Graphs in Machine Learning

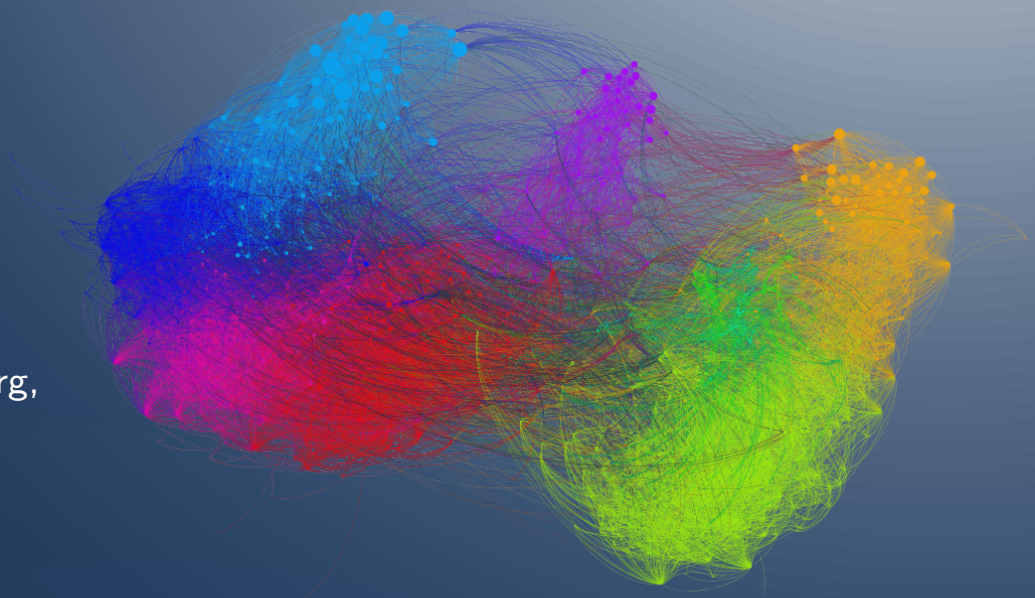
Spectral Clustering: Graph Cuts

MinCut, RatioCut, and NCut

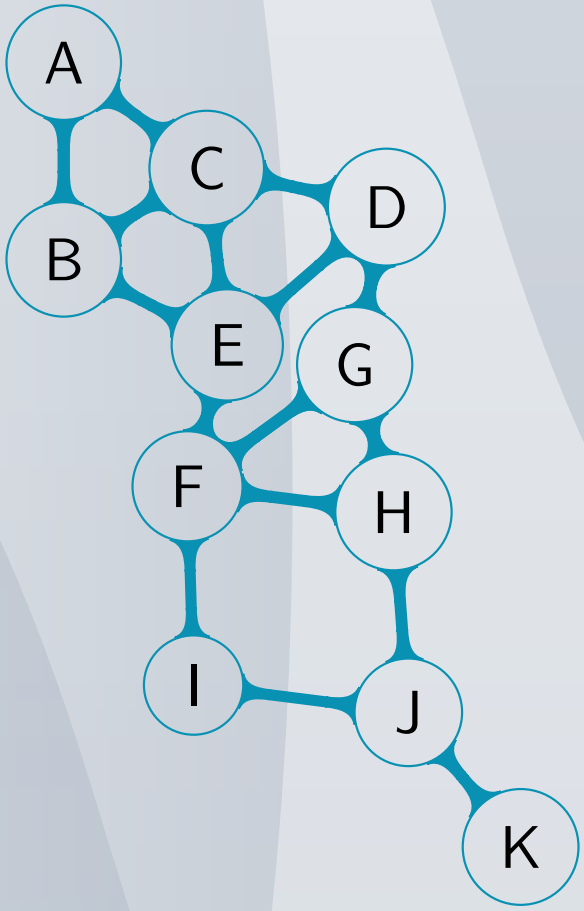
Michal Valko

Inria & ENS Paris-Saclay, MVA

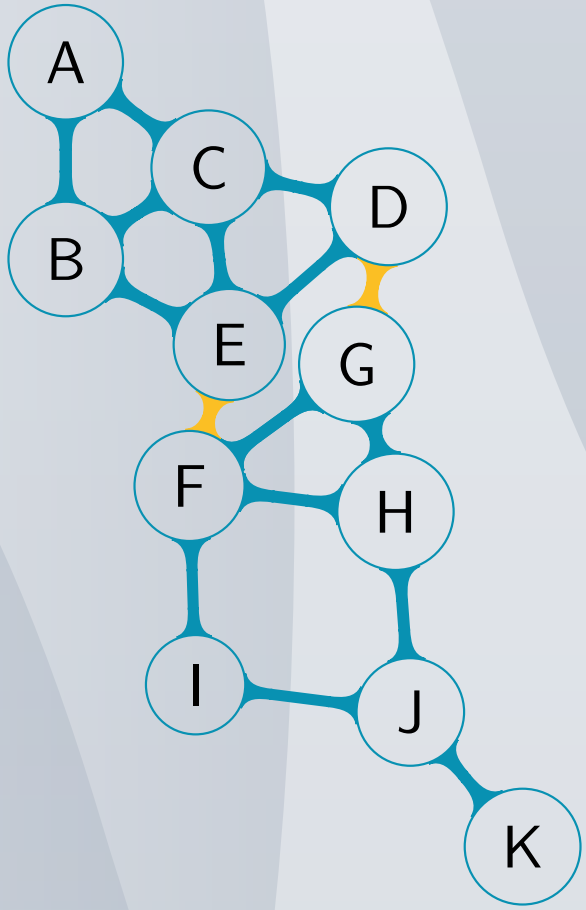
