

Protein Complex Prediction based on Mutually Exclusive Interactions in PPIN

Suk-hoon Jung, Woo-hyuk Jang, Dong-soo Han
Information & Communications University



Contents

- Background
- Method
 - Mutually Exclusive Interaction (MEI)
 - Simultaneous Protein Interaction Cluster (SPIC)
- Evaluation
- Result
- Conclusion



Background

▪ Protein complex

- Molecular component playing crucial roles in cellular process.
- A group of two or more associated proteins formed by protein-protein interaction that is **stable** over time

▪ Protein-Protein Interaction network (PPIN)

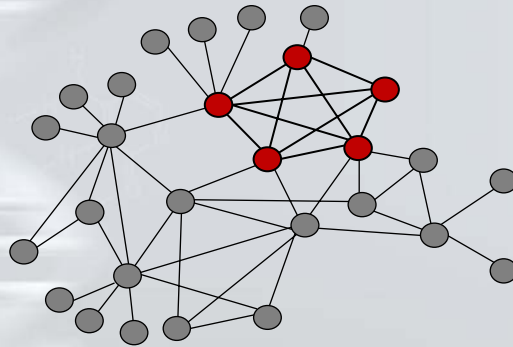
- Static representation of the union of protein-protein interactions
- Visualized with edges (interactions) and nodes (proteins).

▪ Protein complex prediction via PPIN clustering

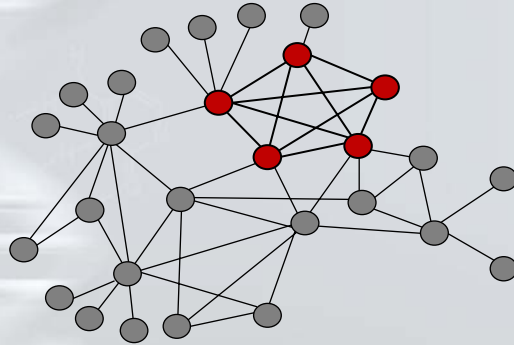
- Protein complexes often correspond to dense regions in PPIN
- MCODE, LCMA, RNSC, STM...



Network clustering on PPIN



Network clustering on PPIN

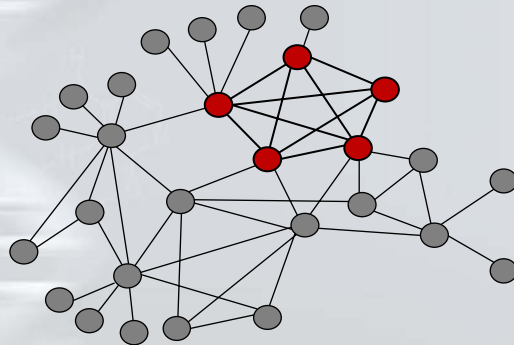


▪ Issues

- Protein complexes are **stable**
- Interactions are mainly dynamic
 - so network clusters may include **dynamic/unstable** interactions



Network clustering on PPIN



▪ Issues

- Protein complexes are **stable**
 - Interactions are mainly dynamic
 - so network clusters may include **dynamic/unstable** interactions
- ← **Mismatched**
- ↓



Network clustering on PPIN

- Clustering must exclude unstable/dynamic interactions to correspond with protein complex
- Elucidating dynamic nature of interactions is difficult, but exclusion of unstable interactions may partly be possible
- Idea: eliminating Mutually Exclusive Interactions (MEI) as a refinement process excluding unstable interactions before clustering



