

Análise Comparativa dos Algoritmos Genéticos Many-Objective em problemas de otimização discreta

Gabriel Augusto Marson
gabrielmarson@live.com

14 de novembro de 2016

Resumo

Esta investigação tem como objetivo mostrar o desempenho de alguns algoritmos genéticos multiobjetivos em problemas discretos clássicos da computação assim como em quais circunstâncias são mais adequados.

1 Introdução

Algoritmos genéticos (AGs) apresentam uma abordagem de busca baseada na teoria da evolução de Darwin que adotam uma sequência de passos para atingir um determinado objetivo. Elas tem em comum o fato de que os indivíduos mais adaptados passarão suas características aos seus descendentes.

Apesar de ter sido constatado o sucesso dos AGs para certos tipos de problemas, como de criptoaritmética, alguns outros permanecem como desafio. Os de otimização discreta com múltiplos objetivos são mais desafiadores pois requerem a maximização ou minimização de mais de um função objetivo.

Consideremos, por exemplo, o problema do Roteamento Multicast (PRM). Como foi dito em Bueno [1] ele pode ser modelado como um grafo direcionado onde os vértices são os hosts e as arestas representam os enlaces de comunicação. Dessa forma, dado o grafo G que representa a rede e um fluxo ϕ , deve-se calcular árvores enraizadas de G para carregar ϕ em G , partindo-se de um vértice que representa o início de uma conexão ou envio de pacote e passando por um subconjunto de vértices de G que tenham os nós folha como destino.

Existe uma série de atributos que podem influenciar o desempenho de uma rede tais como custo, delay, capacidade e tráfego atual. Um AG simples apenas consegue administrar um critério por vez. Os AGs Multi-Objective constituem uma abordagem mais adequada para resolver problemas desse tipo categorizando possíveis melhores respostas em virtude dos critérios avaliados.

Recentemente, uma versão conhecida por Algoritmos Genéticos Multiobjetivos (AGMOs) foi proposta, na qual o processo de avaliação dos indivíduos leve em consideração diferentes objetivos ou critérios.

A fim de considerar todos os objetivos no cálculo de aptidão, os AGMO's utilizam o conceito de fronteira de Pareto [2] que consiste em classificar vários indivíduos por meio de várias funções objetivo. Dessa forma, utilizou-se o conceito de fronteira de Pareto [2], que estabelece uma região do espaço de busca na qual estarão os melhores indivíduos de acordo com várias funções objetivo. O critério para a classificação de qual indivíduo permanece em qual fronteira é a dominância.

Diz-se que $x \preceq y$ (x domina y) se:

1. x para toda função objetivo, x não é pior que y .
2. x é melhor do que y em pelo menos uma função objetivo.

Esse tipo de classificação pode ser melhor compreendida se analisarmos um problema simples como a compra de um carro. Vários critérios podem ser utilizados para comprá-lo como preço e desempenho. Assim, na fronteira de Pareto, estariam os indivíduos com uma combinação de melhor preço e melhor desempenho. Estas duas métricas são normalmente conflitantes, quanto melhor o preço, pior o desempenho. Alguns carros terão bom preço e bom desempenho e portanto estarão mais próximos da fronteira de Pareto. Somente aqueles carros que não tiverem preço e desempenho piores que algum outro estarão na Fronteira de Pareto. Portanto, não existe uma única resposta para esse tipo de problema.

No NSGAII [3], além do agrupamento por fronteiras, ou rank, utiliza-se também a métrica *crowding distance*, como critério de desempate na escolha do indivíduo caso estejam na mesma fronteira de dominância e, também, para impedir que a variabilidade genética da população seja comprometida. O objetivo dessa técnica é que os indivíduos na fronteira de Pareto sejam mais o mais diversificado possível.

2 Objetivos

2.1 Objetivos Gerais

Em virtude de bons resultados gerados por esses algoritmos em problemas discretos, tal qual o PRM, serão analisados métodos de AG multi e *Many-objective* como NSGAII[3], SPEA2[4], SPEA3 e AEMMT aplicados a problemas clássicos da computação como o problema da Mochila, sendo que os dois últimos pertencem a classificação de *Many-objective*.

Além desses, pretende-se investigar uma nova família de AGs conhecida como AG *Many-objectives*, por serem apropriados a problemas com vários objetivos. Dentre esses AGMOs, podemos citar o SPEA3 e o AEMMT.

