Relatório de Iniciação Científica

Algoritmos, Experimentação e Teoria em Otimização Combinatória

Aluna: Juliana Barby Simão
Orientador: José Coelho de Pina Jr.

Período: abril a julho de 2004
Processo n. 04/00580-8

Resumo

Este relatório apresenta um resumo das atividades desenvolvidas durante os primeiros meses da iniciação científica. O projeto visa ao estudo de técnicas recorrentes em otimização combinatória, através da análise, implementação e experimentação de algoritmos e do estudo de teoria. A iniciação está sendo desenvolvida em conjunto com o colega Marcelo Hashimoto (04/0581-4).

Após uma introdução sobre as motivações para o projeto, o documento apresenta um resumo dos objetivos iniciais, seguido pela descrição das atividades realizadas no período considerado. Por fim, o cronograma original e um cronograma revisado para o segundo semestre são apresentados.

Anexo a este relatório, segue o texto que estamos produzindo abrangendo a teoria estudada. Além disso, o projeto possui um sítio na Internet, ainda em desenvolvimento, localizado em: `http://www.ime.usp.br/~coelho/oticomb/`.

1 Introdução

Otimização combinatória é um campo da matemática aplicada que se baseia no uso conjunto de técnicas de combinatória, programação matemática e teoria de desenvolvimento de algoritmos para resolver problemas de otimização formulados sobre estruturas discretas.

Problemas de otimização combinatória têm sido um tópico central para a evolução de algoritmos e da teoria de *complexidade computacional*. Pesquisadores têm apresentado muitas idéias criativas para o projeto de algoritmos eficientes, baseados em conceitos e resultados na área. A sub-área conhecida como *combinatória poliédrica*, por exemplo, apresenta métodos desenvolvidos a partir de conceitos de *programação linear*, como o *primal-dual*, e que têm se mostrado muito úteis no projeto e análise de uma variedade de algoritmos para problemas em outros domínios. Muitas das idéias

inovadoras têm se baseado em um conjunto não muito grande de princípios comuns, como *scaling*, que são simultaneamente simples e poderosos.

Nem sempre, porém, é suficiente restringir a pesquisa de um problema ao estudo teórico de seus aspectos e ao desenvolvimento abstrato de algoritmos que o resolvem. Em diversos casos, existem fatores fundamentais que só podem ser resolvidos, ou mesmo descobertos, durante a implementação propriamente dita. Não é incomum que a obtenção de uma eficiência próxima àquela calculada na teoria, ou mesmo o próprio funcionamento de um algoritmo, dependa do uso de estruturas de dados sofisticadas e técnicas de programação que não são triviais.

Ademais, a complexidade de um algoritmo na teoria é muitas vezes um argumento questionável para classificá-lo na prática: não são incomuns algoritmos teoricamente eficientes mas lentos na prática, e algoritmos cuja análise de pior caso está longe de representar seu comportamento no caso médio. Diante dessa possibilidade, testes e experimentações são ferramentas poderosas e úteis, e que enfatizam ainda mais a importância da implementação.

2 Resumo do plano inicial

Estudar algumas das ferramentas mais fundamentais em otimização combinatória através do estudo e implementação de algoritmos para problemas clássicos da área envolvendo fluxos em redes. A iniciação científica está sendo desenvolvida em conjunto com o colega Marcelo Hashimoto e é um dos objetivos do projeto a nossa preparação para um possível mestrado na área.

A fonte dos algoritmos incluídos no estudo, e também a principal base teórica, é o livro *Network Flows* de Ahuja, Magnanti e Orlin [1]. Também estamos utilizando como apoio secundário o livro *Combinatorial Optimization* de Cook, Cunningham, Pulleyblank e Schrijver [2] e, para os tópicos mais teóricos, livros relacionados de Schrijver [13, 14]. Não nos restringiremos, porém, a seguir somente tais referências: consultas a outras fontes são esperadas e, dependendo do contexto, desejadas.

Como corolário de nosso trabalho, será produzido um texto contendo a descrição de toda a teoria envolvida nos problemas e em cada um dos algoritmos estudados. Também pretendemos montar uma biblioteca de implementações desses algoritmos, que nos permitirá realizar um estudo comparativo de desempenho experimental, cujos resultados também serão incluídos na documentação final.

Pretendemos que a linguagem algorítmica presente no texto final siga os padrões das notas de Feofiloff [6, 7] e nos permita descrever os algoritmos com base em seus invariantes, para destacar de maneira simples a razão pela qual eles funcionam.

Como queremos desenvolver as implementações de modo que elas possam ser facilmente compreendidas e aplicadas, elas serão disponibilizadas na linguagem `CWEB` [12] de programação literária [10]. Pretendemos que a versão final dos programas en-

tregues utilize principalmente as estruturas de dados fornecidas pela plataforma `SGB` para algoritmos em grafos, desenvolvida por Knuth [11].

