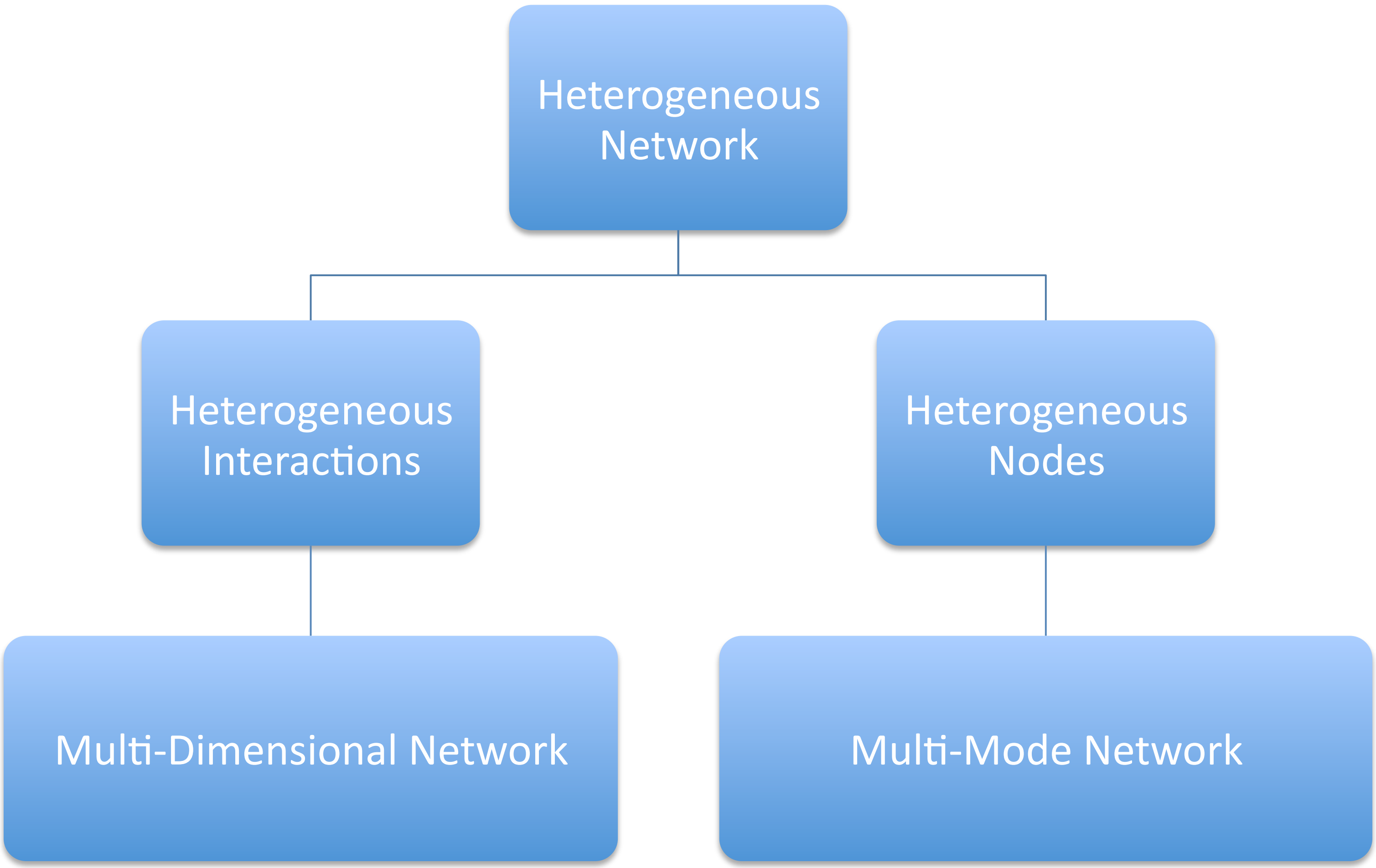
Communities in Heterogeneous Networks

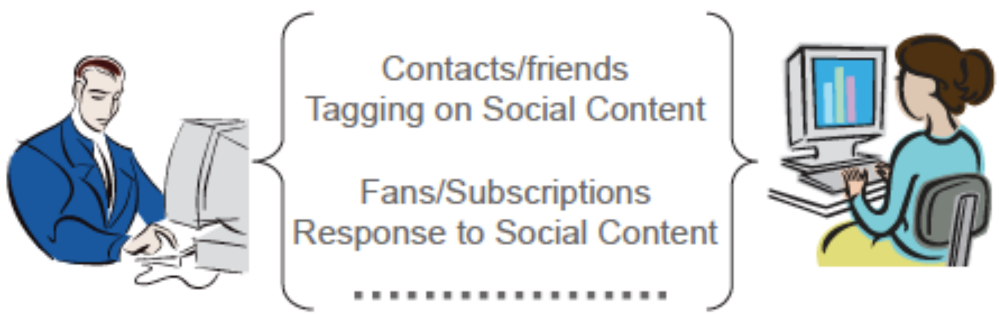
Chapter 4

Heterogeneous Networks



Multi-Dimensional Networks

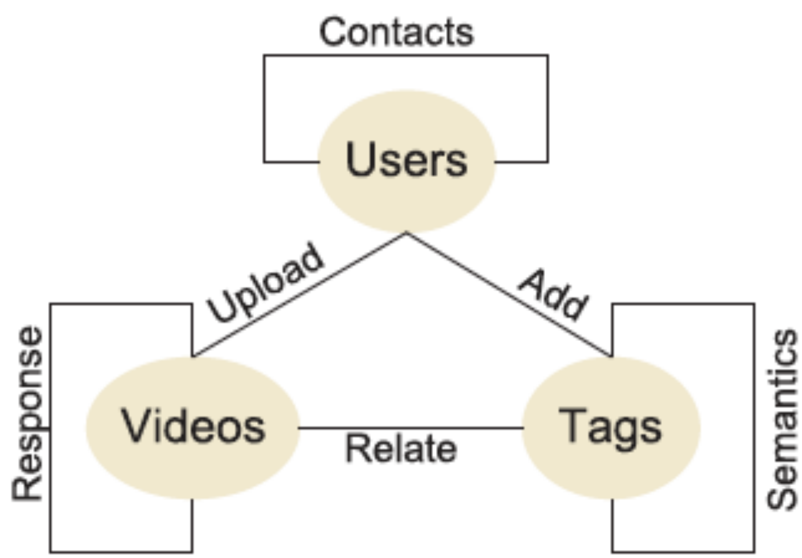
- Communications in social media are multi-dimensional
- Networks often involve heterogeneous connections
  - E.g. at YouTube, two users can be connected through friendship connection, email communications, subscription/Fans, chatter in comments, etc.



- a.k.a. *multi-relational networks*, *multiplex networks*, *labeled graphs*

Multi-Mode Networks

- Interactions in social media may involve heterogeneous types of entities
- Networks involve multiple modes of nodes
  - Within-mode interaction, between-mode interaction
  - Different types of interactions between different modes



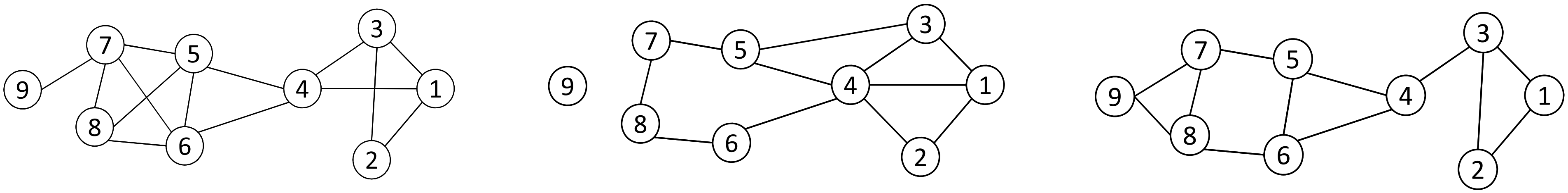
Why Does Heterogeneity Matter

- Social media introduces heterogeneity
- It calls for solutions to community detection in heterogeneous networks
  - Interactions in social media are noisy
  - Interactions in one mode or one dimension might be too noisy to detect meaningful communities
  - Not all users are active in all dimensions or with different modes
- Need integration of interactions at multiple dimensions or modes

COMMUNITIES IN MULTI-DIMENSIONAL NETWORKS

Communities in Multi-Dimensional Networks

- A *p*-dimension network $\mathcal{A} = \{A^{(1)}, A^{(2)}, \cdots, A^{(p)}\}$
- An example of a 3-dimensional network

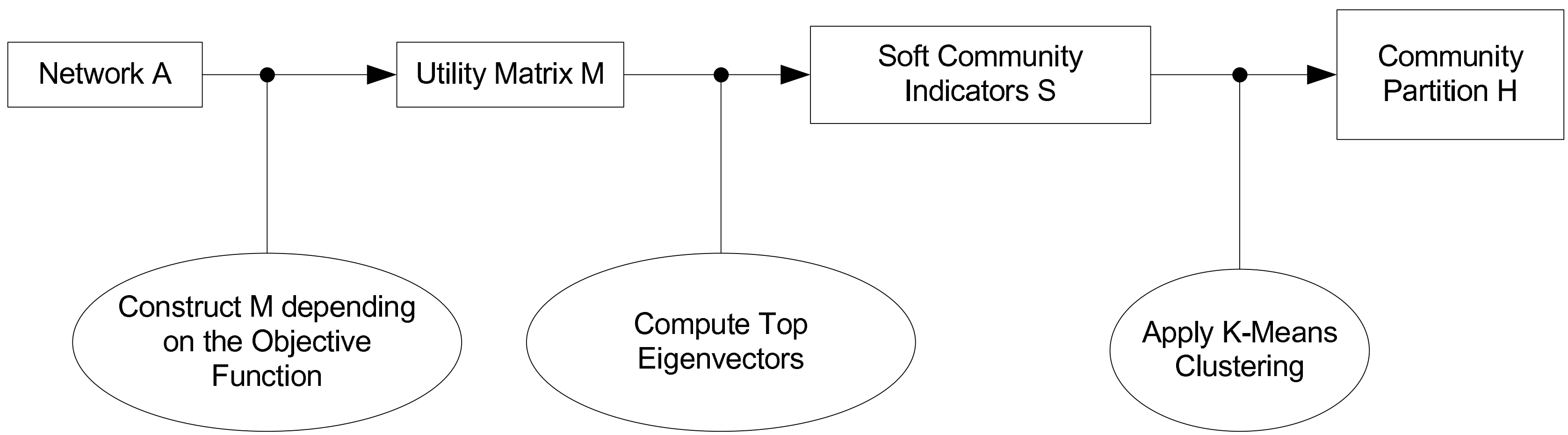


$A^{(1)}$ $A^{(2)}$ $A^{(3)}$

- Goal: integrate interactions at multiple dimensions to find reliable community structures

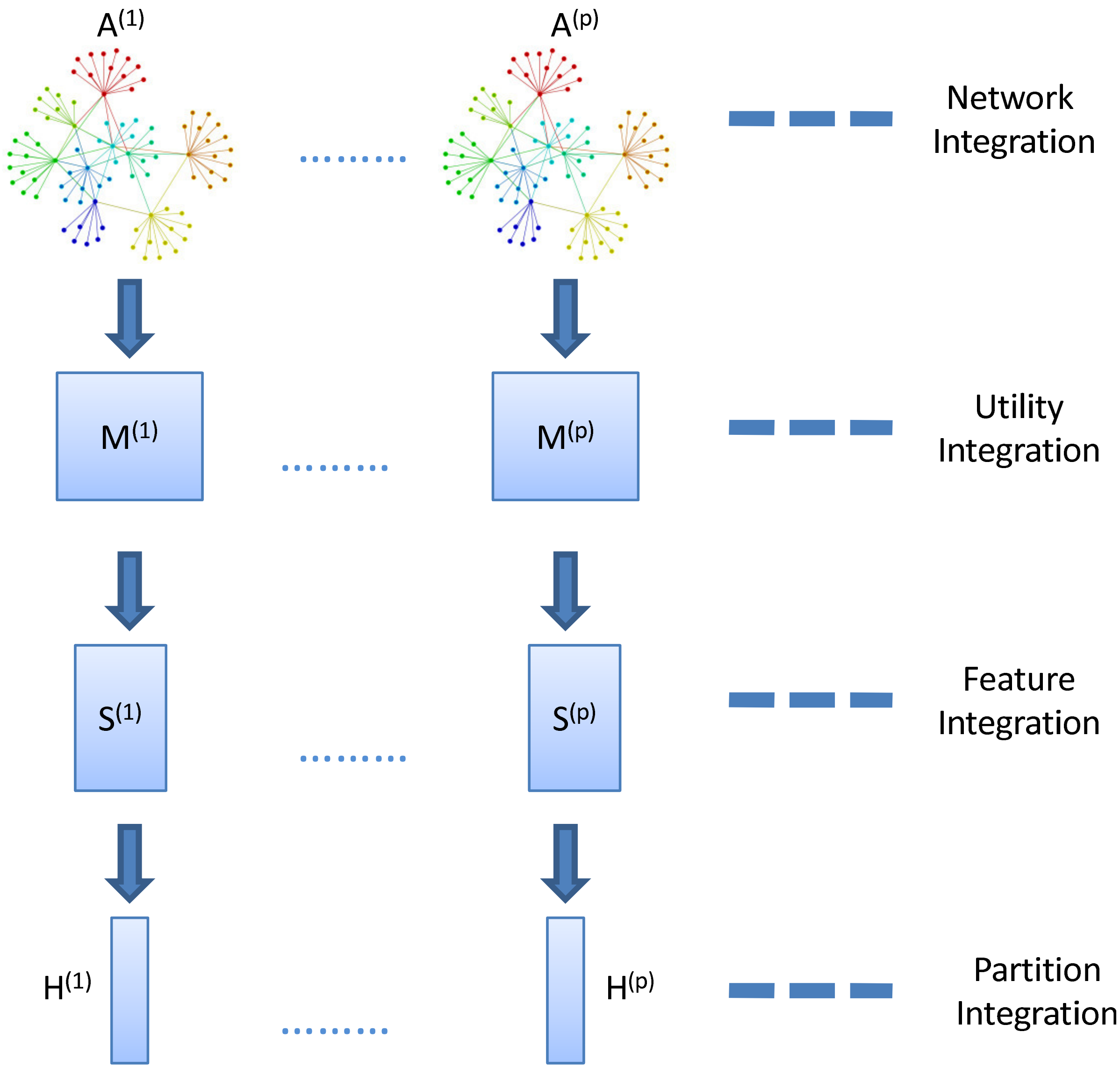
A Unified View for Community Partition (from Chapter 3)

