

IMPLEMENTAÇÃO DE UMA VERSÃO INTERATIVA DO NSGA-III NA FERRAMENTA OPLA-TOOL

Lucas Wolschick (PIC/UEM), Willian Marques Freire (Co-orientador), Aline Maria Malachini Miotto Amaral (Orientadora). E-mail: ammmamaral@uem.br.

Universidade Estadual de Maringá, Centro de Tecnologia, Maringá, PR.

Área e subárea do conhecimento: Ciências Exatas e da Terra / Metodologia e Técnicas da Computação.

Palavras-chave: engenharia de software baseada em busca; algoritmos evolutivos multiobjetivos; otimização em engenharia de software.

RESUMO

Uma Linha de Produto de *Software* é um conjunto de sistemas de *software* com características semelhantes construídos em torno de uma Arquitetura de Linha de Produto de *Software* (ALPS), que fornece uma base de *software* reutilizável e extensível. Projetar uma ALPS é uma tarefa complexa que envolve vários fatores muitas vezes conflitantes. A abordagem MOA4PLA, implementada pela ferramenta OPLA-Tool, descreve um processo de otimização de ALPS que objetiva melhorar a qualidade das arquiteturas. Ela descreve a qualidade de uma ALPS em termos de métricas que representam aspectos de qualidade de *software*, e utiliza algoritmos genéticos, como o NSGA-II, para evoluir um conjunto de soluções de acordo com essas métricas. O arquiteto de *software* é capaz de guiar o processo de otimização dizendo quais métricas devem ser utilizadas antes da execução e também através da avaliação de soluções intermediárias, de modo interativo. Entretanto, o NSGA-II apresenta resultados satisfatórios apenas para até três métricas, a partir do qual o seu desempenho é prejudicado. O NSGA-III é uma evolução do NSGA-II que apresenta melhor desempenho para problemas com vários objetivos. Este trabalho buscou implementar uma versão interativa do algoritmo NSGA-III na ferramenta OPLA-Tool, a fim de avaliar se a abordagem pode oferecer alternativas mais adequadas aos arquitetos durante a otimização de ALPS. O estudo comparou o desempenho dos algoritmos NSGA-II e NSGA-III e atestou que o NSGA-III atingiu desempenho superior ao NSGA-II na otimização de vários objetivos de ALPS.

INTRODUÇÃO

A Engenharia de *Software* Baseada em Busca é uma subárea da Engenharia de *Software* que busca reformular problemas tradicionais da área em problemas de busca, para a aplicação de técnicas de otimização meta-heurísticas (Harman *et al.*, 2001). O projeto de Arquiteturas de Linhas de Produto de *Software* (ALPS), artefatos que descrevem a arquitetura de um conjunto de *softwares* que pertencem a uma mesma linha de produto, é uma tarefa difícil e custosa e requer um alto grau de

experiência e atenção da parte do arquiteto de software responsável. A abordagem MOA4PLA (*Multi-objective Optimization Approach for PLA*) descreve um protocolo de avaliação e otimização de ALPS através do uso de algoritmos genéticos (Colanzi *et al.*, 2014), e é implementada na ferramenta OPLA-Tool, que permite ao arquiteto otimizar uma ALPS de acordo com parâmetros definidos por ele. A otimização pode ser opcionalmente interativa, onde o processo é interrompido e soluções intermediárias são exibidas ao arquiteto para avaliação (Freire *et al.*, 2020).

