

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

SOBRE ALGORITMOS DE OTIMIZAÇÃO ESTOCÁSTICOS:
APLICAÇÕES EM REDESENHO DE REDES DE MONITORAMENTO
E MAPEAMENTO DE QTL

Ramiro Ruiz Cárdenas

2007

SOBRE ALGORITMOS DE OTIMIZAÇÃO ESTOCÁSTICOS:
APLICAÇÕES EM REDESENHO DE REDES DE MONITORAMENTO
E MAPEAMENTO DE QTL

Ramiro Ruiz Cárdenas

Tese de Doutorado submetida ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de Doutor em Ciências Estatísticas.

Orientadores: Alexandra M. Schmidt

Marco A. R. Ferreira

Rio de Janeiro, Julho de 2007.

SOBRE ALGORITMOS DE OTIMIZAÇÃO ESTOCÁSTICOS: APLICAÇÕES
EM REDESENHO DE REDES DE MONITORAMENTO E MAPEAMENTO DE
QTL

Ramiro Ruiz Cárdenas

Orientadores: Alexandra M. Schmidt

Marco A. R. Ferreira

Tese de Doutorado submetida ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de Doutor em Ciências Estatísticas.

Presidente, Prof^a. Alexandra M. Schmidt

IM-UFRJ

Prof. Marco A. R. Ferreira

University of Missouri

Prof^a. Marina S. Paez

IM-UFRJ

Prof. Helio S. Migon

IM-UFRJ

Prof. Ronaldo Dias

IMECC-UNICAMP

Prof. Ricardo S. Ehlers

DEST-UFPR

Rio de Janeiro, Julho de 2007.

Cárdenas, Ramiro Ruiz

Sobre Algoritmos de Otimização Estocásticos: Aplicações em Redesenho de Redes de Monitoramento e Mapeamento de QTL/ Ramiro Ruiz
Cárdenas – Rio de Janeiro:UFRJ/IM, 2007.

vi, 133f.: il, color.; 31cm.

Orientadores: Alexandra Mello Schmidt e Marco Antonio Rosa Ferreira
Tese (Doutorado em Estatística) – UFRJ/IM/Programa de Pós-graduação em Estatística, 2007.

Referências Bibliográficas: f. 99 – 116.

1. Teoria da decisão. 2. Monte Carlo populacional. 3. Amostragem espacial. I. Schmidt, Alexandra M. e Ferreira, Marco A. R. (Orient.).
II. Universidade Federal do Rio de Janeiro, Instituto de Matemática. III.
Título.

RESUMO

SOBRE ALGORITMOS DE OTIMIZAÇÃO ESTOCÁSTICOS: APLICAÇÕES EM REDESENHO DE REDES DE MONITORAMENTO E MAPEAMENTO DE QTL

Ramiro Ruiz Cárdenas

Orientadores: Alexandra M. Schmidt

Marco A. R. Ferreira

Resumo da Tese de Doutorado submetida ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de Doutor em Ciências Estatísticas.

Neste trabalho são propostos alguns algoritmos para a otimização de funções objetivo em espaços de alta dimensão, usando uma abordagem baseada em algoritmos de busca estocásticos e métodos de Monte Carlo populacionais, com ênfase em problemas de redesenho de grandes redes de monitoramento ambiental e mapeamento de QTL.

Inicialmente foi abordado o problema de redução de grandes redes de monitoramento já existentes, quando o critério de otimalidade é baseado nos conceitos de entropia e informação. Um novo algoritmo genético híbrido é proposto para esse problema de otimização, o qual combina o algoritmo genético padrão com um operador de busca local. O novo algoritmo foi comparado com outros algoritmos de busca estocásticos (o *simulated annealing* e o algoritmo genético padrão), mediante simulações, assim como numa aplicação real. Um melhor desempenho do novo algoritmo, em termos de qualidade das soluções, foi verificado para todos os casos considerados.

Posteriormente, baseado num contexto de teoria da decisão, o problema de redesenho de redes foi abordado desde uma perspectiva de simulação. Um novo algoritmo foi proposto para a maximização

de utilidades esperadas que permite fazer uma melhor exploração do espaço dos desenhos. O novo algoritmo combina elementos de um conhecido algoritmo baseado em simulação (Müller et al., 2004), com um algoritmo de Monte Carlo populacional que usa operadores genéticos para atualizar uma população de N cadeias de Markov rodando em paralelo. O algoritmo foi implementado numa aplicação real envolvendo a redução de uma rede de monitoramento de ozônio em larga escala, onde a função de utilidade a ser maximizada engloba vários objetivos conflitantes. O algoritmo proposto superou o desempenho de uma abordagem prévia, baseada numa única cadeia de Markov, e que usa *simulated annealing* como ferramenta de maximização da utilidade esperada.