- Latent space models, block models, spectral clustering, and modularity maximization can be unified as



$$\text{Utility Matrix } M = \begin{cases} \text{modified proximity matrix } \widetilde{P} & \textit{if} \text{ latent space models} \\ \text{adjacency matrix } A & \textit{if} \text{ block models} \\ \text{graph Laplacian } \widetilde{L} & \textit{if} \text{ spectral clustering} \\ \text{modularity maximization } \; B & \textit{if} \text{ modularity maximization} \end{cases}$$

Integration Strategies

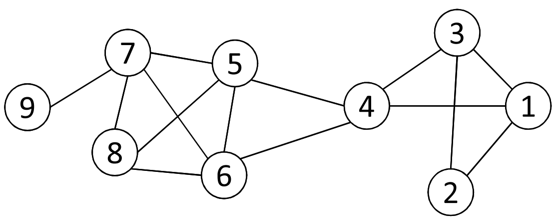


Network Integration

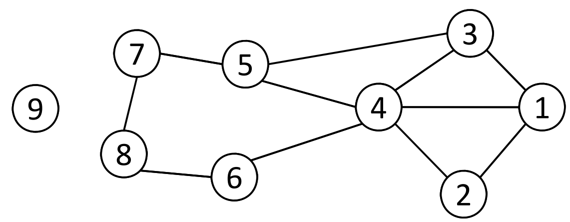
- Convert a multi-dimensional network into a single-dimensional network
- Different types of interaction strengthen one actor's connection
- The average strength is $$\bar{A} = \frac{1}{p}\sum_{i=1}^{p} A^{(i)}$$
- Spectral clustering with a *p*-dimensional network becomes

$$\begin{aligned} \min_{S} \quad & Tr(S^T \bar{L} S) \\ s.t. \quad & S^T S = I \\ \text{where} \quad & \bar{L} = \bar{D}^{-1/2}\bar{A}\bar{D}^{-1/2} \end{aligned}$$

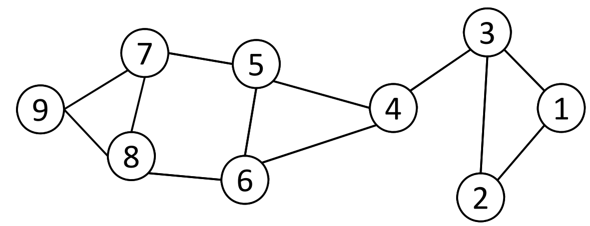
Network Integration Example



$A^{(1)}$



$A^{(2)}$



$A^{(3)}$

$$\bar{A} = \begin{bmatrix} 0 & 1 & 1 & 2/3 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2/3 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2/3 & 0 & 1 & 1/3 & 0 & 0 & 0 & 0 \\ 2/3 & 1/3 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 1 & 0 & 2/3 & 1 & 1/3 & 0 \\ 0 & 0 & 0 & 1 & 2/3 & 0 & 1/3 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1/3 & 0 & 1 & 2/3 \\ 0 & 0 & 0 & 0 & 1/3 & 1 & 1 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2/3 & 1/3 & 0 \end{bmatrix}. \quad S = \begin{bmatrix} -0.33 & -0.44 \\ -0.28 & -0.40 \\ -0.35 & -0.38 \\ -0.40 & -0.18 \\ -0.37 & 0.16 \\ -0.35 & 0.21 \\ -0.35 & 0.41 \\ -0.33 & 0.38 \\ -0.20 & 0.30 \end{bmatrix}.$$

Utility Integration

- Integration by averaging the utility matrix

$$\bar{M} = \frac{1}{p}\sum_{i=1}^{p} M^{(i)}$$

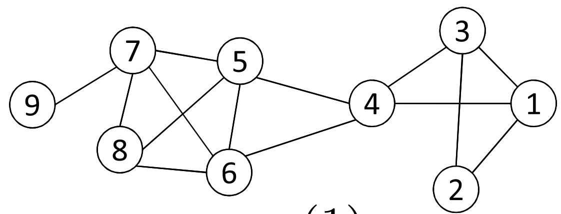
- Equivalent to optimizing average utility function

- For spectral clustering, $\bar{M} = \frac{1}{p}\sum_{i=1}^{p} \widetilde{L}^{(i)}$

- Hence, the objective of spectral clustering becomes

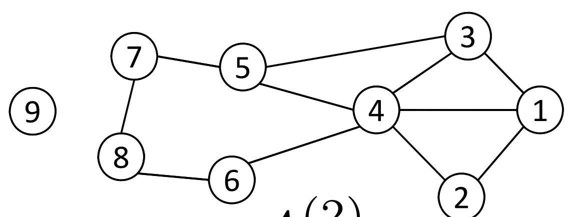
$$\min_S Tr(S^T \bar{M} S) = \min_S \frac{1}{p}\sum_{i=1}^{p} Tr(S^T \widetilde{L}^{(i)} S)$$

Utility Integration Example



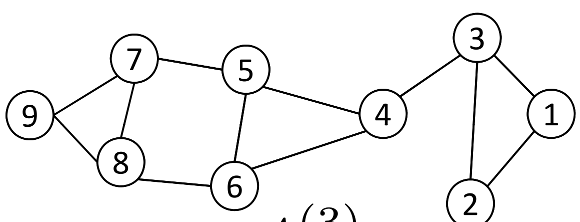
$A^{(1)}$

$$\widetilde{L}^{(1)} = \begin{bmatrix} 1.00 & -0.41 & -0.33 & -0.29 & 0 & 0 & 0 & 0 & 0 \\ -0.41 & 1.00 & -0.41 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.33 & -0.41 & 1.00 & -0.29 & 0 & 0 & 0 & 0 & 0 \\ -0.29 & 0 & -0.29 & 1.00 & -0.25 & -0.25 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.25 & 1.00 & -0.25 & -0.25 & -0.29 & 0 \\ 0 & 0 & 0 & -0.25 & -0.25 & 1.00 & -0.25 & -0.29 & 0 \\ 0 & 0 & 0 & 0 & -0.25 & -0.25 & 1.00 & -0.29 & -0.50 \\ 0 & 0 & 0 & 0 & -0.29 & -0.29 & -0.29 & 1.00 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.50 & 0 & 1.00 \end{bmatrix},$$



$A^{(2)}$

$$\widetilde{L}^{(2)} = \begin{bmatrix} 1.00 & -0.41 & -0.33 & -0.26 & 0 & 0 & 0 & 0 & 0 \\ -0.41 & 1.00 & 0 & -0.32 & 0 & 0 & 0 & 0 & 0 \\ -0.33 & 0 & 1.00 & -0.26 & -0.33 & 0 & 0 & 0 & 0 \\ -0.26 & -0.32 & -0.26 & 1.00 & -0.26 & -0.32 & 0 & 0 & 0 \\ 0 & 0 & -0.33 & -0.26 & 1.00 & 0 & -0.41 & 0 & 0 \\ 0 & 0 & 0 & -0.32 & 0 & 1.00 & 0 & -0.50 & 0 \\ 0 & 0 & 0 & 0 & -0.41 & 0 & 1.00 & -0.50 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.50 & -0.50 & 1.00 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



$A^{(3)}$

$$\widetilde{L}^{(3)} = \begin{bmatrix} 1.00 & -0.50 & -0.41 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.50 & 1.00 & -0.41 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.41 & -0.41 & 1.00 & -0.33 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -0.33 & 1.00 & -0.33 & -0.33 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.33 & 1.00 & -0.33 & -0.33 & 0 & 0 \\ 0 & 0 & 0 & -0.33 & -0.33 & 1.00 & 0 & -0.33 & 0 \\ 0 & 0 & 0 & 0 & -0.33 & 0 & 1.00 & -0.33 & -0.41 \\ 0 & 0 & 0 & 0 & 0 & -0.33 & -0.33 & 1.00 & -0.41 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.41 & -0.41 & 1.00 \end{bmatrix}$$

Utility Integration Example

$$\bar{M} = \left(\widetilde{L}^{(1)} + \widetilde{L}^{(2)} + \widetilde{L}^{(3)}\right)/3$$

$$\bar{M} = \begin{bmatrix} 1.00 & -0.44 & -0.36 & -0.18 & 0 & 0 & 0 & 0 & 0 \\ -0.44 & 1.00 & -0.27 & -0.11 & 0 & 0 & 0 & 0 & 0 \\ -0.36 & -0.27 & 1.00 & -0.29 & -0.11 & 0 & 0 & 0 & 0 \\ -0.18 & -0.11 & -0.29 & 1.00 & -0.28 & -0.30 & 0 & 0 & 0 \\ 0 & 0 & -0.11 & -0.28 & 1.00 & -0.19 & -0.33 & -0.10 & 0 \\ 0 & 0 & 0 & -0.30 & -0.19 & 1.00 & -0.08 & -0.37 & 0 \\ 0 & 0 & 0 & 0 & -0.33 & -0.08 & 1.00 & -0.37 & -0.30 \\ 0 & 0 & 0 & 0 & -0.10 & -0.37 & -0.37 & 1.00 & -0.14 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0.30 & -0.14 & 0.67 \end{bmatrix}$$

Spectral clustering based on utility integration leads to a partition of two communities: {1, 2, 3, 4} and {5, 6, 7, 8, 9}

Feature Integration

- Soft community indicators extracted from each type of interactions are structural features associated with nodes
- Integration can be done at the feature level

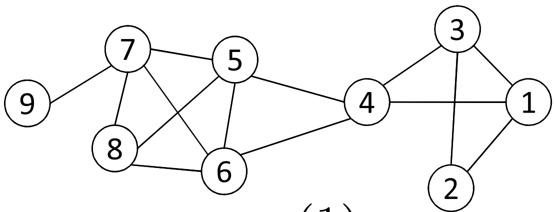
- A straightforward approach: take the average of structural features

$$\bar{S} = \frac{1}{p}\sum_{i=1}^{p} S^{(i)}.$$

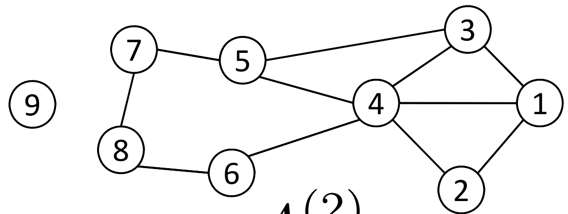
- Direct feature average is not sensible

- Need comparable coordinates among different dimensions

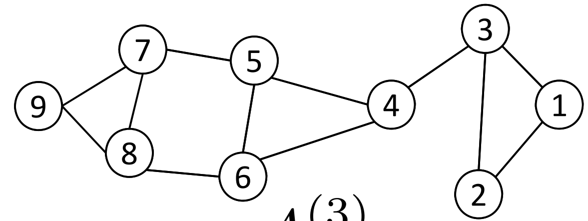
Problem with Direct Feature Average



$A^{(1)}$



$A^{(2)}$



$A^{(3)}$

$$S^{(1)} = \begin{bmatrix} 0.33 & -0.44 & 0.09 \\ 0.27 & -0.43 & 0.22 \\ 0.33 & -0.44 & 0.09 \\ 0.38 & -0.16 & -0.32 \\ 0.38 & 0.24 & -0.30 \\ 0.38 & 0.24 & -0.30 \\ 0.38 & 0.38 & 0.42 \\ 0.33 & 0.30 & -0.16 \\ 0.19 & 0.23 & 0.67 \end{bmatrix}, \quad S^{(2)} = \begin{bmatrix} -0.37 & 0 & 0.39 \\ -0.30 & 0 & 0.33 \\ -0.37 & 0 & 0.23 \\ -0.48 & 0 & 0.21 \\ -0.37 & 0 & -0.08 \\ -0.30 & 0 & -0.31 \\ -0.30 & 0 & -0.46 \\ -0.30 & 0 & -0.56 \\ 0 & 1.00 & 0 \end{bmatrix} \quad S^{(3)} = \begin{bmatrix} 0.29 & 0.47 & 0.21 \\ 0.29 & 0.47 & 0.21 \\ 0.35 & 0.44 & 0.01 \\ 0.35 & 0.04 & -0.50 \\ 0.35 & -0.17 & -0.39 \\ 0.35 & -0.17 & -0.39 \\ 0.35 & -0.33 & 0.28 \\ 0.35 & -0.33 & 0.28 \\ 0.29 & -0.30 & 0.45 \end{bmatrix}$$

$$\bar{S} = \begin{bmatrix} 0.08 & 0.01 & 0.23 \\ 0.08 & 0.01 & 0.26 \\ 0.10 & 0.00 & 0.11 \\ 0.08 & -0.04 & -0.20 \\ 0.12 & 0.02 & -0.26 \\ 0.14 & 0.02 & -0.34 \\ 0.14 & 0.02 & 0.08 \\ 0.13 & -0.01 & -0.15 \\ 0.16 & 0.31 & 0.37 \end{bmatrix}$$

