

CONTROLE DA DIVERSIDADE POPULACIONAL EM ALGORITMOS GENÉTICOS PELA CLUSTERIZAÇÃO k -MÉDIAS E A TRANSFORMADA DISCRETA WAVELET

Elenice da C. Lopes, Felipe A. Calarge e Fabio H. Pereira

*Programa de Mestrado em Engenharia de Produção, Universidade Nove de Julho, Av. Francisco
Matarazzo 612, 05001100 São Paulo, Brasil, mestradoengenharia@uninove.br
<http://www.uninove.br/Paginas/Mestrado/Engenharia/mepHome.aspx>*

Palavras-chave: Algoritmos Genéticos, Diversidade Populacional, Transformada Discreta Wavelet, Clusterização k -médias.

Resumo. Este trabalho propõe a aplicação da transformada discreta wavelet e do algoritmo de clusterização k -médias para controle do tamanho e da diversidade populacional nos Algoritmos Genéticos. A abordagem consiste em trabalhar com uma população inicial relativamente grande, fazendo clusterizações periódicas por meio do k -médias e posterior uso da transformada wavelet discreta para explorar as possíveis correlações entre os indivíduos de um mesmo cluster. A transformada discreta wavelet é aplicada à população após cada clusterização dizimando 50% dos indivíduos. Os indivíduos sobreviventes são avaliados e submetidos às operações genéticas convencionais. Nesse processo tem-se objetivo de eliminar redundâncias dentro dos clusters, preservando a diversidade populacional, reduzir o número de avaliações da função objetivo necessárias para a convergência. Resultados iniciais em problemas testes de otimização irrestrita envolvendo funções com vários mínimos locais mostram o potencial da proposta.

1 INTRODUÇÃO

Devido às dificuldades na obtenção de soluções exatas para problemas de otimização com interesse prático, as técnicas metaheurísticas têm sido apontadas como abordagens promissoras em muitas situações. Apesar de não garantirem a otimalidade de suas soluções, estes métodos fornecem soluções aceitáveis, mesmo para problemas complexos. A facilidade de programação e eficiência computacional dessas técnicas numéricas são suas principais vantagens sobre os métodos determinísticos clássicos.

No entanto, apesar das vantagens da técnica computacional, os métodos baseados em metaheurísticas têm algumas dificuldades, como definir o tamanho da população e preservar a diversidade dessa população. Segundo a literatura (Hiroyasu, *et al.* 2008), uma pequena população pode causar convergência prematura do algoritmo em função da perda de diversidade dos seus indivíduos, enquanto uma população muito grande pode exigir muito tempo de processamento para que uma boa solução seja encontrada.

Muitas pesquisas têm abordado a questão do controle da diversidade populacional nos algoritmos genéticos. Um estudo apresentado em (Zhu, 2003) propõe o controle da diversidade populacional por meio da adaptação das taxas de cruzamento e mutação, obtendo resultados muito competitivos. O autor mede a diversidade populacional por meio da distância de Hamming entre genótipos no ambiente populacional. Outras técnicas para medida da diversidade populacional foram estudadas em (Barker e Martin, 2000; Burke, Gustafson e Graham Kendall, 2002; Morrison e de Jong, 2001).

O uso das estratégias de hipermutação e imigrantes aleatórios para manutenção e aumento da diversidade populacional foi investigada em (Tragante e Tinós, 2009), apresentando resultados satisfatórios em problemas de predição de estruturas em proteínas. Apesar dos bons resultados obtidos, a determinação da taxa de substituição para novos imigrantes não pode ser feita *a priori*, o que pode dificultar a convergência do Algoritmo Genético (AG).

A aplicação de redes neurais artificiais em mecanismos de manutenção de diversidade é objeto de estudo em (Hiroyasu, *et al.* 2008). O mecanismo proposto produz soluções com alta diversidade, mas apresenta elevado custo computacional por envolver o processo de busca, clusterização, treinamento de redes neurais e realocação de indivíduos.

Apesar de existirem várias diferentes abordagens, a solução é frequentemente problema-dependente justificando a realização de modificações nas técnicas básicas do AG que visem aumentar a eficiência computacional do método, evitando convergência prematura para ótimos locais e/ou soluções únicas quando muitas soluções igualmente boas existem.

