Poster K-18

Clustering of Medline abstracts using Particle Swarm Optimisation

Authors:
Ayton Meintjes (*Bioinformatics and Computational Biology Unit, University of Pretoria*)
Fourie Joubert (*Bioinformatics and Computational Biology Unit, University of Pretoria*)

Short Abstract: The growth in the number of biomedical publications has heightened the need for automatic identification of articles of interest to researchers. Particle Swarm Optimisation has previously been shown to be effective in clustering text documents. We propose the application of this technique to clustering Medline abstracts.

Long Abstract:

The growth in the number of biomedical journal articles published each year threatens to overwhelm researchers who need to keep abreast of developments in their field. In response, the application of text-mining techniques to this field has become a popular research theme.

Particle Swarm Optimisation is a stochastic optimisation algorithm proposed by Kennedy and Eberhart in 1995. In short, a population of potential solution 'particles' is initialised and 'flown' through problem space until a stop condition is met.

The technique has been shown to be effective in the clustering of text documents, with each particle in the swarm representing a potential clustering solution.

We propose the application of this algorithm to the clustering of Medline abstracts. Additionally, the initial clusters could be seeded from a researcher's personal collection of publications.

This functionality would form part of a web-based literature management module, which in turn is a component of the Functional Genomics Information Management System being developed at the University of Pretoria.