Finalmente, num outro contexto, o enfoque de teoria da decisão é usado para abordar o problema de mapeamento de QTL (*quantitative trait loci*), isto é, a identificação de regiões no genoma de um organismo que estão diretamente associadas a alguma característica quantitativa de interesse. Geralmente tanto o número, como a localização, desses QTL é desconhecida, levando a um problema de dimensão variável. Uma metodologia é proposta para a estimação do número e localização desses QTL, via a minimização de uma função perda, onde o processo de minimização é feito usando um algoritmo evolucionário desenhado para tal fim.

Palavras-chave : teoria da decisão; amostragem espacial; *simulated annealing*; algoritmos genéticos; Monte Carlo populacional; *quantitative trait loci*.

ABSTRACT

ON STOCHASTIC OPTIMIZATION ALGORITHMS: APPLICATIONS IN MONITORING NETWORK REDESIGN AND QTL MAPPING

Ramiro Ruiz Cárdenas

Advisors: Alexandra M. Schmidt

Marco A. R. Ferreira

Abstract da Tese de Doutorado submetida ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de Doutor em Ciências Estatísticas.

In this thesis some algorithms are proposed for the optimization of objective functions in high dimensional spaces, using an approach based on stochastic search algorithms and population Monte Carlo methods. The focus is on the redesign of large environmental monitoring networks and QTL mapping problems.

Initially, the shrinkage of a large monitoring network already existing was approached, when the optimality criterion is based on entropy and information concepts. A new hybrid genetic algorithm is proposed for this optimization problem, which combines the standard genetic algorithm with a local search operator. The new algorithm was compared with other stochastic search algorithms (simulated annealing and the standard genetic algorithm), through simulations, as well as with a real application. A better performance was verified for the new algorithm in terms of quality of the solutions obtained, for all the cases considered.

Based on a decision theory framework, the redesign problem was afterward approached from a simulation perspective. A new algorithm for expected utility maximization was proposed, which is

able to do a better exploration of the search space. The new algorithm combines elements from a well known simulation based algorithm (Müller et al., 2004), with a population Monte Carlo algorithm, where genetic operators are used to update a population of N Markov chains running in parallel. The algorithm was implemented in a real application involving the reduction of a large-scale ground-level ozone monitoring network, where the utility function to be maximized encodes several competing aims. The proposed algorithm overcome a previous approach based on the simulation of one Markov chain and using the simulated annealing algorithm as the expected utility maximization tool.

Finally, in another context, the decision theory framework was used to approach the QTL (*quantitative trait loci*) mapping problem, i.e., the identification of regions in the genome of an organism which are directly associated with some quantitative trait of interest. In general, both, the number and the locations of these QTL are unknown, leading to a variable dimension problem. One methodology is proposed for the estimation of the locations and the number of these QTL, via the minimization of a loss function, where the minimization process is governed by a genetic algorithm.

Key-words : decision theory; spatial sampling; simulated annealing; genetic algorithms; population Monte Carlo; quantitative trait loci.

AGRADECIMENTOS

O autor expressa seus mais sinceros agradecimentos às seguintes pessoas e entidades por sua valiosa contribuição para a realização deste trabalho:

- Marco A. R. Ferreira
- Alexandra M. Schmidt
- Helio S. Migon
- Dani Gamerman
- Marina Silva Paez
- Roseli Aparecida Leandro
- Edna Afonso Reis
- Alunos do PPG em Estatística da UFRJ, em especial a Romy, Ralph, Esther e Juan.
- CAPES
- FAPERJ

Sumário

Lista de Tabelas	iv
Lista de Figuras	v
Capítulo 1: Introdução	1
Capítulo 2: Revisão de Literatura	4
2.1 O Algoritmo <i>simulated annealing</i>	5
2.1.1 Descrição do algoritmo	6
2.1.2 Esquemas de temperatura	7
2.1.3 A estrutura de vizinhança	10
2.1.4 Extensões	11
2.2 Algoritmos Genéticos	12
2.2.1 Seleção	14
2.2.2 Cruzamento	15
2.2.3 Mutação	16
2.2.4 Controle de parâmetros no algoritmo genético	17
2.2.5 Extensões	18
2.3 Algoritmos de Monte Carlo Populacionais	23
2.3.1 Cadeias Múltiplas Independentes	26

Gracias por visitar este Libro Electrónico

Puedes leer la versión completa de este libro electrónico en diferentes formatos:

- HTML(Gratis / Disponible a todos los usuarios)
- PDF / TXT(Disponible a miembros V.I.P. Los miembros con una membresía básica pueden acceder hasta 5 libros electrónicos en formato PDF/TXT durante el mes.)
- Epub y Mobipocket (Exclusivos para miembros V.I.P.)

Para descargar este libro completo, tan solo seleccione el formato deseado, abajo:

