Spectral Clustering: Going Beyond MinCut

Ira A. Fulton Schools of Engineering
Arizona State University

Objective

Objective

Discuss several graph cut approaches

Objective

Illustrate the algorithm through an example

MinCut

- In MinCut, we used the following objective function:

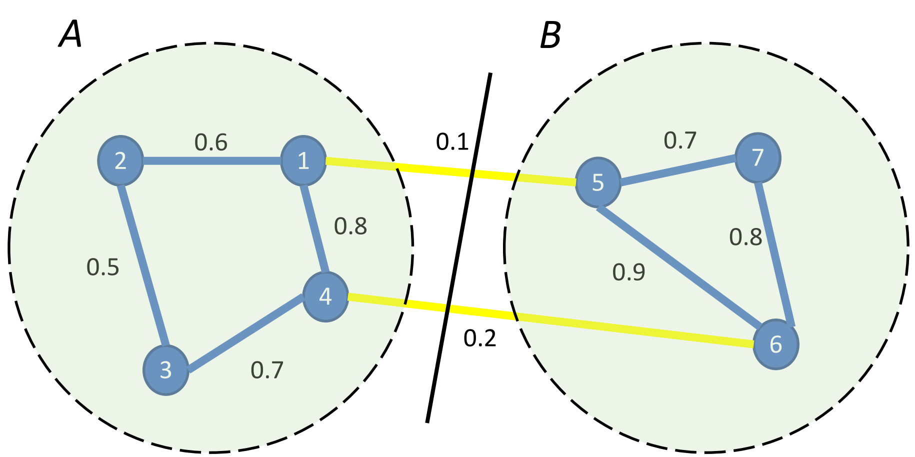
$$J_{MinCut} = Cut(A, B)$$

- We noted one drawback of MinCut: the sizes of the partitions are not considered.

- A few extensions exist.

Characterizing Graph Cut

- $Cut(A,B) = \sum_{i \in A, j \in B} w_{ij}$ $e.g., Cut(A,B) = 0.3$
- $Cut(A,A) = \sum_{i \in A, j \in A} w_{ij}$ $e.g., Cut(A,A) = 2.6$
- $Cut(B,B) = \sum_{i \in B, j \in B} w_{ij}$ $e.g., Cut(B,B) = 2.4$
- $Vol(A) = \sum_{i \in A} \sum_{i=1}^{n} w_{ij}$ $e.g., Vol(A) = 5.5$
- $Vol(B) = \sum_{i \in B} \sum_{i=1}^{n} w_{ij}$ $e.g., Vol(B) = 5.1$
- $|A| = 4, |B| = 3$



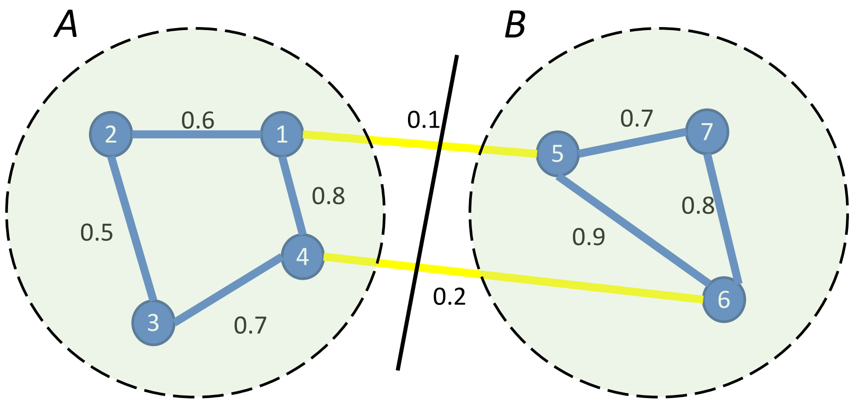
The Ratio Cut Method

- The Objective function:

$$J_{RatioCut}(A,B) = Cut(A,B)\left(\frac{1}{|A|} + \frac{1}{|B|}\right)$$

- Attempts to produce balanced clusters.