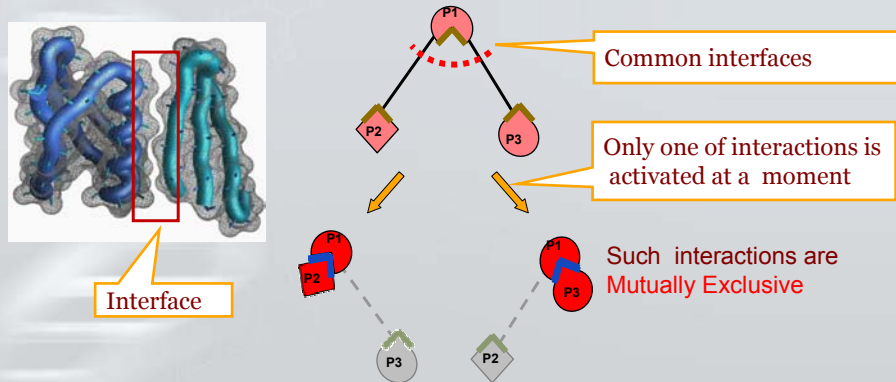
Mutually Exclusive Interactions (MEI)

- Clues
 - Most proteins can interact with several alternative interaction partners mediated by the same interface, and such interactions are likely to be mutually exclusive (*Hu et al, 2005*)
 - Several proteins are reported to have MEI partners
- Only one of MEI can be activated at a time
 - Activation of an interaction guarantees deactivation of other interactions
 - Interactions are in competition for protein complex formation.
- By eliminating MEI, it is able to extract stable clusters which are more likely to be protein complexes



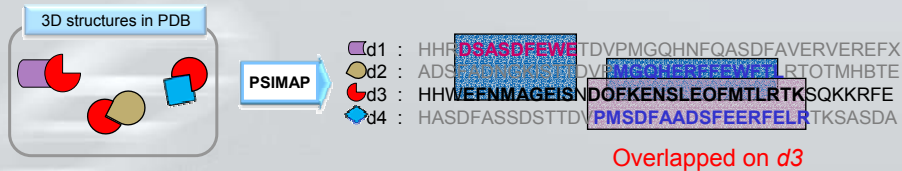
Mutually Exclusive Interactions

- Computational MEI identification using interface data of proteins
 - Interface: contact surface of protein interaction pair

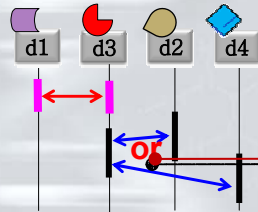


MEI identification

- Interfaces are extracted at the domain level
 - Protein domain is evolutionary conserved unit of functional and structure
 - Domain-domain interactions are regarded as the mediator of PPI
 - Many proteins share a domain, so a domain-domain interface may generate a number of protein-protein interfaces
- By analyzing 3D structures recorded in Protein Data Bank (PDB)
 - PSIMAP (*park et al ,2001*) detects interfacial residues between domains

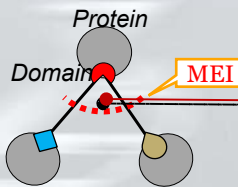


MEI identification



+d2 and d4 are in competition for the interaction with d3

d1 ↔ d3
d2 ↔ d3
d3 ↔ { d1, (d2 or d4) }
d4 ↔ d3



+Many PPIs inherit interactions from a domain-domain interaction

+so such a mutually exclusive domain-domain Interactions may generate many PPIs having MEI feature



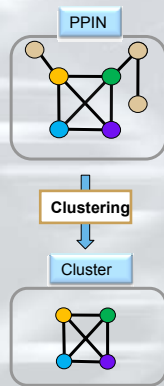
Simultaneous Protein Interaction Cluster (SPIC)

- *The cluster that excludes Mutually Exclusive Interactions*



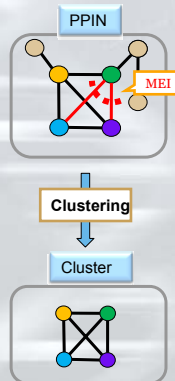
Simultaneous Protein Interaction Cluster (SPIC)

- *The cluster that excludes Mutually Exclusive Interactions*



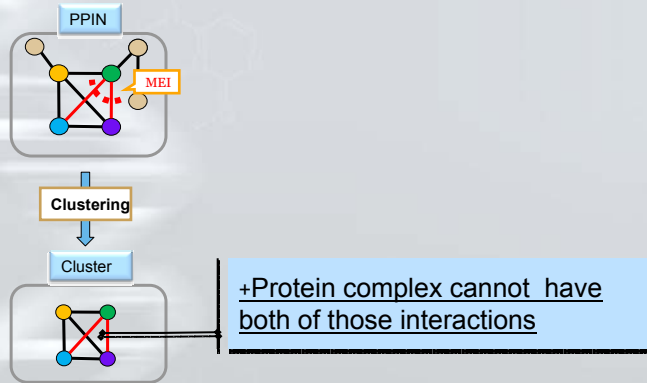
Simultaneous Protein Interaction Cluster (SPIC)

