

Utilização de *Grids* para Paralelização do Processamento de Redes Neurais Artificiais

Mauricio Adami Mariani (Voluntário), Scheila de Avila e Silva, André Martinotto, Gunther Johannes L. Gerhardt, Sergio Echeverrigaray (orientador) - wushubr@gmail.com

Os avanços na genômica possibilitaram o aumento da quantidade de informação genética dos seres vivos. Para realizar a análise desse grande volume de dados, a utilização de computação de alto desempenho se torna necessária. Uma alternativa para a sua realização é o uso de *grids* computacionais. Esses se baseiam na utilização de computadores não dedicados para a obtenção de uma plataforma de alto desempenho, ou seja, visa a utilização de ciclos ociosos de computadores para a solução de problemas que necessitam de um alto poder de processamento. O presente trabalho teve como objetivo paralelizar o método de Redes Neurais Artificiais aplicadas na predição de sequências promotoras de organismos procariotos em um ambiente computacional de alto desempenho, utilizando *grids*. Os *grids* foram escolhidos porque dividem o problema a ser resolvido em partes independentes. A paralelização consistiu em particionar a metodologia de *ten-fold cross-validation* e as repetições necessárias para validar a técnica de Redes Neurais Artificiais estatisticamente. Assim, as 1.600 simulações necessárias para a análise dos dados que consumiam 29 horas de processamento em um único computador passaram a serem realizadas com um tempo de processamento de 39 minutos.

Palavras-chave: computação de alto-desempenho, bioinformática, *grids*.

Apoio: UCS.

XVII Encontro de Jovens Pesquisadores – Setembro de 2009
Pró-Reitoria de Pós-Graduação e Pesquisa
Universidade de Caxias do Sul