2.2 Objetivos Específicos

- Implementar os algoritmos SPEA2[4] e NSGAII[3]

- Realizar uma análise comparativa com as abordagens já existentes para resolução de problemas discretos com multi objetivos
- Estudo de desempenho de algoritmos *many-objective* como SPEA3 e AEMMT

2.3 Justificativas

Existem muitos problemas de otimização discreta multiobjetivo ainda não resolvidos de forma eficiente. O estudo desses problemas, assim como dos algoritmos que os resolvem, constituem prática importante para embasar novas descobertas com o intuito de otimizar as soluções já existentes ou de criar novas.

2.4 Metodologia

O método científico é o método no qual o pesquisador parte de uma premissa inicial do problema até a sua resposta, no final da pesquisa.

Será realizada uma pesquisa descritiva de caráter explicativo com o intuito de construir uma comparação entre os AGMOs e outras abordagens para resolver problemas de otimização discreta multi-objetivo. Pretende-se, também, explicar em quais circunstâncias é indicado o uso dos AGMOs. Para isso, será feita a implementação de alguns AGMOs como SPEA2[4] e NSGAII[3].

Os critérios de comparação serão, inicialmente, a complexidade dos algoritmos segundo a notação O e o desempenho que apresentam em problemas de otimização discreta. Com o objetivo de verificar o desempenho em funções simples, primeiramente, os algoritmos serão testados nos problemas de otimização de funções propostos em NSGAI[5] e NSGA-II[3].

Em seguida, os algoritmos serão aplicados em problemas mais complexos, tidos como clássicos, na computação. São os problemas da mochila e do caixeiro viajante. Espera-se que os AGMOs apresentem desempenho superior aos algoritmos convencionais quando aplicados a problemas de, no máximo, três objetivos.

O desempenho dos AGMOs diminuem consideravelmente a medida que uma maior quantidade de funções é requerida pelo problema, portanto, torna-se necessário o estudo dos algoritmos genéticos *many-objective*. Será feita uma análise descritiva do algoritmo bem como uma comparação de custo e resolução dos algoritmos atuais.

2.5 Cronograma

As atividades principais para desenvolver o artigo são:

- Estudo dos AGMOs(EAGMOs)
- Implementação dos AGMOs(IAGMOs)
- Análise Comparativa(AC)
- Estudo de Algoritmos Many-Objective(EMany-O)

- Dissertação(D)

Sendo que o planejamento para elas é o seguinte:

Tabela 1: Cronograma de Atividades

	EAGMOs	IAGMOs	AC	EMany-O	D
Agosto(2016)	X				
Setembro(2016)	X	X			
Outubro(2016)	X	X			
Novembro(2016)		X	X		
Dezembro(2016)		X	X		
Janeiro(2017)			X		X
Fevereiro(2017)				X	X
Março(2017)				X	X
Abril(2017)				X	X
Mai(2017)					X
Junho(2017)					X

3 Estado da Arte

predicao de estruturas de RNA

Escolha de decisão em mercados imobiliarios de alto risco

Application of multi-objective genetic algorithm to optimize energy efficiency and thermal comfort in building design

<http://www.sciencedirect.com/science/article/pii/S0378778814010305>

4 Trabalhos Correlatos

Trabalhos correlatos constituem artigos que utilizam técnicas e abordagens semelhantes à adotada nesta tese. Os próximos estudos descritos, tem como principal objetivo realizar comparações de AGMOs dado um escopo predefinido de problema.

A análise comparativa proposta nesta tese pretende auxiliar pesquisas futuras no que diz respeito à comparações dos algoritmos genéticos multiobjetivo tradicionais em problemas discretos.

4.1 Comparison Study of SPEA2+, SPEA2, and NSGA-II in Diesel Engine Emissions and Fuel Economy Problem

Nesse artigo[6] é feito o uso de algoritmos como SPEA2 e NSGAII em uma análise comparativa de funções que quantificam as emissões de gases em motores

a diesel(SFC, NOx e Soot). Esse problema é classificado como *Diesel engine fuel emission scheduling problem* e os algoritmos aplicados nesse problema consistem em minimizar a emissão desses tipos de gases de forma simultânea.

Em virtude de calcular a eficácia de cada solução, foi determinada uma métrica de proporção comparada a um conjunto ideal, ou seja, quanto mais próximo os indivíduos não dominados estiverem da solução ideal, melhor será aquele conjunto de soluções. Seja S_1 e S_2 possíveis conjuntos de populações finais para o problema proposto e seja S_u o conjunto união dessas soluções. A partir de S_u serão selecionados os indivíduos não dominados para um conjunto S_p . Quanto mais próximas as soluções estiverem de S_p melhores serão consideradas.

