Decipher the essentials and sub-networks from complex biological networks for drug target selection

複雜生物網路的解析

Reporter: Chia-Hao Chin

Outline

- Prediction of Essential Proteins
  - Hubba
  - CytoHubba

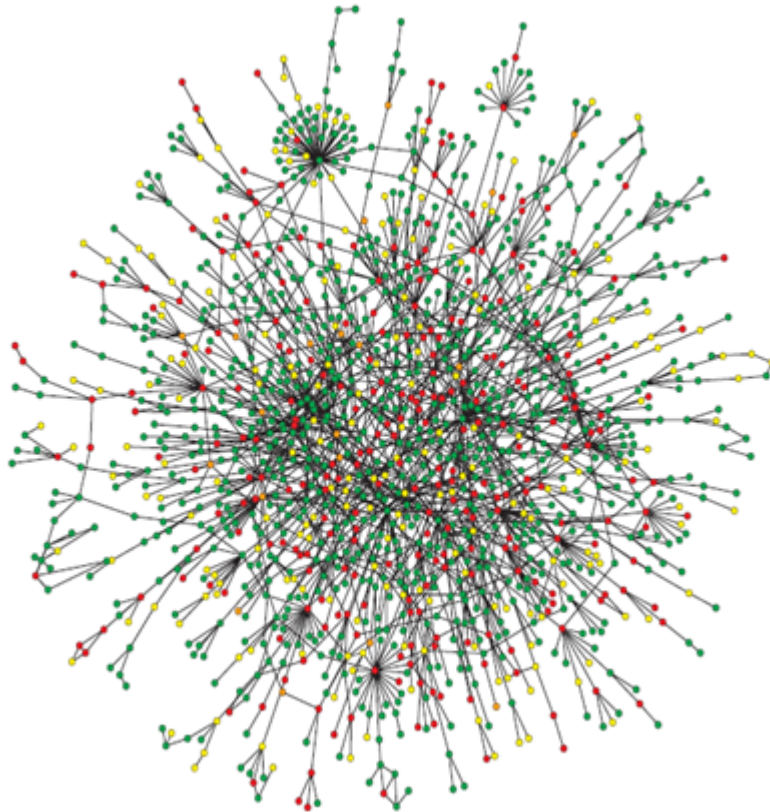- Prediction of Functional Modules
  - SPOTLIGHT

Prediction of Essential Proteins

- A gene (or its associated protein) to be essential if its deletion leads to the loss of cell viability [1].

[1] H. Jeong, Z. N. Oltvai, and A. L. Barabsi, "Prediction of protein essentiality based on genomic data," ComPlexUs, vol. 1, pp. 19-28, 2003.
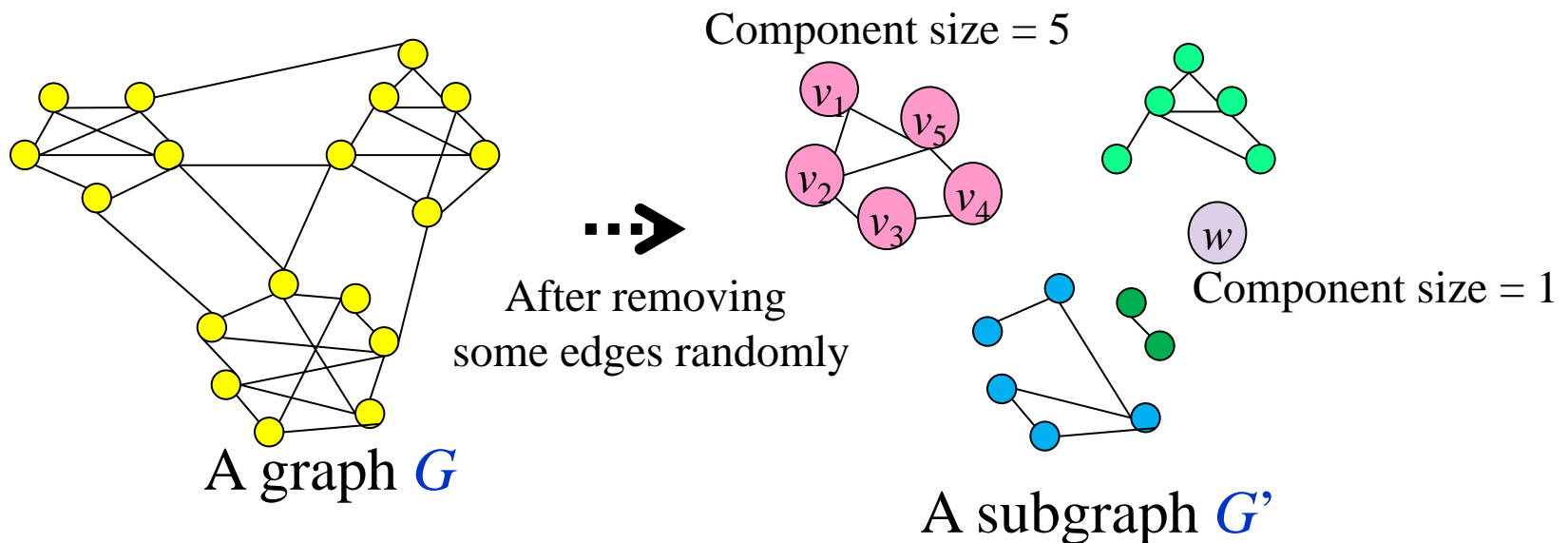
Degree method [2]

A Yeast Protein Interaction Network

Red: **lethal**.

