhos($filhos$)	c_8	$n(2(l * p) + p)$
9 $[melhoresIndividuos\ pioresIndividuos]$ \leftarrow ComparaAptidao ($pais, filhos$)	c_9	$n * 2 * 2$
10 RemovePopulacao ($pioresIndividuos$)	c_{10}	$n * 2$
11 AdicionaPopulacao ($melhoresIndividuos$)	c_{11}	$n * 2$
12 Fim Se		
13 $pm \leftarrow$ CalculaProbabilidadeMutacao()	c_{12}	n
14 Se($pm \leftarrow$ probabilidadeMutacao)	c_{13}	n
15 $pai \leftarrow$ SeleccionaPaiMutacao($populacao$)	c_{14}	n
16 $filho \leftarrow$ RealizaMutacao($populacao$)	c_{15}	n
17 CalculaAptidao($filho$)	c_{16}	$n[(l * p) + p]$
18 SubstituiPaiPeloFilho($pai, filho, populacao$)	c_{17}	n
19 Fim Se		
20 $populacao \leftarrow$ OrdenaCustoDecrescente($populacao$)	c_{18}	$n * p ^2$
21 $melhorSolucao \leftarrow$ ObtemUltimaSolucao($populacao$)	c_{19}	n
22 Fim Para		
23 retornamelhorSolucao		

Segundo Cormen et al. (2009), para calcular $T(n)$, o tempo de execução do algoritmo, somamos o produto do custo em tempo de cada comando pelo número de vezes que cada comando é executado, conforme é mostrado a seguir:

$$T(n) = (c_1 + c_2 + c_3 + c_4 + c_5 + c_6 + 4c_9 + 2c_{10} + 2c_{11} + c_{12} + c_{13} + c_{14} + c_{15} + c_{17} + c_{19}) * n + ((p - 10) * l)c_7 + (2 * (l * p) + p)c_8 + ((l * p) + p)c_{16} * n + n * |p|^2 * c_{18}$$

Assim, podemos expressar o tempo de execução no pior caso como para o algoritmo genético implementado neste trabalho, como:

$$T(n) = a * n + b * n + n * |p|^2 \Rightarrow$$

$$T(n) = O(n + n * |p|^2)$$

Conforme Cormen et al. (2009), a função de ordenação (*Insertion Sort*) tem tempo de execução no pior caso de tempo de $O(n^2)$. Essa função, no pseudo-código do Algoritmo 2, é representada pela função *OrdernaCustoDecrescente(populacao)*, e no nosso caso $n = |p|$. O tempo de execução para o laço {Para i de 1 até n }, linha 1, é $O(n)$. Assim, o tempo de execução no pior caso da Algoritmo 2 é $O(n^3)$, quando o número de gerações for iguais ao número de indivíduos na população.

Como afirmado anteriormente, a codificação genética usa os índices das facilidades. A representação decimal possui menor custo computacional e o algoritmo apresenta melhor em relação à codificação binária que faz com que o algoritmo tenha um desempenho ineficiente mesmo em problemas pequenos (Alp e Erkut, 2003). Assim, neste trabalho, usamos a codificação decimal para os cromossomos.

A Figura 4.3 mostra um cromossomo codificado usado em um problema PMM com 13 facilidades. Os dez primeiros genes (em azul) representam as cidades existentes com uma IES. Os outros genes (em verde) são aqueles que podem ser substituídos.

1	14	17	18	20	34	38	45	50	62	16	48	8
---	----	----	----	----	----	----	----	----	----	----	----	---

Figura 4.3 – Exemplo de um cromossomo codificado com 13 facilidades. O azul indica os genes fixos. Os outros genes (em verde) são aqueles que podem ser substituídos.

O operador de seleção usado é o operador da Roleta (Fernando et al., 2010). Neste estudo, o operador de seleção atribui um valor de probabilidade a cada indivíduo, proporcional ao seu valor de aptidão. Essa probabilidade é dada pela equação (4.11).

$$p(c_i) = \frac{\text{rank}(c_i)}{\sum_{i=1}^n \text{rank}(c_i)} \quad (4.11)$$

Em que:

- c_i é o i -ésimo cromossomo na população;
- $\text{rank}(c_i)$ é a posição do c_i ordenado pelos valores decrescentes de sua função de aptidão;
- $p(c_i)$ é a probabilidade do cromossomo c_i ;

Como temos um problema de minimização, os valores da função de aptidão foram ordenados por valores decrescentes. O valor da função de aptidão não foi utilizado para o cálculo da probabilidade, mas sim pela posição (rank), uma vez que valores grandes da função de aptidão dominam a população, gerando uma convergência prematura do GA (Zapfel et al., 2010).

Um operador de cruzamento de um ponto (Fernando et al., 2010) é usado neste estudo. Esse operador gera aleatoriamente um ponto de referência para permutar genes entre os pais. A probabilidade de cruzamento usada é de 95% (Silva et al., 2019). Para o cromossomo mostrado na Figura 4.3, com 13 genes, o ponto de cruzamento é escolhido aleatoriamente em um ponto entre os genes 11 e 12 ou entre os genes 12 e 13. Esses genes correspondem a cidades sem uma IES.

A Figura 4.4 ilustra a permutação genética realizada pelo operador de cruzamento de um ponto. Para evitar índices repetidos na prole, examinamos os genes de cada filho e substituímos o índice repetido por outro valor selecionado aleatoriamente.

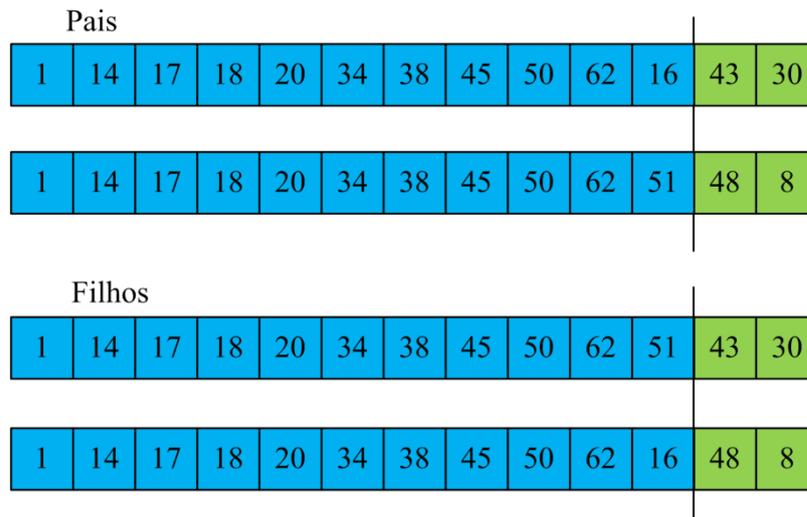


Figura 4.4 – Exemplo de operador de cruzamento de um ponto. Quadrados de cor azul indicam os genes fixos. Os outros genes (em verde) são aqueles que podem ser substituídos.

O operador de mutação seleciona aleatoriamente um gene (Satoglu et al., 2016), com probabilidade de 0,05 (Silva et al., 2019; Shehadeh et al., 2019), e executa uma mutação. A Figura 4.5 ilustra o operador de mutação. Um gene com um valor de índice de 48 é selecionado e substituído por um valor de índice de 32. O valor de substituição é selecionado aleatoriamente. A mutação evitar ótimos locais e permite atingir outros espaços de busca ainda não explorados e evita a convergência prematura (Zapfel et al., 2010).

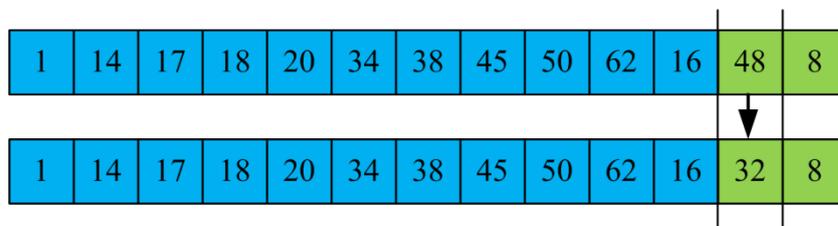


Figura 4.5 – Exemplo de operação de mutação de um gene. Quadrados de cor azul indicam os genes fixos. Os outros genes (em verde) são aqueles que podem ser substituídos.

Os parâmetros usados neste estudo para o GA são descritos na sequência: tamanho da população = 10 (Lee et al., 2009; Roeva et al., 2013) e número máximo de iterações = 600 (Moghadam et al., 2014). A implementação e a simulação para essa heurística foram realizadas no MATLAB R2014a em um computador ASUS Intel® Xeon® com 8 GB e processador de 1,80-2,40 GHz RAM e Windows 10.

4.3 CONSIDERAÇÕES

Neste capítulo, diversos tópicos foram abordados. Inicialmente, mostrou-se como foram adquiridos os materiais, obtidos de instituições brasileiras e de informações derivadas a partir da API do *Google Maps*.

Em seguida, mostrou-se a importância da normalização dos parâmetros utilizados nesse trabalho, impedindo o domínio de um parâmetro em um processo de otimização que utiliza uma expressão da soma dos mesmos.

Abordamos ainda os métodos p-mediana, p-centro e p-dispersão utilizados para a modelagem do problema tratado nesta tese, bem como as métricas propostas para avaliação das soluções. Ressaltamos que as modelagens foram realizadas em atendimento aos objetivos estipulados no capítulo de Introdução.

As adaptações realizadas no método de otimização heurístico Algoritmo Genético aplicado aos métodos p-mediana, p-centro e p-dispersão também foram descritas, bem como os parâmetros e ambiente utilizados para simulação.

No próximo capítulo mostramos os principais resultados alcançados para os métodos propostos e analisamos as soluções obtidas pelas diferentes estratégias propostas.