Esta iniciação é na verdade parte de um projeto maior, no qual estão envolvidos, além de mim e do Marcelo, os alunos Roger Ricardo Flores de Araújo e Nelson Guedes Paulo Júnior. O Roger está atualmente desenvolvendo uma iniciação científica similar à nossa, mas sobre algoritmos para emparelhamentos, e o Nelson está desenvolvendo um programa escrito na linguagem `C++`, que permitirá a visualização gráfica do funcionamento de todos os algoritmos implementados.

Pretendemos reunir todos os textos e implementações em um sítio da Internet, preferencialmente nos moldes da *Network Optimization Library* de Goldberg [9].

3 Atividades já realizadas

Inicialmente, estudamos alguns tópicos de otimização combinatória através dos livros de Cook et al. [2], Schrijver [14] e Cormen et al. [3] para nos familiarizar um pouco mais com a teoria envolvida nos problemas. Em particular, fizemos uma pequena revisão de programação linear, estudamos o *problema dos caminhos mínimos* sob o ponto de vista da combinatória poliédrica e discutimos algumas questões sobre *unimodularidade total*. Partimos, então, para o *problema do fluxo máximo* e, a partir desse ponto, a base teórica de nossos estudos foi quase puramente o livro de Ahuja et al. [1], como previsto no planejamento inicial.

Antes de nos dedicarmos a fundo ao estudo teórico do problema, trabalhamos em cima de uma implementação em `C` para o *método dos caminhos de aumento*, proposto por Ford e Fulkerson [8], que fizemos para a disciplina MAC0328 - Algoritmos em Grafos. A implementação utilizava a plataforma `SGB` e aplicava ao método uma especialização proposta por Edmonds e Karp [5], a qual resultou no algoritmo conhecido como *algoritmo dos caminhos de aumento de comprimento mínimo*. Revisamos nossas implementações e realizamos experimentos com algumas instâncias específicas do problema.

Partindo de minha implementação revisada, desenvolvi uma implementação para uma versão alternativa do método dos caminhos de aumento, conhecida como *algoritmo dos fluxos bloqueadores de aumento*, proposta por Dinits [4]. Enquanto isso, o Marcelo modificou sua versão da implementação e desenvolveu implementações para o *algoritmo dos caminhos de maior aumento* e *capacity scaling*.

Realizamos pequenos testes experimentais com as implementações que fizemos para verificar e comparar seus desempenhos. Os testes iniciais foram feitos sobre instâncias básicas do problema. Para a maioria dos experimentos, porém, utilizamos um gerador de instâncias desenvolvido por Cherkassky e Goldberg [9], capaz de gerar entradas grandes e reconhecidamente difíceis para os algoritmos implementados.

Após a execução e análise dos testes mencionados, começamos a estudar mais a

fundo os conceitos abordados em nosso livro-texto [1] e a discutir algumas dúvidas teóricas entre nós e com o orientador. As principais discussões foram em torno da análise da complexidade dos algoritmos e da modelagem do problema sob o contexto de programação linear e *dualidade*. De uma maneira resumida, estudamos a teoria envolvida nos conceitos básicos do problema do fluxo máximo, particularmente o *teorema do fluxo máximo e corte mínimo*, e na análise dos já mencionados algoritmos de caminhos de aumento. Atualmente, estamos estudando outra classe de algoritmos para fluxo máximo, que utiliza o método conhecido como *pré-fluxo*.

Além das atividades acima descritas, estivémos desenvolvendo um texto que apresenta a teoria estudada. A versão atual do documento está anexada a este relatório, mas ainda não está completa. Tanto o texto como as implementações realizadas podem ser encontrados no sítio: `http://www.ime.usp.br/~coelho/oticomb/`.

De acordo com o mencionado no projeto, o desenvolvimento dessa iniciação científica não é verdadeiramente linear. Por exemplo, antes de iniciarmos o estudo do livro *Network Flows* [1], nos dedicamos a alguns tópicos diversos, mencionados anteriormente, e revisamos conceitos. Além disso, discutimos aspectos de combinatória poliédrica relativos ao problema estudado e também sobre o *problema do emparelhamento máximo em grafos bipartidos*. Juntando a isso o fato de que as disciplinas do primeiro semestre, dentre elas uma de pós-graduação, exigiram bastante dedicação, foi preciso efetuar um reajuste no cronograma.

