

UNIVERSITY OF SOUTH FLORIDA

Defense of a Master's Thesis

Scalable Clustering Using The Dempster-Shafer Theory of Evidence

by

Alireza Chakeri

For the MSCS degree in Computer Science & Engineering

Clustering large data sets has become very important as the amount of available unlabeled data increases. Single Pass Fuzzy C-Means (SPFCM) is useful when memory is too limited to load the whole data set. The main idea is to divide dataset into several chunks and to apply fuzzy c-means (FCM) to each chunk. SPFCM uses the weighted cluster centers of the previous chunk in the next data chunks. However, when the number of chunks is increased, the algorithm shows sensitivity to the order in which the data is processed. Hence, we improved SPFCM by recognizing boundary and noisy data in each chunk and using it to influence clustering in the next chunks. The proposed approach transfers the boundary and noisy data as well as the weighted cluster centers to the next chunks. We show that our proposed approach is significantly less sensitive to the order in which the data is loaded in each chunk.

21st October 2016,

9 AM - 10 AM

ENB 313

THE PUBLIC IS INVITED

Examining Committee

Lawrence O. Hall, Ph.D., Major Professor

Dmitry Goldgof, Ph.D.

Rangachar Kasturi, Ph.D.

Robert Bishop, Ph.D.

Dean, College of Engineering

Dwayne Smith, Ph.D.

Dean, Office of Graduate Studies

Disability Accommodations:

If you require a reasonable accommodation to participate, please contact the Office of Diversity & Equal Opportunity at 813-974-4373 at least five (5) working days prior to the event.