5. RESULTADOS E ANÁLISE DOS RESULTADOS

No capítulo anterior, mostramos a modelagem para o problema p-mediana, p-centro e p-dipersão, e definimos nove estratégias de expansão para os tomadores de decisão. Neste capítulo, são apresentados e discutidos os resultados obtidos para a expansão de IES no estado do Amazonas.

Este capítulo está dividido em seis seções. Na primeira seção, denominada de “Resultado para as métricas do grupo 1”, obtemos resultados para as métricas 1, 2, 3 e 4, respectivamente, para a distância fluvial média entre uma cidade sem IES e uma cidade com IES, distância fluvial média entre uma cidade sem IES e uma cidade com IES, distâncias fluviais máximas entre cidades com IES e a taxa de IES, considerando os três modelos de localização. Na segunda seção, denominada “Resultado para a métrica do grupo 2”, obtemos resultados para a métrica 5, taxa de AEM das mesorregiões. Na terceira seção, denominada “Resultado para as métricas do grupo 3”, obtemos resultados para as métricas 6 e 7, AEM médio e IDH médio, respectivamente.

Na quarta seção é realizada a discussão dos resultados obtidos quanto à combinação modelos-estratégia de localização. Na quinta seção, são realizados comentários a respeito do Algoritmo Genético adaptado. Por fim, na sexta seção, são elencadas considerações a respeito dos resultados obtidos.

5.1 RESULTADOS DAS MÉTRICAS DO GRUPO 1

As Figuras 5.1(a), 5.1(b) e 5.1(c) mostram os resultados para as métricas 1, 2 e 3, respectivamente, para a expansão de 10 a 16 cidades com instituições de ensino superior utilizando o modelo PMM, considerando as três configurações de parâmetros: $(\alpha = 0,5, \beta = 0, \gamma = 0,5)$, $(\alpha = 0,5, \beta = 0,5, \gamma = 0)$ e $(\alpha = 1/3, \beta = 1/3, \gamma = 1/3)$.

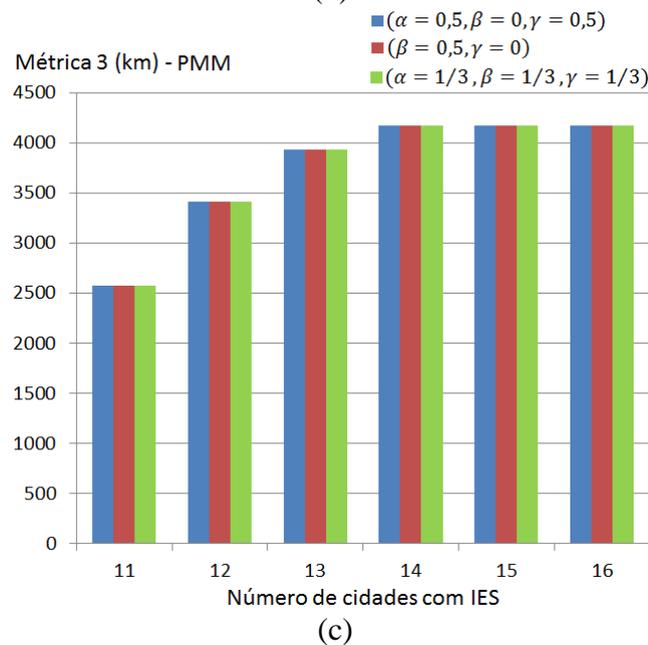
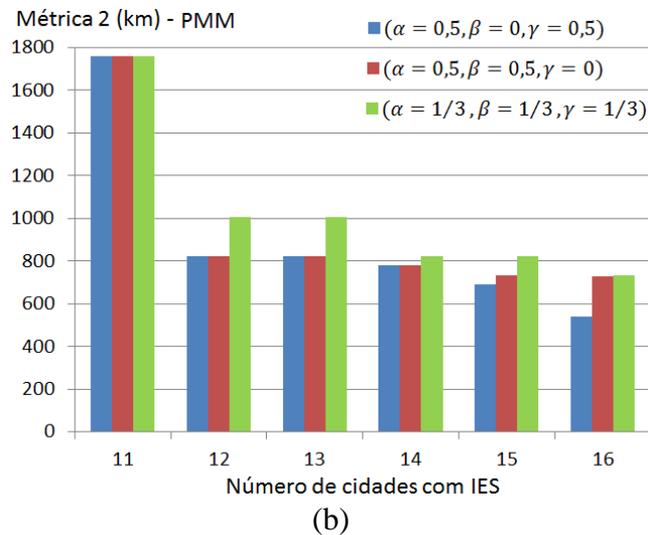
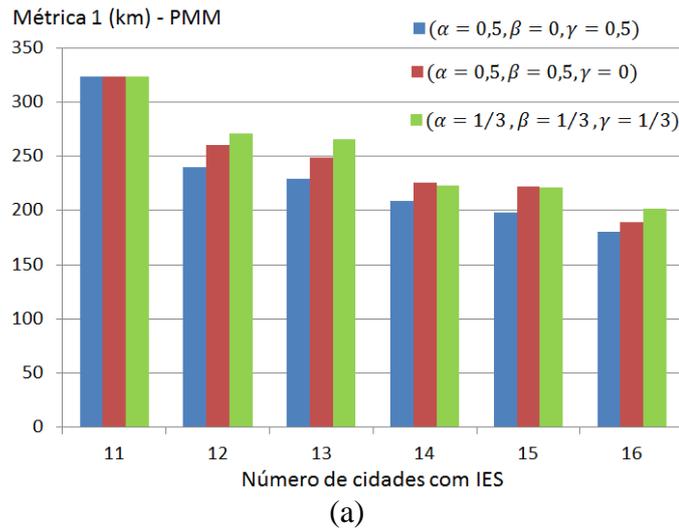


Figura 5.1 – Resultados obtidos usando o modelo PMM para expandir de 10 para 11 a 16 cidades com IES: (a) métrica 1; (b) métrica 2, e; (c) métrica 3. Todas as distâncias estão em quilômetros.

As Figuras 5.2(a), 5.2(b) e 5.2(c) mostram os resultados para as métricas 1, 2 e 3, respectivamente, utilizando o modelo PCM para a expansão de 10 a 16 cidades com IES, levando em conta as três configurações de parâmetros: $(\alpha = 0,5, \beta = 0, \gamma = 0,5)$, $(\alpha = 0,5, \beta = 0,5, \gamma = 0)$ e $(\alpha = 1/3, \beta = 1/3, \gamma = 1/3)$.

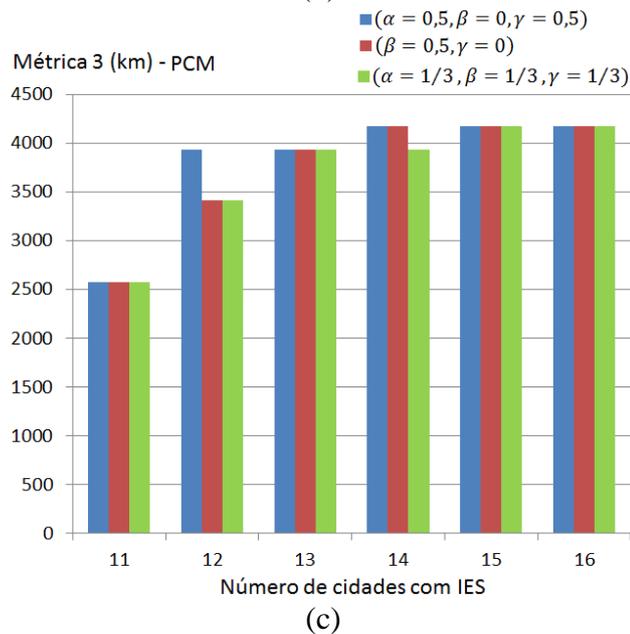
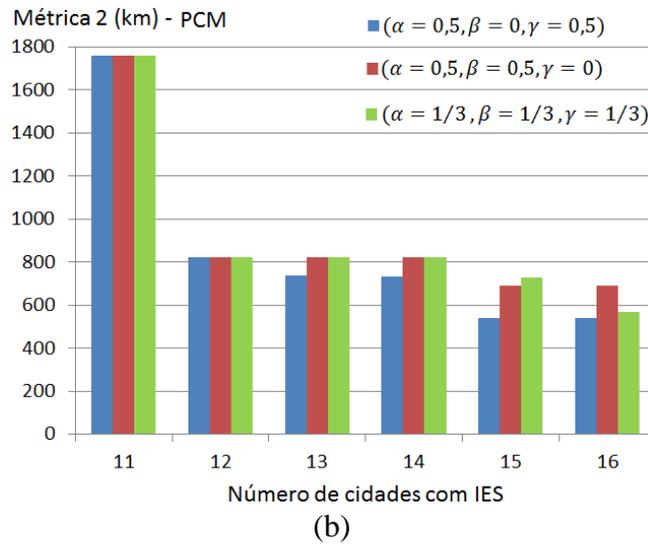
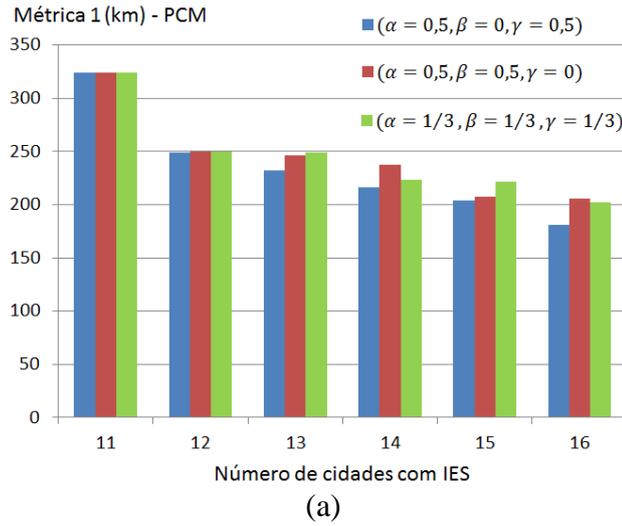


Figura 5.2 – Resultados obtidos da métrica 2 para expandir de 10 para 11 a 16 cidades com IES utilizando os modelos: (a) PMM; (b) PCM, e; (c) PDM. Todas as distâncias estão em quilômetros.

