Some Elements of Spectral Graph Theory

References:

First few chapters of

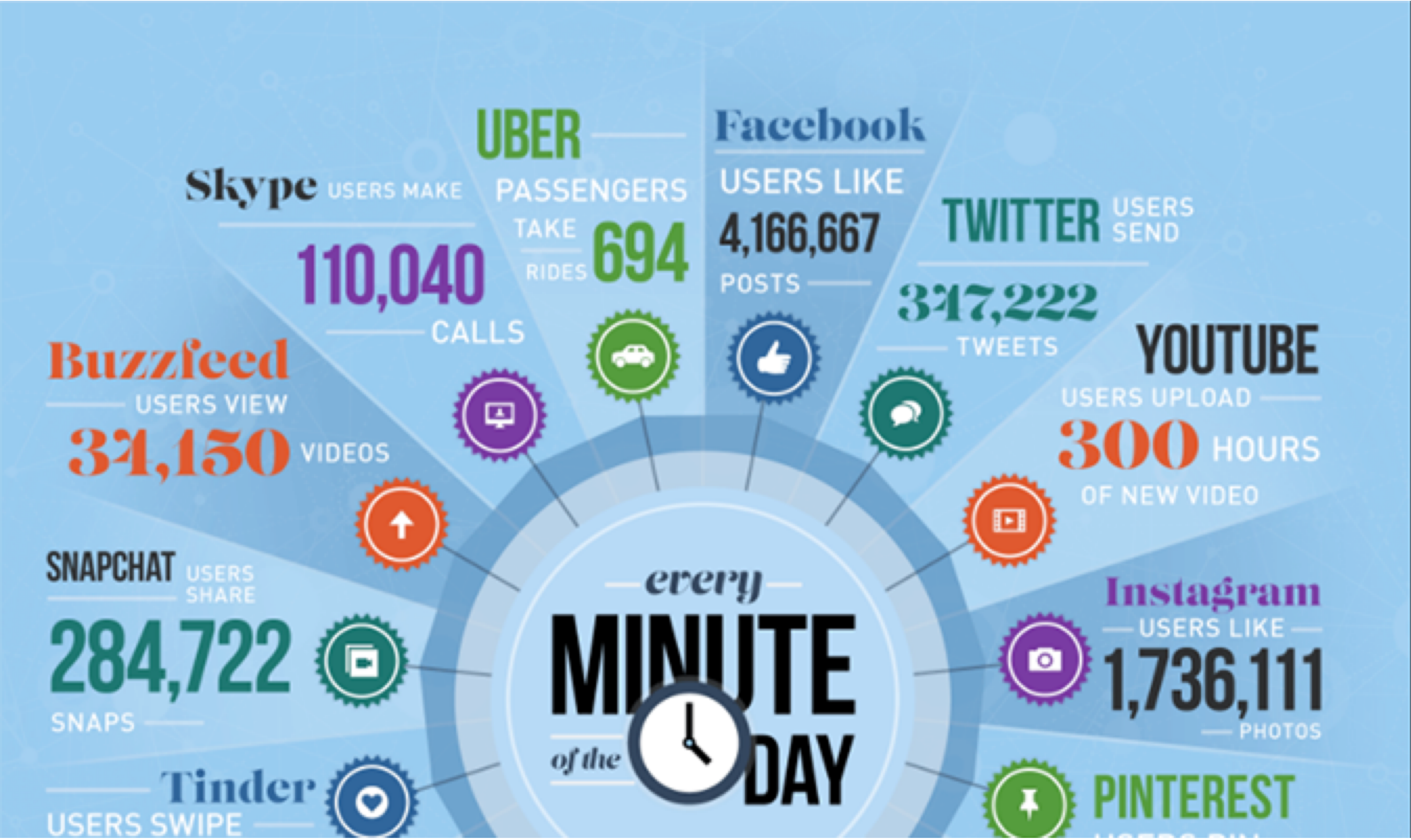
"Spectral graph theory", Fan Chung

LTS EPFL

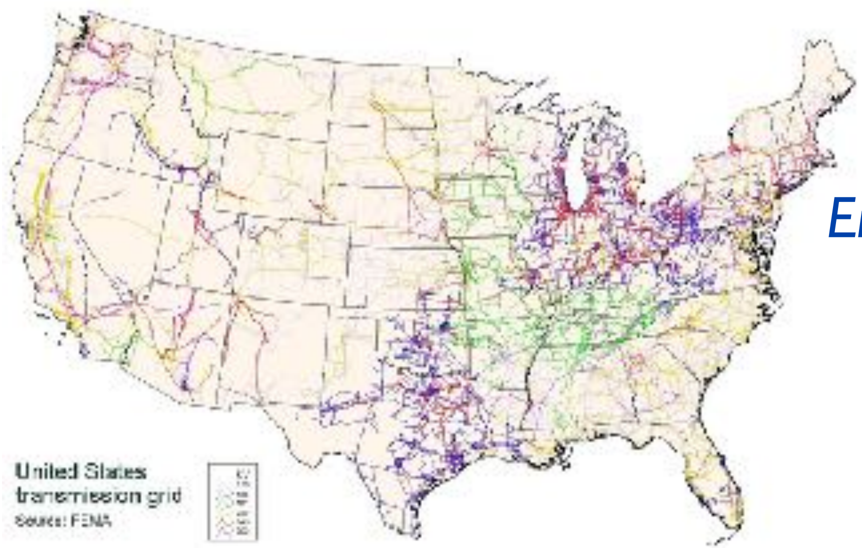
LTS2

EPFL

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

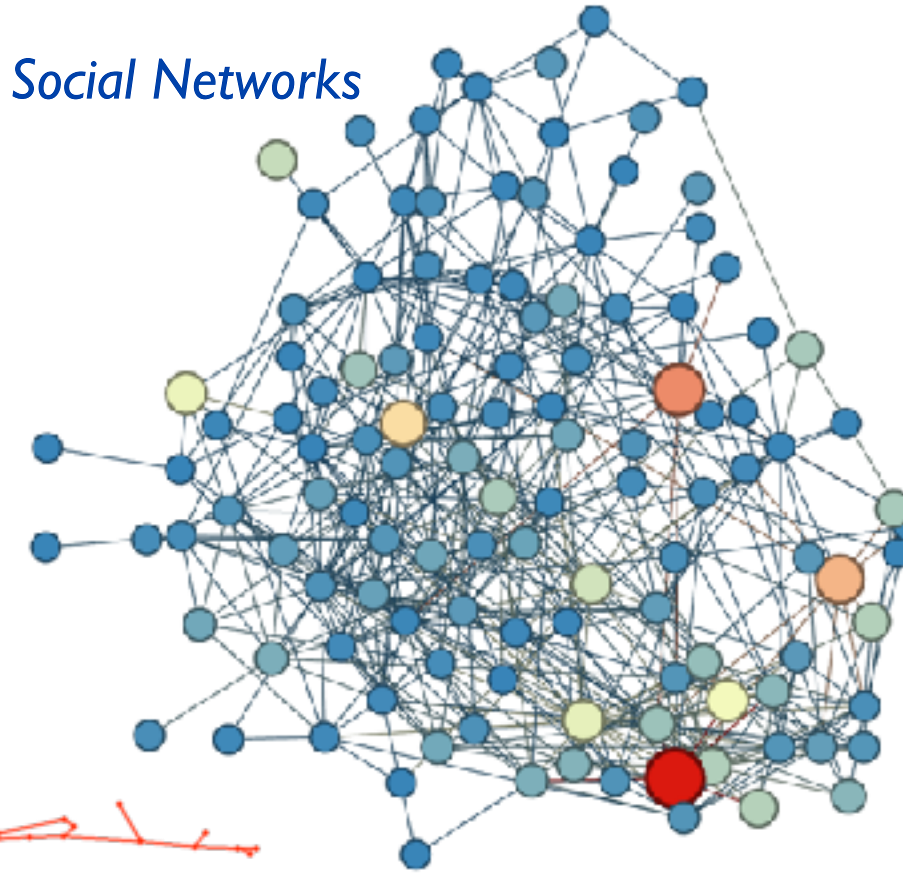

Skype USERS MAKE
110,040
CALLS
UBER
PASSENGERS
TAKE
694
RIDES
Facebook
USERS LIKE
4,166,667
POSTS
TWITTER USERS SEND
347,222
TWEETS
YOUTUBE
USERS UPLOAD
300 HOURS
OF NEW VIDEO
Buzzfeed
USERS VIEW
34,150 VIDEOS
SNAPCHAT USERS SHARE
284,722
SNAPS
every
MINUTE
of the
DAY
Instagram
USERS LIKE
1,736,111
PHOTOS
Tinder
USERS SWIPE
PINTEREST

