



UNIVERSIDADE FEDERAL DE OURO PRETO

Mestrado em Ciência da Computação

Relatório Sobre a Metodologia Utilizada para Escrita da
Dissertação: “*Finding Maximum Clique with a Genetic
Algorithm*”

Autor:

Rodolfo Labiapari Mansur Guimarães - rodolfolabiapari@gmail.com

Professor Orientador:

Gladston Juliano Prates Moreira - gladston.moreira@gmail.com

Ouro Preto - MG

1 de agosto de 2016

Sumário

1	Introdução	1
2	Análise do <i>Abstract</i>	1
3	Análise do Documento em Si	2
4	Conclusão Sobre a Análise	3

1 Introdução

Este trabalho consiste na avaliação da escrita e metodologia utilizada em algum documento confeccionado por algum aluno de Pós-Graduação na área de Ciência da Computação verificando os requisitos que possam tornar este mais claro, coeso como assunto e resultados.

Sendo assim, o documento a ser analisado é a dissertação de Bo Huang da Universidade do Estado da Pensilvânia com o título *Finding Maximum Clique with a Genetic Algorithm*, disponível publicamente em: <https://turing.cs.hbg.psu.edu/mspapers/sources/bo-huang.pdf>.

Serão analisados vários tópicos específicos sendo o estes o *abstract*, a clareza do texto escrito junto com o propósito do trabalho e os dados estatísticos já que este documento possui análise comparativa de algoritmos.

O objetivo deste trabalho não é desqualificar o autor deste, mas sim, analisá-lo em busca de propor melhoria aprendidas na disciplina de Metodologia Científica ministrada pelo professor Moreira.

2 Análise do *Abstract*

O modelo de *abstract* para compor um bom texto, proposto pelo Professor Doutor Valtecir Zucolotto em suas vídeo-aulas *Produção de Artigos de Alto Impacto - Curso de Escrita Científica, Módulo 02 - "Títulos, Autores e Abstract"*, deve possuir os seguintes itens:

1. **Contextualização:** Identificação da grande área de trabalho. Importância da grande área;
2. **GRAP:** Quais ideias dessa área precisam ser trabalhadas/pesquisadas, o que está aberto ou incompreendido;
3. **Propósito:** Conclusão do GRAP;
4. **Metodologia:** Metodologia abordada;
5. **Resultados:** Item obrigatório que aponta os resultados obtidos no trabalho;
6. **Conclusão:** Como os resultados pode ajudar na grande área.

O *abstract* original segue abaixo:

This paper presents a hybrid genetic algorithm for the maximum clique problem. The main features of the algorithm are a strong local optimizer to accelerate the convergence and the adaptive genetic operators to fine tune the algorithm. Our algorithm was tested on DIMACS benchmark graphs of size up to 3,361 vertices and up to 4,619,898 edges. For more than 92% of 64 graphs in the test set, our algorithm finds the maximum clique, and for a small number of graphs, the results achieved by our algorithm are almost as large as the best-known results found by other algorithms. This result is comparable to one of the best existing algorithms for the maximum clique problem.

Após analisado, temos o seguinte resultado:

This paper presents a hybrid genetic algorithm for the maximum clique problem. The main features of the algorithm are a strong local optimizer to accelerate the convergence and the adaptive genetic operators to fine tune the algorithm. Our algorithm was tested on DIMACS benchmark graphs of size up to 3,361 vertices and up to 4,619,898 edges. For more than 92% of 64 graphs in the test set, our algorithm finds the maximum clique, and for a small number of graphs, the results achieved by our algorithm are almost as large as the best-known results found by other algorithms. This result is comparable to one of the best existing algorithms for the maximum clique problem.