As Figuras 5.3(a), 5.3(b) e 5.3(c) mostram os resultados para as métricas 1, 2 e 3, respectivamente, usando o modelo PDM para expandir de 10 até 16 cidades a ser estabelecida uma unidade de uma IES, tendo em vista as três configurações de parâmetros: $(\alpha = 0,5, \beta = 0, \gamma = 0,5)$, $(\alpha = 0,5, \beta = 0,5, \gamma = 0)$ e $(\alpha = 1/3, \beta = 1/3, \gamma = 1/3)$.

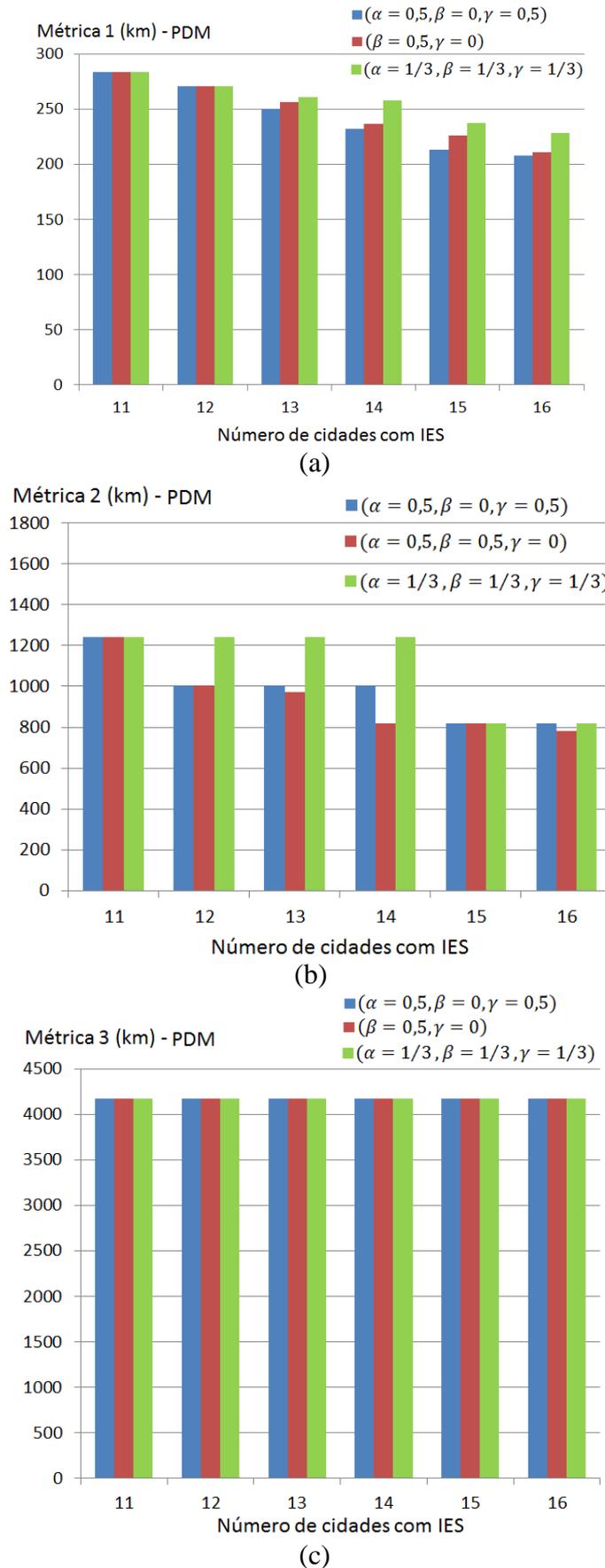


Figura 5.3 – Resultados obtidos usando o modelo PDM para expandir de 10 para 11 a 16 cidades com IES: (a) métrica 1; (b) métrica 2, e; (c) métrica 3. Todas as distâncias estão em quilômetros.

A Tabela 5.1 mostra a quantidade de IES e os valores da métrica 4 para as mesorregiões do estado do Amazonas, para a expansão de 10 a 16 cidade com IESs com os três modelos de localização propostos, considerando as três estratégias: a) estratégia A, ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$); estratégia B, ($\alpha = 0,5$, $\beta = 0,5$, $\gamma = 0$), e; estratégia C, ($\alpha = 1/3$, $\beta = 1/3$, $\gamma = 1/3$).

Tabela 5.1 – Valores da métrica 4 para as três estratégias A, B e C, para expansão de 10 a 16 cidades com IES.

Estratégia	Micro Região	Métodos					
		PMM		PCM		PDM	
		IES	taxa de IES	IES	taxa de IES	IES	taxa de IES
Estratégia A	Região Central	6	0,3750	6	0,3750	6	0,3750
	Norte	1	0,0625	1	0,0625	1	0,0625
	Sudoeste	6	0,3750	6	0,3750	6	0,3750
	Sul	3	0,1875	3	0,1875	3	0,1875
Estratégia B	Região Central	6	0,3750	7	0,4375	8	0,5000
	Norte	1	0,0625	1	0,0625	1	0,0625
	Sudoeste	6	0,3750	5	0,3125	4	0,2500
	Sul	3	0,1875	3	0,1875	3	0,1875
Estratégia C	Região Central	6	0,3750	7	0,4375	7	0,4375
	Norte	1	0,0625	1	0,0625	1	0,0625
	Sudoeste	6	0,3750	5	0,3125	6	0,3750
	Sul	3	0,1875	3	0,1875	2	0,1250

5.2 DISCUSSÃO DOS RESULTADOS PARA AS MÉTRICAS DO GRUPO 1

Nesta seção serão discutidos e analisados os resultados obtidos para as métricas 1, 2, 3 e 4, descritos anteriormente.

5.2.1 Avaliação se as soluções obtidas pela combinação de modelos e estratégias de localização citadas na Seção 1 fornecem uma melhor distribuição das IES

Conforme observado na Figura 5.1(a), a combinação de PMM com a configuração de parâmetros ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$) resulta em valores menores para a métrica 1, ou seja, menor distância fluvial média entre uma cidade sem e com IES. O pior desempenho é obtido com a combinação de PMM com a configuração de parâmetros ($\alpha = 0,5$, $\beta = 0,5$, $\gamma = 0$), para as expansões para 12, 13 e 16 cidades com IES. Nas expansões apresentadas, a distância fluvial média diminuiu com o aumento do número de cidades com IES. Na Figura 5.1(b), o PMM com a configuração de parâmetros ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$) e o PMM com a configuração de parâmetros ($\alpha = 0,5$, $\beta = 0,5$, $\gamma = 0$) resultaram em valores cada vez mais

baixos para a métrica 2, para expansões de 11 a 15 IES e para 16 IES, respectivamente. Na Figura 5.1(c), o PMM com quaisquer uma das configurações de parâmetros ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$), ($\alpha = 0,5$, $\beta = 0,5$, $\gamma = 0$) e ($\alpha = 1/3$, $\beta = 1/3$, $\gamma = 1/3$) resultaram em valores semelhantes de distância fluvial máxima entre as cidades com IES, para a métrica 3.

Conforme observado na Figura 5.2(a), o PCM com a configuração de parâmetros ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$) resulta em valores menores para a métrica 1, menor distância fluvial média entre localidades sem e com IES. Conforme observado na Figura 5.2(a), o PCM com a configuração de parâmetros ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$) resulta em valores menores para a métrica 1, menor distância fluvial média entre localidades sem e com IES. O piores desempenho são obtido para o PCM com a configuração de parâmetro ($\alpha = 0,5$, $\beta = 0,5$, $\gamma = 0$) e ($\alpha = 1/3$, $\beta = 1/3$, $\gamma = 1/3$), para as expansões para 14 e 15 cidades com IES, respectivamente. Nas expansões apresentadas, a distância fluvial média diminuiu com o aumento do número de cidades com IES. Na Figura 5.2(b), o PCM com a configuração de parâmetros ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$) resulta em valores mais baixos para a métrica 2. Na Figura 5.2(c), o PCM com a configuração de parâmetros ($\alpha = 1/3$, $\beta = 1/3$, $\gamma = 1/3$) resulta em valores mínimos para a métrica 3 e os piores resultados são obtidos para o PCM com ajuste de parâmetros ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$).

Conforme observado na Figura 5.3(a), o PDM com configuração de parâmetro ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$) resulta em valores menores para a métrica 1. Os piores desempenhos foram obtidos com o PDM com a configuração de parâmetro ($\alpha = 1/3$, $\beta = 1/3$, $\gamma = 1/3$). Nas expansões apresentadas, a distância média fluvial diminuiu com o aumento do número de cidades com IES. Na Figura 5.3(b), o PDM com configuração de parâmetro ($\alpha = 0,5$, $\beta = 0,5$, $\gamma = 0$) resultou em valores mais baixos para a métrica 2. Na Figura 5.3(c), todas as configurações de parâmetros resultam no mesmo valor mínimo para a métrica 3.

Na Tabela 5.1 mostramos os resultados para métrica 4 (taxa de IES) obtidos para cada um dos três modelos combinados com as estratégias A, B e C. Para a estratégia A, priorizando cidades com menor IDH, os modelos PMM, PCM e PDM resultaram na mesma proporção de IES. Para a estratégia B, em que são priorizadas cidades com população mais alta, a melhor taxa de IES foi obtida pelo modelo PMM. Para essa estratégia, a pior taxa de IES foi alcançada pelo PDM. Para a estratégia C, o modelo PMM obteve os melhores valores de taxa de IES para as mesorregiões Sudoeste e Sul. O PMM prioriza cidades nas mesorregiões Sudoeste e Sul, ou seja, cidades com menor valor de IDH e que estão distantes dos maiores

centros populacionais do Estado do Amazonas, tais como Manaus, Itacoatiara e Parintins. As piores taxas de IES foram alcançadas pelos modelos PCM e PDM para as mesorregiões Sudoeste e Sul, respectivamente. O modelo PMM prioriza cidades nas mesorregiões Sudoeste e Sul.