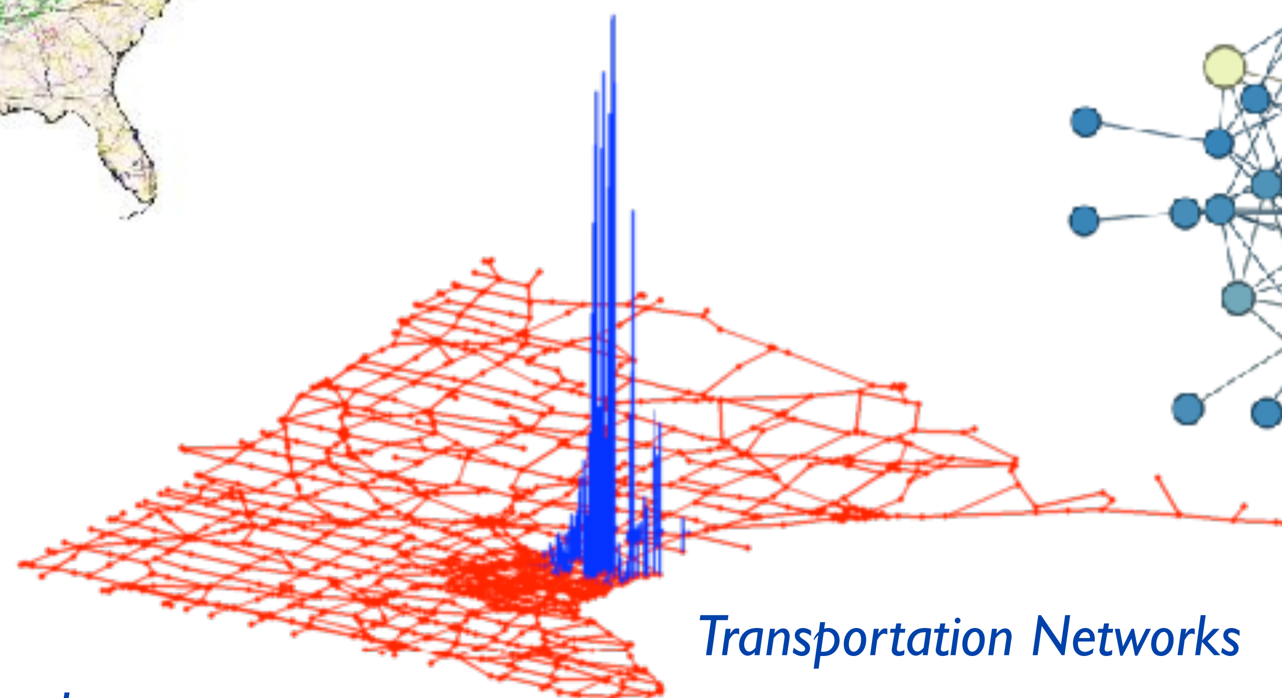
Processing Data on/with Graphs



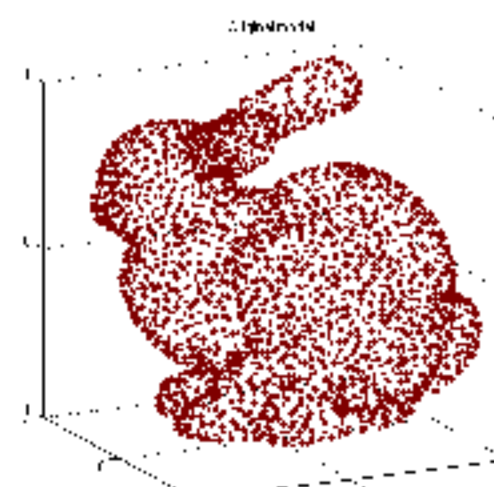
Energy Networks

Social Networks

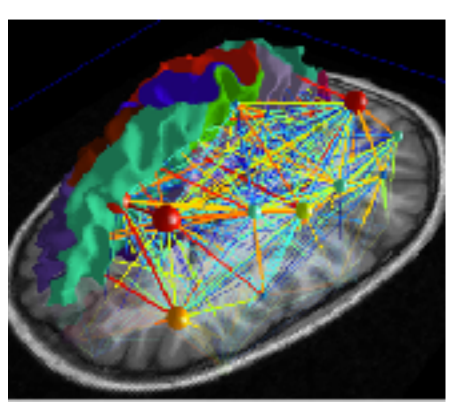




Transportation Networks



Biological Networks

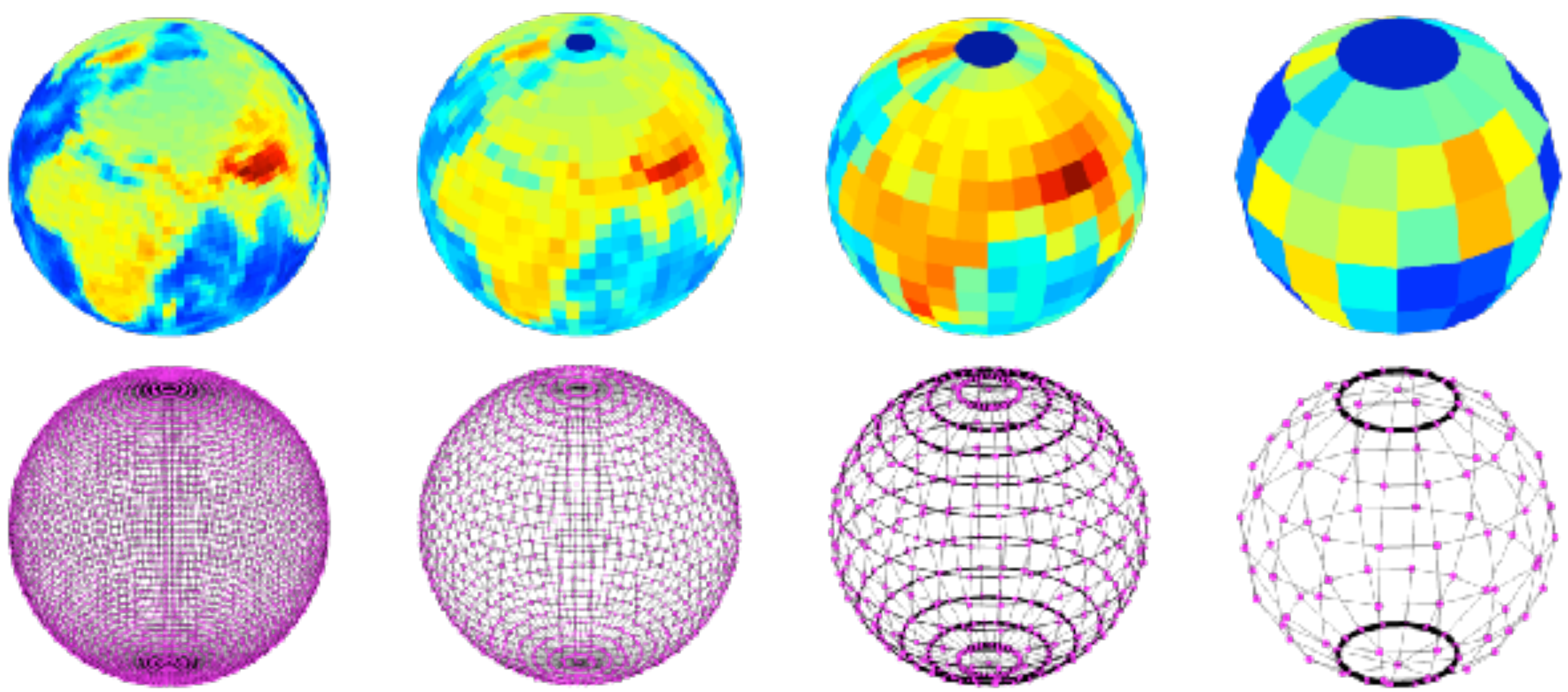


Irregular Data Domains

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

