2.4 Community Detection

Definition:

Community Detection is the process of identifying groups of nodes (communities) in a graph such that:

- Nodes within the same group are densely connected
- Nodes between different groups are sparsely connected

Why Community Detection?

- Identify social groups
- Discover clusters in networks
- Improve recommendations

Common Methods:

- Modularity maximization
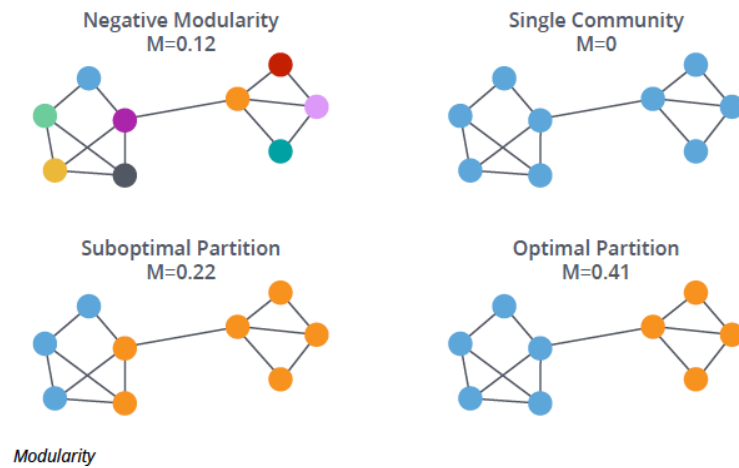- Spectral clustering
- Graph partitioning

Basic Concepts

- **Graph**: Nodes (vertices) + edges (connections)
- **Community / Cluster / Module**: A tightly connected subgraph
- **Intra-community edges**: Connections inside a group
- **Inter-community edges**: Connections between groups

Real-World Examples

- **Social Networks** → friend groups, interest groups
- **Biology** → protein interaction modules
- **Web Networks** → related web pages
- **E-commerce** → customer segments
- **Citation Networks** → research communities

Modularity-Based Methods

Negative Modularity
M=0.12

Single Community
M=0

Suboptimal Partition
M=0.22

Optimal Partition
M=0.41

Modularity

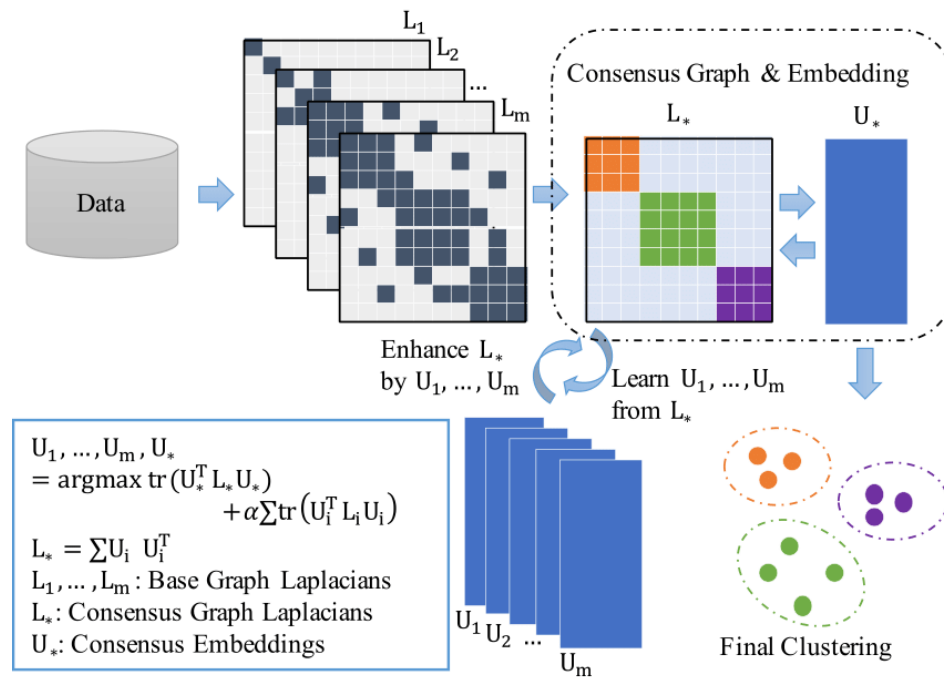## Modularity (Q)

Measures quality of community structure

$$Q = \frac{1}{2m}\sum(A_{ij} - \frac{k_i k_j}{2m})\delta(c_i, c_j)$$

✓ High modularity → strong communities

Popular Algorithms

- Louvain Algorithm
- Girvan–Newman

Spectral Clustering

$U_1, \ldots, U_m, U_*$
$= \text{argmax } \text{tr}(U_*^T L_* U_*)$
$\qquad\qquad + \alpha \sum \text{tr}(U_i^T L_i U_i)$
$L_* = \sum U_i \ U_i^T$
$L_1, \ldots, L_m$: Base Graph Laplacians
L_*: Consensus Graph Laplacians
U_*: Consensus Embeddings

Idea

- Uses eigenvalues of graph Laplacian

- Converts graph into vector space

Used For

- Image segmentation

- Network analysis