# **Modularity-Based Graph Clustering Algorithm**

## Abstract

Modularity is a measure of the structure of networks or graphs. The network is divided into groups, clusters or communities. Modularity is used to measure the strength of division of a network.

Modularity-based graph clustering algorithms are applied to various applications related to Artificial Intelligence and Web Communities. However, existing algorithms (K-means, Hierarchal clustering algorithms) are not applied to large graphs because they have to scan all vertices/edges iteratively.

This graph clustering algorithm compute clusters with high modularity from extremely large graphs with more than a few billion edges. The main part of the solution is to compute clusters by incrementally pruning unnecessary vertices/edges and optimizing the order of vertex selections.

The algorithm is based on three ideas. First, it incrementally aggregates vertices, which are placed in same cluster, into a single vertex to eliminate unnecessary vertices/edges from the graph. Second, it incrementally prunes vertices whose clusters can be obtained without modularity computing. Last, it optimized the order of vertex selections to reduce the number of modularity computations in the clustering process and dynamically selects the vertex with the smallest degree.

The advantages of the approach are that it can extract clusters with quite-small computational cost for complex networks (Newman 2003). With the high clustering coefficients and power-law degree distributions, this algorithm runs faster on large size complex networks. Because of pruning methods and prevention of imbalanced clustering results, this algorithm can produce results with high modularity.

**Reference:**

Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence

Modularity-Based Graph Clustering Algorithm by

Hiroaki Shiokawa, Yasuhiro Fujiwara and Makoto Onizuka