



ALGORITMO GENÉTICO PARA RESOLUÇÃO DO PROBLEMA DA LOCALIZAÇÃO DE CENTROS PÚBLICOS DE EDUCAÇÃO INFANTIL

Kellen Dayelle Endler

Universidade Federal do Paraná - UFPR
Avenida Coronel Francisco H. dos Santos, 100, Curitiba - PR
kellen_dayelle@hotmail.com

Cassius Tadeu Scarpin

Universidade Federal do Paraná - UFPR
Avenida Coronel Francisco H. dos Santos, 100, Curitiba - PR
cassiusts@gmail.com

Maria Teresinha Arns Steiner

Pontifícia Universidade Católica do Paraná - PUC-PR
Rua Imaculada Conceição, 1155, Curitiba - PR
maria.steiner@pucpr.br

RESUMO

A educação e o cuidado na primeira infância vêm sendo tratados, cada vez mais, como assuntos prioritários em âmbito nacional. O objetivo deste artigo é apresentar uma estratégia de solução do problema da localização das creches públicas para o município de Curitiba-PR baseada em Algoritmo Genético (AG). Utilizou-se para avaliar a distribuição espacial dessas unidades, o modelo das p -medianas não capacitado. Tal problema consiste num aspecto importante da administração pública, podendo atenuar problemas de desigualdades observadas no ensino público, em termos de acessibilidade e disponibilidade de vagas, por exemplo. O modelo de programação matemática foi gerado utilizando-se a linguagem de modelagem VB.NET[®] e resolvido pelo solver CPLEX[®]. Os resultados de validação do AG indicaram que o método proposto neste trabalho apresenta soluções satisfatórias. O algoritmo proposto destaca-se pela qualidade de soluções para instâncias de maior porte (com 3282 vértices), por apresentar *gaps* consideravelmente inferiores às meta-heurísticas da literatura.

PALAVRAS CHAVE. Algoritmo Genético, Localização, Rede pública de educação.

Tópicos (PO na Administração Pública, Meta-heurísticas)

ABSTRACT

Education and care in early childhood have increasingly been treated as priority issues at the national level. The purpose of this article is to present a strategy to solve the location problem of the public day care centers in Curitiba-PR based on Genetic Algorithm (GA). In order to evaluate the spatial distribution of these units, the model of non-trained p -medians is used. Such a problem is an important aspect of public administration and can alleviate inequalities observed in public education, in terms of accessibility and availability of places, for example. The mathematical programming model was generated using the VB.NET[®] modeling language and solved by the CPLEX[®] solver. The validation results of the GA indicated that the method proposed in this work presents satisfactory solutions. The proposed algorithm stands out for the quality of solutions for larger instances (with 3282 vertices), because it has *gaps* considerably lower than the meta-heuristics in the literature.

KEYWORDS. Genetic Algorithm, Location, Public education network.

Paper topics (OR in Public Administration, Metaheuristics)



1. Introdução

A educação infantil é amplamente reconhecida como a base para a formação integral do cidadão, “em seus aspectos físico, psicológico, intelectual e social, complementando a ação da família e da comunidade” (artigo 29 da LDB – Lei de Diretrizes e Bases da Educação Nacional). A educação é duplamente protegida pela Constituição Federal Brasileira de 1988: tanto é direito das crianças como é direito dos trabalhadores em relação a seus filhos e dependentes.

O PNE (Plano Nacional de Educação), aprovado pela Lei nº 13.005/2014, visa a ampliação da oferta de Educação Infantil em Creches de forma a atender, no mínimo, 50% das crianças de até 3 anos até o final de 2024. Um indicador da Pesquisa Nacional por Amostra de Domicílios do Instituto Brasileiro de Geografia e Estatística (PNAD/IBGE), mostra que desde 2005 há um crescimento constante na porcentagem dessas crianças na Educação Infantil, atingindo a marca de 29,6% em 2014. Em números absolutos, isso significa que mais de 3,5 milhões de crianças estão em creches no Brasil. No entanto, essa melhora não pode ser considerada suficiente para o cumprimento da meta considerando que a porcentagem continue crescendo no ritmo dos últimos anos avaliados.

Observando a porcentagem de atendimento de crianças em creches atingida pelas regiões e Estados brasileiros, o Estado de Santa Catarina obteve o melhor resultado (44,6%), seguido por São Paulo (40,2%) e Paraná (35,2%). Especificamente o município de Curitiba, a partir dos dados do Censo Demográfico [2010], apresentou uma taxa de 39,5% de atendimento das crianças de 0 a 3 anos. Diante disso, uma importante decisão estratégica competente ao poder público diz respeito à melhor localização de novas instalações, considerando as já existentes, visando garantir o número de vagas necessário, assim como facilitar a acessibilidade do local.

O objetivo deste artigo é apresentar uma estratégia de solução do problema da localização das creches públicas de Curitiba-PR baseada em Algoritmo Genético (AG). Utilizou-se, para avaliar a distribuição espacial dessas unidades, o modelo das p -medianas não capacitado. A análise da localização ideal implica na análise da realocação das creches existentes em suas localizações ideais para comparação com o cenário existente. Esta análise é apresentada e utilizada por outros trabalhos, como os mencionados em Pizzolato, et al. [2004].

O estudo promove racionalização da tomada de decisão para a avaliação locacional da rede pública de educação. Assim, outros municípios podem se beneficiar da aplicação da metodologia apresentada. Principalmente os municípios que apresentarem um grande desequilíbrio entre oferta e demanda de vagas e que precisam, portanto, avaliar as melhores localizações por meio de estratégias de ampliação ou criação de novas vagas.

Na sequência, o artigo está organizado como segue: a Seção 2 apresenta os passos do AG e trabalhos relacionados e a metodologia proposta. A Seção 3 apresenta a descrição do AG proposto e os resultados obtidos com sua aplicação em Curitiba-PR, e, por fim, a Seção 4 apresenta as considerações finais e sugestões para futuros trabalhos.