Example: $J_{RatioCut}(A,B) = \frac{7}{40}$



The Ratio Cut Method (cont'd)

- Similar to MinCut, the solution can be found by the following generalized eigenvalue problem:

$$(\mathbf{D} - \mathbf{W})\mathbf{q} = \lambda \mathbf{D}\mathbf{q}$$

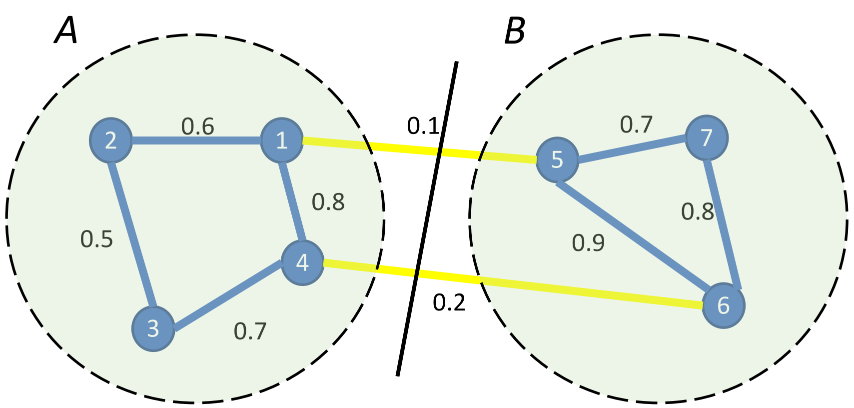
$$\mathbf{L}\mathbf{q} = \lambda \mathbf{D}\mathbf{q}$$

Normalized Cut (NCut)

- In Ratio Cut, the balance of the partitions is defined based on the number of vertices.
- We may consider the “size” of a set based on weights of its edges ➔ Ncut
- The objective function is:

$$J_{NCut}(A,B) = Cut(A,B)\left(\frac{1}{Vol(A)} + \frac{1}{Vol(B)}\right)$$

Example: $J_{NCut}(A,B) = 0.1134$



Additional Considerations

- In clustering, we should also consider **within-cluster** connections.
- A good partition should consider
  - Inter-cluster connections, and
  - Intra-cluster connections.

MinMaxCut

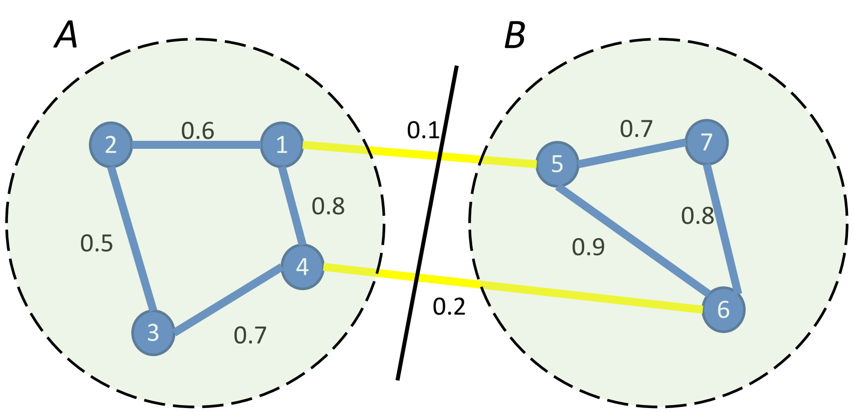
- 1st constrant: inter-connection should be minimized: $MinCut(A,B)$

- 2nd constraint: intra-connection should be maximized : $MaxCut(A,A)\ and\ MaxCut(B,B)$

- These requirements may be simultaneously satisfied by minimizing the objective function:

$$J_{MinMaxCut}(A,B) = Cut(A,B)\left(\frac{1}{Cut(A,A)} + \frac{1}{Cut(B,B)}\right)$$