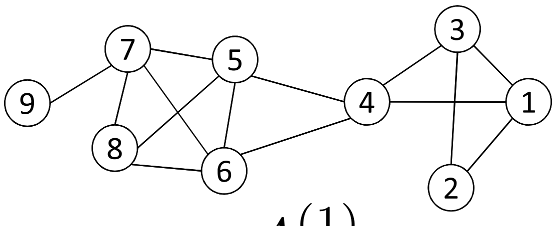
Two communities:
{1, 2, 3, 7, 9}
{4, 5, 6, 8}

Proper way of Feature Integration

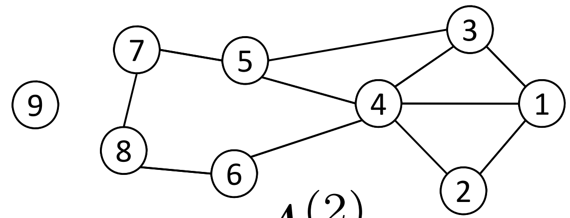
- Structural features of different dimensions are highly correlated after a certain transformation
- Multi-dimensional integration can be conducted after we map the structural features into the same coordinates
  - Find the transformation by maximizing pairwise correlation
  - Suppose the transformation associated with dimension (i) is $\mathbf{w}^{(i)}$
  - The average of structural features is $\frac{1}{p}\sum_{i=1}^{p} S^{(i)}\mathbf{w}^{(i)}$
  - The average is shown to be proportional to the top left singular vector of data X by concatenating structural features of each dimension

$$X = \left[S^{(1)}, S^{(2)}, \cdots, S^{(p)}\right]$$

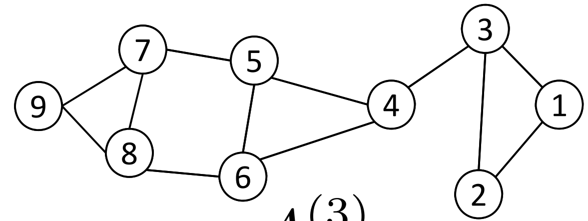
Feature Integration Example



$A^{(1)}$



$A^{(2)}$



$A^{(3)}$

$$X = [S^{(1)}, S^{(2)}, S^{(3)}] = \begin{bmatrix} 0.33 & -0.44 & 0.09 & -0.37 & 0 & 0.39 & 0.29 & 0.47 & 0.21 \\ 0.27 & -0.43 & 0.22 & -0.30 & 0 & 0.33 & 0.29 & 0.47 & 0.21 \\ 0.33 & -0.44 & 0.09 & -0.37 & 0 & 0.23 & 0.35 & 0.44 & 0.01 \\ 0.38 & -0.16 & -0.32 & -0.48 & 0 & 0.21 & 0.35 & 0.04 & -0.50 \\ 0.38 & 0.24 & -0.30 & -0.37 & 0 & -0.08 & 0.35 & -0.17 & -0.39 \\ 0.38 & 0.24 & -0.30 & -0.30 & 0 & -0.31 & 0.35 & -0.17 & -0.39 \\ 0.38 & 0.38 & 0.42 & -0.30 & 0 & -0.46 & 0.35 & -0.33 & 0.28 \\ 0.33 & 0.30 & -0.16 & -0.30 & 0 & -0.56 & 0.35 & -0.33 & 0.28 \\ 0.19 & 0.23 & 0.67 & 0 & 1.00 & 0 & 0.29 & -0.30 & 0.45 \end{bmatrix}$$

The top 2 left singular vectors of X are

$$\bar{S} = \begin{bmatrix} -0.30 & 0.42 \\ -0.26 & 0.38 \\ -0.33 & 0.38 \\ -0.39 & 0.24 \\ -0.37 & -0.07 \\ -0.36 & -0.14 \\ -0.36 & -0.40 \\ -0.35 & -0.36 \\ -0.23 & -0.40 \end{bmatrix}$$

Two Communities:
{1, 2, 3, 4}
{5, 6, 7, 8, 9}

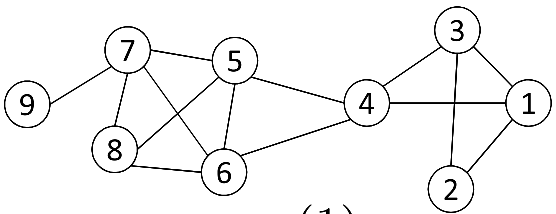
Partition Integration

- Combine the community partitions obtained from each type of interaction
  - a.k.a. *cluster ensemble*
- *Cluster-based Similarity Partitioning Algorithm (CPSA)*
  - Similarity is 1 is two objects belong to the same group, 0 otherwise
  - The similarity between nodes is computed as

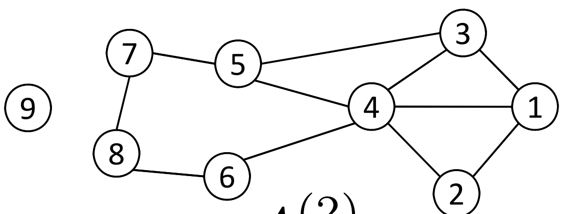
$$\frac{1}{p}\sum_{i=1}^{p} H^{(i)}(H^{(i)})^T = \frac{1}{p}\sum_{i=1}^{p} YY^T \quad \text{where} \quad Y = \left[H^{(1)}, H^{(2)}, \cdots, H^{(p)}\right]$$

  - The entry is essentially the probability that two nodes are assigned into the same community
  - Then apply similarity-based community detection methods to find clusters