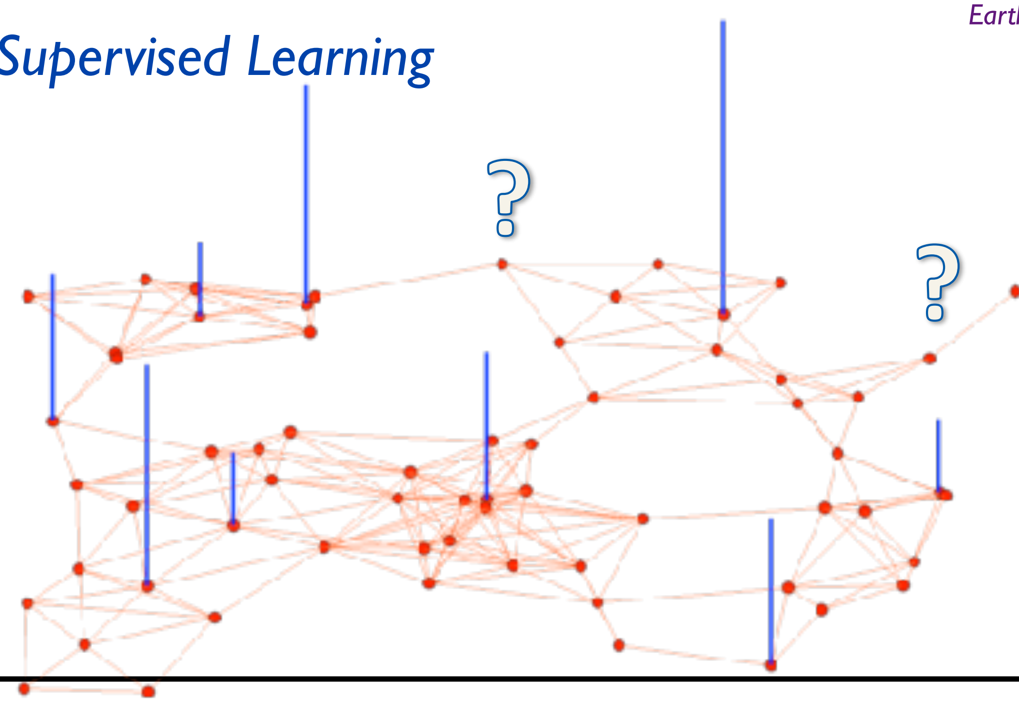
Some Typical Processing Problems

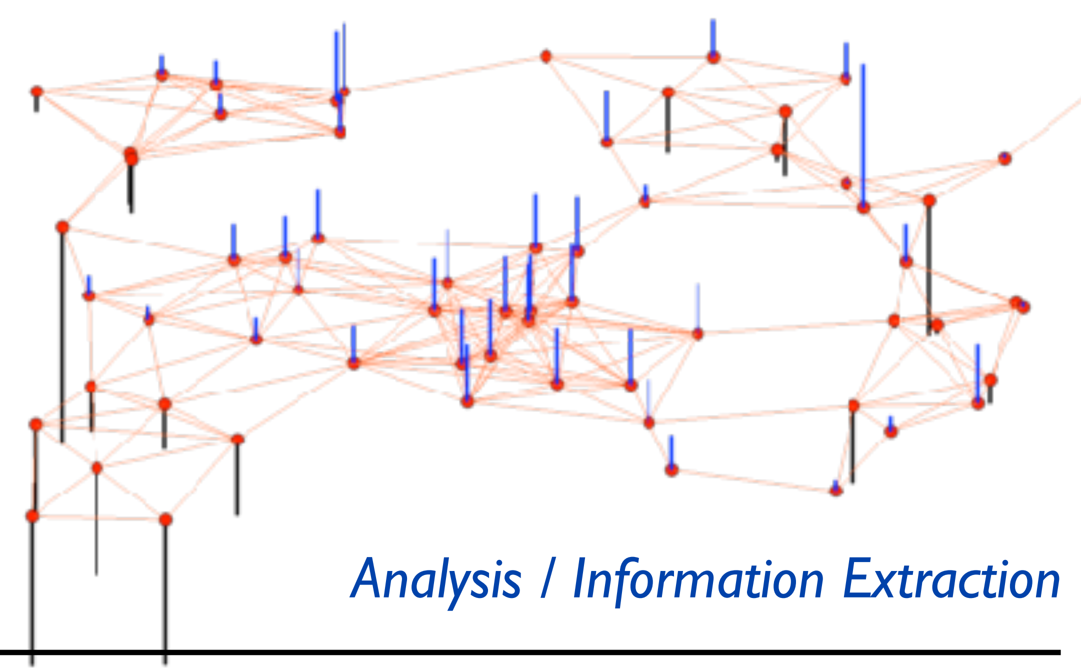
Compression / Visualization



Earth data source: Frederik Simons

Semi-Supervised Learning





Analysis / Information Extraction

Outline

- A (very short) introduction to spectral graph theory
  - Laplacian eigenvalues, eigenvectors, Cuts and Cheeger's inequality
- Unsupervised Learning: Spectral Clustering
  - several views, including smooth partition functions
- Embedding and Visualization: Laplacian Eigenmaps
  - Connections with smooth partition functions
- Signal Processing on Graphs
  - Signals, filters, algorithms and applications in ML