4 Plano de trabalho para o próximo período

O cronograma inicial foi construído em função de nossa expectativa de estudo de teoria e algoritmos contidos nos seguintes capítulos do livro *Network Flows* [1]:

6 Maximum flows: basic ideas

7 Maximum flows: polynomial algorithms

9 Minimum cost flows: basic algorithms

10 Minimum cost flows: polynomial algorithms

11 Minimum cost flows: network simplex

13 Minimum spanning trees

14 Convex cost flows

15 Generalized flows

16 Lagrangian relaxation and network optimization

18 Computational testing of algorithms

A idéia original era, na primeira parte da iniciação, o estudo conjunto dos capítulos $6, 7, 9, 10, 11$ e 18 do livro. Nos meses finais, eu passaria a estudar os capítulos 13 e 14 e o Marcelo os capítulos 15 e 16.

A tabela abaixo apresenta o cronograma original de estudos:

Atividade (capítulo)	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ	JAN	FEV
6	√	√										
7		√	√									
9			√	√								
10				√	√							
11					√	√	√					
13							√	√	√			
14										√	√	√
18	√	√										

Entretanto, conseguimos cumprir apenas o referente aos três primeiros meses da tabela, pois acabamos dedicando algum tempo extra aos tópicos adicionais e às disciplinas do semestre, conforme anteriormente explicado. Cabe mencionar também que o desenvolvimento de uma primeira versão do texto exigiu mais tempo do que o esperado. Dedicamos praticamente um mês inteiro a isso. Percebemos que tarefas como a definição de uma notação a ser seguida, bem como o estilo de texto a ser produzido, são bastante delicadas e também que escrever um texto em grupo dispende mais tempo do que escrever um texto individualmente.

De qualquer forma, fizemos um novo planejamento e continuamos com a mesma expectativa de conteúdo de estudos a ser cumprido. A única ressalva é o fato de que inicialmente prevíamos o término do projeto para fevereiro do próximo ano, quando, na verdade, o último relatório deve ser encaminhado no início do próximo mês de dezembro. O meu planejamento de estudos para o próximo período pode ser visualizado através da seguinte tabela:

Atividade (capítulo)	AGO	SET	OUT	NOV	DEZ	JAN
7	√					
9	√	√				
10		√	√			
11			√	√		
13				√	√	√

O texto envolvendo toda a teoria estudada deverá ser continuamente incrementado e revisado no decorrer do semestre, na medida em que cada tópico for sendo estudado. Além disso, é nossa intenção que as próximas implementações sejam realizadas diretamente em `CWEB` para já serem armazenadas no sítio da biblioteca.

Cabe observar que, no próximo semestre, tanto o Marcelo como eu estaremos cursando a disciplina de pós-graduação MAC5781 - Otimização Combinatória. Uma vez que muitos dos assuntos tratados durante a disciplina serão comuns aos que estaremos estudando em nossa iniciação científica, é provável que algum trabalho seja adiantado, de forma a tornar possível nossa dedicação a tópicos complementares que podem eventualmente ser incluídos no projeto.

Referências

[1] R.K. Ahuja, T.L. Magnanti, and J. Orlin, *Network flows: Theory, algorithms, and applications*, Practice Hall, 1993.

[2] W.J. Cook, W.H. Cunningham, W.R. Pulleyblank, and A. Schrijver, *Combinatorial otimization*, Wiley-Interscience series in discrete mathematics and optimization, John Wiley & Son's, New York, 1998.

[3] T.H. Cormen, C.E. Leiserson, R.L. Rivest, and C. Stein, *Introduction to algorithms*, 2nd. ed., The MIT Press and McGraw-Hill, 2001.

[4] E.A. Dinic, *Algorithm for solution of a problem of maximum flow in a network with power estimation*, Sov. Math. Dokl. **11** (1970), no. 5, 1277–1280.

[5] J. Edmonds and R.M. Karp, *Theoretical improvements in algorithmic efficiency for network flow problems*, J. ACM **19** (1972), no. 2, 248–264.

[6] P. Feofiloff, *Algoritmos de programação linear*, EDUSP, 2000.

[7] ———, *Notas de aula de MAC 5781 otimização combinatória*, `"http://www.ime.usp.br/~pf/"`, 2002.

[8] L.R. Ford and D.R. Fulkerson, *Maximal flow through a network*, CJM **8** (1956), 399–404.

[9] A.V. Goldberg, *Network otimization library*, http://www.avglab.com/andrew/soft.html.

[10] D.E. Knuth, *Literate programming*, Center for the study of Language and Information (CSLI), 1992.

[11] ———, *The Stanford GraphBase: A plataform for combinatorial computing*, ACM Press, 1993.

[12] D.E. Knuth and S. Levy, *The CWEB System of Structured Documentation*, Addison-Wesley, 1994.

[13] A. Schrijver, *Theory of integer and linear programming*, John Wiley & Son's, New York, 1986.

[14] _____, *Combinatorial optimization: Polyhedra and efficiency*, Algorithms and Combinatorics, vol. 24, Springer, 2003.