2. Descrição do Problema e Metodologia

Os AG foram introduzidos por Holland [1975] através de seu livro *Adaptation in Natural and Artificial Systems*. Eles representam técnicas de busca baseadas na Teoria da Evolução por meio da seleção natural, apresentada por Charles Darwin em 1858. Goldberg [1989] publicou *Genetic Algorithms in Search, Optimization and Machine Learning* que, juntamente com a obra de Holland [1975], são consideradas as obras pioneiras sobre este tema.

A meta-heurística AG associa-se aos mecanismos de evolução e representação gênica de indivíduos para a resolução de problemas combinatórios de otimização. Os genes são as unidades básicas dos cromossomos e estabelecem as características dos indivíduos. No contexto dos AG, “Aptidão” (ou *Fitness*) é um atributo utilizado na literatura para qualificar os indivíduos e definir a sua permanência ou exclusão da população entre gerações sucessivas [Corrêa et al., 2004]. Em geral, esse termo está associado à expressão matemática da função objetivo do modelo que representa o problema. O algoritmo, proposto por Holland, é conhecido na literatura como *Simple*



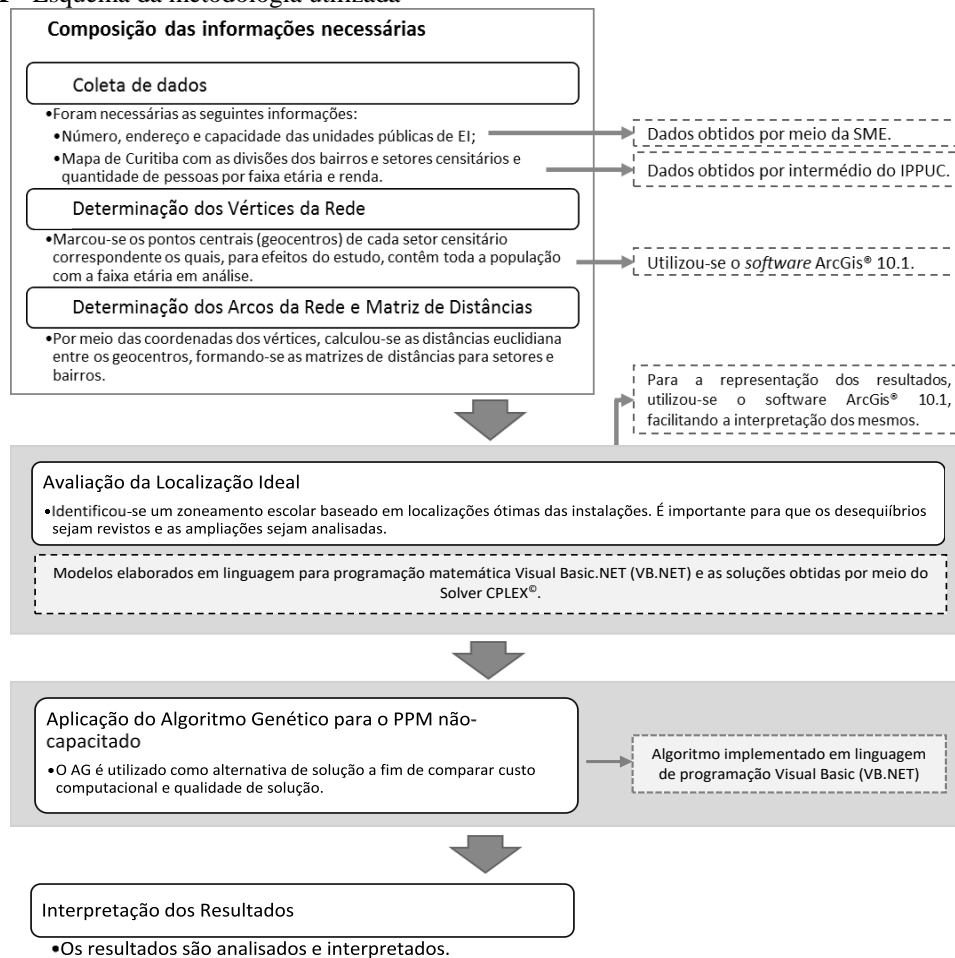
Genetic Algorithm ou *Standard Genetic Algorithm* ou simplesmente pela sigla SGA. Pode-se descrever o algoritmo, sucintamente, em seis passos [Davis 1991]:

- Passo 1.** Inicie uma população, de tamanho N , com cromossomos gerados aleatoriamente;
- Passo 2.** Aplique a função de adequação em cada cromossomo desta população;
- Passo 3.** Crie novos cromossomos através de cruzamentos de cromossomos selecionados desta população. Aplique recombinação e mutação nestes cromossomos;
- Passo 4.** Elimine membros da antiga população, de modo a ter espaço para inserir estes novos cromossomos, mantendo a população com o mesmo número N de cromossomos;
- Passo 5.** Aplique a função de adequação nestes cromossomos e insira-os na população;
- Passo 6.** Se a solução ideal for encontrada ou, se o tempo (ou número de gerações) se esgotou, retorne o cromossomo com a melhor adequação. Caso contrário, volte ao Passo 3.

Esta simulação do processo evolutivo pretende produzir, à medida que as gerações forem se sucedendo, cromossomos cada vez mais bem adaptados, isto é, com melhor valor da função de adequação, de maneira que, no final, obtém-se uma solução (cromossomo) com alto grau de adequação ao problema proposto.

A metodologia utilizada neste trabalho está dividida em 7 etapas ilustradas de forma detalhada na Figura 1.

Figura 1 - Esquema da metodologia utilizada



Os modelos matemáticos de localização citados foram implementados em um computador Core i5-2400, 8GB de RAM, usando linguagem Visual Basic.NET. Os resultados foram obtidos através do solver CPLEX® 6.5.1. Para facilitar a interpretação dos resultados, utilizou-se o software ArcGis® 10.1 na representação gráfica dos mesmos.