A partir da Tabela 5.1, extraímos uma conclusão importante deste trabalho: todos os métodos de localização (PMM, PCM e PDM) obtiveram, para a expansão de 10 a 16 cidades com IES, uma melhor taxa de IES para as mesorregiões Norte, Sudeste e Sul comparada com os valores mostrados na Tabela 1.1. A única exceção é a mesorregião Central. De fato, na Tabela 1, temos que a taxa IES da mesorregião Central, a mais populosa, é de 0,8, enquanto na Tabela 5.1, as taxas IES da mesorregião Central são menores ou iguais a 0,5. Portanto, os modelos de localização aumentam o número de cidades com instituições de ensino superior nas mesorregiões com menor proporção de alunos.

5.3 RESULTADOS DA MÉTRICA DO GRUPO 2

A Tabela 5.4 mostra a quantidade de IES, o número de AEM e os valores da métrica 5 para as mesorregiões do Estado do Amazonas, para a expansão de 10 a 16 cidade com IESs com os três modelos de localização propostos, considerando as três estratégias: a) estratégia A, ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$); estratégia B, ($\alpha = 0,5$, $\beta = 0,5$, $\gamma = 0$), e; estratégia C, ($\alpha = 1/3$, $\beta = 1/3$, $\gamma = 1/3$).

Tabela 5. 2 – Valores da métrica 5 com as estratégias A, B e C, para expansão de 10 a 16 cidades com IES.

Estratégia	Micro Região	Métodos								
		PMM			PCM			PDM		
		IES	AEM	taxa de AEM	IES	AEM	taxa de AEM	IES	AEM	taxa de AEM
Estratégia A	Central	6	129.914	0,8664	6	129.914	0,8723	6	129.015	0,8792
	Norte	1	2.707	0,0181	1	2.707	0,0182	1	2.707	0,0184
	Sudoeste	6	11.281	0,0752	6	10.283	0,0690	6	10.219	0,0696
	Sul	3	6.037	0,0403	3	6.037	0,0405	3	4.799	0,0327
Estratégia B	Central	6	131.581	0,8679	7	132.195	0,8796	8	133.981	0,8873
	Norte	1	2.707	0,0179	1	2.707	0,0180	1	2.707	0,0179
	Sudoeste	6	11.281	0,0744	5	9.345	0,0622	4	8.279	0,0548
	Sul	3	6.037	0,0398	3	6.037	0,0402	3	6.037	0,0400
Estratégia C	Central	6	131.641	0,8704	7	130.377	0,8756	7	134.027	0,8829
	Norte	1	2.707	0,0179	1	2.707	0,0182	1	2.707	0,0178
	Sudoeste	6	10.849	0,0717	5	9.783	0,0657	6	10.651	0,0702
	Sul	3	6.037	0,0399	3	6.037	0,0405	2	4.414	0,0291

5.4 DISCUSSÃO DOS RESULTADOS PARA AS MÉTRICAS DO GRUPO 2

Nesta seção serão discutidos e analisados os resultados obtidos para a métrica 5, descritos anteriormente.

5.4.1 Avaliação se as soluções obtidas combinando modelos de localização e estratégias citadas na Seção 1 fornecem melhor equalização da taxa de AEM de todas as mesorregiões

Para a métrica 5, a Tabela 5.4 mostra os valores da taxa de AEM para os três modelos PMM, PCM e PDM com as estratégias A, B e C, respectivamente. Como podemos observar, para a mesorregião Central, as IES existentes em Manaus, Parintins, Itacoatiara, Coari e Tefé representam o total de 128.501 AEM e a nova IES em Maués, 3.140 AEM, como mostra a Figura 5.6. As outras mesorregiões têm 19.593 AEM. Assim, o número de AEM e a taxa de AEM no primeiro são maiores que o anterior. Essa taxa de AEM da mesorregião Central diminui com o aumento das IESs nas outras mesorregiões.

As melhores taxas de AEM obtidas, conforme mostrado na Tabela 5.4, foram: a) para a mesorregião Norte, PDM com a configuração de parâmetros ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$), com taxa de AEM de 0,0184; b) para a mesorregião Sudoeste, PMM com as estratégias A, B, ou C, com taxas de AEM de 0,0752, 0,0744 e 0,0717, respectivamente; c) para a mesorregião Sul, o PCM com as estratégias A, B e C, respectivamente, com taxas de AEM e 0,0405, 0,0402 e 0,0405. Quando analisamos a população de AEM atendidas nas mesorregiões Sudoeste, Norte e Sul observamos que o modelo PMM obtém os melhores resultados. Por

exemplo, para a estratégia C, a população total atendida dessas mesorregiões é, respectivamente, de 19.593, 18.527 e 17.772, para os modelos PMM, PCM e PDM.

É importante notar que, para aumentar a taxa de AES das mesorregiões Sul, Sudoeste e Norte em relação à mesorregião Central, é necessário não apenas aumentar o número de IES nessas regiões, mas também aumentar o número de estudantes nas instituições de ensino superior localizadas nelas. No entanto, com cidades pequenas, para aumentar o número de AES, é altamente recomendável aumentar o número de cidades com IES. Aumentar o número de IES implica diminuir as distâncias entre cidades sem IES e cidades com IES, porque facilita o acesso dos estudantes. De fato, em todas as locações com 11, 12, 13, 14, 15 e 16 cidades, temos uma diminuição da métrica 1, como mostram as Figuras 5.1(a), 5.2(a) e 5.3(a) e uma diminuição da métrica 2 também, conforme mostrado nas Figuras 5.1(b), 5.2(b) e 5.3(b), entre uma cidade sem uma IES e uma cidade com uma IES.

5.5 RESULTADOS DAS MÉTRICAS DO GRUPO 3

As Figuras 5.4 e 5.5 mostram os valores das métricas 6 e 7, respectivamente, obtidos para os modelos PMM, PCM e PDM, para expansão das cidades com IES de 10 para 11, 12, 13, 14, 15 e 16 cidades, considerando as três configurações de parâmetros: $(\alpha = 0,5, \beta = 0, \gamma = 0,5)$, $(\alpha = 0,5, \beta = 0,5, \gamma = 0)$ e $(\alpha = 1/3, \beta = 1/3, \gamma = 1/3)$.

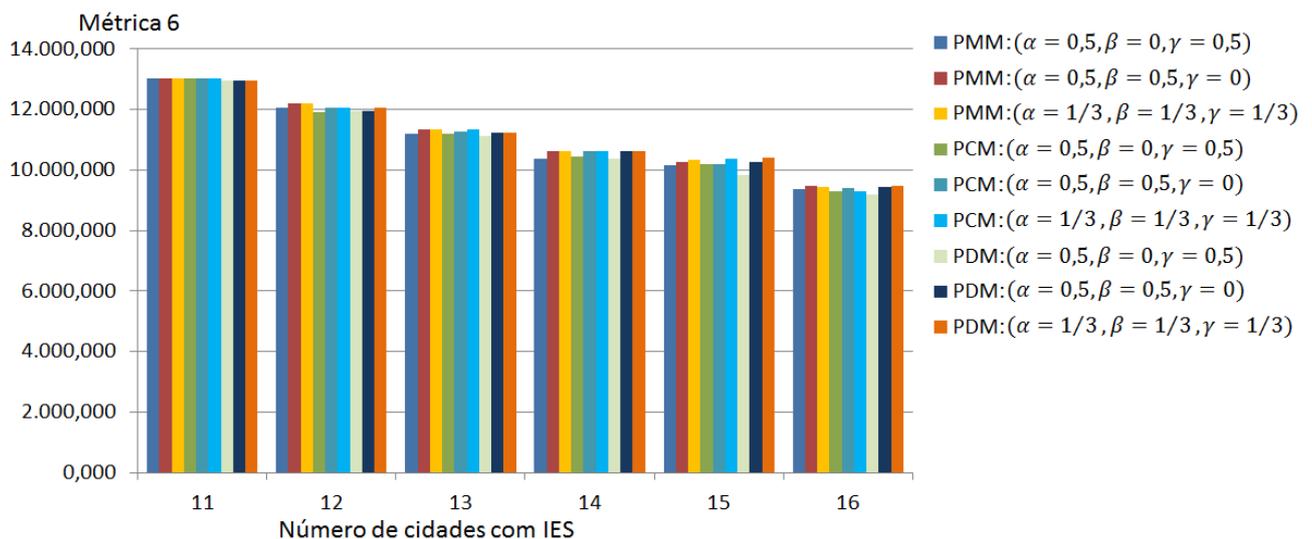


Figura 5. 4 – Valores para a métrica 6 obtidos para os modelos PMM, PCM e PDM, com a expansão de 11 a 16 cidades com IES.