- *The cluster that excludes Mutually Exclusive Interactions*



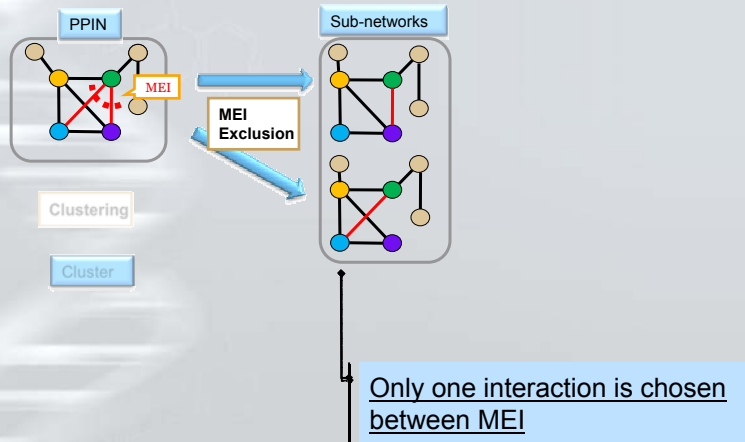
Simultaneous Protein Interaction Cluster (SPIC)

- *The cluster that excludes Mutually Exclusive Interactions*



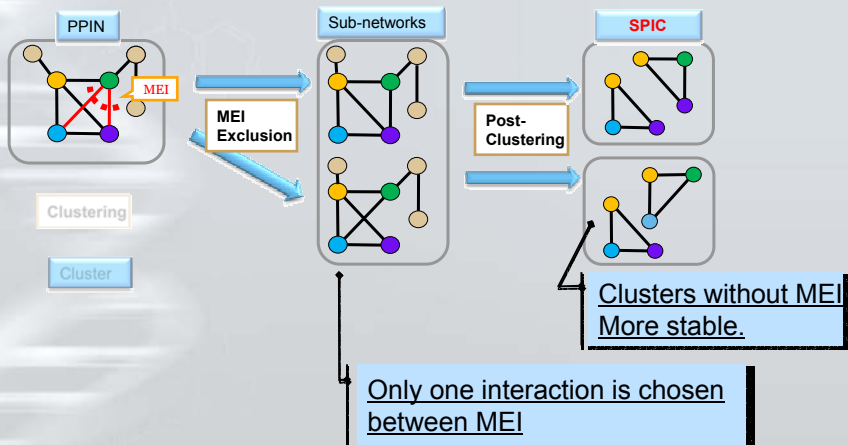
Simultaneous Protein Interaction Cluster (SPIC)

- *The cluster that excludes Mutually Exclusive Interactions*

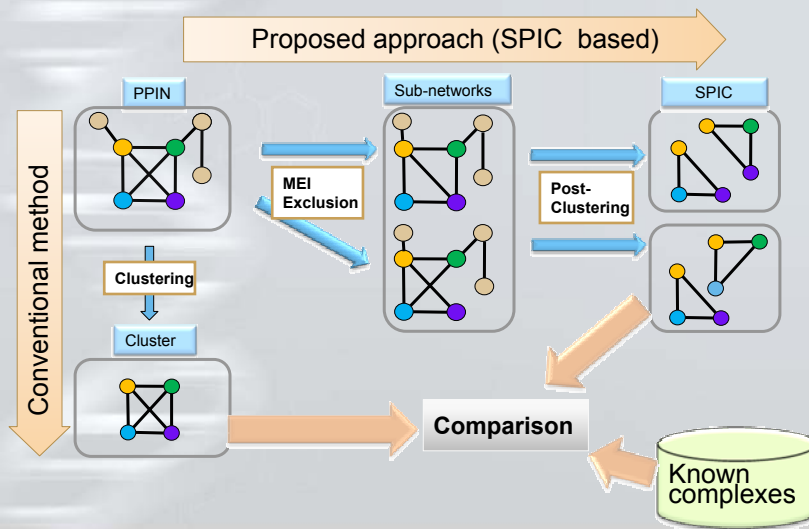


Simultaneous Protein Interaction Cluster (SPIC)