3. Resultados

3.1 Composição das Informações Necessárias

Para o estudo da distribuição espacial das creches públicas, foi preciso discretizar a distribuição da população em microrregiões chamados setores censitários. Esta discretização possibilitou os estudos normativos sobre a mesma. Portanto, o mapa da região a ser estudada deve estar dividido nos setores censitários (referenciados a partir deste ponto como setor ou setores), que são pequenas divisões definidas pelo IBGE (Instituto Brasileiro de Geografia e Estatística), que facilitam a realização do Censo. Os dados necessários para a condução deste estudo foram obtidos a partir das seguintes fontes:

- Mapa de Curitiba-PR com as divisões dos bairros e setores censitários – dados obtidos através do Instituto de Pesquisa e Planejamento Urbano de Curitiba (IPPUC);
- Quantidade de crianças de 0 a 3 anos em cada setor censitário – obtida pelo Censo do IBGE [2010].
- Endereço e capacidade das creches públicas – dados obtidos por meio da Secretaria Municipal de Educação (SME).

Para cada setor censitário é preciso localizar seu respectivo centro geométrico, ou seja, um ponto representativo do centro demográfico do setor. Local no qual assume-se que toda a população se encontra, já que não há como conhecer o exato posicionamento de cada aluno.

Um importante fator a ser considerado devido ao seu impacto na utilização da rede pública de educação consiste no número instituições privadas existentes. Por ocasião da pesquisa, o Ministério da Educação estimava que as unidades privadas representavam 54,60% do total de unidades existentes. Deste modo, foi adotada a seguinte simplificação: assume-se que a rede privada de ensino contribui com 50% do atendimento das vagas necessárias. O ajuste no valor da demanda mencionado é considerado nas análises consequentes, resultando em uma demanda a ser suprida de 43.647 crianças.

3.2 Avaliação da Localização Ideal

Esta fase de análise da localização ideal ou ótima é obtida pela solução do modelo das p -medianas apresentado de (1) a (5). Sob a terminologia da Teoria dos Grafos, [Corrêa et al. 2004] associam o Problema das p -medianas Capacitado (PPMC) a um grafo não direcionado $G=(V,A)$ cujo conjunto de vértices V contém os n pontos de demanda interligados pelos arcos do conjunto A . O objetivo do problema é identificar um conjunto de vértices V_p ($V_p \in V$) selecionados como medianas, de modo que a soma dos custos entre os vértices de $\{V-V_p\}$ e os vértices correspondentes (ou associados) de $\{V_p\}$ seja mínima. Assim, o problema das p -medianas pode ser formulado de acordo com o seguinte modelo de programação linear binária [Pizzolato et al. 2004]:

$$PPM \quad \text{Min} \sum_{j \in N} \sum_{i \in N} d_{ij} \cdot dem_i \cdot x_{ij} \quad (1)$$

$$s. a. \quad \sum_{j \in N} x_{ij} = 1 \quad \forall i \in N \quad (2)$$

$$\sum_{j \in N} x_{jj} = p \quad (3)$$

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

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

Em que, como parâmetros do modelo, tem-se: N é o número total de vértices (pontos de demanda) do grafo G ; p é o número de pontos de demanda a serem escolhidos como medianas; d_{ij} é a matriz de distâncias euclidianas entre i e j ; dem_i refere-se à demanda do vértice i , sempre que a



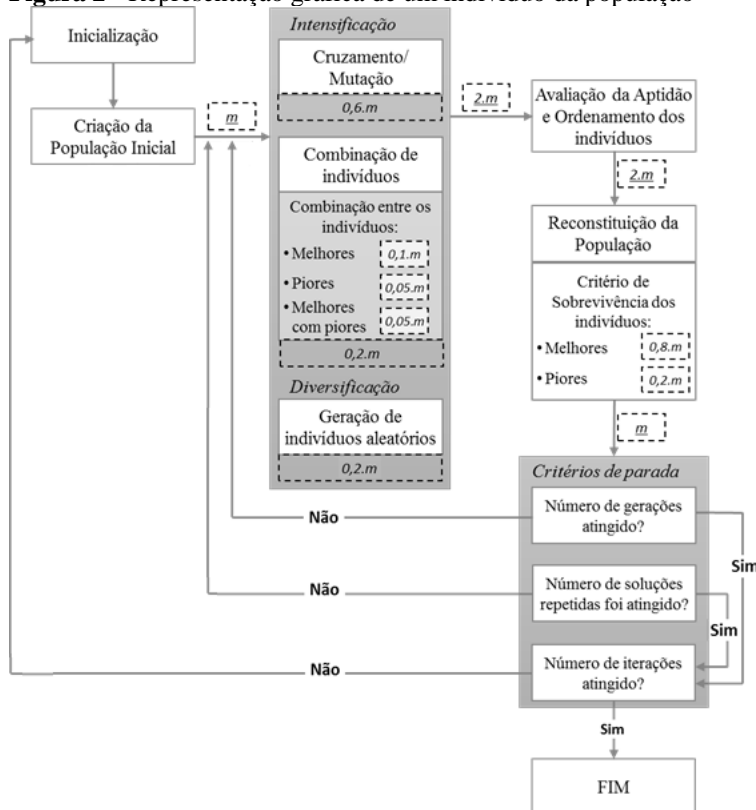
demanda do vértice i for nula, assume-se que $dem_i=1$ no modelo matemático. Como variáveis de decisão do modelo, tem-se: $[x_{ij}]_{N \times N}$ é uma matriz de alocação, com $x_{ij}=1$ se o vértice i está alocado à mediana j , e $x_{ij}=0$, caso contrário; $x_{jj}=1$ se o vértice j é uma mediana e $x_{jj}=0$, caso contrário.

O objetivo do problema, em (1), é minimizar a soma das distâncias ponderadas de cada vértice à mediana mais próxima, ou seja, minimizar custos entre os vértices de $\{V-V_p\}$ alocados às medianas de $\{V_p\}$. As restrições representadas em (2) e em (4) garantem que cada vértice i (ponto de demanda de $\{V-V_p\}$) seja alocado a apenas um vértice j o qual deve ser uma mediana. A restrição em (3) garante que o conjunto $\{V_p\}$ contenha exatamente p elementos; e as restrições apresentadas em (5) definem as variáveis de decisão x_{ij} como binárias.