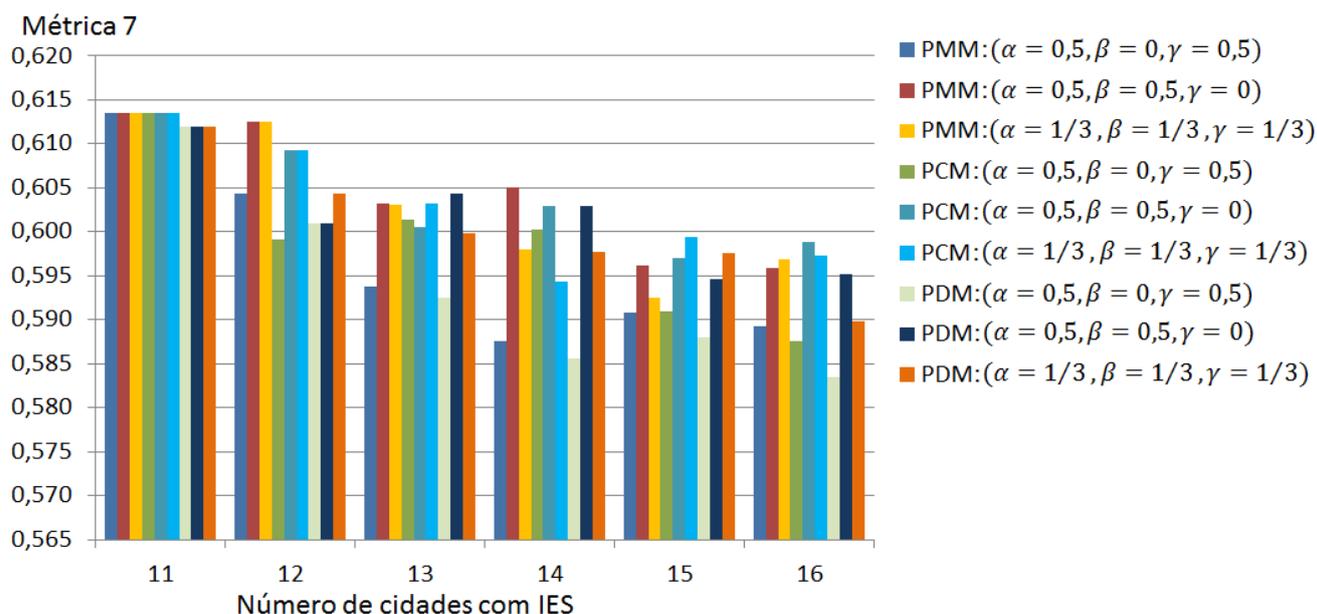


Figura 5. 5 – Valores da métrica 7 obtidos para os modelos PMM, PCM e PDM, com expansão de 11 para 16 cidades com IES.

5.6 DISCUSSÃO DOS RESULTADOS DAS MÉTRICAS DO GRUPO 3

Nesta seção serão discutidos e analisados os resultados obtidos para as métricas 6 e 7, descritos anteriormente.

5.6.1 Avaliação se as soluções obtidas combinando modelos de localização e algumas estratégias citadas na Seção 1 favorecem cidades com menor IDH ou cidades com maior número de AEM

A Figura 5.4 mostra os valores da métrica 6 obtidos para o PMM, PCM e PDM, para a expansão do número de cidades com IES de 10 para 11 a 16 cidades. Conforme mostrado nessa figura, se o tomador de decisão escolher uma das estratégias A, B ou C e um dos métodos de otimização, PMM, PDM ou PCM, os valores da métrica 6 serão quase os mesmos.

A Figura 5.5 mostra os valores da métrica 7 obtidos para o PMM, PCM e PDM, para a expansão do número de cidades com IES de 10 para 11 a 16 cidades. Conforme mostrado nessa figura, em algumas expansões, para obter valores mínimos da métrica 7, os tomadores de decisão devem selecionar um modelo PMM combinado com a estratégia A, enquanto em outras expansões devem selecionar o modelo PDM combinado com a estratégia C.

A Figura 5.6 mostra as cidades com IES para a expansão de 10 para 16 cidades com IES obtidas para o modelo PMM combinado com a estratégia C. Como podemos observar nesta figura, as novas cidades com IES são representadas por um símbolo +. Quatro novas

idades com instituições de ensino superior estão na mesorregião Sudoeste. Uma nova cidade com uma IES está na mesorregião Central e a outra na mesorregião Sul. A escolha da cidade de Maués, na mesorregião Central, é afetada pela equalização espacial, pois está próxima de grandes polos populacionais como Manaus, Itacoatiara e Parintins do Estado do Amazonas.

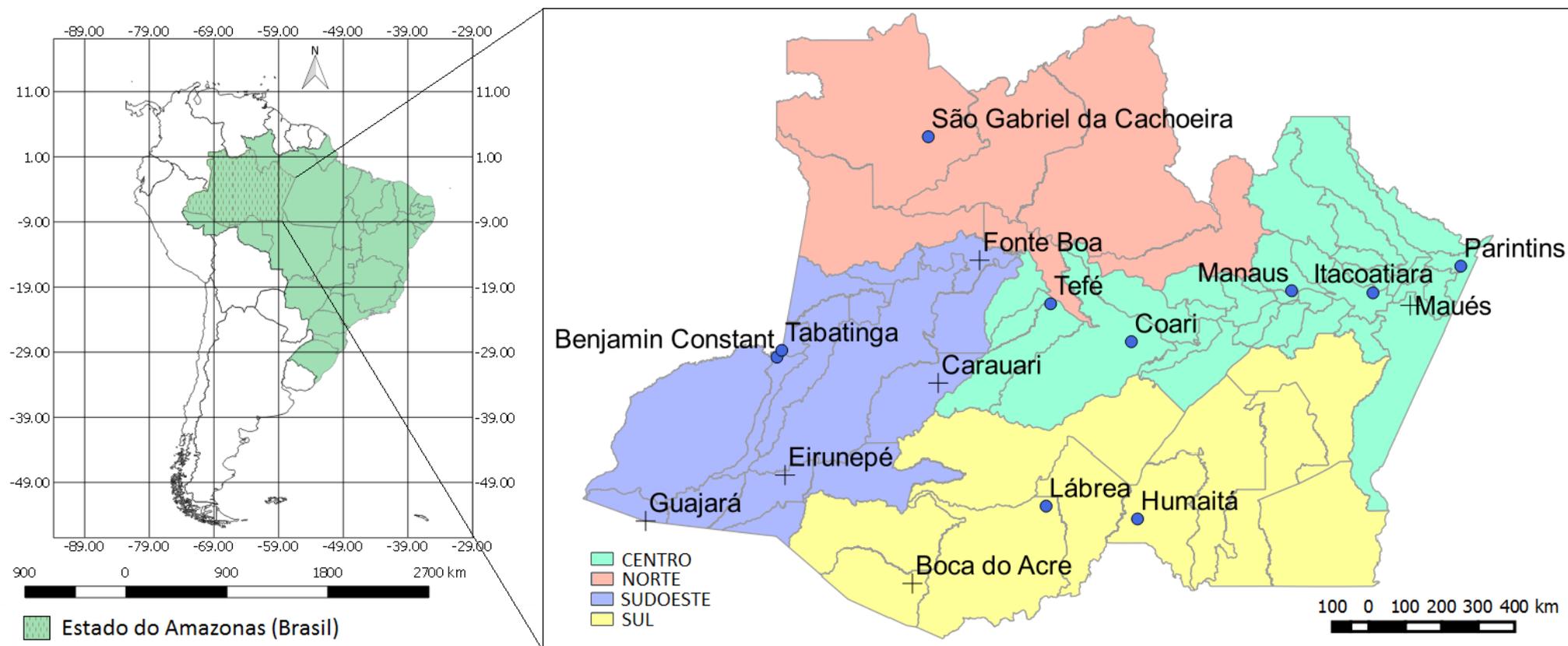


Figura 5. 6 – Resultados obtidos usando o modelo PMM combinado com a estratégia C para a expansão de 10 a 16 cidades com IES. *Shapefiles* – arquivos com dados geospaciais – obtidos através do IBGE (2016a), IBGE (2016b), IBGE (2019) e FORESTGIS (2020), respectivamente, para o mapa do Estado do Amazonas, mapa do Brasil, mapa das mesorregiões do Estado do Amazonas e mapa da América do Sul.

A partir da Figura 5.5, extraímos outra conclusão importante deste trabalho: todos os métodos de alocação usados neste trabalho, PMM, PCM e PDM, obtiveram, para a expansão de 10 para 11 a 16 cidades, uma diminuição dos valores da métrica 7, ressaltando mais municípios com o IDH médio. De fato, da Tabela 1.2, observamos que a distribuição existente apresenta um valor médio de IDH de 0,619. Todas as expansões mostradas na Figura 5.5 apresentam valores médios mais baixos de IDH em comparação com 0,619. Os melhores desempenhos neste trabalho foram obtidos com a parametrização ($\alpha = 0,5$, $\beta = 0$, $\gamma = 0,5$), estratégia A, ou com a parametrização ($\alpha = 1/3$, $\beta = 1/3$, $\gamma = 1/3$), estratégia C. O pior desempenho foi novamente obtido com o ajuste de parâmetros ($\alpha = 0,5$, $\beta = 0,5$, $\gamma = 0$), estratégia B.

Por fim, observamos que, para todas as expansões, de 10 para 11 a 16 cidades, há uma diminuição nos valores da métrica 6. Na Tabela 1.2, percebemos que, para a distribuição existente, o número médio de AEM é 14.170,40. Nas expansões de 10 para 11 a 16 cidades, o número médio de AEM variou de 13.018,91 a 9.171,25, para os três modelos. Portanto, embora o modelo propicie uma expansão para cidades com populações mais altas, houve uma diminuição no número médio de AEM, quando o número de cidades com IES aumenta. Com relação a esta questão, a seguinte consideração pode ser levada em conta: a distribuição de IES existente já favorece cidades com um número maior de AEM. Portanto, a inclusão de mais cidades com instituições de ensino superior tende a diminuir o número médio de AEM de cidades com instituições de ensino superior.

5.7 RESULTADOS OBTIDOS PELO ALGORITMO GENÉTICO

A respeito do GA empregado neste trabalho, os tempos de simulação usando os três modelos, para expansões de 10 para 11, 12, 13, 14, 15 e 16 cidade com IES, são mostrados na Tabela 5.3. Enfatizamos que, neste trabalho, usamos uma versão adaptada da solução do estado-da-arte do GA proposto por Silva et al. (2019).