Partially based on material by: Ulrike von Luxburg, Gary Miller, Doyle & Schnell, Daniel Spielman



Spectral Clustering: Cuts on graphs

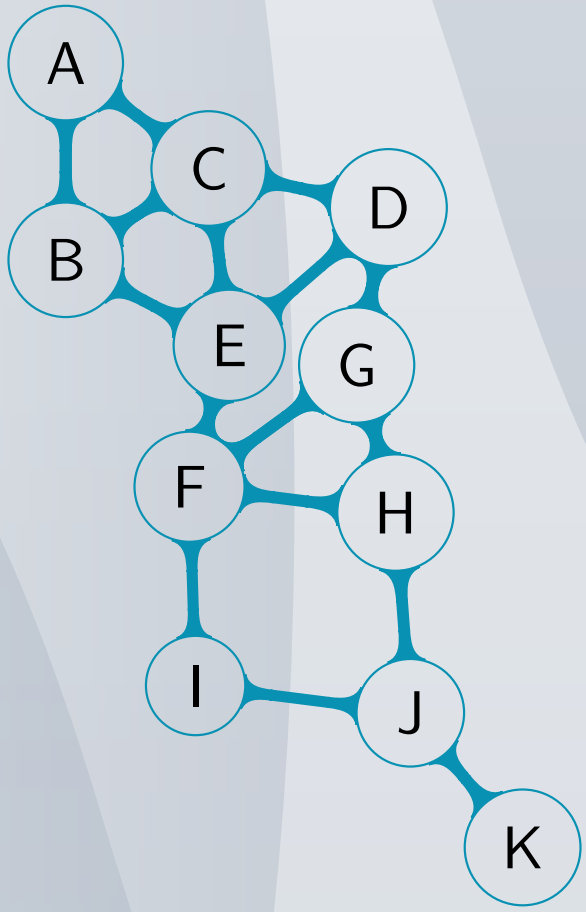


Spectral Clustering: Cuts on graphs



Defining the cut objective we get the clustering!