3.3 Estratégia de Solução baseada em Algoritmo Genético

Um AG, em geral, envolve etapas de: avaliação da Aptidão dos indivíduos de uma geração; seleção dos melhores indivíduos; criação de uma nova população através dos operadores de cruzamento e mutação; e atualização da população. O processo de reprodução e atualização da população é repetido até que um critério de parada seja atingido como, por exemplo, um número fixo de gerações. A representação gráfica da sequência de procedimentos realizados no método proposto para resolução do PPM não-capacitado através de um AG é apresentada na Figura 2, com detalhamento de cada etapa nas próximas subseções.

Figura 2 - Representação gráfica de um indivíduo da população



Inicialização

Quanto aos dados de entrada, no método de solução proposto, a soma das distâncias euclidianas entre os vértices e as respectivas medianas e a demanda de crianças de cada vértice, são utilizados nos cálculos da Aptidão de cada indivíduo nas gerações; a função de Aptidão, que corresponde ao *Fitness*, deverá ser minimizada. Além disso, é necessário que sejam informados o número de medianas (p).



Segundo Alp et al. [2003], em geral, o tamanho da população em um AG deve ser suficientemente grande para que a sua diversificação garanta a exploração das potenciais regiões factíveis que contenham soluções de boa qualidade. Por outro lado, populações excessivamente grandes incorrem em tempos de processamento elevados, de modo que uma possível perda de eficiência não compense o ganho de qualidade da solução.

O AG proposto no presente trabalho considera uma população com número de indivíduos baseado em uma formulação empírica parametrizada pelo número de vértices (n), sugerida por [Isler et al. 2012]. Seja $H=\{h_1, h_2, \dots, h_f\}$ o conjunto de indivíduos, o tamanho da população (f), proposto por estes autores é dado por:

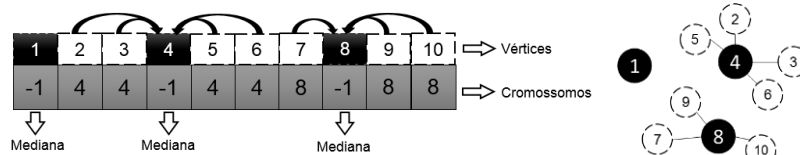
$$f = PAR[17,5 * \ln(n)] \quad (6)$$

O valor considerado como o tamanho da população (m), neste trabalho é o múltiplo de 10 mais próximo de f definido por (6), a fim de facilitar a combinação de indivíduos.

Em relação a representação do indivíduo, o cromossomo de um indivíduo é definido pelo conjunto $\{V_p\}=\{v_1, v_2, \dots, v_p\}$ ($p \times 1$) contendo os índices dos vértices de demanda escolhidos como medianas. A cada indivíduo está associado um conjunto $\{S_n\}=\{s_1, s_2, \dots, s_n\}$ ($n \times 1$), com a i -ésima posição equivalente ao índice do vértice e s_i a mediana a ela associada, tal que o valor inteiro “-1” é atribuído à posição do vértice escolhido como mediana. Assim, $s_i = -1$ para $i = v_1, v_4$ e v_8 , conforme exemplificado na Figura 3, em que $\{V_p\}=\{v_1, v_2, v_8\}=\{1, 4, 8\}$ e, portanto, $\{S_n\}=\{-1, 4, 4, -1, 4, 4, 8, -1, 8, 8\}$.

Ainda, cada indivíduo possui o atributo Aptidão, que representa a soma das distâncias entre as medianas (genes) e os respectivos centroides designados, ponderadas pela demanda de cada vértice i , de modo que os indivíduos mais aptos a permanecerem na população entre gerações sucessivas são aqueles com menor valor desse atributo, em consonância com a função objetivo estabelecida em (1).

Figura 3 - Representação gráfica de um indivíduo da população



Criação da população inicial

Para cada indivíduo da população inicial – com tamanho m – procede-se à escolha das p medianas do PPM não-capacitado. A fim de garantir a variabilidade da população inicial do AG, optou-se pela escolha aleatória dos vértices que integram o conjunto de medianas de cada indivíduo. Dessa forma, o processo de escolha das medianas do conjunto $\{V_p\}$ do primeiro indivíduo da população inicial é realizado aleatoriamente, segundo uma distribuição de probabilidade uniforme, sob a condição de não repetição de um mesmo vértice no conjunto de medianas do indivíduo.

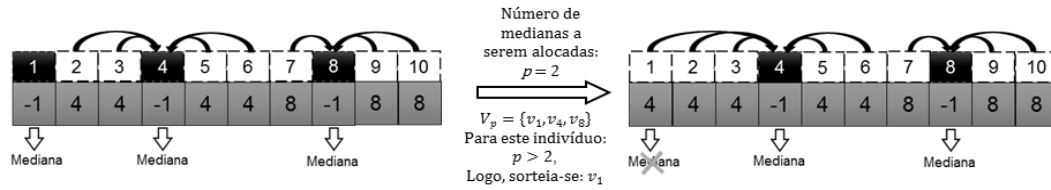
A escolha das medianas dos demais indivíduos é análoga a esta estratégia, sem a imposição de restrições que impeçam a existência de mais de um indivíduo com a mesma configuração do conjunto de medianas. Após a definição do conjunto de medianas de todos os indivíduos da população, o AG é direcionado à atribuição dos vértices de $\{V-V_p\}$, conforme o procedimento descrito a seguir. Os passos apresentados a seguir são atribuídos a todos os indivíduos na etapa de criação da população inicial do AG e após a execução da combinação de indivíduos.