Xavier et al. (2017) modelaram um GA para solucionar um modelo PMM multi-objetivo para localizar novas IES no estado do Amazonas que consistia de um cromossomo composto por 62 genes binários, tamanho da população igual a 300, variando o número de iterações. Esses autores alcançaram o melhor tempo de simulação para 10 IES com 275 segundos. O melhor tempo de simulação, neste trabalho, foi alcançado para o modelo PCM com 11 escolas, com um tempo de 0,006 segundos, conforme mostrado na Tabela 5.7. Como

podemos observar, os melhores tempos de simulação foram obtidos para o modelo PMM. Os piores tempos de simulação foram obtidos para o modelo PDM.

Tabela 5.3 – Tempos de simulação obtidos com a configuração de parâmetros ($\alpha = 1/3$, $\beta = 1/3$, $\gamma = 1/3$) para a expansão de 10 para 11, 12, 13, 14, 15 e 16 cidades com IES.

Método	Tempo de Simulação (Segundos)					
	Número de cidades com IES					
	11	12	13	14	15	16
PMM	0,0492	0,0284	0,027	0,6382	0,0873	0,2691
PCM	0,0663	0,0927	0,2523	0,1068	0,0887	0,4832
PDM	0,1252	1,477	2,0401	0,2337	0,2084	0,2395

Aplicando o teste t para avaliar as diferenças entre a distância fluvial, e os critérios AEM e IDH obtidos para a simulação de GA com tamanho inicial da população de 10, 20 e 30, com os três modelos de localização, concluímos que essas diferenças não são estatisticamente significativas com um nível de confiança de 90% (Com conjuntos de amostra diferentes).

5.7.1 Comentários sobre os Resultados do Algoritmo Genético

A Tabela 5.3 mostra os tempos de simulação usando os três modelos, para expansões de 10 a 11, 12, 13, 14, 15 e 16 cidades com IES, para a definição de parâmetros ($\alpha = 1/3$, $\beta = 1/3$, $\gamma = 1/3$). Como pode ser visto, os melhores tempos de simulação foram obtidos com o modelo PMM, expandindo para 13 IES em 0,027 segundos. Os piores tempos de simulação foram obtidos com o modelo PDM.

5.8 CONSIDERAÇÕES FINAIS

Neste capítulo mostramos e analisamos os resultados para a expansão de unidades de ensino superior no estado do Amazonas. A seguir mostramos uma síntese do que foi alcançado neste capítulo.

Para os resultados das métricas do grupo 1, foram discutidas as combinações dos métodos de localização e estratégias de expansão, relacionadas às métricas 1, 2 e 3, explanando ao tomador de decisão as melhores opções. Para a métrica 4, foi constatado que todos os métodos de localização obtém uma melhor taxa de IES para as mesorregiões Norte, Sudoeste e Sul, exceto para a mesorregião Central, comparada à taxa de IES existente.

Em relação aos resultados para a métrica do grupo 2, vimos que, para aumentar a taxa de AES, representada pela métrica 5, das mesorregiões Sul, Sudoeste e Norte em relação à mesorregião Central, é necessário aumentar o número de estudantes nas instituições de ensino superior localizadas nelas. No entanto, considerando cidades pequenas, é altamente recomendável aumentar o número de cidades com IES.

Quanto aos resultados para as métricas do grupo 3, constatamos que, para todas as expansões, de 10 para 11 a 16 cidades, há uma diminuição nos valores da métrica 6 e também uma diminuição dos valores da métrica 7, ressaltando mais municípios com o IDH médio.

Em relação aos resultados do Algoritmo Genético, observamos que os melhores tempos de simulação foram obtidos para o modelo PCM, e os piores tempos de simulação foram obtidos para o modelo PDM.

No próximo capítulo, serão apresentadas as conclusões sobre a pesquisa realizada neste trabalho, traçando um paralelo entre a mesma e os objetivos específicos propostos.

6. CONCLUSÕES E TRABALHOS FUTUROS

No capítulo anterior apresentamos os resultados obtidos, bem como uma interpretação dos principais conhecimentos derivados do desenvolvimento deste trabalho. Neste capítulo, vamos mostrar as limitações do nosso trabalho, contribuições mais significativas no que tange ao atendimento aos objetivos estipulados, bem como as opções para guiar a tomada de decisão.

Estruturalmente, este capítulo é constituído por duas seções. Na primeira, traçaremos as conclusões pertinentes a cada um dos objetivos propostos neste trabalho. Na segunda, avançamos em termos dos próximos passos com vistas ao prosseguimento deste trabalho.

6.1 CONCLUSÕES

A seguir discorreremos sobre as principais contribuições deste trabalho de pesquisa.

Em relação ao grupo 1, que inclui as métricas 1, 2, 3 e 4, os principais resultados para guiar o tomador de decisão são elencados a seguir. Quando se quer priorizar uma menor distância fluvial média entre uma cidade sem IES e uma cidade com IES (métrica 1), o tomador de decisão pode escolher entre os modelos PMM, PCM ou PDM combinados com a estratégia A. Quando se quer priorizar uma menor distância fluvial máxima entre uma cidade sem IES e uma cidade com IES (métrica 2), as melhores opções são os modelos PMM ou PCM combinados com a estratégia A ou PDM combinado com a estratégia B. Quando se quer priorizar uma maior distância fluvial máxima entre cidades com IES (métrica 3), pode ser usada a combinação do método PMM com as estratégias A, B ou C, ou a combinação do método PCM com a estratégia C, ou a combinação do método PDM com as estratégias A, B ou C. Quando se quer priorizar uma melhor taxa de IES (métrica 4), as melhores opções são o modelo PMM combinado com qualquer uma das estratégias A, B ou C, ou os modelos PCM ou PDM combinados com a estratégia A.

Quanto ao grupo 2, quando se quer priorizar uma melhor taxa de AEM (métrica 5), a melhor opção para o tomador de decisão é o modelo PMM combinado com qualquer uma das estratégias A, B ou C.

Considerando os resultados para ao grupo 3, para otimização das métricas 6 e 7, o tomador de decisão pode optar pelas seguintes escolhas. Quando se quer priorizar o AEM médio (métrica 6), o tomador de decisão poder selecionar um dos modelos PMM, PCM ou PDM combinado com as estratégias A, B ou C, uma vez que os valores da métrica 6 são semelhantes. Quando se quer priorizar o IDH médio (métrica 7), o tomador de decisão pode selecionar o modelo PMM combinado com a estratégia A, para algumas expansões, ou o modelo PDM combinado com a estratégia C, para outras expansões. Esses modelos de localização foram selecionados em atendimento ao Objetivo Específico 1. E as métricas para avaliar as soluções obtidas foram propostas em atendimento ao Objetivo Específico 2.

Com respeito ao tempo para obtenção das soluções de locação obtidas pelo GA – em atendimento ao Objetivo Específico 3 – para cada um dos métodos, o tomador de decisão pode escolher como melhor opção o método PMM. É importante enfatizar que utilizando diferentes tamanhos iniciais da população, não foi constatada diferença estatisticamente significativa com um nível de confiança de 90%.

A metodologia desenvolvida neste trabalho pode ser aplicada a qualquer problema de localização de recursos públicos – em atendimento ao Objetivo Geral desta pesquisa. Embora o problema de localizar recursos do governo, em geral, envolva uma dimensão política, isso não invalida uma análise técnica do problema, como fizemos neste trabalho. A análise aqui apresentada tem como principal objetivo contribuir para a avaliação de uma solução para localização de IES no estado do Amazonas, Brasil.

Gostaríamos também de salientar que este trabalho difere de outros trabalhos já publicados na literatura, porque o foco aqui é a expansão de um sistema existente de IES. Acreditamos que este estudo é novo no sentido de que é o único estudo que compara modelos clássicos de localização de instalações com múltiplos critérios.

Por fim, gostaríamos de salientar que a principal contribuição deste estudo ressalta a necessidade de desenvolvimento social das cidades com valores de IDH mais baixos e propomos uma expansão futura do sistema de ensino superior do Amazonas que favorece mesorregiões com pequenas concentrações de estudantes, impedindo um número enorme de estudantes em uma mesorregião com uma população maior.

Algumas limitações de nosso trabalho são as seguintes:

1. Um conjunto ótimo de pesos não é proposto para cada termo nas equações (4.4), (4.5) e (4.6);
2. Não é proposto nenhum tratamento de fronteira para cidades que poderiam ser atendidas por instituições de estados vizinhos;
3. Cidades com populações altas podem gerar agregação espacial (Erkut e Bozkaya, 1999);

Considerando a primeira limitação, mecanismos de pesos que procuram definir a importância relativa de um critério são categorizados em três grupos: 1) métodos subjetivos, em que o papel de atribuir a importância relativa é uma escolha do tomador de decisão; 2) métodos objetivos, em que o tomador de decisão não tem nenhum papel na determinação da importância do critério, e; 3) mecanismos de pesos combinando as duas abordagens anteriores (Jahan et al., 2016). Delmelle et al. (2018) utilizaram um mecanismo de peso como uma mera questão de escolha do tomador de decisão, variando os valores entre 0 e 1.

Experimentos realizados, ponderando os termos da medida de proximidade de Xavier et al. (2017), mostraram que altos valores de peso não melhoraram a qualidade da solução, com as melhores soluções obtidas para pesos iguais. No entanto, são necessários estudos mais detalhados para uma melhor conclusão a esse respeito. Assim, pretendemos fazer estudos sobre as principais técnicas utilizadas nos mecanismos de pesos na literatura.

Quanto à segunda limitação, considera-se o tratamento das condições de fronteira nos casos em quem as cidades possam se atendidas por instituições de ensino superior dos estados vizinhos. Essa abordagem foge ao escopo desta pesquisa. É importante ressaltar que o tipo dos cursos ofertados pela IES não é considerado nessa pesquisa, sendo uma decisão política e administrativa.