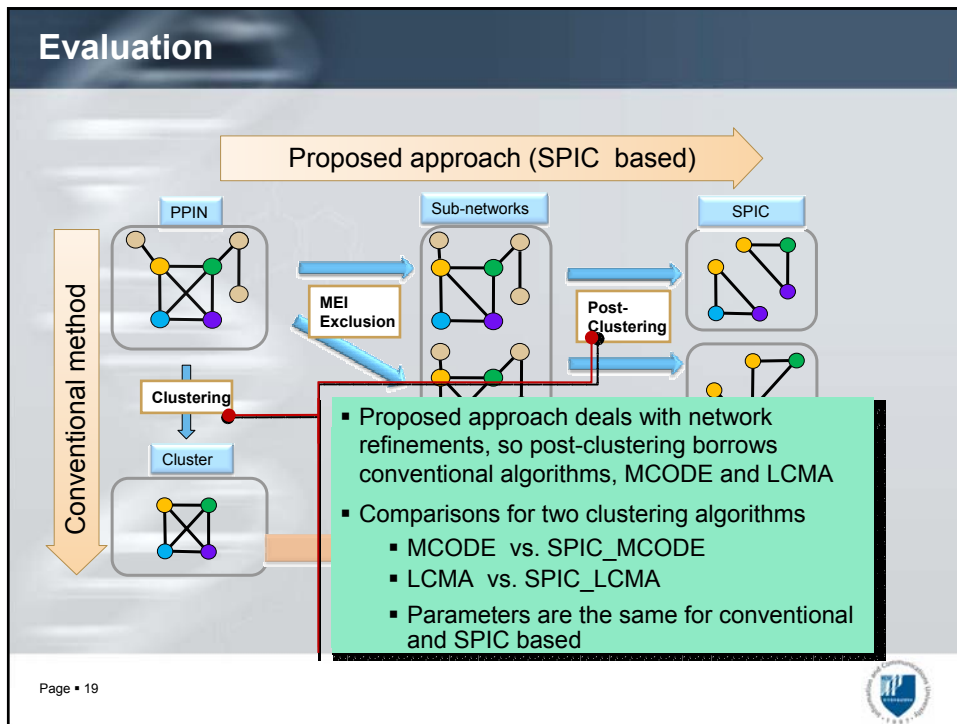
- *The cluster that excludes Mutually Exclusive Interactions*