É possível ver que na primeira parte vermelha temos a Contextualização (1), como sugerido pelo Professor Zucolotto. Em seguida, na parte azul, temos o Propósito do trabalho (3), pulando o item GRAP (2). Já pode-se concluir que este *abstract* não é tão rico em informações tal como deveria ser. Continuando, na segunda parte em vermelho temos os Resultados (5) pulando novamente outra etapa que no caso é a de Metodologia (4) no qual especificaria como o trabalho foi desenvolvido. E por fim, o resumo também não conclui nada após os resultados (6).

O pesquisador que realizará a leitura deste *abstract* e não possuirá nenhuma informação concreta do documento sendo assim ser obrigado a ler o documento por completo para assim decidir se este é ou não algo que lhe ajudará em sua pesquisa.

3 Análise do Documento em Si

Como pontos positivos do trabalho é possível ver que a dissertação descreve muito bem o problema a ser abordado e a natureza do algoritmo utilizado para a resolução deste. Descreve várias aplicações deste problema mostrando claramente sua boa utilidade além da sua classificação como NP-Completo¹ justificando o uso de uma Metaheurística para tal e não um algoritmo determinístico.

Ele também descreve várias formas de abordar este problema mostrando os algoritmos e sua estrutura de funcionamento e ao final descreve como o algoritmo híbrido, algoritmo proposto no trabalho, funciona. Para os testes, o autor também fornece os dados do computador utilizado para execução dos testes o que permite que outros possam reproduzir o mesmo experimento comparando cada um dos resultados.

O experimento é realizado com várias instâncias diferentes, comparando com três algoritmos, mostrando a diversidade dos seus testes, mas com poucas comparações de qualidade.

Entretanto, por mais que a dissertação deste autor seja clara o suficiente para pessoas de vários níveis de conhecimento consigam entender a proposta apresentada, poucos resultados são exibidos mostrando realmente o potencial do algoritmo em comparação com outros concorrentes.

Como é descrito pelo autor, realizou-se capturas do tempo de execução de cada algoritmo em cada instância e ao final é feita uma média de cada um destes e comparando-os. É possível perceber a pobreza nas comparações pois este realiza somente uma média dos valores sem exibir nenhuma informação a mais das execuções. Fica claro que ele poderia expor todas as suas execuções numa tabela para

¹NP significa Tempo Polinomial Não-Determinístico.

que outros pesquisadores possam reproduzir os resultados a fim de comparar estes com outros algoritmos.

Também foi possível que o autor não utilizou nenhuma forma de visualização de resultado como gráficos ficando somente pela visualização de dados por meio de tabelas. Entretanto, tais tabelas são insuficientes no quesito de informações a serem passadas para o leitor.

Exibindo os resultados obtidos de forma mais científica mostrando qual distribuição os resultados do seu algoritmo pertence, a taxa de desvio padrão e variância dos seus resultados, confiabilidade além de todos os cálculos estatísticos, tornaria seu documento mais rico em informações úteis que podem servir de conclusão para quem põe-se a ler o texto. Sem estas informações, é impossível comprovar de fato os resultados de seu novo algoritmo para que a sociedade científica possa interpretar realmente o desempenho deste e utilizá-lo como fundamentos para futuros outros.

4 Conclusão Sobre a Análise

O documento possui uma linguagem de fácil leitura além de boa fundamentação para qualquer pessoa que deseja aprender sobre o problema de Clique-Máximo ou sobre Meta-Heurísticas. O documento também termina mostrando que o algoritmo proposto possui uma eficácia de 92% dos testes realizados além de sugestões de problemas futuros.

Todavia, seu texto pode ser classificado mais como texto ‘didático’ que um texto ‘científico’ já que a sua apresentação (*abstract*) e representação de resultados não exibem realmente os resultados claros e concisos obtidos pelo autor. O autor necessitaria de realizar testes estatísticos e exibi-los no documento para que os leitores tenham mais informações sobre o resultado obtido. Ao final, o autor não especifica quais são as qualidades boas e ruins do seu algoritmo deixando detalhes importantes obscuros à visão do leitor.