



guidebook

16° WORKSHOP DE

WORCAP



COMPUTAÇÃO

2016
PLICADA

25 e 26 de outubro de 2016

Auditório Fernando de Mendonça - LIT
Instituto Nacional de Pesquisas Espaciais

www.lac.inpe.br
/worcap2016

 Fototerra



/worcap2016



Livro de Resumos
16 Workshop de Computação Aplicada
Pós-Graduação em Computação Aplicada (CAP)

25 e 26 de outubro de 2016

Auditório Fernando de Mendonça, Laboratório de Integração e Testes (LIT)
Instituto Nacional de Pesquisas Espaciais (INPE)
São José dos Campos – SP
<http://www.lac.inpe.br/worcap2016/>

turbulence, ionospheric plasma turbulence associated to equatorial plasma bubbles can be described by power laws and some information regarding their generation mechanism may be inferred from the slopes of the power spectra. Different data analysis techniques have been used to study ionospheric data and results show deviations from K41 theory for homogeneous turbulence. Mandelbrot surmised that turbulent dissipation can be modeled by a fractal and that part of fractal analysis of turbulence can be seen as the geometric counterpart of the analytic analysis of correlations and spectra. Detrended Fluctuation Analysis (DFA) technique is being recently used on in situ ionospheric measurements and results show a wide range of spectral index variations indicating lack of universality of class, which can be probably attributed to the non-homogeneity in energy cascade. In this work, we propose to use Multi-Fractal DFA (MFDFA) technique on in situ ionospheric data to try to analyze non-homogeneity in the energy cascade. In addition, it may also help in inferring the nature of turbulence ie whether it is hydrodynamic or more complex.

UM MÉTODO DE CLASSIFICAÇÃO HIERÁRQUICA UTILIZANDO UMA FUNÇÃO SVM MULTI-KERNEL

Michelle Parreira, Luciano Dutra, Eliana Pantaleão, Sherfis Ruwer

Esse trabalho propõe um método de classificação baseado em máquinas de vetores suporte (SVM) que usa múltiplos kernels e atribui os rótulos às classes de modo hierárquico. Uma árvore binária é criada automaticamente pelo algoritmo e cada nó realiza a classificação entre duas partições do conjunto de classes pré-classificado pelo nó superior. A classificação é realizada pelo classificador SVM com múltiplos kernels combinados aproveitando as diferentes características de cada kernel. A escolha pelas classes que compõem cada partição em cada nó é feita por otimização junto com os parâmetros dos kernels e os coeficientes da combinação linear entre eles. Para isso é empregado o algoritmo Infestação por Ervas Daninhas (Invasive Weed Optimization, IWO). Foi também realizado um experimento Monte Carlo para definir os parâmetros do algoritmo IWO que afetam tanto a acurácia da classificação como o tempo de processamento. Esse novo método consegue separar hierarquicamente as classes com melhor separabilidade segundo um classificador SVM multi-kernel otimizado para cada classificação binária. Os resultados foram comparados com outros métodos já estabelecidos na literatura, e demonstraram significativa melhora na acurácia da classificação.

EVOLUÇÃO DE SOFTWARE: UMA REVISÃO SISTEMÁTICA DE LITERATURA

Rogério Marínke, Valdivino Santiago Junior, Nandamudi Vijaykumar

É comum o uso de software legado pelas grandes corporações. Embora existam muitos esforços para que eles sejam substituídos totalmente, isso é uma tarefa complicada que envolve custos, profissionais e fatores humanos. Portanto, pesquisas que tentam estimar como um determinado produto de software pode evoluir baseando-se em diversos aspectos técnicos e sociais, são muito importantes pois podem otimizar os esforços financeiros e de pessoal para evoluir um determinado produto. Esse ponto também é válido para desenvolvimento de software espacial. Este trabalho tem como objetivo apresentar uma revisão sistemática de literatura sobre as diversas abordagens que tratam de prever como um determinado produto de software pode evoluir.

UM EXPERIMENTO CONTROLADO PARA TESTE COMBINATORIAL

Juliana Balera, Valdivino Santiago Junior

Exhaustive execution of tests is impractical, this is a fact. Considering complex systems, such as software developed for space applications like satellite, stratospheric balloons and rockets, this fact is even more true where defects in software can cause disasters in which large financial loss, or damage to the environment occur. For this reason it is very important to have available methods/software testing techniques with great potential for the development of defects. One of the activities of the most studied software testing process is the generation/selection of test cases. However, select test cases that show a greater number of defects is complex, given the large number of entries that the software can process, and even because of the different characteristics of software products in various fields. Combinatorial designs come calling attention of the software testing community to generate smaller test cases sets (lower cost to run) and efficient (ability to find defects in software). Such techniques have been effective to detect faults due to the interaction of various input variables. Therefore, this proposed master's dissertation aims to present a new algorithm, called T-Tuple Reallocation (TTR) to generate instances of software testing via combinatorial designs specifically via the Mixed-Level Covering Arrays (MCAs). A literature review is presented with the main algorithms to generate combinatorial designs, pointing out their strengths and weaknesses that led to the creation of the TTR. A preliminary comparison between the TTR and four other algorithms/tools found in the literature is presented, where the TTR achieved a better performance considering the size of the set of test cases generated.

DETECÇÃO DE BORDAS EM IMAGENS POLARIMÉTRICAS USANDO APRENDIZAGEM PROFUNDA E DEEP LEARNING

Gilberto Siva Junior, Marcos Quiles, Fabio Cappabianco

A detecção de bordas é uma das operações mais importantes para a análise visual e automática de imagens. É a abordagem mais usada para detecção de descontinuidades nos valores ou atributos de uma imagem.