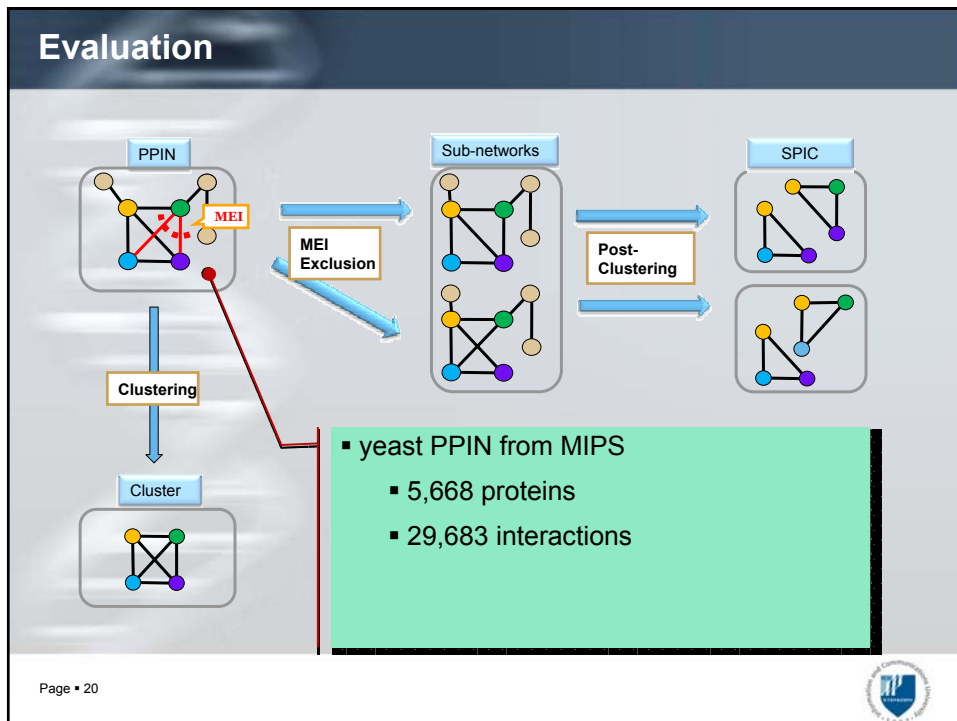
Evaluation



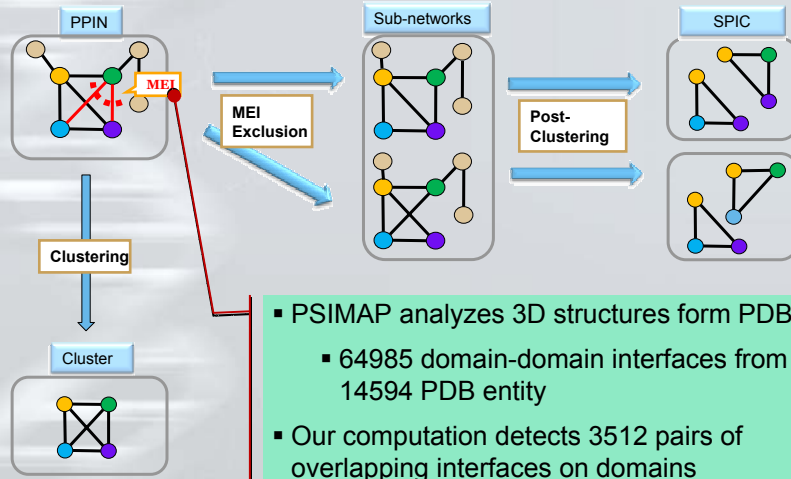
Evaluation



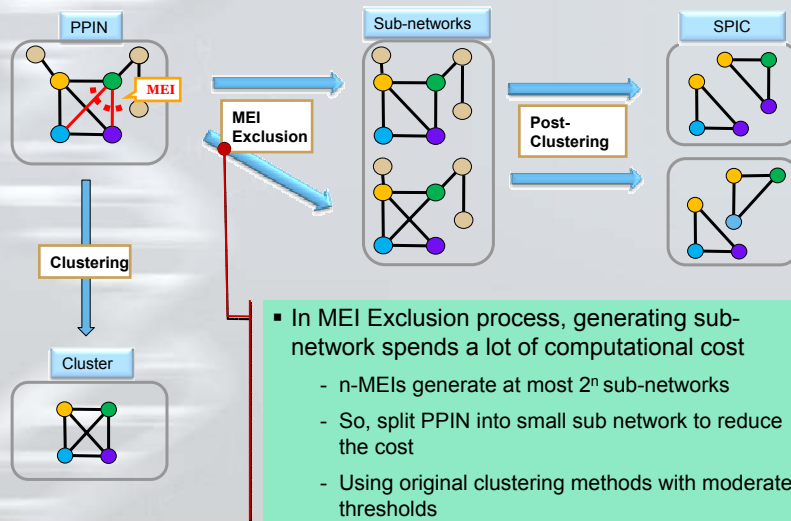
Evaluation



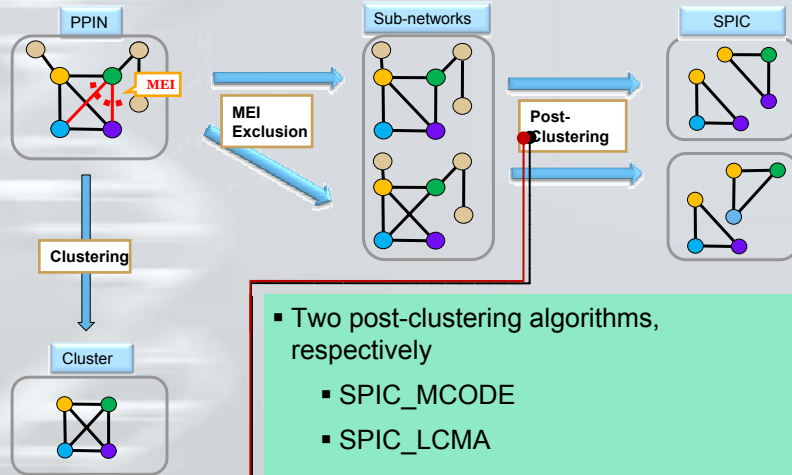
Evaluation



Evaluation



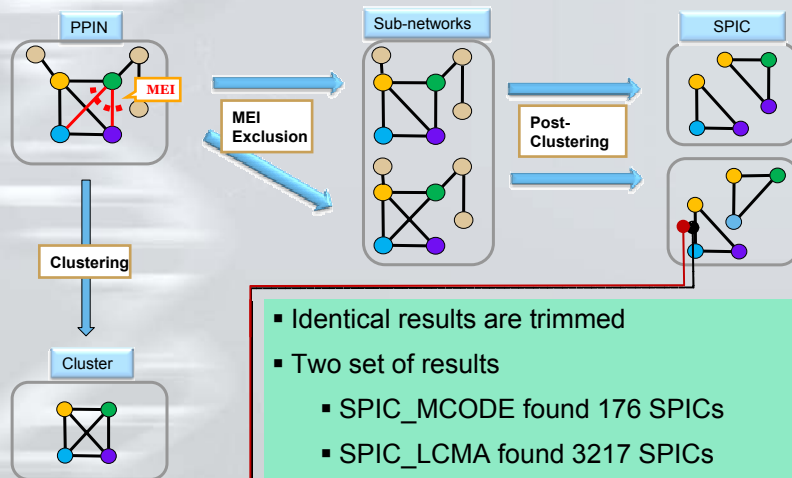
Evaluation



- Two post-clustering algorithms, respectively
 - SPIC_MCODE
 - SPIC_LCMA