Em relação à terceira limitação, gostaríamos de lembrar, conforme já expresso, que dois termos nas equações (4.4), (4.5) e (4.6) são conflitantes: população e IDH. O conflito surge porque cidades muito pequenas têm IDH mais baixos, no caso do estado do Amazonas. Esse conflito minimiza o problema de agregação espacial observado em outros trabalhos (Erkut e Bozkaya, 1999; Francis et al. 1999). Em outras palavras, minimiza o perigo de que as medianas locais nas soluções p-medianas sejam localizadas onde quer que existam centros específicos, com suas populações com mais habitantes que outros centros.

Especificamente, a agregação é a representação por um único ponto de um conjunto de pontos de demanda que estão próximos um do outro espacialmente. Certamente, isso leva a uma redução do problema. Contudo, também introduz erros no modelo devido à inerente

perda de informação. Para uma determinada solução, o valor da função objetivo (distância total percorrida) é calculado incorretamente devido à agregação de vários pontos de demanda em um ponto. Esses erros podem resultar na seleção de uma solução não ótima como na de uma solução ótima (Erkut e Bozkaya, 1999).

6.2 TRABALHOS FUTUROS

Como trabalhos futuros, vislumbramos três linhas de continuidade:

- 1) Estudo para definição de um conjunto ótimo de pesos para cada termo das equações (4.4), (4.5) e (4.6);
- 2) Avaliação ou acréscimo de outros critérios nas equações (4.4), (4.5) e (4.6), respectivamente, em substituição ao IDH.
- 3) Expansão do trabalho ora desenvolvido, para inclusão de um tratamento de fronteira para as cidades que poderiam ser atendidas por instituições de estados vizinhos;
- 4) Estudo de uma modelagem que propicie um tratamento das cidades com populações altas, que podem gerar agregação espacial.

REFERÊNCIAS BIBLIOGRÁFICAS

AMAYA, J.; PEETERS, D.; URIBE, P.; and VALENZUELA, J. P. **Optimization Modeling for Resource Allocation in the Chilean Public Education System**. International Regional Science Review April 2016, vol. 39, no. 2, pp. 155-176, first published on February 24, 2015.

AGÊNCIA NACIONAL DE ÁGUAS (ANA): **HidroWeb – Sistemas de Informações Hidrológicas. Shapefiles Bacia Amazônica**. Disponível em: <<http://hidroweb.ana.gov.br/baixar/mapa/Bacia1.zip>>. Acessado em: 05 de Maio de 2016. Brasil, 2016.

ALP, O.; ERKUT, E.; and DREZNER, Z. **An Efficient Genetic Algorithm for the p-Median Problem**. Annals of Operations Research. 122. pp.21-42, 2003.

AGÊNCIA NACIONAL DE TRANSPORTES AQUAVIÁRIOS (ANTAQ): **Caracterização da oferta e da demanda do transporte fluvial de passageiros da região amazônica**. Brasília, Brasil, 2013.

BEASLEY, J.E. (1990). **OR-Library – Distributing Test Problems by Electronic Mail**. Journal of the Operational Research Society 41(11), 1069–1072.

BELTRAN, C.; TADONKI, C. and VIAL, J. **Solving the p-median Problem with a Semi-Lagrangian Relaxation**, 2004.

BRUNO, G.; GENOVESE, A.; PICCOLO, C; and STERLE, C. **A Location Model for the Reorganization of a School System: The Italian Case Study**. Procedia - Social and Behavioral Sciences, Volume 108, 8 January 2014, Pages 96-105.

BUSSAB, W. O., & MORETTIN, P. A. (2009). **Estatística Básica**. Editora Saraiva, 6ª Edição. (Seção 3.1: Medidas de Posição), 2009.

CHEN, M.; THILL, J. C.; and DELMELLE, E. **iGLASS: An Open Source SDSS for Public School Location-Allocation**. In: Thill JC., Dragicevic S. (eds) GeoComputational Analysis and Modeling of Regional Systems. Advances in Geographic Information Science. Springer, Cham, 2018.

CHONG, E. K. P.; and ZAK, S. H. (2013). **An Introduction to Optimization**. (4th Ed.). Canada. John Wiley & Sons. (Chapter 14: Global Search Algorithms).

CHIYOSHI, F.; and GALVÃO, R. D. **A Statistical Analysis of Simulated Annealing Applied to the p-Median Problem**. Annals of Operations Research 96, 61–74, 2000.

CORMEN, T. H.; LEISERSON, C. E.; RIVEST, R. L.; and STEIN, C. **Introduction to Algorithms**. Third Edition. The MIT Press, Massachusetts Institute of Technology, USA, 2009.

DASKIN, M. **Network and Discrete Location: Models, Algorithms, and Applications**, Second Edition. John Wiley & Sons, 2013.

DASKIN, M. S.; MAASS, K. L. **The p-Median Problem**, Chapter Location Science, pp 21-45, 2015.

DEB, K. **Multi-Objective Optimization Using Evolutionary Algorithms**. John Wiley & Sons, New Yor, USA, 2001.

DELMELLE, E.; DESJARDINS, M. R.; and DENG, J. **Designing Spatially Cohesive Nature Reserves with Backup Coverage**, International Journal of Geographical Information Science, 31:12, 2505-2523, 2017.

DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES (DNIT): **Shapefiles Mapas Multimodais Rodovias**. Disponível em: <<https://189.9.128.64/mapas-multimodais/shapefiles/shapefiles.rar>>. Acessado em: 02 de maio de 2016. Brasil, 2016.

DREZNER, Z. and HAMACHER, H. W. **Facility Location: Applications and Theory**. Springer, 2002.

DREZNER, Z.; BRINBERG, J., MLADENOVIC, N. and SALHI, S. **New Heuristic Algorithms for Solving the Planar p-median Problem**. Computational Operations Research, 62:296–304, 2015.

ERKUT, E.; and BOZKAYA, B. **Analysis of Aggregation Errors for the p-median Problem**. Computers & Operations Research, 26:1075–1096, 1999.

FERNANDO, P. R.; KATKOORI, S.; KEYMEULEN, D.; ZEBULUM, R.; STOICA, A. **Customizable FPGA IP Core Implementation of a General-Purpose Genetic Algorithm Engine**. IEEE Transactions of Evolutional Computation, 14:133–149, 2010.

FORESTGIS – BASE DE DADOS SHAPEFILE DO BRASIL. Mapa da América do Sul. Disponível em: <https://sites.google.com/site/joabelb/Home/America_do_Sul.zip?attredirects=0&d=1>. Acessado em 26 de março de 2020.

FOTHERINGHAM, A. S.; DENSHAM, P. J. and CURTIS, A. **The Zone Definition Problem in Location-Allocation Modeling**. Geographical Analysis, Volume 27, issue 1, pages 60–77, January 1995.

FRANCIS, R. L.; LOWE, T. J.; RUSHTON, G.; and RAYCO, M. A. **Synthesis of Aggregation Methods for Multifacility Location Problems: strategies for containing error**. Geographical Analysis,, 31:67–87, 1999.

GAC, I.; MARTÍNEZ, F.; and WEINTRAUB, A. **A deterministic linear optimization model for allocating schools to zones**. Journal of the Operational Research Society, Volume 60, pp.895-905, July 2009.

HAKIMI, S. **Optimal location of switching centers and the absolute centers and medians of a graph**. Operations Research, 12, 450-459, 1964.

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

HOLLAND, J. H. 1975. **Adaptation in Natural and Artificial Systems**. University of Michigan Press. (Second edition: MIT Press, 1992.)

INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo da Educação Superior 2018**. Disponível em: <http://download.inep.gov.br/educacao_superior/centso_superior/documentos/2019/apresentacao_censo_superior2018.pdf>, pg. 6. Acessado em 01 de fevereiro de 2020.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE): **XII Censo Demográfico**. Brasil, 2010.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE): **Shapefiles Unidades da Federação**. Disponível em: <ftp://geoftp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_municipais/municipio_2015/Brasil/BR/br_unidades_da_federacao.zip>. Acessado em: 28 de março de 2020. Brasil, 2016a.

_____. **Shapefiles Malhas Municipais do Estado Amazonas**. Disponível em: <ftp://geoftp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_municipais/municipio_2014/AM/am_municipios.zip>. Acessado em: 02 de maio de 2016. Brasil, 2016b.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE): **Alunos no Ensino Médio**. Disponível em: <https://cidades.ibge.gov.br/brasil/am/panorama>. Acessado em: 06 julho de 2019.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE): **Shapefiles Malhas Municipais do Estado Amazonas**. <ftp://geoftp.ibge.gov.br/organizacao_do_territorio/malhas_territoriais/malhas_municipais/municipio_2015/UFs/AM/am_microrregioes.zip>. Acessado em: 10 de maio de 2019. Brasil, 2019.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE): **Pesquisa Nacional por Amostra de Domicílios Contínua - PNAD Contínua**. <ftp://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Nacional_por_Amostra_de_Domicilios_continua/Anual/Educacao_2018/PNAD_Continua_2018_Educacao.xls>. Acessado em: 15 de março de 2020. Brasil, 2018.

JAHAN, A.; EDWARDS, K.; BAHRAMINASAB, M. **Multi-criteria Decision Analysis for Supporting the Selection of Engineering Materials in Product Design**. Elsevier, Chapter 5. Multi-attribute decision-making for ranking of candidate materials, pg. 81-124, 2016.

KARATAS, M.; N. RAZI, N.; and TOZAN, H. **A Multi-Criteria Assessment of the p-Median, Maximal Coverage and p-Center Location Models**. Tehnicki Vjesnik, 2:399–407, 2017.

KARIV, O.; and HAKIMI, S. L. **An algorithmic approach to network location problems. II: the p -medians**. *SIAM J Appl Math* 37:539–560, 1979a.

KARIV, O.; and HAKIMI, S. L. **An Algorithmic Approach to Network Location Problems. I: The p -Centers**. *SIAM Journal on Applied Mathematics*, 37, 513–538, 1979b.