Por outro lado, quando é aplicada em um conjunto de dados as funções *wavelets* criam uma aproximação esparsa desses dados, eliminando redundâncias entre os valores. A aproximação esparsa criada pela transformada wavelet carrega consigo a maior parte da informação contida nos dados originais. Portanto, as *wavelets* podem ser aplicadas a uma população do Algoritmo Genético a fim de selecionar os indivíduos mais adequados, como é proposto neste trabalho. A abordagem deste artigo consiste em trabalhar no AG com uma população inicial relativamente grande, fazendo clusterizações periódicas por meio da função *k*-médias e posterior uso da transformada wavelet discreta para explorar as possíveis correlações entre os indivíduos de um mesmo cluster, visando manter a diversidade populacional e evitar a convergência prematura do processo busca.

A transformada wavelet é aplicada à população após a clusterização dizimando 50% dos indivíduos. Os indivíduos sobreviventes são avaliados e submetidos às operações convencionais do AG. Nesse processo tem-se objetivo de eliminar redundâncias dentro dos clusters e, conseqüentemente, reduzir o número de avaliações da função objetivo necessárias

para a convergência.

O método foi aplicado a problemas testes de otimização irrestrita envolvendo funções com vários mínimos locais, e várias combinações de tamanho da população e número de gerações foram testadas. O número médio de avaliações da função objetivo e o desempenho da abordagem proposta foram analisados e comparados aos resultados do AG convencional. Valores de eficiência, medidos pela taxa de convergência em 20 simulações, são apresentados de acordo com o número de avaliações da função objetivo. Os resultados iniciais do uso da clusterização e transformada wavelet no AG mostraram-se viáveis para os casos testes, apresentando resultados semelhantes com um menor número de avaliações da função. O comportamento menos oscilatório na convergência mostra que a nova abordagem é menos sensível à variação do tamanho da população, indicando uma maior robustez do método.

2 ALGORITMOS GENÉTICOS

Os algoritmos genéticos (AGs) são uma família de modelos computacionais inspirados no processo de seleção natural, com base no princípio da sobrevivência do mais forte, conforme definição darwinista dos fenômenos naturais e da herança genética nas lutas pela sobrevivência (Russell, Shi and Kennedy, 2001; Mitchell, 1997).

Em vez de trabalhar com uma única solução por iteração, um AG trabalha com uma série de potenciais soluções conhecido como população, que é geralmente escolhida aleatoriamente na inicialização (Mitchell, 1996). Estas soluções potenciais são codificadas como uma seqüência de números (*bits*) de comprimento fixo, chamado de cromossomo. Assim, em cada iteração (também chamada geração) uma avaliação desses cromossomos é realizada por meio da função objetivo para atribuir um valor de aptidão (*fitness*) para cada solução. O valor de aptidão deve representar a qualidade de uma solução, que normalmente é definido em relação à população atual, e é utilizado para classificar os membros da população a fim de proporcionar uma espécie de limiar de probabilidade para a operação de reprodução.

Se, por exemplo, o objetivo da otimização com o GA é maximizar uma função então o indivíduo com valor de *fitness* relativamente mais alto devem ter uma maior probabilidade de reprodução. Em seguida, após a definição desse limiar de probabilidade, a população é atualizada por meio de operações de seleção, cruzamento e mutação e o processo é repetido até que um critério de parada, que pode ser definido como atingir um limite de tempo pré-definido (ou número de gerações ou de convergência da população) seja satisfeito.

Os operadores de seleção, cruzamento e mutação podem ser resumidamente definidos da seguinte forma (Sean, 2010; Weise, 2009):

Seleção é um processo em que p indivíduos são selecionados com uma probabilidade proporcional à sua aptidão para serem pais. *Cruzamento* é um processo no qual dois pais diferentes são iterativamente selecionados a partir do conjunto de p pais para trocar informações entre si com a finalidade de gerarem dois novos indivíduos (filhos). Isto é feito escolhendo aleatoriamente um ponto de ruptura para os pais e trocando-se posições entre eles. *Mutação* é a alteração dos *bits* de um indivíduo com uma pequena probabilidade pré-determinada, às vezes conhecida como coeficiente de mutação (*cm*). Estes novos indivíduos alterados compõem a nova população (Nemati *et al.* 2009).

2.1 O problema da perda de diversidade populacional

Os AGs operam em uma população formada por indivíduos que diferem uns dos outros por meio de um valor de aptidão e uma identificação individual satisfazendo a critério pré-definido para operação de seleção. Os indivíduos selecionados são combinados em cada ciclo

e essa combinação ocorre pela troca de parte dos indivíduos. Portanto, para criar novos indivíduos, os indivíduos combinados devem ser diferentes e a porção de diferenças na população é chamada de diversidade (Mitchell, 1997).