Green: non-lethal.

Orange: slow growth.

Yellow: unknown..

[2] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai, "Lethality and centrality in protein networks," Nature, 2001.

Edge Percolation Component [3]

- Threshold p: a given percolation probability
- Generate a subgraph G'
  - For each edge $e = (i, j)$
    - assign a random number p_{ij}.
   - Remove the edge $e = (i, j)$ if $p_{ij} > p$.

Component size = 5

After removing
some edges randomly

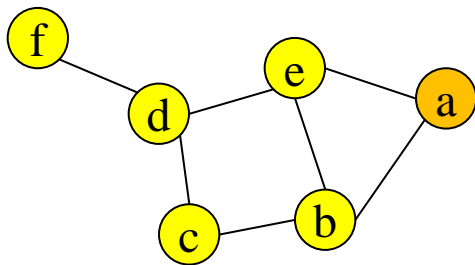A graph G

A subgraph G'

Component size = 1

[3] C. S. Chin and M. P. Samanta, "Global snapshot of a protein interaction network - a percolation based approach," Bioinformatics, 2003.
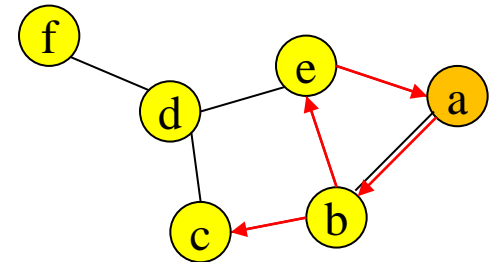
Subgraph Centrality [4]

- For each node v
  - $u_k(v)$:the number of closed walks of v of length k.
  - The subgraph centrality of v

$$SC(v) = \sum_{k=1}^{\infty} \frac{\mu_k(v)}{k!}$$

A example graph
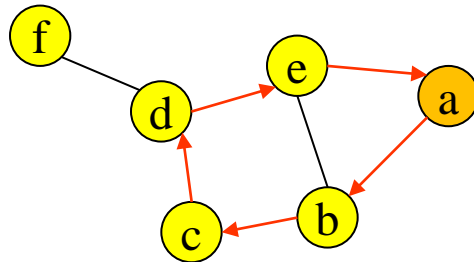
(a, b, c, d, e, a) is a
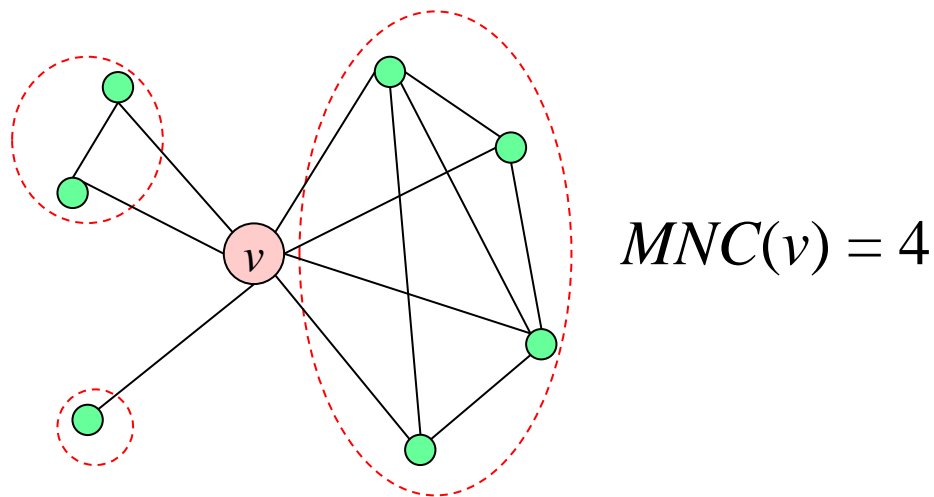closed walk of length 5.

(a, b, e, a, b, c) is an
open walk with length 5

[4] E. Estrada and J. A. Rodriguez-Velazquez, "Subgraph centrality in complex networks," Physical Review E, 2005.

Maximum Neighborhood Component [5]

- The score of node *v*, *MNC(v)*, is defined to be the size of the maximum connected component of the subgraph induced by the neighbors of *v*.

$MNC(v) = 4$

[5] C. Y. Lin, C. H. Chin, H. H. Wu, S. H. Chen, C. W. Ho, and M. T. Ko, "Hubba: hub objects analyzer - a framework of interactome hubs identification for network biology," Nucleic Acids Research, 2008.