KUBY, M. (1987). **Programming Models for Facility Dispersion: The p -Dispersion and Maxisum Dispersion Problems**. *Geographical Analysis*, 19(4), 315-329.
<https://doi.org/10.1111/j.1538-4632.1987.tb00133.x>

LEE, J.; SHALABY, A.; and ABDULHAI, B. **Optimized Strategy for Integrated Traffic and Transit Signal Control**. In Sayed, T, Masad, E, Al-Qadi, I L, & Alnuaimi, N A (Eds.) *Efficient Transportation and Pavement Systems: Characterization, Mechanisms, Simulation, and Modeling*, 2009.

LOTFIAN, R.; and NAJAFI, M. **Optimal Location of Emergency Stations in Underground Mine Networks using a Multiobjective Mathematical Model**. *Injury prevention: journal of the International Society for Child and Adolescent Injury Prevention*, 2018.

LOPEZ, I. C; and OSPINA, H. A. L. **School location and capacity modification considering the existence of externalities in students school choice** *Journal Computers and Industrial Engineering archive*, Volume 80 Issue C, February 2015, Pages 284-294.

MENEZES, R. F.; and PIZZOLATO, N. D. **Locating public schools in fast expanding areas: application of the capacitated p -median and maximal covering location models**. *Pesqui. Oper. Rio de Janeiro*, v. 34, n. 2, p. 301-317, 2014.

MITCHEL, M. **An Introduction to Genetic Algorithms**. A Bradford Book The MIT Press Cambridge, Massachusetts. London, England Fifth printing, 1999.

MOGHADAM, A. M.; PIROOZFARD, H.; MAARAM, A. B.; and MIRZAPOUR, S. A. **Solving a Capacitated p -Median Location Allocation Problem Using Genetic Algorithm: a Case Study**. In *Advanced Materials Research*, Vol. 845, pp. 569-573, 2014.

MÜLLER, S.; HAASE, K.; and KLESS, S. **A Multiperiod School Location Planning Approach with Free School Choice**. *Environment and Planning A* December 2009 41: 2929-2945,

NDIAYE, F.; NDIAYE, B.; and LY, I. **Application of the p -Median Problem in School Allocation**. *American Journal of Operations Research*, Vol. 2 No. 2, 2012, pp. 253-259.
OECD. **Education at a Glance 2018: OECD Indicators**. OECD Publishing, Paris,
<https://doi.org/10.1787/19991487> , 2018.

PASCHOS, V. T. **Paradigms of Combinatorial Optimization: Problems and New Approaches**, vol. 2, pp. 75, 2013.

- PLASTRIA, F. **Static competitive facility location: An overview of optimisation approaches**, European Journal of Operational Research, Volume 129, Issue 3, 2001, Pages 461-470.
- RAO, S. S. **Engineering Optimization**. Theory and Practice, 4th Edition, John Wiley & Sons, Inc.
- RAVI, S.S.; ROSENKRANTZ, D. J.; TAYI, G. K. **Facility dispersion problems: Heuristics and special cases**. In: Dehne F., Sack JR., Santoro N. (eds) Algorithms and Data Structures. WADS 1991. Lecture Notes in Computer Science, vol 519. Springer, Berlin, Heidelberg, 1991.
- REID, R. D.; and SANDERS, N. R. **Operations Management. An Integrated Approach Fourth Edition**, John Wiley & Sons, Inc, 2011.
- REINELT, G. **TSLIB Traveling Salesman Library**. ORSAJ Comput. 1991; 3: pp. 376–384.
- REVELLE, C. S; and SWAIN, R. W. **Geographical Analysis**, Volume 2, Issue 1, pages 30–42, January 1970.
- ROEVA, O.; FIDANOVA, S.; PAPRZYCKI, M. **Influence of the Population Size on the Genetic Algorithm Performance in Case of Cultivation Process Modelling**. Federated Conference on Computer Science and Information Systems, FedCSIS 2013. pp.371-376, 2013.
- SALHI, S. **Defining Tabu List Size and Aspiration Criterion within Tabu Search Methods**. Computers & Operations Research, 29:67–86, 2002.
- SASTRY, K.; GOLDBERG, D.; and KENDALL, G. **Genetic algorithms**. In Edmund K. Burke and Graham Kendall, editors, Search Methodologies - Introductory Tutorials in Optimization and Decision Support Techniques, chapter 4, pages 97–125. Springer Science, Business Media, Inc., 2005.
- SATOGLU, S.; OKSUZ, M.; KAYAKUTLU, G.; and BUYUKOZKAN, K. A. **Genetic Algorithm for the p-Median Facility Location Problem**. GJCI2016 – Global Joint Conference on Industrial Engineering, Istanbul, 2016.
- SEMESP - Sindicato das Mantenedoras de Ensino Superior do Estado de São Paulo. **Mapa da Educação Superior do Brasil**, São Paulo, 2018.
- SHEHADEH, H.; AHMEDY, I.; and IDRIS, M. **Empirical Study of Sperm Swarm Optimization Algorithm**. Volume 869 of the Advances in Intelligent Systems and Computing series, Edition: Volume 2, Chapter: K. Arai et al. (Eds.): IntelliSys 2018, AISC 869, 2019., Springer International Publishing Switzerland, pp. 1082-1104, 2019.
- SILVA, S. D. de S., COSTA, M. G. F. and COSTA FILHO, C. F. F. **Customized Genetic Algorithm for Facility Allocation using p-median**, 2019 Federated Conference on Computer Science and Information Systems (FedCSIS), Leipzig, Germany, 2019, pp. 165-169. doi: 10.15439/2019F158

SPIEGEL, M. R.. **Estatística**. Makron Books, 3^a Edição, São Paulo, 1994.

SRINIVAS, N.; and DEB, K. **Muiltiobjective Optimization Using Nondominated Sorting in Genetic Algorithms**. in *Evolutionary Computation*, vol. 2, no. 3, pp. 221-248, Sept. 1994.

TEIXEIRA, J. C.; ANTUNES, A. P; and PEETERS, D. **An Optimization-Based Study on the Redeployment of a Secondary School Network**, Environment and Planning B: Planning and Design April 2007 34: 296-315, doi:10.1068/b31173.

TEIXEIRA, J. C.; and ANTUNES, A. P. **A hierarchical location model for public facility planning**, European Journal of Operational Research, Volume 185, Issue 1, 16 February 2008, Pages 92-104, ISSN 0377-2217, <http://dx.doi.org/10.1016/j.ejor.2006.12.027>.

TURKOGLU, D. C; and GENEVOIS, M. D. **A Comparative Survey of Service Facility Location Problems**. Ann Oper Res (2019), <https://doi.org/10.1007/s10479-019-03385-x>.

TUTUNCHI, G. K.; and FATHI, Y. **Effective Methods for Solving the Bicriteria p-Center and p-Dispersion Problem**. Comput. Oper. Res., vol. 101, pp. 43-54, Jan. 2019.

UNDP – Indicadores e Índice de Desenvolvimento Humano.: *Statistical update*, UN, New York, <https://doi.org/10.18356/656a3808-en>, pg. 11, 2018.

UNIVERSIDADE DO ESTADO DO AMAZONAS (UEA). **Plano de Desenvolvimento Institucional 2012-2016**. Disponível em: <<http://www.pdi.uea.edu.br/data/area/c33/download/1-1.pdf>>. Acessado em: 10 de Agosto de 2016.

UNIVERSIDADE FEDERAL DO ESTADO DO AMAZONAS (UFAM). **Plano de Desenvolvimento Institucional 2006-2015**. Disponível em: <http://www.proplan.ufam.edu.br/PDI_06a15.pdf>. Acessado em: 08 de Agosto de 2016.

XAVIER, C. M.; COSTA, M. G. F. FILHO, C. F. F. C. **Optimizing Universities Allocation with Integer Programming using a New Distance Measure**. in Proc. 5th Int. Conf. Eng. Optim. (EngOpt), e-Papers, Rio de Janeiro, vol. 1, pp. 1–10, 2016.

XAVIER, C. M; COSTA, M. G.F; FILHO, C. F. F. C. **A New Multi Objective Approach for Optimizing p-median Modeling in School Allocation using Genetic Algorithm**. In: SEMINÁRIO INTEGRADO DE SOFTWARE E HARDWARE (SEMISH), 44. , 2017, São Paulo. Anais do XLIV Seminário Integrado de Software e Hardware. Porto Alegre: Sociedade Brasileira de Computação, july 2017 . ISSN 2595-6205.

XU, R.; WUNSCH, D. C. **Clustering**. IEEE Press Series on Computational Intelligence, 2008.

WHITAKER, R. A. **Fast Algorithm for the Greedy Interchange for Large-Scale Clustering and Median Location Problems**. INFOR, 21:95–108, 1983.

ZAPFEL, G.; BRAUNE, R.; BOGL, M. **Metaheuristic Search Concepts. A Tutorial with Applications to Productions and Logistics**. Springer, New York, 2010.

APÊNDICE A**LISTA DE EVENTOS E PUBLICAÇÕES**

A.1 ARTIGO EM PERIÓDICO

1. **“Combining Facility-Location Approaches for Public Schools Expansion”**, publicado no periódico *IEEE Access* – Qualis A1 (Fator de impacto de 4,098 – 2018) (Xavier et al., 2020);

A.2 CONFERÊNCIA NACIONAL

1. **“A New Multi Objective Approach for Optimizing p-median Modeling in School Allocation using Genetic Algorithm”**, trabalho apresentado no XLIV Seminário Integrado de Software e Hardware – Sociedade Brasileira de Computação 2017 – Qualis B4 (de 02 a 04 de julho de 2017, São Paulo, SP - Brasil) (Xavier et al., 2017);

A.3 CONFERÊNCIA INTERNACIONAL

1. **“Optimizing Universities Allocation with Integer Programming Using a New Distance Measure”**, resumo e artigos aceitos e o trabalho apresentado na 5th *International Conference on Engineering Optimization* – Qualis B4 (de 19 a 23 de junho de 2016, Foz do Iguaçu, PR - Brasil) (Xavier et al., 2016).