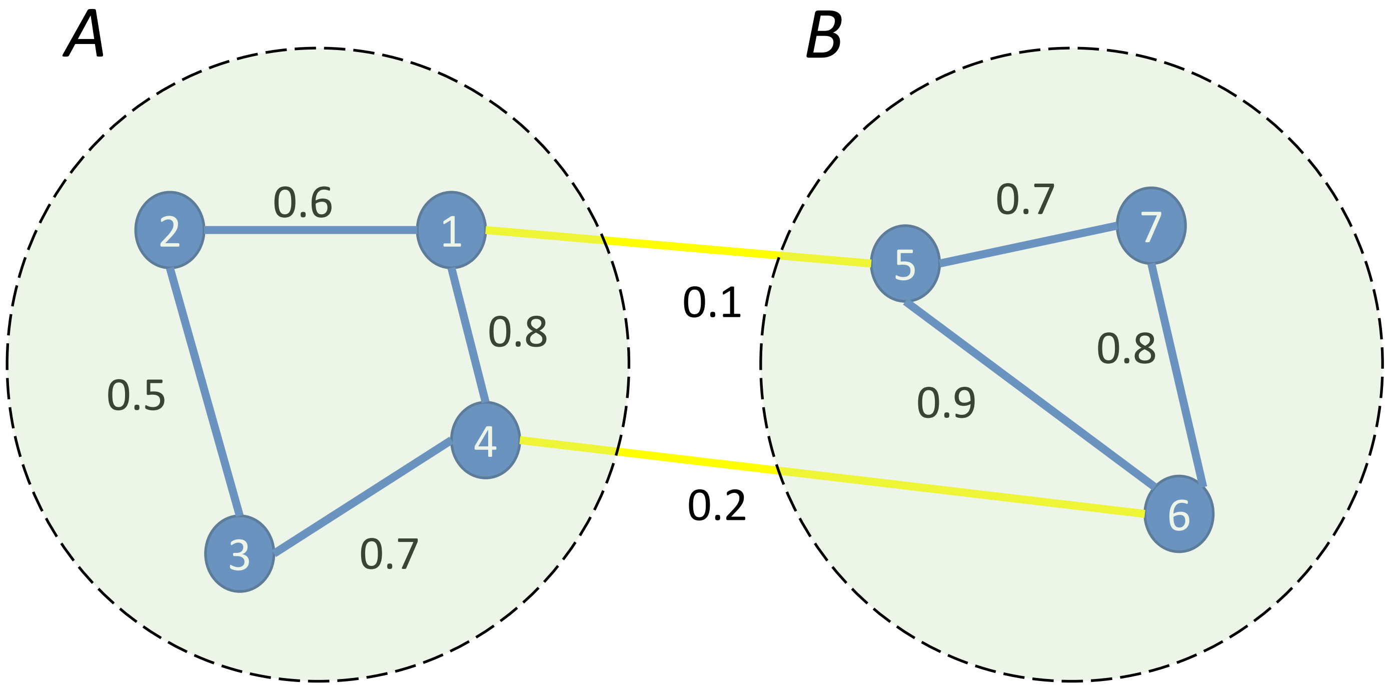
Example: $J_{MinMaxCut}(A,B) = 0.240$



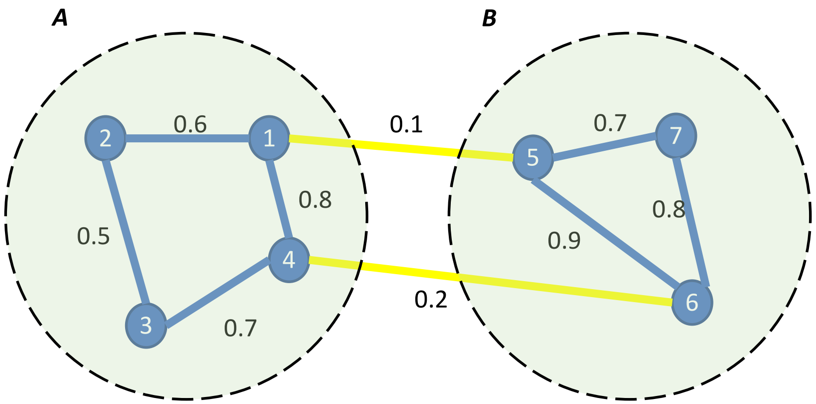
Normalized and MinMaxCut methods

- Similar to before, we may relax the indicator vector **q** to real values.
- For both NCut and MinMaxCut, the solution may be found by solving generalized eigenvalue problems.