Density of Maximum Neighborhood Component [5]

- For a node *v*, let *N* be the number of nodes and *E* be the number of edges of the subgraph induced by the neighbors of *v*. The score of node *v*, *DMNC(v)*, is defined to be E/N^{ε} for some $1 \leq \varepsilon \leq 2$.
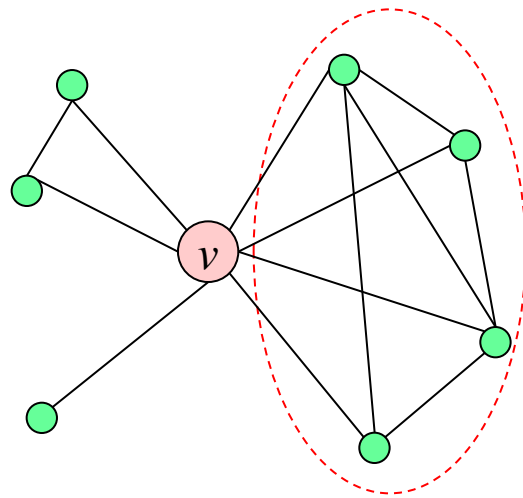
$$DMNC(v) = \frac{5}{4^{\varepsilon}}$$

[5] C. Y. Lin, C. H. Chin, H. H. Wu, S. H. Chen, C. W. Ho, and M. T. Ko, "Hubba: hub objects analyzer - a framework of interactome hubs identification for network biology," Nucleic Acids Research, 2008.

Maximal Clique Centrality

- Given a vertex v

  - $MCC(v) = \displaystyle\sum_{C \in S(v)} |C|!$

  - $S(v) = \{ C \mid C$ is a maximal clique in the subgraph induced by the neighbors of $v \}$

  An example
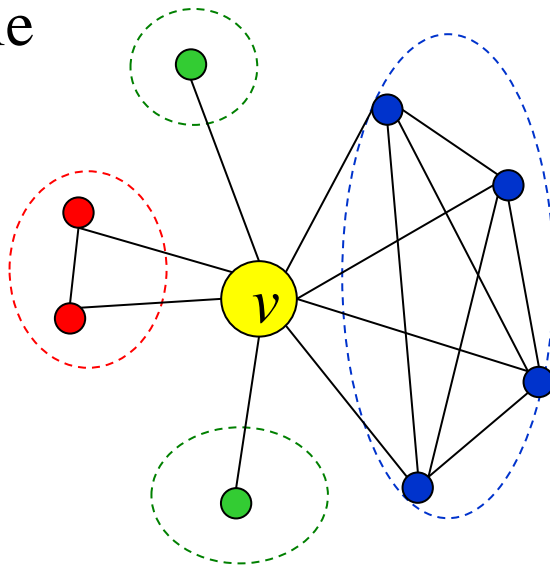
  $MCC(v) = 2*1! + 2! + 4!$
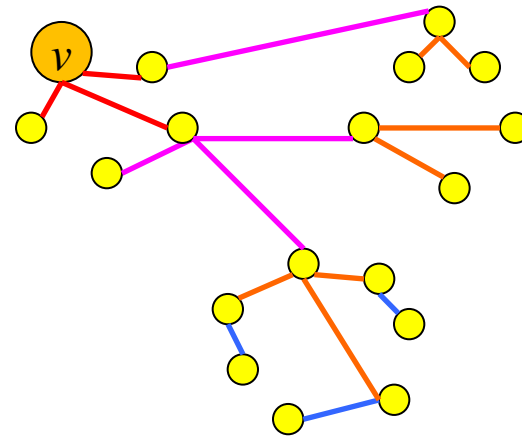 $= 2*(1) + (1*2) + (1*2*3*4)$

BottleNeck method [6]

- For each node v in the undirected PPI graph
  - T_v : a shortest pathtree rooted at v.
    - n_v: the size of T_v.
    - Bottleneck node w: at least $n_v/4$ paths of T_v "meet" at w.
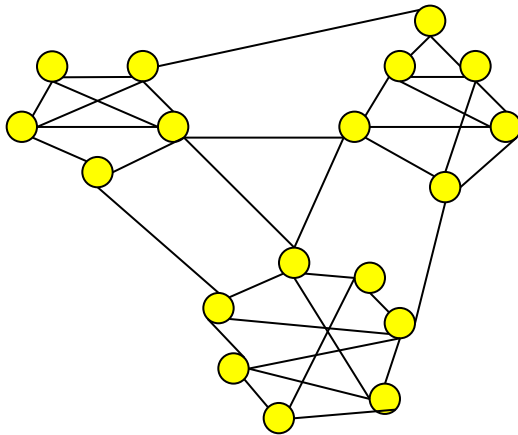
A graph G

T_v: a shortest path tree rooted at v.

[6] N. Przulj, D. A. Wigle, and I. Jurisica, "Functional topology in a network of protein interactions," Bioinformatics, 2004.