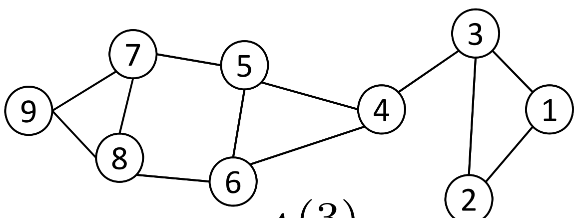
CPSA Example



$A^{(1)}$



$A^{(2)}$



$A^{(3)}$

$$H^{(1)} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad H^{(2)} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad H^{(3)} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

$$\frac{1}{p}\sum_{i=1}^{p} H^{(i)}(H^{(i)})^T = \begin{bmatrix} 1.00 & 1.00 & 1.00 & 0.67 & 0.33 & 0.33 & 0.33 & 0.33 & 0 \\ 1.00 & 1.00 & 1.00 & 0.67 & 0.33 & 0.33 & 0.33 & 0.33 & 0 \\ 1.00 & 1.00 & 1.00 & 0.67 & 0.33 & 0.33 & 0.33 & 0.33 & 0 \\ 0.67 & 0.67 & 0.67 & 1.00 & 0.67 & 0.67 & 0.67 & 0.67 & 0.33 \\ 0.33 & 0.33 & 0.33 & 0.67 & 1.00 & 1.00 & 1.00 & 1.00 & 0.67 \\ 0.33 & 0.33 & 0.33 & 0.67 & 1.00 & 1.00 & 1.00 & 1.00 & 0.67 \\ 0.33 & 0.33 & 0.33 & 0.67 & 1.00 & 1.00 & 1.00 & 1.00 & 0.67 \\ 0.33 & 0.33 & 0.33 & 0.67 & 1.00 & 1.00 & 1.00 & 1.00 & 0.67 \\ 0 & 0 & 0 & 0.33 & 0.6667 & 0.67 & 0.67 & 0.67 & 1.00 \end{bmatrix}$$

Applying spectral clustering to the above matrix results in two communities: {1, 2, 3, 4} and { 5, 6, 7, 8, 9}

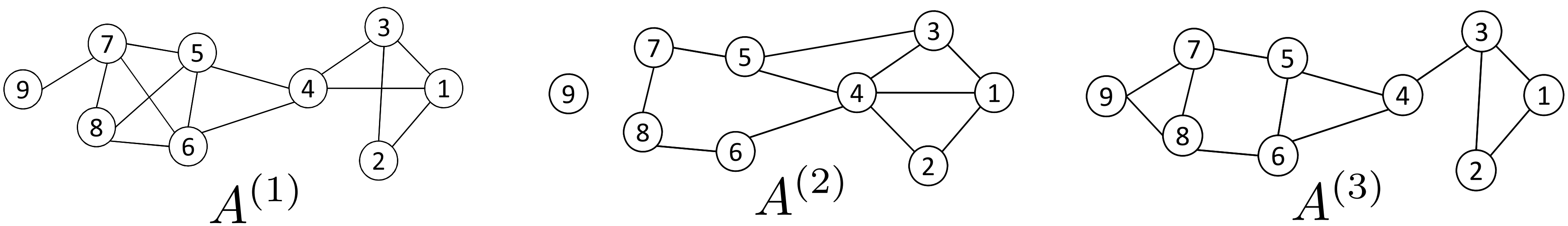
More Efficient Partition Integration

- CPSA requires the computation of a *dense* similarity matrix
  - Not scalable
- An alternative approach: Partition Feature Integration
  - Consider partition information as features
  - Apply a similar procedure as in feature integration
- A detailed procedure:
  - Given partitions of each dimension $H^{(1)}, H^{(2)}, \cdots, H^{(p)}$
  - Construct a *sparse* partition feature matrix

$$Y = \left[H^{(1)}, H^{(2)}, \cdots, H^{(p)}\right]$$

  - Take the top left singular vectors of Y as soft community indicator
  - Apply k-means to the singular vectors to find community partition

Partition Integration Example



$$Y = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \xrightarrow{\text{SVD}} \bar{H} = \begin{bmatrix} -0.27 & -0.47 \\ -0.27 & -0.47 \\ -0.27 & -0.47 \\ -0.35 & -0.14 \\ -0.39 & 0.22 \\ -0.39 & 0.22 \\ -0.39 & 0.22 \\ -0.39 & 0.22 \\ -0.24 & 0.35 \end{bmatrix} \xrightarrow{\text{k-means}}$$

{1, 2, 3, 4}
{5, 6, 7, 8, 9}

Y is sparse

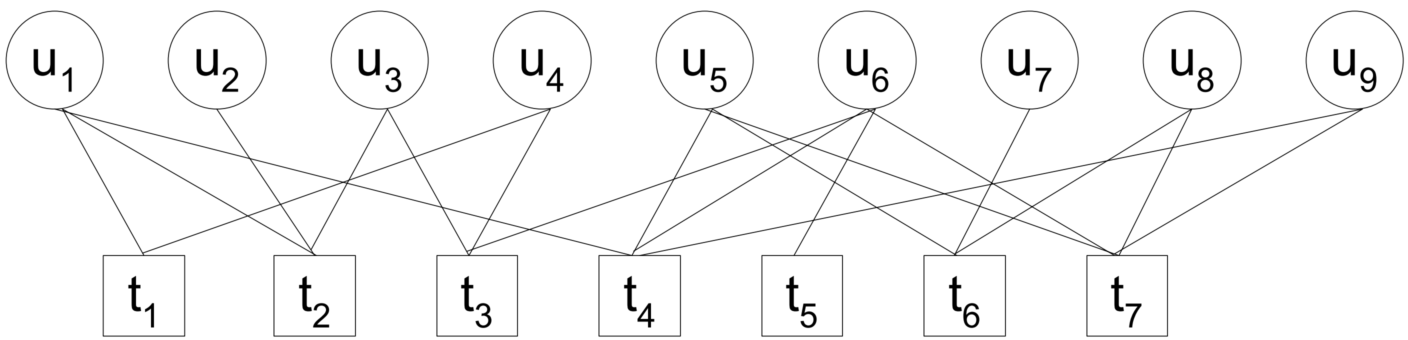
Comparison of Multi-Dimensional Integration Strategies

	Network Integration	Utility Integration	Feature Integration	Partition Integration
Tuning weights for different types of interactions	X	X	X	X
Sensitivity to noise	Yes	OK	Robust	Yes
Clustering quality	bad	Good	Good	OK
Computational cost	Low	Low	High	Expensive

COMMUNITIES IN MULTI-MODE NETWORKS

Co-clustering on 2-mode Networks

- Multi-mode networks involve multiple types of entities
- A 2-mode network is a simple form of multi-mode network
  - E.g., user-tag network in social media
  - A.k.a., *affiliation network*
- The graph of a 2-mode network is a bipartite



  - All edges are between users and tags
  - No edges between users or between tags