Primeiramente, conta-se o número de medianas do conjunto $\{V\}$ do indivíduo gerado. Se esse número for maior que o número de medianas que se deseja alocar, escolhe-se de forma aleatória dentre a lista de medianas existentes, uma quantidade suficiente de medianas a serem fechadas de forma a viabilizar o indivíduo gerado. Da mesma forma, se o número de medianas do indivíduo gerado exceder o número de medianas desejado, então, sorteia-se dentre os vértices



não-medianas, aqueles que se tornarão medianas. Esse procedimento é exemplificado na Figura 4.

Figura 4 - Ilustração do procedimento de alocação na etapa de criação dos indivíduos da população



Na Figura 5 avalia-se a distribuição da localização ideal das medianas (p) ao longo dos vértices. Neste exemplo, o número de unidades de educação a serem alocadas é dois, assim, como o indivíduo gerado apresentou uma mediana além do valor desejado, escolheu-se de forma aleatória a mediana do v_i para ser excluído da lista de medianas existentes.

Os demais vértices são alocados às medianas, enfim determinadas, de acordo com a proximidade dos mesmos às medianas. Ou seja, para cada vértice do conjunto $\{V - V_p\}$, calcula-se a distância euclidiana deste com os membros do conjunto $\{V_p\}$, sendo que a mediana que apresentar menor distância, será alocada ao vértice. Esse procedimento de alocação dos vértices às medianas é atribuído também para os operadores de cruzamento e mutação caracterizados nas seções subsequentes.

Cruzamento/Mutação

Para obtenção das gerações sucessivas do AG, o operador de cruzamento é executado após o agrupamento dos indivíduos aos pares, escolhidos aleatoriamente segundo uma distribuição de probabilidade uniforme. Este cruzamento ocorre pela troca de medianas entre os pares indivíduos e posterior realocação dos vértices remanescentes. O procedimento é realizado a partir de um cruzamento único em uma posição aleatória.

Figura 5 - Operador de cruzamento



Após o cruzamento, o operador de mutação é responsável pela inserção de pequenas mudanças aleatórias nos cromossomos dos filhos de modo a torná-los viáveis. Assim, para cada um dos filhos gerados, calcula-se o número de genes que devem ser transformados em medianas ou o número de medianas a serem fechadas, de modo que o indivíduo seja factível. Como ilustrado na Figura 6, são escolhidos aleatoriamente, segundo uma distribuição de probabilidade uniforme, dentre o conjunto de não-medianas, aquelas que serão abertas, ou dentre o conjunto de medianas as que serão fechadas.

Combinação de indivíduos

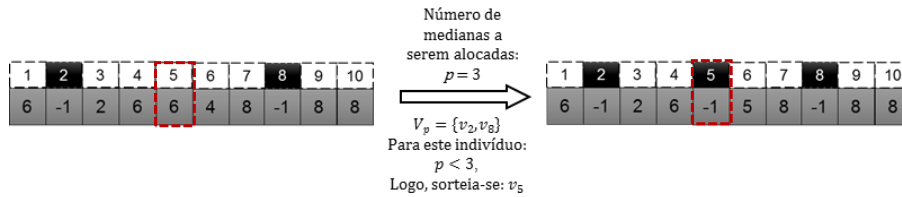
Essa etapa, juntamente com a etapa anterior de Cruzamento e Mutação, é caracterizada por explorar o espaço de busca a partir de informações de pontos anteriormente visitados, de modo a encontrar melhores soluções. Ou seja, adota-se aqui a estratégias de **intensificação** na busca.

A partir do cálculo do *fitness* da população inicial criada, faz-se o ordenamento do mesmo, de modo que os primeiros indivíduos são os que apresentam o melhor valor (menor no caso do problema abordado) para a função de aptidão. Assim, são realizadas as combinações entre os indivíduos. Primeiramente, selecionam-se os dois melhores indivíduos e através da comparação de seus genes, cria-se um novo indivíduo através da repetição dos genes em comum



dos dois. A sequência se dá através do segundo melhor indivíduo com o terceiro, e assim sucessivamente. Dessa análise gera-se 10% do tamanho da população inicial em novos indivíduos.

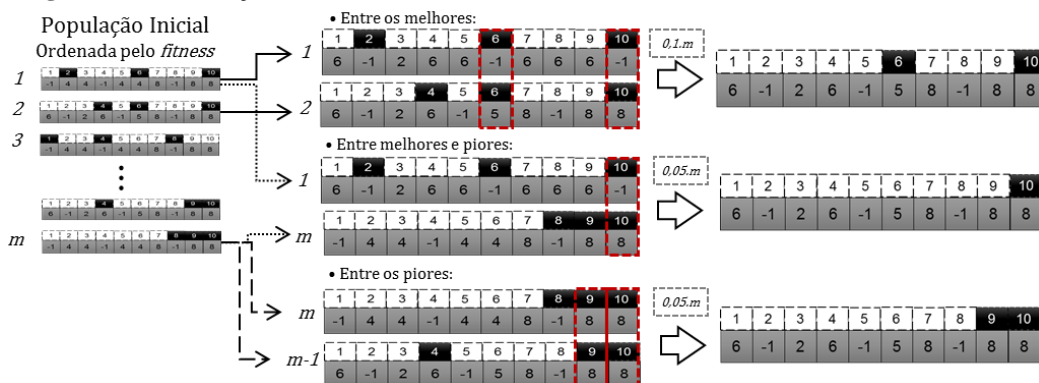
Figura 6 - Operador de mutação



O procedimento é repetido, agora com uma estratégia de **diversificação** da população, a partir dos indivíduos já criados, comparando-se os melhores com os piores indivíduos. Assim, cria-se um novo indivíduo, repetindo-se os cromossomos em comum. Como os indivíduos da população estão ordenados em relação ao *fitness*, compara-se o primeiro com o último, o segundo com o penúltimo e assim consecutivamente. Dessa análise gera-se 5% do tamanho da população inicial em novos indivíduos.

Analogamente, o procedimento é realizado entre os piores indivíduos. São comparados o último com o penúltimo indivíduo, mantendo-se os cromossomos comuns. Em seguida, analisa-se o penúltimo com o antepenúltimo e assim por diante. Dessa análise gera-se 5% do tamanho da população inicial em novos indivíduos. O procedimento encontra-se ilustrado na Figura 7. Desse procedimento é gerado, no total, uma quantidade de indivíduos igual a 20% da população inicial (m).