Betweenness centrality [7]

- The number of shortest paths in the graph that pass through the node divided by the total number of shortest paths.

$$BC(k) = \sum_i \sum_j \frac{\rho(i,k,j)}{\rho(i,j)}, \ i \neq j \neq k \text{ and } i < j$$
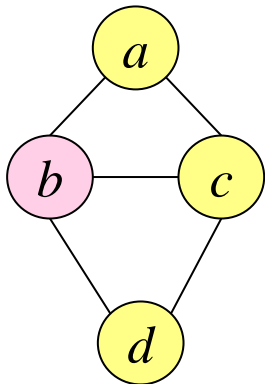
- Shortest paths: *ab, ac, a**b**d , acd, bc, bd, cd.*

$\rho(a, b, d) = 1$
$\rho(a, d) = 2$
$BC(b) = 0.5$

[7] M. P. Joy, A. Brock, D. E. Ingber, and S. Huang, "High-betweenness proteins in the yeast protein interaction network," Journal of Biomedicine and Biotechnology, 2005.

Hubba (http://hub.iis.sinica.edu.tw/Hubba/)

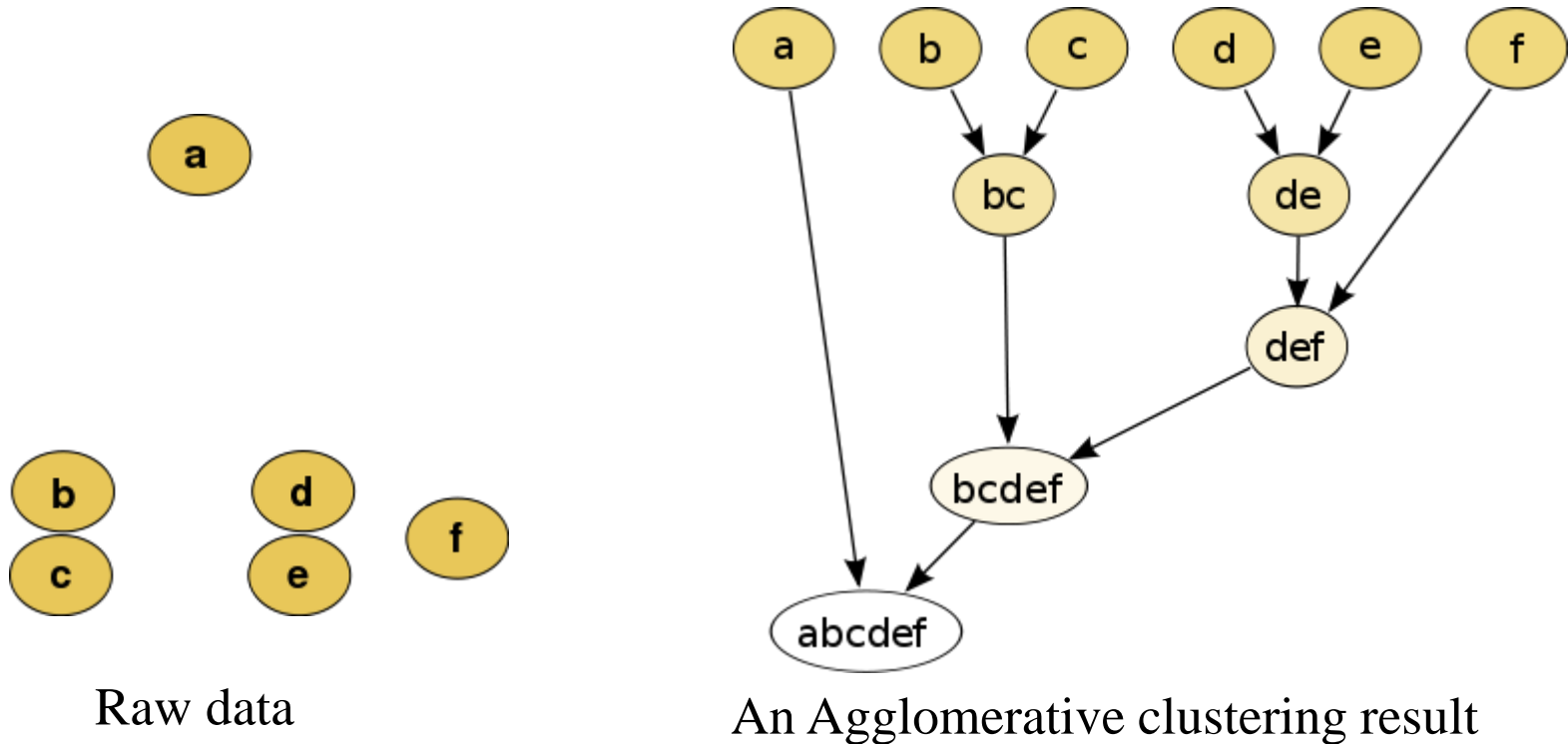# cytoHubba (http://hub.iis.sinica.edu.tw/cytoHubba/)

Prediction of Functional Modules

- A functional module is a discrete entity whose function is separable from those of other modules [8].
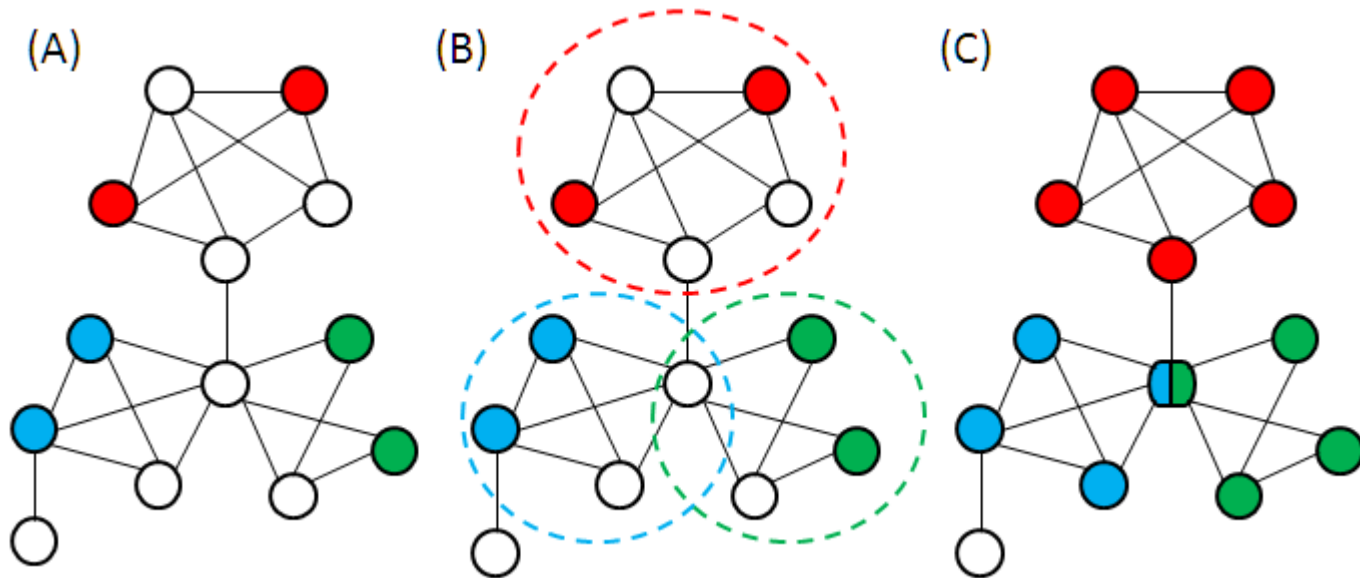
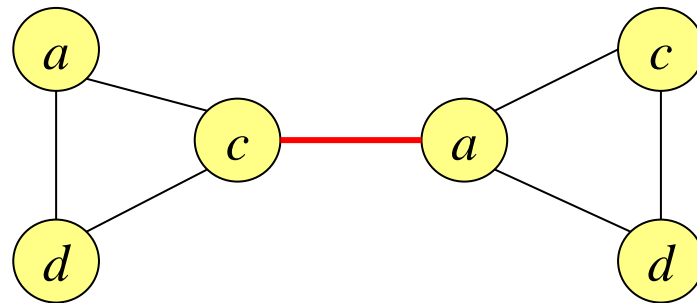[8] L. H. Hartwell, J. J. Hopfield, S. Leibler, and A. W. Murray, "From molecular to modular cell biology," Nature, 1999.

Hierarchical clustering

Raw data

An Agglomerative clustering result
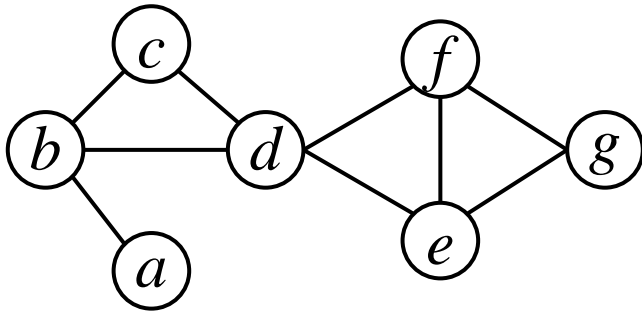
Edge-Betweenness Clustering Method [9]

Remove an edge with the highest betweenness iteratively.

[9] R. Dunn, F. Dudbridge, C. M. Sanderson, "The use of edge-betweenness clustering to investigate biological function in protein interaction networks, "BMC Bioinformatics, 2005.
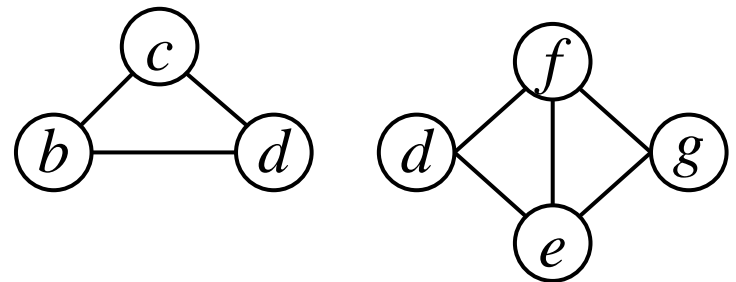
Clique Percolation Method [10]

(a)

(b)

	a, b	b, c, d	d, e, f	e, f, g
a, b	2	1	0	0
b, c, d	1	3	1	0
d, e, f	0	1	3	2
e, f, g	0	0	2	3

$k = 3$

(c)

	a, b	b, c, d	d, e, f	e, f, g
a, b	0	0	0	0
b, c, d	0	1	0	0
d, e, f	0	0	1	1
e, f, g	0	0	1	1

(d)

[10] I. Derényi, G. Palla, and T. Vicsek, "Clique percolation in random networks," Physical Review Letters, 2005.
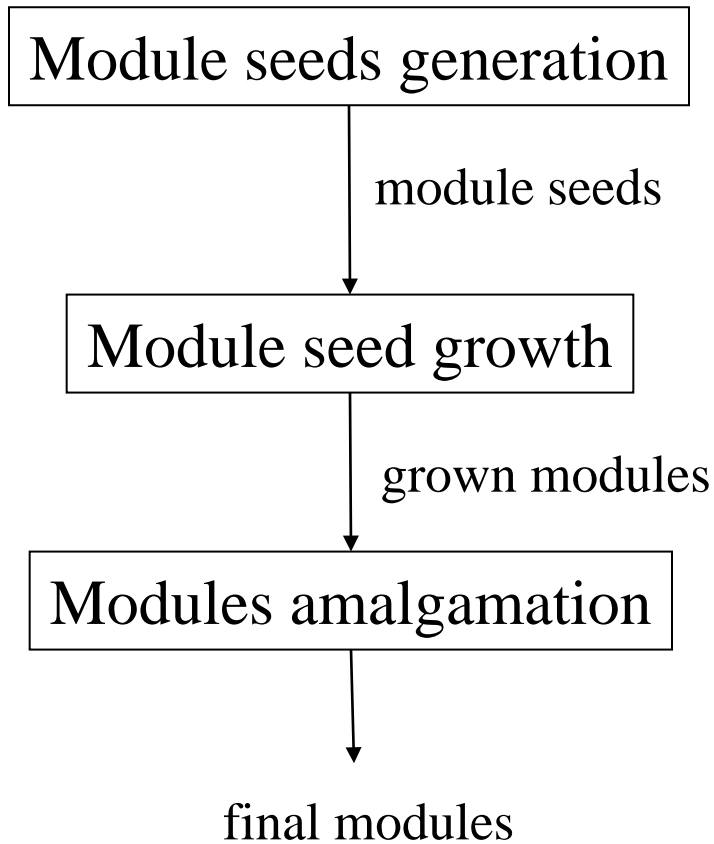
InfoMap [11]

[11] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," PNAS, 2008.

Consensus Graph [12]
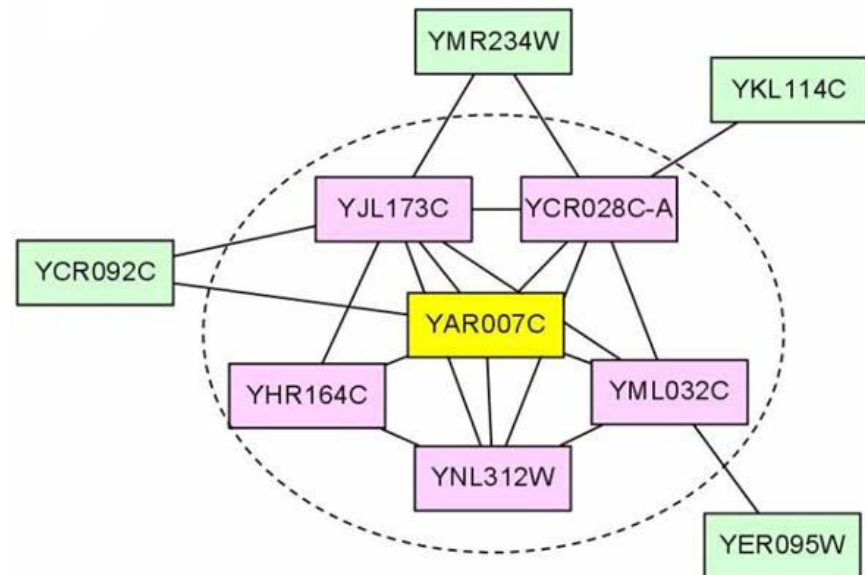

Original Graph

4 partitions

Consensus Graph

$D_{ij} = 1$
$D_{ij} = 3/4$
$D_{ij} = 2/4$
$D_{ij} = 1/4$
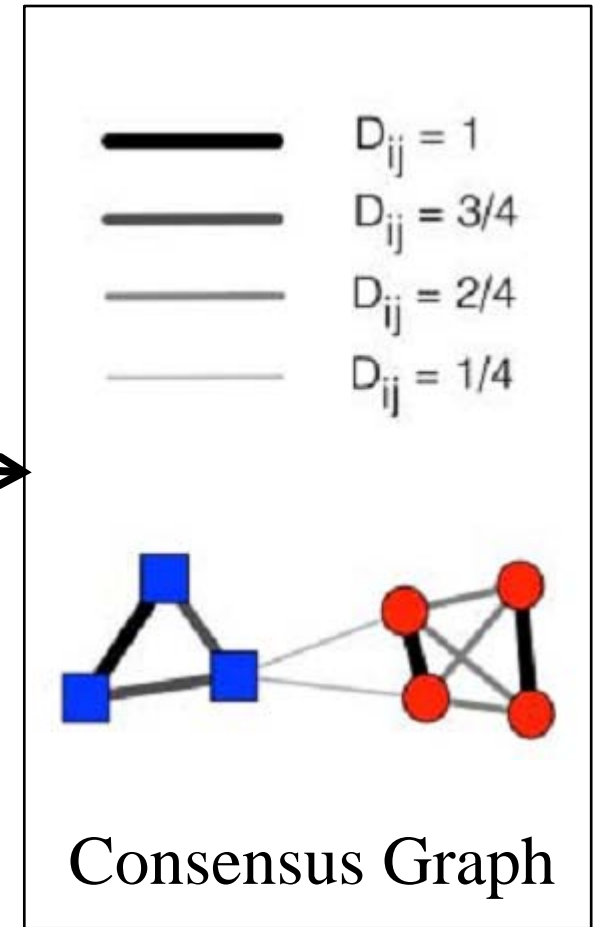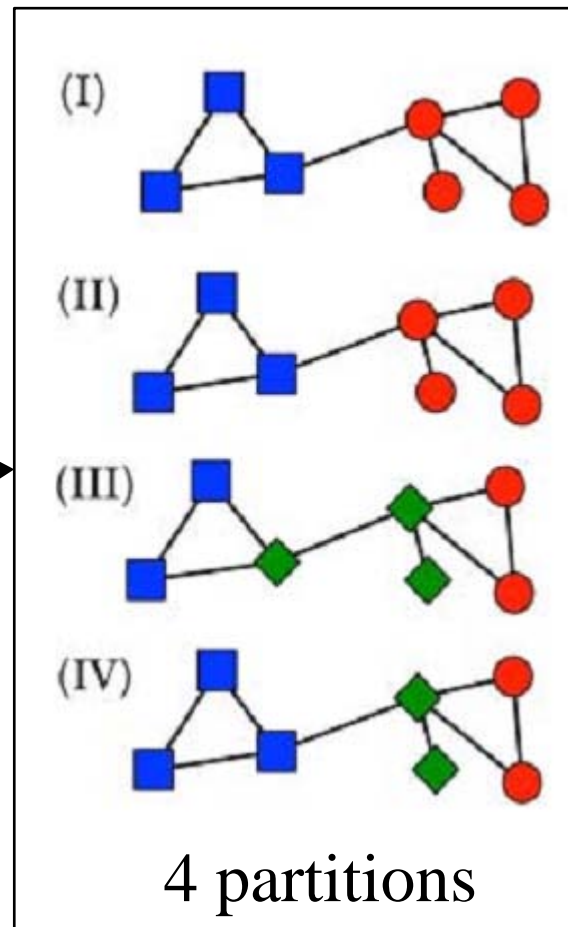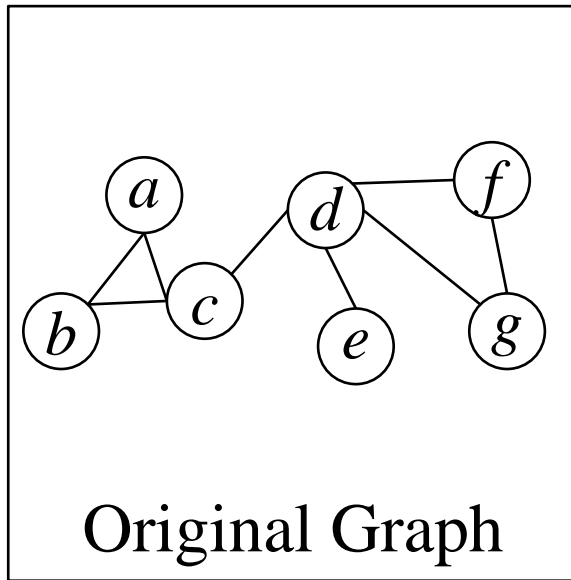
[12] A. Lancichinetti, S.Fortunato, "Consensus clustering in complex networks," Scientific Reports, 2012.

The Overview of HUNTER [13]

Module seeds generation

    module seeds

↓

Module seed growth

    grown modules
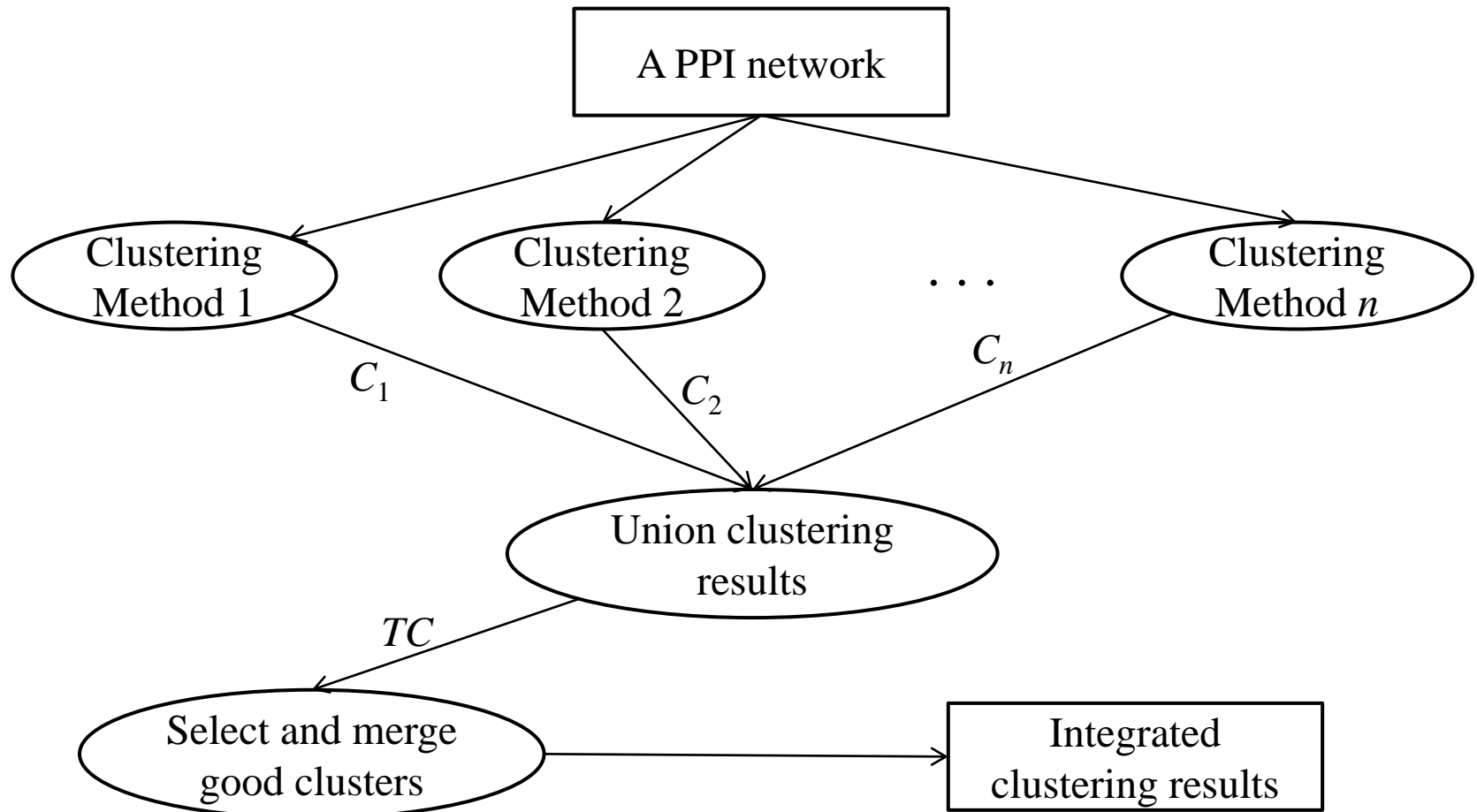
↓

Modules amalgamation

↓

final modules

An Example

[13] C. H. Chin, S. H. Chen, C. W. Ho, M. T. Ko, and C. Y. Lin, "A hub-attachment based method to detect functional modules from confidence-scored protein interactions and expression profiles," BMC Bioinformatics, 2010.

An Integration Method [12]

[12] C. H. Chin, S. H. Chen, C. Y. Chen, C. A Hsiung, C. W. Ho, M. T. Ko and C. Y. Lin, "Spotlight: Assembly of Protein Complexes by Integrating Graph Clustering Methods," to be appeared in Gene.

Spotlight(http://hub.iis.sinica.edu.tw/spotlight/)

Please input your data and other related information.

Job ID	(string with character 0~9, a~z, A~Z)
Input format	● Tab ○ Data with weight value ○ PSI
Data input	☐ **Example File** Shortcut to the Result more examples Clear Input Data
Or load it from file	Choose File No file chosen
Job note	
Email for communication	

Submit   Reset

Thank you for your attention