Por outro lado, um dos principais problemas na aplicação dos algoritmos genéticos em problemas do mundo real é como reduzir o grande número de avaliações da função objetivo. A abordagem mais simples é uma busca com uma população com um número pequeno de indivíduos. No entanto, a diversidade de soluções é muitas vezes perdida com essa estratégia de solução (Hiroyasu et al., 2008), comprometendo uma varredura completa do espaço de busca e provocando, em alguns casos, a convergência prematura do método para mínimos (máximos) locais.

3 TRANSFORMADA DISCRETA WAVELET

Wavelets são funções $\psi_{a,b}(t)$ obtidas por meio de translações e dilatações (contrações ou escalamento) de uma função $\psi(t)$ chamada *wavelet mãe*. Introduzindo as dilatações por um fator de escalamento a e as translações por meio de uma variável b , a forma geral de uma família dessas funções é dada por:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad (1)$$

com $a > 0$ e $b \in \mathbb{R}$. O termo $1/\sqrt{a}$ é introduzido para preservar a norma da função no espaço das funções quadraticamente integráveis $L^2(\mathbb{R})$.

Uma importante característica das funções definidas em (4.1) é que elas formam uma base para o espaço $L^2(\mathbb{R})$. Assim, uma função $f(t)$ nesse espaço pode ser aproximada por uma combinação das wavelets usando coeficientes $c_{a,b}$ dados por (2) (Frasier, 1999; Sarkar, 2002),

$$c_{a,b} = \langle f, \psi_{a,b} \rangle, \quad (2)$$

na qual $\langle \cdot, \cdot \rangle$ denota o produto interno em $L^2(\mathbb{R})$.

O termo wavelet é uma tradução para o inglês das palavras francesas *ondelettes* ou *petites ondes*, ou seja, pequenas ondas. No português, recebem também o nome de ondaletas (Morettin, 1999). Na verdade, os termos wavelet e ondaletas são usados para sugerir que se tratam de funções que são localizadas em frequência em torno de um valor central.

A transformada wavelet usa uma técnica de multiresolução para analisar diferentes frequências em diferentes resoluções. Como as funções wavelets possuem localização temporal e, além disso, são funções com dois parâmetros (translações e escalamento ou contração), os componentes de alta frequência do sinal podem ser representados por wavelets de curta duração (maior resolução temporal). Por outro lado, os componentes de baixa frequência são analisados em períodos de tempo mais longos (Daubechies, 1988; Frasier, 1999; Sarkar, 2002).

Formalmente, a Análise em Multiresolução (AMR) com wavelets pode ser definida da seguinte forma:

Definição 1: uma análise em multiresolução é uma seqüência $V_j (j \in \mathbb{Z})$ de subespaços fechados aninhados,

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots \quad (3)$$

tais que a união é densa no espaço das funções quadraticamente integráveis $L^2(\mathbb{R})$. Além disso, tem-se que

$$\bigcap_{j \in \mathbb{Z}} V_j = 0; \quad (4)$$

$$f(x) \in V_j \Leftrightarrow f(2^j x) \in V_0; \quad (5)$$

$$f(x) \in V_0 \Rightarrow f(x-n) \in V_0, \text{ para todo } n \in \mathbb{Z}; \quad (6)$$

Outras importantes características desses subespaços são:

a) V_0 é gerado por um conjunto ortonormal obtido por meio de translações inteiras de uma única função $\phi \in L^2(\mathbb{R})$, chamada *função escaladora*,

$$V_0 = \{(\phi(x-n))\}_{n \in \mathbb{Z}} \quad (7)$$

Para cada inteiro j , V_j é gerado por uma base ortonormal obtida da forma,

$$V_j = \{\phi(2^{-j}x-k)\}_{k \in \mathbb{Z}} \quad (8)$$

b) Para cada inteiro j , $V_j = V_{j+1} \oplus W_{j+1}$, na qual o símbolo \oplus significa a soma direta entre os espaços e cada W_{j+1} é gerado por um conjunto ortonormal obtido por meio de translações e escalamentos de uma única função wavelet,

$$W_j = \{\psi(2^{-j}x-k)\}_{k \in \mathbb{Z}} \quad (9)$$

c) O conjunto obtido da forma

$$\{\psi(2^{-j}x-k)\}_{j,k \in \mathbb{Z}} \quad (10)$$

é uma base de $L^2(\mathbb{R})$.