Uma meta-heurística utilizada em problemas de otimização multiobjetivo é o NSGA-II, um algoritmo genético, que é oferecido pela ferramenta OPLA-Tool para o processo de otimização. Entretanto, este algoritmo não possui bom desempenho para otimização de quatro ou mais objetivos, não conseguindo gerar conjuntos de soluções de boa qualidade. Pensando nisso, o algoritmo NSGA-III, contendo estratégias de preservação da diversidade das soluções, foi proposto como melhoria ao NSGA-II. O NSGA-III apresenta melhor desempenho em otimização com vários objetivos em relação ao NSGA-II. O objetivo desta pesquisa foi avaliar comparativamente o desempenho do algoritmo NSGA-III em relação ao NSGA-II no problema de otimização de ALPS sob a abordagem MOA4PLA (Deb, Jain, 2013). Para isso, implementamos o NSGA-III na ferramenta OPLA-Tool e executamos uma série de experimentos comparando os dois algoritmos.

MATERIAIS E MÉTODOS

ALPS podem ser avaliadas de acordo com critérios objetivos que descrevem características de qualidade de software, neste trabalho denominados métricas. Foram utilizadas seis métricas ao todo: acoplamento de classes, coesão, modularização orientada a *features*, função convencional, extensibilidade e elegância. Para avaliar o desempenho do processo de otimização, são utilizados indicadores como a distância euclidiana (ED) e o *hypervolume* (HV), que analisam as soluções individualmente e em agregado, respectivamente (Bechikh *et al.*, 2015), e utilizam os valores das métricas nos seus cálculos. EDs baixas indicam soluções mais otimizadas e HVs altos indicam alta diversidade de soluções.

Implementação do NSGA-III

A ferramenta OPLA-Tool utiliza a versão 4 da biblioteca *jMetal* como uma base estrutural para o processo de otimização de ALPS. A biblioteca fornece implementações de várias meta-heurísticas, mas o algoritmo NSGA-III apenas estava disponível na versão 6 da biblioteca. Após analisarmos de alternativas, realizamos o *backport* do algoritmo NSGA-III para a versão 4 da biblioteca *jMetal*, e adequamos o algoritmo para utilizar a infraestrutura de interação já existente na ferramenta. Além disso, o algoritmo foi modificado para assumir um tamanho de população especificado pelo usuário, pois na implementação original esse número era restrito ao número de pontos de referência, conforme Deb e Jain (2013).

Validação da Implementação e Análise de Desempenho

Após a implementação do NSGA-III, elencamos protocolos de testes já utilizados em trabalhos anteriores como base para a análise comparativa do desempenho com o NSGA-II. Elaboramos um protocolo de teste comum para os experimentos, fixando parâmetros e métricas em comum. Os experimentos foram realizados sobre três ALPS utilizadas em trabalhos anteriores. Para a análise, utilizamos os valores das métricas das soluções geradas, os quais foram normalizados e utilizados no cálculo do HV de cada execução. Os HV resultantes de cada experimento foram comparados para verificar se havia diferença estatística entre os resultados dos dois algoritmos. Os resultados foram resumidos em gráficos box-plot. A análise estatística se deu com a linguagem de programação R. Devido ao tempo restrito, não conseguimos realizar uma comparação efetiva do desempenho do NSGA-III interativo com o NSGA-II, mas isto está planejado para estudos futuros.

RESULTADOS E DISCUSSÃO

Dez experimentos foram realizados ao todo. Desses, três utilizaram três métricas, três utilizaram quatro, três utilizaram cinco e um utilizou seis. Alguns experimentos (3 e 6) buscaram medir se as mudanças na implementação do NSGA-III afetaram negativamente o seu desempenho, e em ambos os casos concluiu-se que não. A tabela 1 contém dados de desempenho sumarizados dos experimentos realizados após a implementação do NSGA-III na ferramenta OPLA-Tool. A tabela descreve o número do experimento, as medianas da ED e do HV dos dois algoritmos (NSGA-II e NSGA-III respectivamente), o valor-p do teste de diferença estatística dos HVs e a medida de tamanho de efeito.