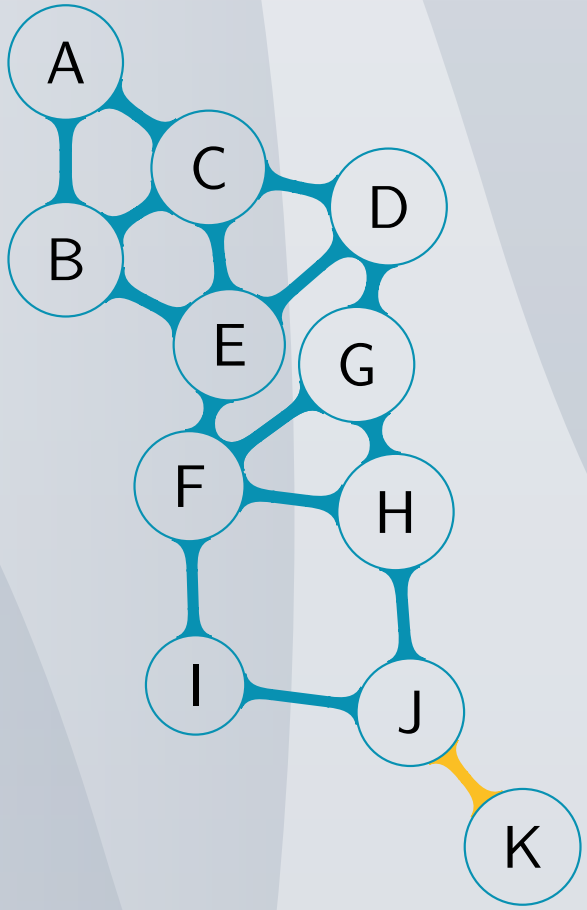
Spectral Clustering: Cuts on graphs



MinCut: $\mathrm{cut}(A, B) = \sum_{i \in A, j \in B} w_{ij}$

Are we done?

Spectral Clustering: Cuts on graphs



MinCut: $\mathrm{cut}(A, B) = \sum_{i \in A, j \in B} w_{ij}$

Are we done?

Can be solved efficiently, but maybe not what we want . . . .

Spectral Clustering: Balanced Cuts

Let's balance the cuts!

Spectral Clustering: Balanced Cuts

Let's balance the cuts!

MinCut

$$\mathrm{cut}(A, B) = \sum_{i \in A, j \in B} w_{ij}$$

Spectral Clustering: Balanced Cuts

Let's balance the cuts!

MinCut

$$\mathrm{cut}(A, B) = \sum_{i \in A, j \in B} w_{ij}$$

RatioCut

$$\mathrm{RatioCut}(A, B) = \sum_{i \in A, j \in B} w_{ij} \left(\frac{1}{|A|} + \frac{1}{|B|} \right)$$

Spectral Clustering: Balanced Cuts

Let's balance the cuts!

MinCut

$$\mathrm{cut}(A, B) = \sum_{i \in A, j \in B} w_{ij}$$

RatioCut

$$\mathrm{RatioCut}(A, B) = \sum_{i \in A, j \in B} w_{ij} \left(\frac{1}{|A|} + \frac{1}{|B|} \right)$$

Normalized Cut

$$\mathrm{NCut}(A, B) = \sum_{i \in A, j \in B} w_{ij} \left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)} \right)$$

Spectral Clustering: Balanced Cuts

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

$$\mathrm{NCut}(A, B) = \mathrm{cut}(A, B)\left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)}\right)$$

Easily generalizable to $k \geq 2$

Spectral Clustering: Balanced Cuts

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

$$\mathrm{NCut}(A, B) = \mathrm{cut}(A, B)\left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)}\right)$$

Easily generalizable to $k \geq 2$

Can we compute this?

Spectral Clustering: Balanced Cuts

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

$$\mathrm{NCut}(A, B) = \mathrm{cut}(A, B)\left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)}\right)$$

Easily generalizable to $k \geq 2$

Can we compute this? RatioCut and NCut are NP hard :(

Spectral Clustering: Balanced Cuts

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

$$\mathrm{NCut}(A, B) = \mathrm{cut}(A, B)\left(\frac{1}{\mathrm{vol}(A)} + \frac{1}{\mathrm{vol}(B)}\right)$$

Easily generalizable to $k \geq 2$

Can we compute this? RatioCut and NCut are NP hard :(

Approximate!

Michal Valko

michal.valko@inria.fr

Inria & ENS Paris-Saclay, MVA

https://misovalko.github.io/mva-ml-graphs.html