Na prática, a análise em multiresolução é empregada em problemas discretos que estão definidos, ou são projetados, no subespaço V_0 que representa o mais alto nível de resolução em AMR. Assim, dada um vetor (sinal discreto) $\mathbf{u} \in V_0$, os coeficientes $c_{0,k}$ da expansão de \mathbf{u} em termos da base (7) são obtidos usando a ortogonalidade de ϕ , e ψ ,

$$\mathbf{u} = \sum_k c_{0,k} \phi(x-k). \quad (11)$$

A expressão (11) pode ser considerada como a representação de \mathbf{u} .

Outros coeficientes $c_{1,k}$ e $d_{1,k}$ podem ser encontrados para representar \mathbf{u} em V_1 e W_1 da forma:

$$\mathbf{u} = \sum_k c_{1,k} \phi\left(\frac{x}{2}-k\right) + d_{1,k} \psi\left(\frac{x}{2}-k\right). \quad (12)$$

Segundo a teoria de multiresolução, o resultado da decomposição de V_0 em uma soma de subespaços V_1 e W_1 , é que a maior parte da variação (energia da função) fica contida em V_1 e a representação dessa função em W_1 é praticamente nula. Assim, os coeficientes $c_{1,k}$ podem ser considerados como uma aproximação de \mathbf{u} no espaço V_1 , e o processo pode continuar até atingir o nível mais grosseiro, sempre decompondo cada representação de \mathbf{u} em V_j no próximo par de subespaços V_{j+1} e W_{j+1} . A decomposição de um espaço vetorial V_n em

subespaços mutuamente ortogonais na forma

$$V_n = V_l \oplus W_l \oplus W_{l+1} \oplus \dots \oplus W_{n+1} \quad (13)$$

é denominada Transformada Discreta Wavelet (DWT).

Na prática, a DWT é realizada usando filtros digitais passa-banda que capturam uma aproximação (coeficientes $c_{j,k}$) do sinal e os coeficientes de detalhes ($d_{j,k}$). Mais detalhes sobre a relação entre a transformada wavelet e banco de filtros pode ser encontrado em (Sarkar, 2002).

Uma DWT com três níveis é ilustrada esquematicamente na Figura 1. Na Figura, os blocos G e H representam, respectivamente, os filtros passa-alta e passa-baixa e o símbolo ($\downarrow 2$) representa uma operação de decimação (sub-amostragem) que consiste em eliminar as amostras de índice ímpar do sinal de entrada. A decimação é feita para preservar a dimensão do sinal.

Na DWT tradicional, muito empregada no processamento de sinais e imagens digitais, os coeficientes de detalhes capturados pelos filtros passa-alta G são utilizados na reconstrução do sinal decomposto. Entretanto, como no contexto do algoritmo genético não existe a necessidade de reconstruir o sinal digital (população), a nova abordagem apresentada aqui consiste no uso apenas dos filtros passa-baixa H, capturando as aproximações necessárias para representar a população em um nível com menos refinamento. Essa operação é ilustrada pelo ramo direito da árvore apresentada na Figura 1.