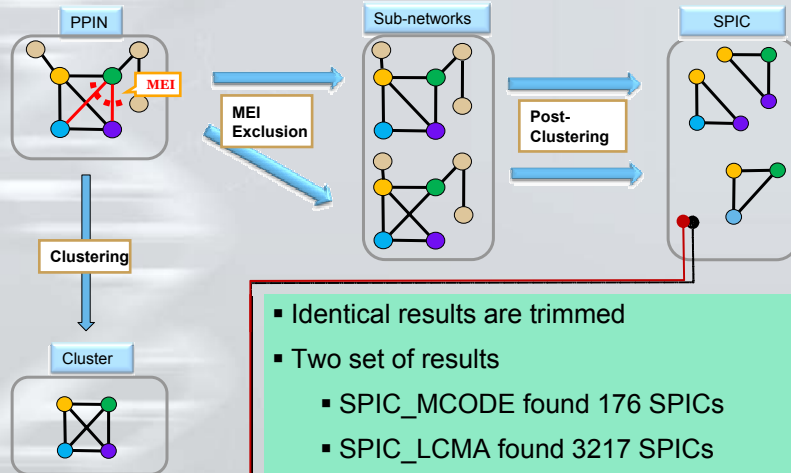
Evaluation



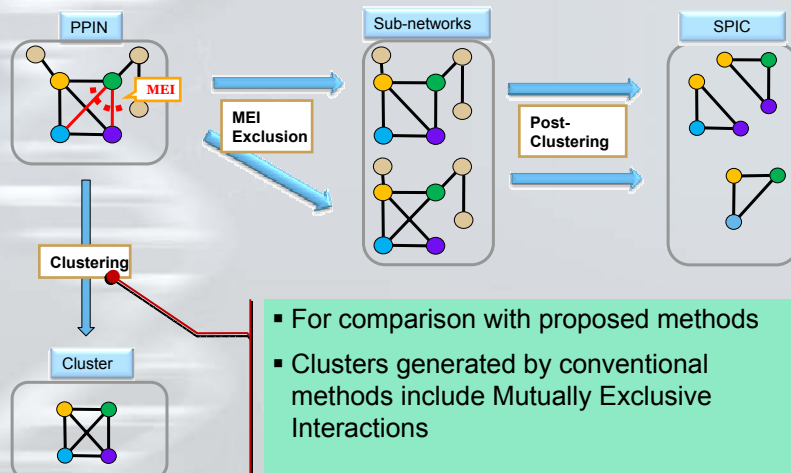
- Identical results are trimmed
- Two set of results
 - SPIC_MCODE found 176 SPICs
 - SPIC_LCMA found 3217 SPICs



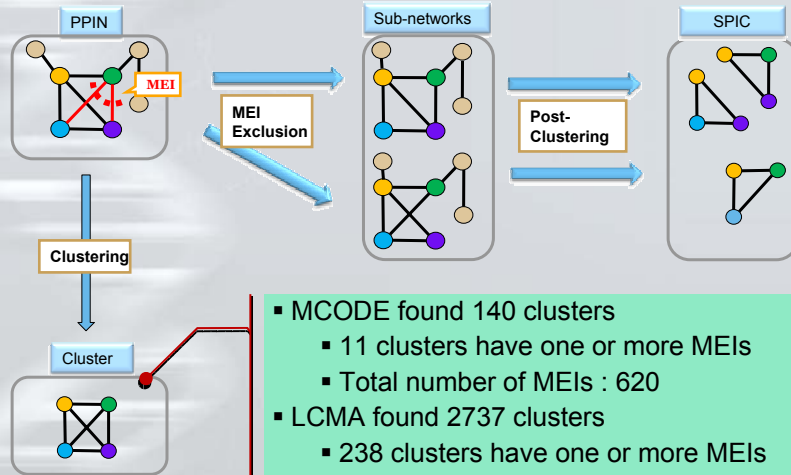
Evaluation



Evaluation



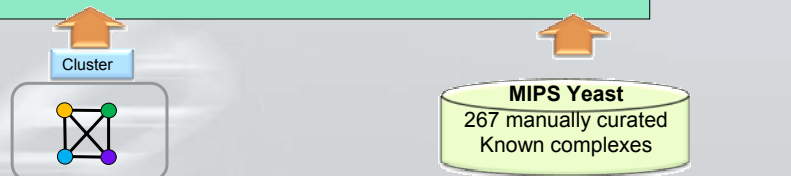
Evaluation



Evaluation

- Results are compared with 267 known yeast complexes from MIPS
 - MCODE vs. SPIC_MCODE
 - LCMA vs. SPIC_LCMA
- *Overlapping Score (OS)* determines how effectively a predicted cluster (P_c) matches a known complex (K_c)

$$OS(P_c, K_c) = \frac{|P_c \cap K_c|^2}{|P_c| \cdot |K_c|}$$



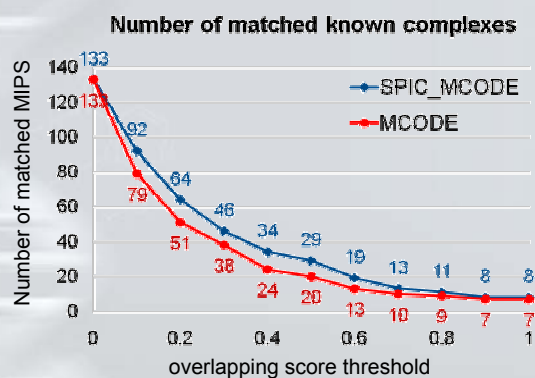
Result: Conventional vs. SPIC_based methods