O artigo conclui que o algoritmo SPEA2+ [7] é ligeiramente mais eficiente do que os tradicionais SPEA2 pois encontra soluções mais diversificadas. Além disso, por meio dos gráficos apresentados conclui-se que a capacidade de mostrar soluções diversificadas, para o problema em questão, do algoritmo NSGAII é consideravelmente inferior ao SPEA2.

4.2 Multi-Objective Optimization of Vehicle Passive Suspension System using NSGA-II, SPEA2 and PESA-II

No âmbito engenharia mecatrônica, pesquisadores aplicaram algoritmos genéticos multiobjetivo com ênfase na suspensão de automóveis. De acordo com o artigo[8], o objetivo era resolver o conflito entre proporcionar mais conforto para os passageiros e motorista ou aumentar a capacidade de controle do automóvel. Como esses dois critérios são conflitantes, foi utilizada uma abordagem de testes e comparações com AGMOs.

O estudo demonstrou que as fronteiras de Pareto obtida pelo NSGAII foi considerada de rendimento extremo, sendo superior ao SPEA2 e ao PESAI[9]. No entanto, NSGAII perde para os dois outros métodos em termos de diversidade de soluções na fronteira de Pareto.

5 Referencial Teórico

5.1 Algoritmos Genéticos

Publicada em 1859, a obra **Na Origem das Espécies – Sob o Conhecimento da Seleção Natural**, de Charles Darwin, foi alvo de críticas pois explicitava que os seres de um ambiente mudam de uma determinada geração para outra e são selecionados de acordo com essas mudanças. A partir disso, Darwin escreveu que todos os seres vivos possuem um ancestral comum, contrastando com a visão criacionista da época.

Essa teoria foi imprescindível para a criação dos algoritmos genéticos por John Henry Holland. Físico formado pelo MIT (Massachusetts Institute of Technology), Holland elaborou um algoritmo não determinístico com foco em resolução de problemas de custo muito alto.

O AG consiste em dar uma interpretação para um conjunto de dados e evoluí-los para uma solução ótima. Todos os dados seriam uma população e cada indivíduo representa um dado, além de ser uma solução para o problema. O algoritmo avalia cada membro usando uma métrica de aptidão. Isso consiste em dar uma nota para cada indivíduo que simboliza as chances dele transmitir suas características para a próxima geração.

Alguns métodos são utilizados para selecionar os indivíduos. Dado que o método gera dois filhos por cruzamento de um par de pais, serão escolhidos N pares de pais, onde N é o tamanho da população inicial. O processo de troca de material genético chamado *crossover* retorna um ou mais filhos sendo que, geralmente, existe uma pequena chance de ocorrer uma mutação no material genético dos filhos, em virtude de aumentar a variabilidade na população.

Após o cruzamento, os filhos resultantes são reinseridos na população. Com o intuito de manter o custo computacional não crescente e de eliminar indivíduos fracos, utilizam-se métodos de reinserção para diminuir a população.

Este ciclo se repete por um número determinado gerações ou pode ser parado quando todos os indivíduos possuem o mesmo material genético (é improvável que a mutação nas próximas gerações produza um indivíduo mais apto).

5.2 Fronteira de Pareto

FALAR COISAS AKI

5.3 Non-dominated Sorting Genetic Algorithm II

O Non-dominated Sorting Genetic Algorithm(NSGA-II)[3] é um algoritmo baseado em uma ordenação elitista por dominância. Tem como objetivo classificar os indivíduos de um conjunto M por fronteiras F sendo que na fronteira $F1$ estariam os melhores indivíduos por critérios de dominância de todo o conjunto M . Na fronteira $F2$, estariam todos os indivíduos não dominados de $M - F1$ e assim sucessivamente. A classificação dos indivíduos ocorre utilizando-se o conceito de dominância que pode ser explicado usando os seguintes critérios.

1. ndi o número de soluções que dominam a solução i .
2. Ui o conjunto de soluções que são dominadas por i . Indivíduos com menor ndi possível ($ndi = 0$) estão na fronteira de Pareto. As demais fronteiras são estabelecidas pelos indivíduos com ndi subsequente.

