

Kiatichai Treerattanapitak 2012: Automatic Exponential Fuzzy Clustering with Outlier Detection. Doctor of Philosophy (Computer Science), Major Field: Computer Science, Department of Computer Science. Thesis Advisor: Associate Professor Chuleerat Jaruskulchai, Ph.D. 111 pages.

The degree of membership is the key element to achieve high quality of Fuzzy Clustering algorithm. Traditional algorithm like Fuzzy C-Means (FCM) does not produce the degree of membership to reflect the actual level of belonging in some situations. In addition, performing clustering on the dataset containing noise and outliers leads to inaccurate clustering result due to cluster centroids are influenced and shift away from their actual positions and sometimes generates coincident clusters. Furthermore, parameters setting is difficult for inexperienced users to operate the clustering algorithm. In this thesis, Exponential Fuzzy Clustering (XFCM) is proposed based on the three types of degree of membership concept to improve its representation. Additionally, the problem of noise and outlier are handled by combining the Possibilistic approach with Exponential Fuzzy Clustering. To solve the problem of setting number of clusters and estimating fuzzifier, the Agglomerative Fuzzy Clustering (AFC) is proposed with a single parameter. Various experiments were setup to validate XFCM, PXFCM and AFC performance. The experiments for XFCM were carried out to measure the prediction of errors by Mean Absolute Error (MAE) on Collaborative Filtering. The results showed that XFCM outperforms FCM by 5.2-9.8%, FCME by 1.0-6.1%, the Item-based method by 2.7-6.9% and SVD by 1.0-3.0% for 100K and 1M MovieLens dataset. PXFCM produced minimum centroid errors comparing to other algorithms and did not generate coincidence clusters. In the outlier detection perspective, the XOF that calculated based on the residual distance yielded the better result than other outlier detection algorithms. AFC also selected the right value of fuzzifier and number of cluster parameters for fuzzy clustering. This method can be used to automate the algorithm and it is easy to operate by novice.

---

Student's signature

---

Thesis Advisor's signature

---

\_\_\_\_/\_\_\_\_/\_\_\_\_