Diferentes técnicas de detecção de bordas geram diferentes resultados e não é possível afirmar que uma única borda seja consistentemente melhor em todos os pontos para todos os problemas.

A aprendizagem profunda (Deep Learning) envolve desenvolver sistemas para reconhecer padrões, muitas vezes complexos e abstratos, alimentando grandes quantidades de dados através de redes sucessivas de neurônios artificiais e refinar a forma como essas redes respondem à entrada. Essa abordagem tem se mostrado muito eficaz para o reconhecimento de fala ou outros tipos de áudio, ou ainda para classificar informações visuais.

A nossa proposta objetiva, criar um arcabouço baseado em Aprendizagem Profunda que permita extrair informações das imagens de radar polarimétricas produzindo uma borda mais completa oriunda dessas informações. O desafio consiste em ser capaz de fazer essa extração de informações ponderando de forma quantitativa e qualitativa a contribuição de cada fonte produzindo um sistema amigável e adaptado à prática dos usuários de imagens de sensoriamento remoto.

SENTINEL - UM AMBIENTE VIRTUAL PARA GERENCIAMENTO E MONITORAMENTO DE APLICAÇÕES CIENTÍFICAS

Bruno Leonor, Walter dos Santos, Reinaldo Rosa

A computação vem revolucionando o modo como diversas áreas científicas tem conduzido suas pesquisas. Para realizar simulações/pesquisas os sistemas utilizados lidam com uma enorme quantidade de variáveis que podem exigir uma grande quantidade de cálculos, o que tem contribuído com o aumento na demanda por recursos computacionais cada vez mais poderosos. O crescente número de projetos científicos e um explosivo aumento do volume de dados tem motivado a adoção de abordagens em eScience em diversas partes do mundo como apoio na resolução de problemas práticos e de pesquisa. A proposta deste trabalho é apresentar uma solução de eScience para apoiar cientistas no acompanhamento de suas pesquisas em um ambiente de nuvem, possibilitando acesso as suas aplicações computacionais onde poderão parametrizar, executar, visualizar e compartilhar os resultados obtidos transparentemente através de qualquer dispositivo que ofereça acesso a web. A infraestrutura computacional utilizada para demonstração deste projeto é constituída basicamente de microcomputadores Raspberry PI devido a seu baixo custo e recursos oferecidos, podendo ser associados outros recursos de hardware sempre que necessário. Cada uma das aplicações é executada em um ambiente isolado através de um container (Docker). Para armazenamento dos dados (entrada e saída) das aplicações optamos pelo uso do banco de dados MongoDB, um banco de dados orientado a documentos. Como estudo de caso abordaremos o programa EMBRACE (Estudo e Monitoramento Brasileiro do Clima Espacial), pois o conhecimento e previsão de diferentes fenômenos que afetam diretamente as atividades humanas, como a atividade solar por exemplo, têm grande importância para o bem estar da sociedade. Como um protótipo para a implementação, a técnica DFA (detrended fluctuation analysis) foi aplicada na análise das medições de radiação solar de 1978 a 2012. Discutimos também como novas aplicações podem ser adicionadas ao ambiente considerando o novo paradigma do Data Science/Big Data.

MOBILIDADE URBANA EM UM DIA TÍPICO - SÉRIE TEMPORAL DE PROPRIEDADES ESTATÍSTICAS

Jessica Santos, Leonardo Bacelar, Marcos Quiles, Elbert Macau

Um sistema complexo pode ser modelado através do conceito de rede complexa que, por meio de grafos, retrata um conjunto de conexões com características topológicas não completamente regulares nem mesmo completamente aleatórias. Na sociologia, na biologia, na mobilidade urbana e em muitas outras áreas, essas características apresentam semelhanças permitindo a reutilização das soluções a partir da generalização dos modelos. A modelagem utilizada neste estudo visa utilizar ferramentas tradicionalmente referentes a áreas como Física Estatística e Sistemas Complexos para analisar dados de mobilidade urbana em cenário de crescimento urbano não planejado que pode gerar impactos negativos, como congestionamentos, poluição do ar e falhas de infraestruturas. A cidade de São José dos Campos/SP detém um estudo para melhorar sua infraestrutura a partir de dados reais de mobilidade obtidos em um dia típico, pesquisa efetuada pelo IPPLAN, Instituto de Pesquisa, Administração e Planejamento. Neste estudo a cidade é dividida em 55 zonas de tráfego. Um grafo de origem e destino foi gerado com representação matricial em 3 dimensões, no qual cada vértice representa um centroide da zona de tráfego e as arestas são ponderadas pelo “sampling”, fluxo de pessoas, com variação no tempo. Para a análise estatística os dados foram agrupados numa janela de definida pelo pesquisador. Neste trabalho foi considerado um agrupamento a cada 1h. Os resultados - séries temporais de propriedades estatísticas - mostram que o maior pico de fluxo ocorre entre 12h e 13h, os valores mínimos no período das 2h e 6h, e os horários que representam a maior média de fluxo entre zonas de tráfegos ocorrem, respectivamente, entre 7h e 8h, 13h e 14h e 17h e 18h. As variáveis analisadas são ainda espacializadas, com base em Séries Temporais em Bancos de Dados Geográficos.

PROJETO E DESENVOLVIMENTOS DE UM BANCO DE DADOS DE AMOSTRAS PARA USO EM PROJETOS DE MAPEAMENTO DE USO E COBERTURA DA TERRA