An Illustrative Example

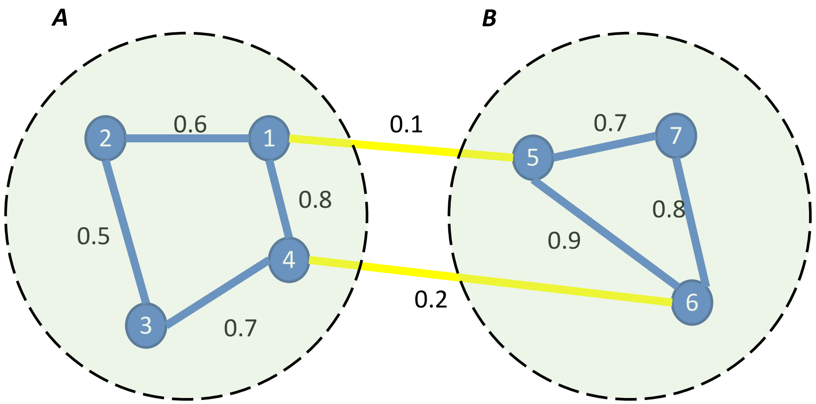


Graph and Similarity Matrix



	x1	x2	x3	x4	x5	x6	x7
x1	0	0.6	0	0.8	0.1	0	0
x2	0.6	0	0.5	0	0	0	0
x3	0	0.5	0	0.7	0	0	0
x4	0.8	0	0.7	0	0	0.2	0
x5	0.1	0	0	0	0	0.9	0.7
x6	0	0	0	0.2	0.9	0	0.8
x7	0	0	0	0	0.7	0.8	0

Graph and Laplacian Matrix

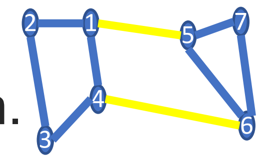


	x1	x2	x3	x4	x5	x6	x7
x1	1.5	-0.6	0	-0.8	-0.1	0	0
x2	-0.6	1.1	-0.5	0	0	0	0
x3	0	-0.5	1.2	-0.7	0	0	0
x4	-0.8	0	-0.7	1.7	0	-0.2	0
x5	-0.1	0	0	0	1.7	-0.9	-0.7
x6	0	0	0	-0.2	-0.9	1.9	-0.8
x7	0	0	0	0	-0.7	-0.8	1.5

Solve Eigen Problem

Pre-processing

- Build Laplacian matrix L of the graph.



Λ =

0
0.1588
1.2705
1.3692
2.2751
2.6238
2.9027

X =

0.378	-0.2962	0.3027	-0.6041	0.0429	0.3638	-0.4226
0.378	-0.3805	0.6392	0.4487	-0.0125	-0.233	0.2192
0.378	-0.3608	-0.5812	0.4834	-0.0221	0.2736	-0.2832
0.378	-0.2649	-0.398	-0.4373	0.0429	-0.3899	0.5323
0.378	0.4298	0.0443	0.0159	-0.6004	0.4291	0.3544
0.378	0.406	-0.0317	0.0012	-0.2174	-0.6116	-0.5196
0.378	0.4665	0.0247	0.0923	0.7667	0.1681	0.1195

Find

- Eigenvalues Λ and eigenvectors **x** of matrix **L**.
- Map vertices to the corresponding components of the 2nd eigenvector.

x1	-0.2962
x2	-0.3805
x3	-0.3608
x4	-0.2649
x5	0.4298
x6	0.406
x7	0.4665

Spectral Clustering

x1	-0.2962
x2	-0.3805
x3	-0.3608
x4	-0.2649
x5	0.4298
x6	0.406
x7	0.4665

Split at value 0
Cluster A: Negative points
Cluster B: Positive Points

x1	-0.2962
x2	-0.3805
x3	-0.3608
x4	-0.2649

X5	0.4298
X6	0.406
X7	0.4665