O NSGAII trabalha com duas populações P e Q sendo que P é a população inicial da geração corrente e Q são os filhos de P resultantes de um processo de cruzamento após a seleção. Ao fim de cada geração, essas duas populações convergem para uma só e ocorrerá um processo de seleção de todos os indivíduos para formar uma nova população P que será a inicial do próximo ciclo. Serão escolhidos, portanto, os indivíduos que estão nas melhores fronteiras até o número total de P ser preenchido. Como critério de desempate para o caso de a última fronteira ser maior do que o restante de elementos faltantes em P , utiliza-se o

critério de desempate *Crowding Distance* em que são priorizados os indivíduos com maior variabilidade genética.

O cálculo dessa métrica é feito obtendo-se a soma da média da distância das duas soluções adjacentes a cada indivíduo para cada objetivo referente a ele. Dessa forma, são selecionados os indivíduos que estão mais distantes um dos outros na fronteira.

A aptidão de cada solução é determinada usando-se os seguintes critérios.

1. *Rank*(ou fronteira) em que o indivíduo está.
2. *Crowding Distance* como desempate para indivíduos na mesma fronteira.

Ao fim desse processo espera-se vários, ou todos, os indivíduos na fronteira de Pareto com os variabilidade elevada.

5.4 Strength Pareto Evolutionary Algorithm 2

O Strength Pareto Evolutionary Algorithm 2[4] é um algoritmo que trabalha com duas populações P e Q sendo que P é a população inicial e em Q serão apenas armazenadas as soluções não dominadas encontradas pelo algoritmo. Serão denotados por P_t e Q_t as populações para a geração $t = 1, 2, \dots, N_{iter}$.

Inicia-se estabelecendo uma população inicial aleatória para P_1 e classificando todos os indivíduos de $R_t = P_t \cup Q_t$. Um dos critérios que irá compor a função de aptidão(*fitness*) é o *strength* de cada indivíduo que pode ser estabelecido por:

$$strength_i = |\{j, j \in R_t, |i \leq j\}|$$

Como foi dito em Waldo e Alexandre, [10] a interpretação para o valor *strength* é a quantidade de elementos em R_t que são dominados por i . Calcula-se, também o valor de *raw fitness* denotado na equação como:

$$raw_i = |\{\sum_{j \in R_t, j \leq i} strength_j\}|$$

Sendo assim, raw_i é o somatório dos $strengths_j$ tal que $j \in R_t$ e j domina i . As soluções com $raw_i = 0$ não são dominadas por nenhum outro indivíduo.

Referências

- [1] M. L. d. P. Bueno, “Heurísticas e algoritmos evolutivos para formulações mono e multiobjetivo do problema do roteamento multicast,” 2010.
- [2] D. A. Van Veldhuizen and G. B. Lamont, “Evolutionary computation and convergence to a pareto front,” in *Late breaking papers at the genetic programming 1998 conference*, pp. 221–228, Citeseer, 1998.
- [3] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multi-objective genetic algorithm: Nsga-ii,” *IEEE Transactions on Evolutionary Computation*, vol. 6, pp. 182–197, Apr 2002.
- [4] E. Zitzler, M. Laumanns, L. Thiele, *et al.*, “Spear2: Improving the strength pareto evolutionary algorithm,” in *Eurogen*, vol. 3242, pp. 95–100, 2001.
- [5] N. Srinivas and K. Deb, “Multiobjective optimization using nondominated sorting in genetic algorithms,” *Evolutionary computation*, vol. 2, no. 3, pp. 221–248, 1994.
- [6] T. Hiroyasu, S. Nakayama, and M. Miki, “Comparison study of spear2+, spear2, and nsga-ii in diesel engine emissions and fuel economy problem,” in *2005 IEEE congress on evolutionary computation*, vol. 1, pp. 236–242, IEEE, 2005.
- [7] M. Kim, T. Hiroyasu, M. Miki, and S. Watanabe, “Spear2+: Improving the performance of the strength pareto evolutionary algorithm 2,” in *International Conference on Parallel Problem Solving from Nature*, pp. 742–751, Springer, 2004.
- [8] B. Gadhvi, V. Savsani, and V. Patel, “Multi-objective optimization of vehicle passive suspension system using nsga-ii, spear2 and pesa-ii,” *Procedia Technology*, vol. 23, pp. 361–368, 2016.
- [9] D. W. Corne, N. R. Jerram, J. D. Knowles, M. J. Oates, and M. J., “Pesa-ii: Region-based selection in evolutionary multiobjective optimization,” in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO’2001)*, pp. 283–290, Morgan Kaufmann Publishers, 2001.
- [10] W. G. C. Ticona and A. C. B. Delbém, “Algoritmos evolutivos para otimização multi-objetivo,” Aug 2008.