

# Enhanced Hierarchical Clustering for Gene Expression data

Geetha.T

Department of Computer Applications  
National Institute of Technology  
Trichirappalli, Tamil Nadu, India

Michael Arock

Department of Computer Applications  
National Institute of Technology  
Trichirappalli, Tamil Nadu, India

## ABSTRACT

Micro arrays are used to assess the transcriptome of many biological systems that has generated an enormous amount of data. Cluster analysis is a technique used to group and analyze micro array data. Identification of groups of genes that manifest similar expression patterns is a key step in the analysis of gene expression data. Hierarchical clustering is the one of the clustering techniques used for this purpose. In this paper, we design an enhanced hierarchical clustering algorithm which scans the dataset and calculates distance matrix only once unlike other papers, (up to authors' knowledge). Our main contribution is to reduce time, even when a large database is analyzed. Also, the results of hierarchical clustering are represented as a binary tree which gives clarity in grouping and further helps to find clustered objects easily. Our algorithm is able to retrieve number of clusters with the help of cut distance and measures the quality with validation index in order to obtain the best one; does not require initial parameter like number of clusters.

The full text of the article is not available in the cache. Kindly refer the IJCA digital library at [www.ijcaonline.org](http://www.ijcaonline.org) for the complete article. In case, you face problems while downloading the full-text, please send a mail to editor at [editor@ijcaonline.org](mailto:editor@ijcaonline.org)