Figura 7 - Combinação de indivíduos



Geração de indivíduos aleatórios

A geração de indivíduos aleatórios é realizada para a diversificação da população, com o objetivo de direcionar a busca para novas regiões, de forma a atingir o maior espaço de soluções possíveis, evitando que o processo estabilize em um ótimo local. Dessa forma, soluções inteiramente novas são avaliadas no espaço de busca. Os indivíduos são definidos pelo mesmo processo de criação da população inicial. São gerados por esse procedimento, no total, uma quantidade de indivíduos igual a 20% da população inicial (m).

Reconstituição da População

A atualização da população é dada pela criação de um conjunto único com o dobro do número de indivíduos da população inicial (ou seja, com tamanho $2.m$) contendo os indivíduos da população resultante dos cruzamentos ($0,6.m$), combinações ($0,2.m$), indivíduos aleatórios ($0,2.m$) além daqueles que originaram esta população (m). A população obtida ao final do processo é aquela formada pelos $0,8.m$ primeiros indivíduos com os melhores (menores) valores para a função de Aptidão neste conjunto e $0,2.m$ piores (maiores) valores para a função de Aptidão. Esta estratégia garante que os melhores indivíduos permaneçam candidatos à solução ótima do PPM,



sem que parte das piores soluções seja descartada, a fim de que se evite a estabilização em um ótimo local precocemente.

Critérios de parada

Existem diferentes critérios que podem ser considerados como regra de encerramento do AG: número de gerações, como em Corrêa et al. [2004]; porcentagem dos indivíduos da população com o mesmo valor de Aptidão, como sugerido por Alp et al. [2003]; ou ainda a combinação de critérios em função do método de solução abordado para resolução de um problema. Os critérios de parada do AG para resolução do PPM foram estabelecidos empiricamente com base na avaliação do seu desempenho em relação à qualidade das soluções e tempo computacional. Assim, foi estabelecido como critério de término do algoritmo a execução da operação de reconstituição da população até que 1000 gerações fossem atingidas; ou até que $m/2$ soluções repetidas fossem encontradas. Todo esse procedimento foi repetido por 100 iterações, a fim de que se pudesse calcular os valores médios e erros das soluções e seus tempos computacionais.

3.4 Experimentos computacionais para validação do algoritmo

A avaliação do desempenho do AG proposto para a resolução do PPM foi realizada com instâncias que representam dados reais da área central da cidade de São José dos Campos, com diferentes quantidades de vértices e medianas (pmedian324, pmedian818 e pmedian3282). As informações estão disponíveis em <http://www.lac.inpe.br/~lorena/instancias.html>. O algoritmo foi avaliado em termos da qualidade das soluções geradas. Esse algoritmo foi implementado em Visual Basic.NET e executado em computador com processador Intel Core i5, 8GB de RAM e sistema operacional Microsoft Windows 8[®]. Os resultados da aplicação do AG às instâncias indicadas, apresentados na Tabela 1, foram comparadas às seguintes meta-heurísticas utilizadas na literatura por Vasconcelos et al. [2010]:

- *Jumping Frog Optimization* (JFO) ou Algoritmo de Otimização por Saltos de Rãs [Garcia, Perez 2008];
- Algoritmo Genético (AG);
- Algoritmo Genético Híbrido, ou Algoritmo Genético com a Busca Local (AG-BL).

No AG-BL aplica-se o algoritmo de Gillet e Johnson (G&J) para a geração dos *clusters*, ou seja, das regiões de atendimento, juntamente com a Heurística de Localização-Alocação (HLA). Na Tabela 1 estão em negrito os valores em que o algoritmo proposto apresentou qualidade da solução inferior ao apresentado na literatura. É possível observar que o algoritmo proposto neste trabalho identifica soluções melhores em todas as instâncias quando comparado ao AG da literatura. Da mesma forma, em relação ao AG-BL apresentou-se as melhores soluções exceto pela instância pmedian324 com 20 medianas. O JFO apresenta resultados melhores que o algoritmo proposto para instâncias de menor porte, ou seja, para número de vértices igual a 324. Embora, não tenha superado as meta-heurísticas da literatura em todas as instâncias, o método descrito neste trabalho fornece resultados próximos dos valores ótimos. O algoritmo proposto destaca-se pela qualidade de soluções para instâncias de maior porte (pmedian3282), por apresentar *gaps* consideravelmente inferiores às meta-heurísticas da literatura.

Ainda, este método pode ser considerado robusto, pois quando utilizado para instâncias de maior porte para o problema real analisado (apresentado na Tabela 2 da seção subsequente), com número de vértices igual a 2.395 setores censitários, este fornece soluções com pequenos desvios em relação ao valor ótimo, dado os baixos valores de erro médio e desvio padrão decorrentes da execução de 100 replicações para cada conjunto de instâncias. Além de tempos de processamento muito inferiores aos obtidos com a solução do modelo de forma otimizada com o CPLEX[®]. O tempo computacional necessário para obtenção destas soluções foi significativamente menor do que os valores apresentados na literatura. No entanto, esses valores não foram considerados por se tratar de sistemas operacionais e processadores diferentes.

A respeito dos baixos tempos computacionais, testes indicaram que o AG não converge prematuramente para as soluções identificadas, pois o uso de diferentes valores para os critérios



de parada não resultou em soluções de melhor qualidade. Constatou-se, através de testes empíricos, que alterações de percentual entre os indivíduos gerados por operadores de mutação, cruzamento e combinação de indivíduos na reconstituição da população não trazem melhoras significativas no tempo de processamento ou na qualidade das soluções que justifiquem a alteração dos valores destes parâmetros.