Adjacency Matrix of 2-Mode Network

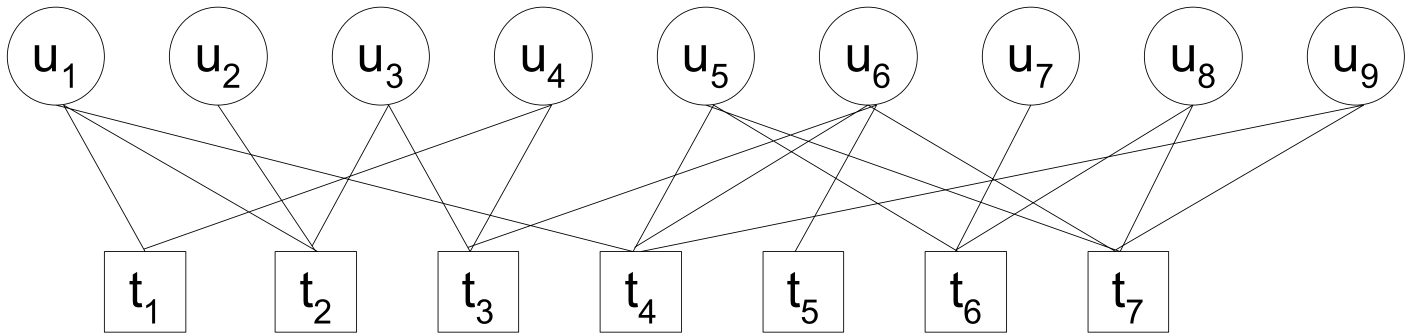


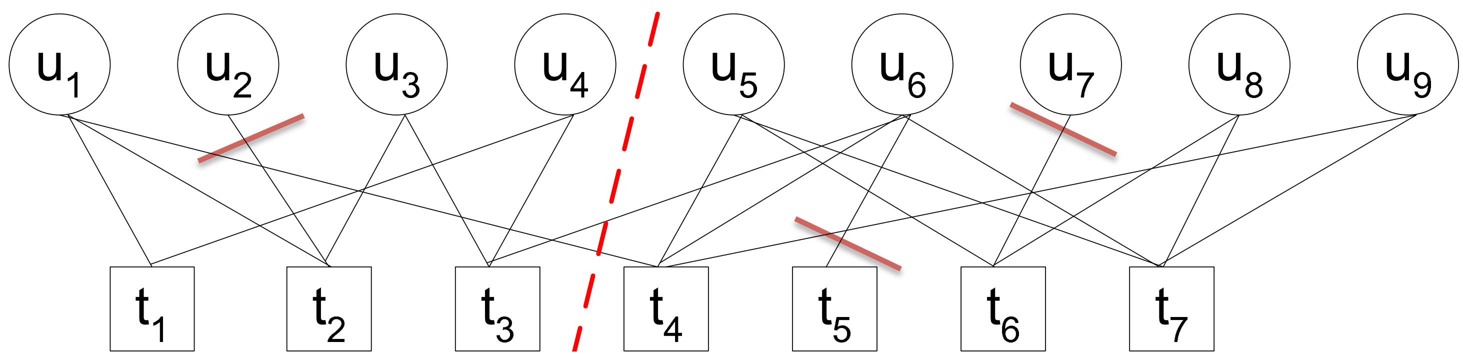
Table 4.2: Adjacency Matrix

	t_1	t_2	t_3	t_4	t_5	t_6	t_7
u_1	1	1	0	1	0	0	0
u_2	0	1	0	0	0	0	0
u_3	0	1	1	0	0	0	0
u_4	1	0	1	0	0	0	0
u_5	0	0	0	1	0	1	1
u_6	0	0	1	1	1	0	1
u_7	0	0	0	0	0	1	0
u_8	0	0	0	0	0	1	1
u_9	0	0	0	1	0	0	1

Each mode represents one type of entity; not necessarily a square matrix

Co-Clustering

- Co-clustering: find communities in two modes simultaneously
  - a.k.a. *biclustering*
  - Output both *communities of users and communities of tags* for a user-tag network
- A straightforward Approach: Minimize the cut in the graph



- The minimum cut is 1; a trivial solution is not desirable
- Need to consider the size of communities

Spectral Co-Clustering

- Minimize the normalized cut in a bipartite graph
  - Similar as spectral clustering for undirected graph
- Compute normalized adjacency matrix

$$\widetilde{A} = D_u^{-1/2} A D_t^{-1/2}$$

$$D_u = diag(d_{u_1}, d_{u_2}, \cdots, d_{u_m}), \quad D_t = diag(d_{t_1}, d_{t_2}, \cdots, d_{t_n})$$

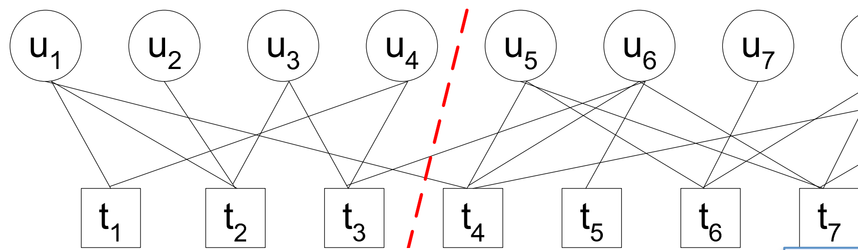
- Compute the top singular vectors of the normalized adjacency matrix

$$\widetilde{A} \approx S^{(u)} \Sigma_k S^{(t)}$$

- Apply k-means to the joint community indicator Z to obtain communities in user mode and tag mode, respectively

$$Z = \begin{bmatrix} D_u^{-1/2} S^{(u)} \\ D_t^{-1/2} S^{(t)} \end{bmatrix}$$

Spectral Co-Clustering Example



Two communities:
{ u_1, u_2, u_3, u_4, t_1, t_2, t_3 }
{ u_5, u_6, u_7, u_8, u_9, t_4, t_5, t_6, t_7}

$$\widetilde{A} = D_u^{-1/2} A D_t^{-1/2} = \begin{bmatrix} 0.41 & 0.33 & 0 & 0.29 & 0 & 0 & 0 \\ 0 & 0.58 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.41 & 0.41 & 0 & 0 & 0 & 0 \\ 0.50 & 0 & 0.41 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.29 & 0 & 0.33 & 0.29 \\ 0 & 0 & 0.29 & 0.25 & 0.50 & 0 & 0.25 \\ 0 & 0 & 0 & 0 & 0 & 0.58 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.41 & 0.36 \\ 0 & 0 & 0 & 0.36 & 0 & 0 & 0.36 \end{bmatrix}$$