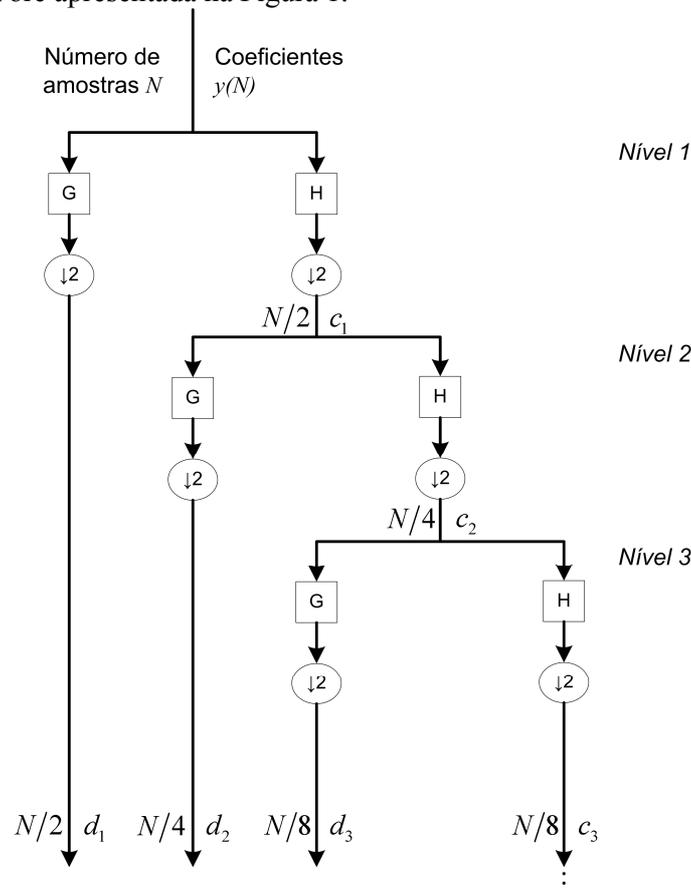


Figura 1: Transformada Discreta Wavelet com três níveis

4 O ALGORITMO DE CLUSTERIZAÇÃO *K*-MÉDIAS

O Algoritmo *k*-médias (MacQueen, 1967) é um dos mais simples algoritmos de aprendizado não supervisionado que resolvem conhecido problema de agrupamento. O procedimento segue uma maneira simples e fácil de classificar um determinado conjunto de dados por meio de certo número pré-estabelecido de clusters (*k* clusters).

A idéia principal do método é definir *k* centróides iniciais, um para cada cluster, e associar cada ponto do conjunto de dados ao centróide mais próximo, definindo assim um agrupamento inicial. Na sequência, *k* novos centróides são calculados como os baricentros dos clusters resultantes da etapa anterior. De posse dos novos centróides, o processo de associação dos pontos ao centróide mais próximo se repete gerando um processo iterativo. Como resultado desse ciclo, pode-se notar que os centróides mudam a sua localização passo a passo até que não haja mais mudanças a serem feitas. Em outras palavras, os centróides não se movem mais.

Finalmente, este algoritmo visa minimizar a função objetivo de erro quadrático dada por (14),

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (14)$$

na qual, $\|x_i^{(j)} - c_j\|^2$ é uma medida de distância entre um ponto x_i do conjunto de dados e o centro do cluster, c_j .

A função objetivo (14) fornece um indicador da distância dos *n* pontos do conjunto de dados aos centros de seus respectivos clusters.

5 A ABORDAGEM PROPOSTA

Deve ser observado que cada cromossomo, chamado de indivíduo no AG, corresponde a um ponto no espaço de soluções do problema de otimização. Além disso, os AGs trabalham com uma população de pontos e cada ponto é um vetor representando um potencial candidato para a solução do problema. Portanto a população é considerada apenas como um conjunto de vetores, e uma matriz populacional pode ser definida de forma que suas colunas sejam formadas pelos vetores que representam cada cromossomo.

A abordagem deste artigo consiste em trabalhar no AG com uma população inicial relativamente grande e usar transformada wavelet discreta para explorar as possíveis correlações entre os indivíduos da população. O processo descrito na seção 3 é aplicado às linhas da matriz populacional dizimando 50% dos indivíduos. Os indivíduos sobreviventes são avaliados e submetidos às operações convencionais do AG. Nesse processo tem-se objetivo de eliminar redundâncias dentro da população e, conseqüentemente, reduzir o número de avaliações da função objetivo necessárias para a convergência. Nesse processo, apenas os filtros passa-baixa são utilizados, capturando as aproximações necessárias para representar a população em um nível com menos refinamento.

O método foi aplicado a problemas testes de otimização irrestrita envolvendo funções com vários mínimos locais, e várias combinações de tamanho da população e número de gerações foram testadas. O número médio de avaliações da função objetivo e o desempenho da abordagem proposta foram analisados e comparados aos resultados do AG convencional. Valores de eficiência, medidos pela taxa de convergência em 20 simulações, são apresentados de acordo com o número de avaliações da função objetivo.

Considerando o caráter local da ação da transformada discreta wavelet, o algoritmo de

clusterização k -médias deve ser aplicado periodicamente ao longo do processo iterativo e a transformada aplicada de forma independente dentro de cada cluster. Entretanto, nos resultados apresentados a seguir o k -médias não foi utilizado. Novos testes considerando a utilização do algoritmo de clusterização estão sendo realizados e serão apresentados durante o evento e em futuras publicações.