	MCODE		LCMA	
	MCODE	SPIC_MCODE	LCMA	SPIC_LCMA
# clusters found	140	176	2737	3217
# identical clusters	139		2499	
# unique clusters	11	37	238	718
# clusters with MEI	11	0	238	0

- Since SPIC based approaches borrow conventional algorithms with the same parameters for post-clustering
 - Each SPIC cluster is a sub-cluster of one of the conventional results.
 - Many SPIC clusters are identical with conventional results.
 - Only conventional clusters having MEI are split into SPIC clusters, so they are refined
 - Splitting leads to improvement in SPIC_based prediction



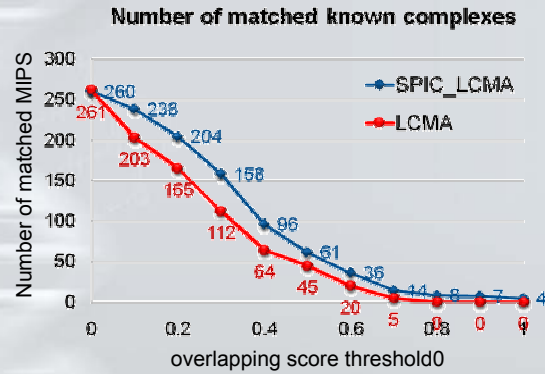
Results: MCODE vs. SPIC_MCODE



- Overall, SPIC_MCODE outperforms original MCODE
- All refined clusters have higher overlapping scores than corresponding conventional clusters



Results: LCMA vs. SPIC_LCMA



- Overall, SPIC_LCMA outperforms original LCMA
- Refined clusters have higher overlapping scores than corresponding conventional clusters except one case.



Refinement effect

- Trimming/ splitting over-predicted clusters
 - 16 cases for MCODE
 - 74 case for LCMA

~~YER111C, YDL074C, YGR104C, YOL122C, YEL036C, YGL167C, YDL246C, YJR142W, YLR342W, YDR414C, YBR200W, YGL215W, YMR306W, YPL075W, YNL271C, YLR425W, YGR078C, YLR371W, YJR075W, YJL101C, YOL021W, YPL217W, YML115C, YHL028W, YCR081W, YOL119C, YML122C, YCR073C, YNL199C, YDR150W, YPL161C, YLR360W, YBR260C, YCR006W, YDR456W, YOR008C, YGR157W, YDR207C, YNL098W, YNL090C, YDL095W, YJR159W, YOR014W, YGR161C, YER096W, YLR372W, YMR212C, YNL201C, YBR081C, YHR063C, YLR200W, YDR473C, YER155C, YDR245W, YHR030C, YGL015C, YER049W, YPL253C, YMR307W, YNL323W, YGR229C~~

YDL179W, YER059W, YHR071W, YIL050W,

YNL289W, YOL001W, YDL127W, YGL134W, YPL219W,

YOR140W

YDR443C, YPL042C,

YNL025C, YCR081W

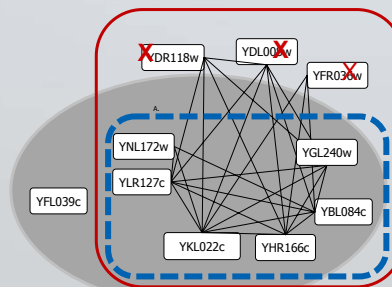
YHR012W, YJL154C,

YJL053W

known Complex

LCMA predicted

SPIC_LCMA predicted



Conclusion

- We propose the SPIC based approach for protein complex prediction
 - To eliminate unstable interactions in PPI network clusters using Mutually Exclusive Interactions (MEI) .
 - MEIs are identified by 3D analysis on PDB entity.
- Comparison with two conventional algorithms, MCODE and LCMA, showed that proposed approach
 - outperforms conventional methods.
 - rarely has adverse effects
 - correctly predicts what conventional method does and additional complexes more.
- The more stable a cluster is, the more likely it is to correspond with protein complexes.