k-means

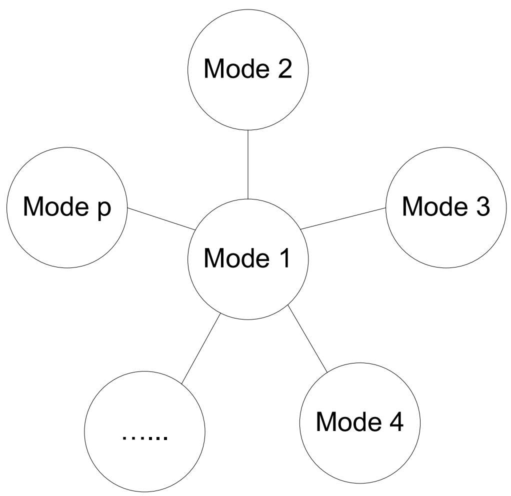
$$S^{(u)} = \begin{bmatrix} -0.39 & 0.33 \\ -0.22 & 0.35 \\ -0.32 & 0.40 \\ -0.32 & 0.35 \\ -0.39 & -0.37 \\ -0.45 & -0.04 \\ -0.22 & -0.38 \\ -0.32 & -0.42 \\ -0.32 & -0.19 \end{bmatrix}, \quad S^{(t)} = \begin{bmatrix} -0.32 & 0.35 \\ -0.39 & 0.35 \\ -0.39 & 0.33 \\ -0.45 & -0.10 \\ -0.22 & -0.02 \\ -0.39 & -0.58 \\ -0.45 & -0.38 \end{bmatrix}, \quad Z = \begin{array}{c} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \\ u_8 \\ u_9 \\ t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \\ t_7 \end{array} \begin{bmatrix} -0.22 & 0.19 \\ -0.22 & 0.35 \\ -0.22 & 0.28 \\ -0.22 & 0.25 \\ -0.22 & -0.22 \\ -0.22 & -0.02 \\ -0.22 & -0.38 \\ -0.22 & -0.29 \\ -0.22 & -0.14 \\ -0.22 & 0.25 \\ -0.22 & 0.32 \\ -0.22 & 0.19 \\ -0.22 & -0.05 \\ -0.22 & -0.02 \\ -0.22 & -0.33 \\ -0.22 & -0.19 \end{bmatrix}.$$

Generalization to A Star Structure

- Spectral co-clustering can be interpreted as a *block model approximation* to normalized adjacency matrix

$$\min_{S^{(1)},S^{(2)},\Sigma} \|\widetilde{A} - S^{(1)}\Sigma S^{(2)}\|_F^2, \quad s.t. \quad (S^{(1)})^T S^{(1)} = I_k, \ (S^{(2)})^T S^{(2)} = I_k$$

generalize to
a star structure



$$\min \sum_{q=2}^{p} \|\widetilde{A}^{(q)} - S^{(1)}\Sigma^{(q)}(S^{(q)})^T\|_F^2$$
$$s.t. (S^{(1)})^T S^{(1)} = I_k,$$
$$(S^{(q)})^T S^{(q)} = I_k, \quad q = 2, \cdots, p.$$

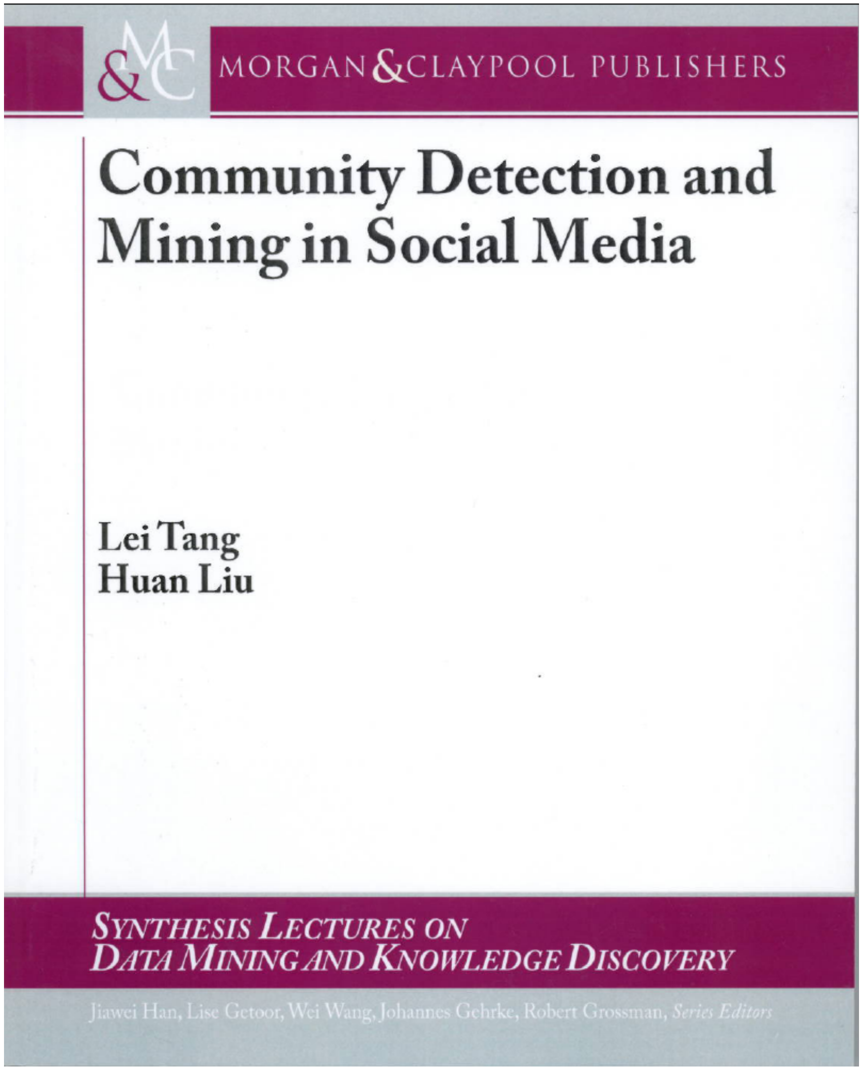
$S^{(1)}$ corresponds to the top left singular vectors of the following matrix

$$X = \left[\widetilde{A}^{(2)} S^{(2)}, \widetilde{A}^{(3)} S^{(3)}, \cdots, \widetilde{A}^{(p)} S^{(p)}\right]$$

Generalization to Multi-Mode Networks

- For a multi-mode network, compute the soft community indicator of each mode one by one
- It becomes a star structure when looking at one mode vs. other modes

- Community Detection in Multi-Mode Networks
  - Normalize interaction matrix
  - Iteratively update community indicator as the top left singular vectors
  - Apply k-means to the community indicators to find partitions in each mode



Book Available at

- Morgan & claypool Publishers
- Amazon

If you have any comments, please feel free to contact:

- **Lei Tang**, Yahoo! Labs, ltang@yahoo-inc.com
- **Huan Liu**, ASU huanliu@asu.edu