Tabela 1 - Resultados da aplicação do AG e comparação com instâncias da literatura

Instâncias	N	P	Ótimo	Gap (%) ¹			Algoritmo proposto			
				AG	AG-BL	JFO	Gap (%) ¹	Tempo (s) ²	Erro médio (%) ³	Tempo médio (s) ³
Pmedian 324	324	5	122518	4,409	3,722	3,431	0,542	1,866	1,841	2,504
	324	10	7956,36	4,541	4,026	1,522	2,393	4,786	3,715	4,107
	324	20	54533,11	10,225	1,907	1,641	2,978	7,840	4,904	6,274
	324	50	32101,52	20,421	4,344	4,032	4,254	15,442	6,332	9,610
Pmedian 818	818	5	605856,1	5,285	3,777	2,292	1,026	3,565	2,283	3,764
	818	10	379473	23,540	21,944	14,480	2,731	11,070	4,493	7,868
	818	20	2517718	-	-	-	4,028	12,801	5,441	11,311
	818	50	1308957	-	-	-	3,471	43,047	5,169	37,138
	818	100	98992,31	-	-	-	5,151	129,397	5,950	101,139
Pmedian 3282	3282	5	6381120	4,510	2,943	1,777	0,4259	29,484	1,2801	30,1165
	3282	10	3914250	38,311	20,688	17,346	2,0043	55,825	4,3811	50,5572
	3282	20	2350502	37,460	21,243	15,070	5,0944	98,314	6,9766	75,135
	3282	50	1308957	61,163	32,921	23,654	6,7461	192,836	8,1902	181,8247
	3282	100	841380,8	62,288	35,206	21,340	8,7585	727,156	10,4278	615,0153
	3282	500	332954,8	13,048	11,836	10,414	4,3865	4681,709	4,9317	4716,565
	3282	1000	194813,5	38,21	35,302	24,094	3,6206	7422,922	4,1625	7215,947

Nota: ¹Gap=100*(Best find - Optimal)/Optimal, ou seja, representa a diferença relativa percentual entre o melhor valor encontrado e o valor ótimo; ²Tempo computacional de processamento da melhor solução encontrada; ³Referente à média para 10 replicações. Os campos preenchidos com “-” representam instâncias não apresentadas.

3.5 Interpretação dos Resultados

A fim de se determinar uma proposta de zoneamento baseada em localizações ótimas, o modelo das *p*-medianas não-capacitado foi aplicado, usando método exato de solução. As creches estão distribuídas nos 75 bairros e 2.395 setores de Curitiba-PR. A conciliação da localização ideal com o fato de existirem unidades operando no município pode ser feita com um raciocínio simples. De fato, as unidades existentes não podem ser fisicamente removíveis. Entretanto, a experiência sugere que, além do interesse acadêmico, o estudo da localização ótima justifica-se quando os desequilíbrios são graves ou quando há uma necessidade suplementar de construir-se novas instalações. É importante para que os desequilíbrios sejam revistos e as ampliações necessárias sejam analisadas.

Para efeitos gerenciais, a proposta de realocação inclui procedimentos para conciliá-la com a rede existente, não ignorando as unidades já existentes, com seus terrenos e instalações. As localizações propostas dividem a área em regiões e cada creche existente pertencerá a uma delas, agregando para ela a sua capacidade. Dessa forma, embora as atuais instalações possam não coincidir com as localizações propostas, elas oferecerão capacidade de atendimento às regiões apontadas pelo modelo. Assim, para cada região proposta houve uma demanda e oferta de vagas das creches existentes.

Tabela 2 - Resultado do AG para a avaliação da localização ideal e método exato

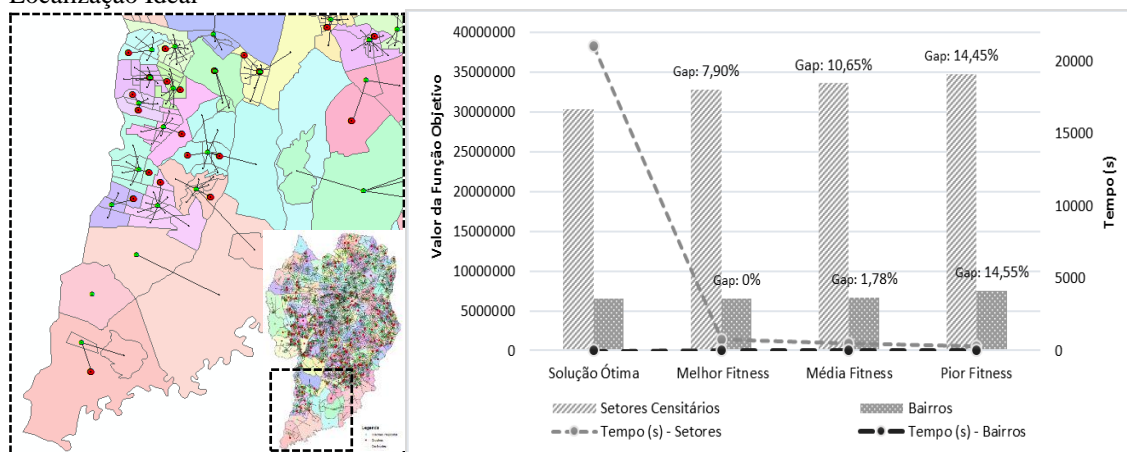
Problema	Método Exato				Algoritmo Genético (100 iterações)					
	n	P	Ótimo	Tempo (s)	Melhor Solução	Gap (%) ¹	Tempo (s) ²	Erro Médio (%) ³	Tempo Médio (s) ³	Desvio Padrão (%) ³
Bairros (0-3 anos)	75	52	6587254	0,19	6.587.254	0%	0,447	1,78%	0,57s	2,55%
Setor (0-3 anos)	2.395	258	30.354.553	14.900,4	32.751.643	7,90%	786,0	10,65%	530,55	0,98%

Nota: ¹Representa a diferença relativa percentual entre o melhor valor encontrado e o valor ótimo; ²Tempo computacional de processamento da melhor solução encontrada; ³Em 100 replicações.

A Tabela 2 apresenta resultados para o AG e as respectivas comparações em relação a qualidade das soluções e tempo computacional gasto com os valores ótimos obtidos. Os resultados dessa avaliação comparativa estão presentes na Figura 8.