6 TESTES NUMÉRICOS E RESULTADOS PRELIMINARES

O método proposto foi aplicado em um problema de otimização irrestrita da bem conhecida função *Rastrigin* a qual é frequentemente usada para testar o algoritmo genético. Essa função teste possui vários mínimos locais e um único mínimo global no ponto $[0,0]$ do plano $x - y$, conforme apresentado nas Figuras 2 e 3.

Para duas variáveis independentes, a função *Rastrigin* é definida como em (15)

$$Ras(x, y) = 20 + x^2 + y^2 - 10(\cos 2\pi x + \cos 2\pi y) \quad (15)$$

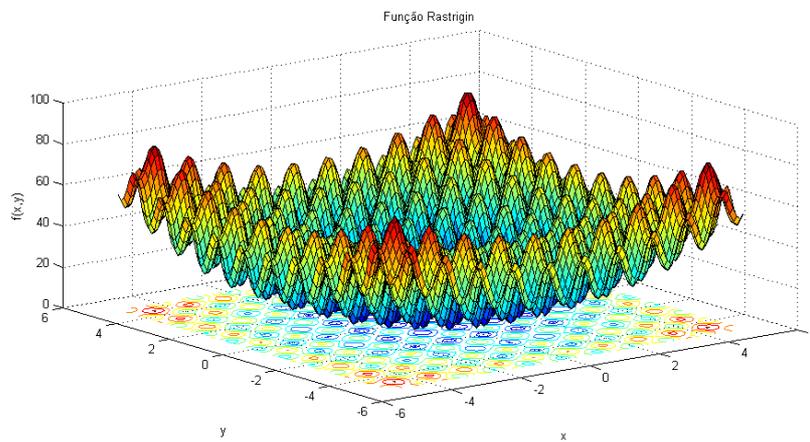


Figura 2: Função teste *Rastrigin*

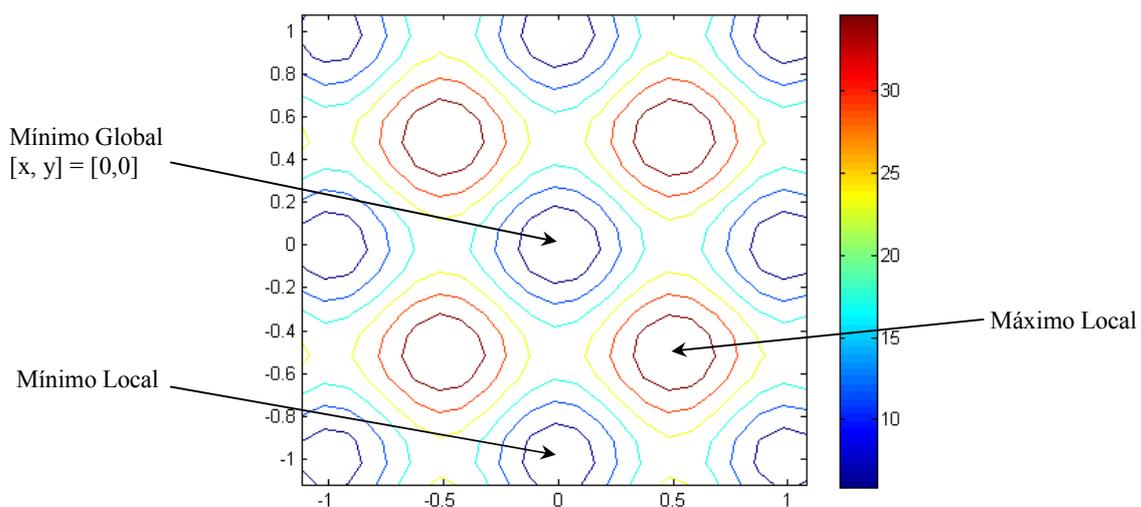


Figura 3: Curvas de nível da função teste

Foram testadas diversas combinações de tamanho da população e números de gerações. O número médio de avaliações da função objetivo e o desempenho após 20 simulações foram analisados e comparados aos resultados do AG tradicional.

Foram realizadas 20 simulações para cada conjunto de parâmetros (tamanho da população e número de gerações). O número médio de avaliações da função objetivo e a porcentagem de convergência (eficiência) foram avaliados, e os resultados estão apresentados resumidamente na Tabela 1 e, graficamente, na Figura 4.