Tabela 1

Id	M _{DE} (A)	M _{DE} (B)	M _{hV} (A)	M _{hV} (B)	valor-p	VD-A
1	82.079	44.702	0.392	0.592	0.0001286	0.2122222 (grande)
2	61.515	39.193	0.453	0.487	0.1242000	0.3844444 (pequeno)
3	39.193	52.307	0.351	0.487	0.0000042	0.8311111 (grande)
4	111.311	68.198	0.321	0.497	0.000000027	0.0822222 (grande)
5	114.033	64.753	0.276	0.320	0.002439	0.2722222 (médio)
6	34.423	89.578	0.487	0.338	0.009267	0.6955556 (médio)
7	119.034	93.087	0.517	0.496	0.332	0.5933333 (pequeno)
8	200.029	141.032	0.340	0.521	2.00E-16	0.0322222 (grande)
9	360.125	349.405	0.505	0.497	0.14	0.6033333 (pequeno)
10	141.789	382.467	0.333	0.491	2.00E-16	0.133333333 (grande)

Observa-se que no geral, o NSGA-III obteve um desempenho melhor que o NSGA-II para os casos de otimização com vários objetivos, e obteve desempenho semelhante ou ligeiramente melhor nos casos multiobjetivos. A figura 1 representa alguns box-plots selecionados dos experimentos, e indicam uma diferença significativa entre os dois algoritmos. Os experimentos 4, 8 e 10 foram executados com cinco métricas e o experimento 5, com seis.

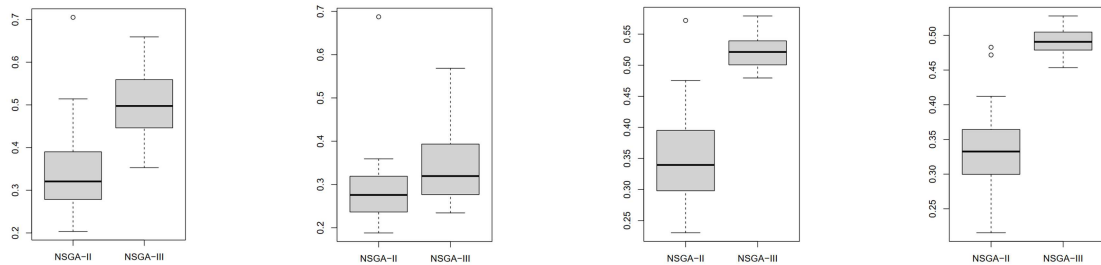


Figura 1 – Box-plots dos HV dos experimentos 4, 5, 8 e 10.

CONCLUSÕES

Este trabalho visou verificar se o algoritmo NSGA-III oferece uma vantagem de desempenho sobre o NSGA-II em um contexto de otimização de vários objetivos de ALPS na abordagem MOA4PLA. Através da execução de experimentos de acordo com um protocolo formal, foi possível constatar que o NSGA-III evidentemente apresenta melhor desempenho sobre o NSGA-II, superando-o na métrica HV em quase todos os experimentos, com tempo de execução comparável. Em trabalhos futuros pretendemos analisar o impacto da interatividade no algoritmo NSGA-III mais profundamente, envolvendo mais arquitetos, através de um método comparativo assim como neste trabalho.

REFERÊNCIAS

BECHIKH, S. et al. Preference incorporation in evolutionary multiobjective optimization: A survey of the state-of-the-art. **Advances in Computers**. Elsevier, 2015. p. 141-207.

COLANZI, T. E. et al. A search-based approach for software product line design. In: **Proceedings of the 18th International Software Product Line Conference**, v. 1. p. 237-241, set. 2014.

DEB, K.; JAIN, H. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: solving problems with box constraints. **IEEE transactions on evolutionary computation**, v. 18, n. 4, p. 577-601, set. 2013.

FREIRE, W. M. et al. Opla-tool v2.0: a tool for product line architecture design optimization. In: **Proceedings of the XXXIV Brazilian Symposium on Software Engineering**, p. 818-823, out. 2020.

HARMAN, M; JONES, B. F. Search-based software engineering. **Information and software Technology**, v. 43, n. 14, p. 833-839, dez. 2001.