Figura 8 - Solução do modelo não-capacitado para setores censitários e Resultados do AG para análise da Localização Ideal



O que se pode observar é que para os problemas com número de vértices igual a 75 bairros, o método proposto identificou como melhor solução a solução ótima. Porém o tempo computacional foi superior quando comparado ao obtido pelo método exato. Para os problemas com número de vértices igual a 2395, a melhor solução apresentou 7,90% de erro em relação ao valor ótimo. O tempo computacional, no entanto, para essa análise foi substancialmente reduzido em relação ao realizado pelo método exato. Por exemplo, o tempo para o modelo exato foi de 14900,45s, ou seja, 4h08min. Com o método apresentado, o tempo para obtenção da melhor solução foi 786,057 ou 13h10min. Destaca-se o fato de que o desvio-padrão obtido para esta análise de maior porte, foi bastante reduzido (com no máximo 1,03% de desvio).

A Figura 8, também apresenta um comparativo dos valores da melhor e pior solução encontrada, a média das soluções dentre 100 iterações e os respectivos *gaps* em relação ao valor ótimo obtido com o modelo exato. Para esta análise de bairros, o método proposto encontra a solução ideal como melhor solução, sendo que em média o erro obtido é de 1,78%, enquanto para setores é de 10,65%.

4. Conclusões

Os resultados de validação do AG indicam que o método proposto neste estudo apresenta soluções satisfatórias para o problema de p -medianas. O algoritmo proposto destacou-se pela qualidade de soluções para instâncias de maior porte (com 3282 vértices), por apresentar *gaps* consideravelmente inferiores às meta-heurísticas da literatura. Apesar de não terem sido obtidas as soluções ótimas em todas as instâncias de validação do AG, o objetivo de estabelecer um método de solução para o problema real de avaliação da localização de centros de Educação Infantil foi atingido. Além disso, o AG proposto pode ser utilizado na análise do planejamento da localização de unidades de educação infantil para cidades maiores, ou com um número elevado instituições para análise, nas quais a aplicação do modelo exato pode tornar-se inviável.

A estratégia da atualização da população através dos operadores de cruzamento, mutação, e combinação de indivíduos buscou garantir a variabilidade da população entre gerações de modo a evitar a estagnação do algoritmo em soluções ótimas locais. Enquanto, o procedimento de escolha aleatória dos indivíduos e genes a serem cruzados e mutados foi implementado para impedir que o algoritmo permanecesse estagnado em uma solução de mínimo local, de forma a proporcionar menores tempos computacionais para obtenção de soluções de boa qualidade.

Ainda, em relação aos critérios de parada, a variação dos parâmetros do AG indica que o tempo computacional é sensível ao número de iterações em que são executados os operadores de cruzamento e de mutação e a reconstituição da população, e empiricamente, o algoritmo proposto apresentou resultados satisfatórios em termos do *trade off* entre qualidade da solução e tempo



computacional. A adaptação do modelo ao problema de Localização de unidades de Educação Infantil indica que o algoritmo proposto pode ser utilizado como ferramenta para a localização dessas instalações em outros municípios.

Do ponto de vista metodológico, como estímulo para trabalhos futuros sugere-se a identificação de melhores soluções fornecidas pelo AG por meio da implementação de métodos de buscas locais ou de um método heurístico alternativo para geração de soluções de melhor qualidade (por exemplo, o Gillet e Johnson ou o *Greedy Randomized Adaptive Search Procedure* – GRASP) inserido na etapa de criação e reconstituição da população.

Enfim, a utilização dos diferentes modelos em paralelo para fins deste estudo revelou uma vantagem, devido à possibilidade de comparar as soluções obtidas. O estudo demonstrou que, de forma geral, existe a necessidade de planejamento imediato para expansão do ensino público infantil em Curitiba-PR, inclusive para com o cumprimento efetivo das metas do PNE. Da mesma forma, tendo em vista uma evidente necessidade de grandes investimentos nesse campo, os resultados deste trabalho podem contribuir inclusive, para uma melhor estratégia de investimentos em Centros de Educação Infantil.

Referências

Alp, O. E., Erkut E. e Drezner Z. (2003). An Efficient Genetic Algorithm for the p -Median Problem. *Annals of Operations Research*, 122:21–42.

Censo Demográfico (2010). Web page. <http://www.ibge.gov.br/home/estatistica/populacao/censo2010/>. Acessado: 2016-08-07.

Corrêa, E. S.; Steiner M. T.; Freitas A. A.; Carnieri. C. (2004) A genetic algorithm for solving a capacitated p -median problem. *Numerical Algorithms*, 35:373–388.

Davis, L. *Handbook of Genetic Algorithms*. (1991). New York: Van Nostrand Reinhold.

Garcia, F. J. M., Perez, J. A. M. (2008) Jumping frogs optimization: a new swarm method for discrete optimization. *Documentos de Trabajo del DEIOC*, 3.

Goldberg, D. E. (1989) *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison-Wesley.

Holland, J. H. (1975) *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. Ann Arbor: University of Michigan Press.

Isler, C. A., Bonassa, A. C. e Cunha, C. B. (2012) Algoritmo genético para resolução do problema de p -medianas capacitado associado à distribuição de peças automotivas. *Transportes*, 20(2):5-14.

Pizzolato, N. D., Barcelos, F. B., Lorena, N., e Antonio L. (2004) School location methodology in urban areas of developing countries. *International Transactions in Operational Research*, 11(6):667-681.

Smith, S. e Jones, J. (2002). A paper on operations research. *Pesquisa Operacional*, 32:5–44.

Vasconcelos, A. M., Souza, S. R., Vitor, J. F. A. e Bezerra, S. N. (2010) Jumping frog optimization e algoritmo genético aplicados à solução do problema das p -medianas. In: *Anais do XXXI Iberian Latin American Congress on Computational Methods in Engineering*, p. 9645-9657. Buenos Aires, Argentina.