AG Convencional				AG Wavelet			
População	Gerações	Avaliações (nº médio)	%	População	Gerações	Avaliações (nº médio)	%
50	200	8075	0,50	4000	6	5437	0,70
	400	9822	0,70		7	5084	0,75
100	200	11800	0,80	5000	6	7437	0,60
	400	9385	0,95		7	7222	0,70
150	200	12787	0,95	6000	7	8306	0,80
	400	12112	1,00		8	7996	0,80
200	200	6080	1,00	7000	8	9655	0,75
	400	17420	1,00				

Tabela 1: Número de avaliações e eficiência das abordagens testadas

Considerando o ponto ótimo global da função, o critério de convergência foi definido por meio da norma euclidiana do erro obtido por cada método nas diferentes combinações. As soluções obtidas em cada caso são apresentadas na Tabela 2.

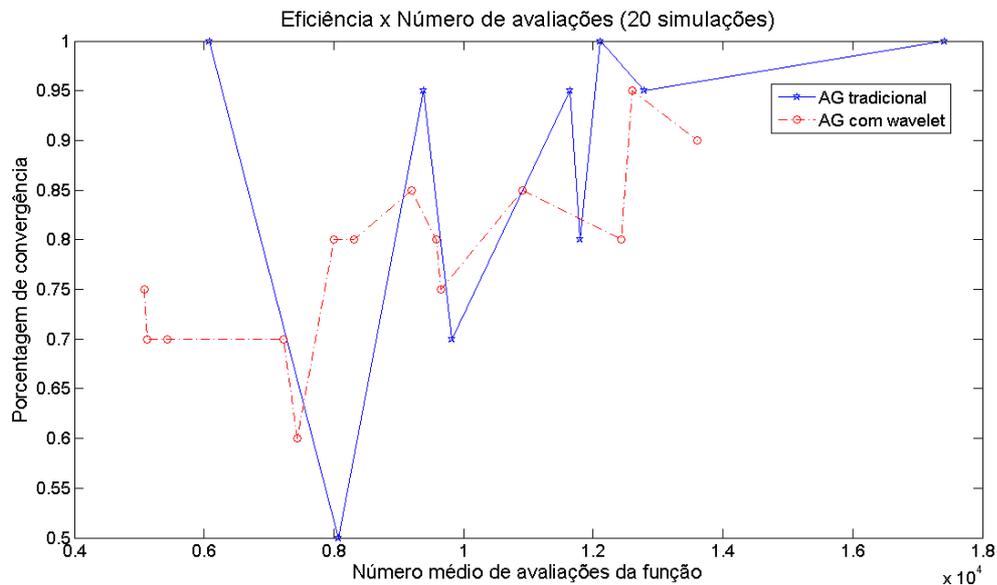


Figura 4: Eficiência versus Custo para os métodos testados

Método	População	Gerações	Melhor x	Melhor y	Norma	Melhor f(x,y)
AG Convencional	50	200	0,005234	-0,010235	0,0115	0,026209
		400	-0,004297	-0,00961	0,01053	0,021977
		600	-0,003203	-0,000234	0,00321	0,002046
	100	200	-0,000391	0,01211	0,01212	0,029109
		400	0,003672	0,010235	0,01087	0,023448
	150	200	-0,001484	0,003672	0,00396	0,003112
		400	-0,01211	0,009766	0,01556	0,047993
	200	200	-0,002734	0,002891	0,00398	0,003141
		400	0,000547	-0,008047	0,00807	0,012903
	AG com transformada wavelet	14000	8	-0,005703	-0,010391	0,01185
12000		7	-0,004453	-0,001953	0,00486	0,004691
		8	-0,010703	0,007891	0,0133	0,035069
10000		6	0,006953	0,004141	0,00809	0,012992
		7	-0,000391	-0,001328	0,00138	0,00038
8000		5	0,012266	-0,007266	0,01426	0,040305
		6	0,003047	-0,006172	0,00688	0,009398
		7	-0,001641	-0,012578	0,01268	0,031906

Tabela 2: Melhores soluções obtidas em 20 replicações

7 CONCLUSÕES

O presente trabalho propõe o uso da transformada discreta wavelet para explorar as possíveis correlações entre os indivíduos da população nos algoritmos genéticos, visando manter a diversidade populacional e evitar a convergência prematura do processo busca.

O uso da transformada wavelet da forma como está sendo proposta mostrou-se viável para o caso testado. Apesar de não atingir 100% de eficiência, como visto na Fig. 4, o AG com wavelet apresentou, em geral, os melhores resultados para número médio de avaliações menores (tabela 1). O comportamento menos oscilatório apresentado no gráfico de eficiência mostra que a nova abordagem pode ser menos sensível à variação do tamanho da população, indicando maior robustez do método.

Vale destacar que o trabalho está em fase inicial e, conseqüentemente, os resultados apresentados são bastante embrionários. Nesse sentido, diversos aspectos importantes estão sob investigação. Considerando o caráter local da atuação da transformada discreta wavelet, por exemplo, a transformação direta da população pode envolver nos cálculos de uma aproximação dois indivíduos que, de fato, não são espacialmente próximos no espaço de busca, gerando um novo elemento que não será uma boa aproximação dos indivíduos usados. Nesse caso, a proposta é aplicar o algoritmo de clusterização k -médias periodicamente ao longo do processo iterativo e então realizar a transformada wavelet de forma independente dentro de cada cluster.

Novos testes considerando a utilização do algoritmo de clusterização estão sendo realizados e serão apresentados durante o evento e em futuras publicações. Adicionalmente, testes envolvendo a aplicação da abordagem proposta em outras funções testes, bem como em problemas reais, e o uso de outras funções wavelets estão sendo realizados.

REFERENCES

- Barker, A.L., e Martin, W.N., Dynamics of a distance-based population diversity measure. In: *Proceedings of the 2000 Congress on Evolutionary Computation*, Piscataway, NJ, 1:1002-1009, 2000.
- Burke, E., Gustafson, S., e Kendall, G., A survey and analysis of diversity measures in genetic programming. In: *GECCO 2002: Proceedings of the Genetic and Evolutionary Computation Conference*. Morgan Kaufmann Publishers, 1:716-723, 2002.
- Daubechies, I., Orthonormal bases of compactly supported wavelets, *Comm. Pure and Appl. Math.*, 41:909-996, 1988.
- Frasier, M.W., *An Introduction to Wavelets Through Linear Algebra*. Springer-Verlag, New York, 1999.
- Hiroyasu, T., Kobayashi, K., Nishioka, M., Miki, M., Diversity Maintenance Mechanism for Multi-Objective Genetic Algorithms Using Clustering and Network Inversion. In: *Proceedings of the 10th international conference on Parallel Problem Solving from Nature*, Dortmund, Germany, 1:722-732, 2008.
- MacQueen, J.B. Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*", Berkeley, University of California Press, 1:281-297, 1967.
- Mitchell, M., *An Introduction to Genetic Algorithms*. MIT Press, 1996.
- Mitchell, T., *Machine Learning*. McGraw Hill, 1997.
- Morettin, P.A., *Ondas e Ondaletas: da Análise de Fourier à Análise de Ondaletas*. São Paulo, Edusp, 1999.
- Morrison, R.W., De Jong, K.A.. Measurement of population diversity. In: *5th International Conference, Evolution Artificielle*, EA 2001, Le Creusot, France, 1:31-41. Springer, 2001.
- Nemati, S., Basiri, M.E., Ghasem-Aghae, N., e Aghdam, M.H., A novel ACO-GA hybrid algorithm for feature selection in protein function prediction. *Expert Syst. Appl.* 36(10):12086-12094, 2009.
- Sarkar, T.K., Salazar-Palma, M., Michael, C.W., *Wavelet Applications in Engineering Electromagnetics*. Artech House, Boston, 2002.
- Sean L., *Essentials of Metaheuristics*, online version, March, 2010. Disponível em <http://cs.gmu.edu/~sean/book/metaheuristics/>
- Tragante, V., Tinós, R., Diversity Control in Genetic Algorithms for Protein Structure Prediction. In: *VII Encontro Nacional de Inteligência Artificial (ENIA'2009)*, 2009, Bento Gonçalves. Anais do XXIX Congresso da Sociedade Brasileira de Computação (CSBC 2009), 1:727-737, 2009.
- Weise, T., *Global Optimization Algorithms: Theory and Application*, 2nd ed., online version, 2009. Disponível em <http://www.it-weise.de/>
- Zhu, K.Q. A Diversity-Controlling Adaptive Genetic Algorithm for the Vehicle Routing Problem with Time Windows. In: *Proceedings of the 15th IEEE international Conference on Tools with Artificial intelligence*. ICTAI. IEEE Computer Society, Washington, DC, 176, 2003.