Orientation-agnostic definitions

$$V = \{v_1, \ldots, v_N\} \qquad E = \{e_1, \ldots, e_M\}$$

Degrees and Degree Matrix:

$$d(v) = \left|\{u \in V \text{ s.t. } (u,v) \in E \text{ or } (v,u) \in E\}\right|$$

$$\mathbf{D}(G) = \text{diag}(d_1, \ldots d_N)$$

Incidence Matrix:

$$\mathbf{S}(i,j) = \begin{cases} +1 & \text{if } e_j = (v_i, v_k) \text{ for some } k \\ -1 & \text{if } e_j = (v_k, v_i) \text{ for some } k \\ 0 & \text{otherwise} \end{cases}$$

Orientation-agnostic definitions

$$V = \{v_1, \ldots, v_N\} \qquad E = \{e_1, \ldots, e_M\}$$

Adjacency matrix

$$\mathbf{A}(i,j) = \begin{cases} +1 & \text{if there is an edge } (v_i, v_j) \text{ or } (v_j, v_i) \in E \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbf{A}(i,j) = \begin{cases} +1 & \text{if there is an edge } \{v_i, v_j\} \in E \\ 0 & \text{otherwise} \end{cases}$$

Extensions to weighted graphs

$$V = \{v_1, \ldots, v_N\} \qquad E = \{e_1, \ldots, e_M\}$$

Weight Matrix:

A *symmetric* N-by-N matrix $\mathbf{W}$

$$\mathbf{W}(i, j) \geqslant 0 \qquad \mathbf{W}(i, i) = 0$$

$\mathbf{W}(i,j)$ is the weight ("strength") of the edge between i,j (if any)

Degrees:

$$d(v_i) = \sum_{j \sim i} \mathbf{W}(i, j)$$

Orientation-agnostic definitions

With these definitions we have:

$$\mathbf{S}\mathbf{S}^T = \mathbf{D} - \mathbf{A}$$

$\mathbf{L} = \mathbf{D}$ - $\mathbf{A}$ is called unnormalized Laplacian of G

$\mathbf{L}$ does not depend on the orientation (so OK for undirected)

For a weighted graph we have $\mathbf{L} = \mathbf{D}$ - $\mathbf{W}$ (attention to degrees)

$\mathbf{L}$ is a symmetric, positive semi-definite matrix

Graph Laplacian

Proposition: L is positive semi-definite

For any N-by-N weight matrix $\mathbf{W}$, if $\mathbf{L} = \mathbf{D}$-$\mathbf{W}$ where $\mathbf{D}$ is the degree matrix of $\mathbf{W}$, then

$$x^T \mathbf{L} x = \frac{1}{2} \sum_{i,j} \mathbf{W}(i,j)(x[i] - x[j])^2 \geqslant 0 \quad \forall x \in \mathbb{R}^N$$

Rem: to ease notations we will sometimes use $w_{ij} = \mathbf{W}(i,j)$

Since $\mathbf{L}$ is real, symmetric and PSD:

- It has an eigendecomposition into real eigenvalues and eigenvectors λ_i, u_i
- The eigenvalues are non-negative

$$0 = \lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_N$$

$$\Updownarrow$$

$$\mathbf{L1} = 0$$

What can be learned from eigenvectors and eigenvalues ?

Some examples

Path graph DCT II transform

$$\begin{bmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & & & \\ & 2 & & & & \\ & & 2 & & & \\ & & & \ddots & & \\ & & & & 2 & \\ & & & & & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 & & & & \\ 1 & 0 & 1 & & & \\ & 1 & 0 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & 0 & 1 \\ & & & & 1 & 0 \end{bmatrix}$$

$$\lambda_k = 2 - 2\cos\frac{\pi k}{N} = 4\sin^2\frac{\pi k}{2N},\ k = 0, ..., N-1$$

$$u_k[\ell] = \cos\left(\pi k(\ell + \frac{1}{2})/N\right),\ \ell = 0, ..., N-1$$

Some examples

Ring graph

$$\begin{pmatrix} 2 & -1 & & & -1 \\ -1 & 2 & -1 & & \\ & & \ddots & & \\ -1 & & & -1 & 2 \end{pmatrix}$$

DCT transform

$$\lambda_k = 2 - 2\cos\frac{\pi k}{N} = 4\sin^2\frac{\pi k}{2N},\ k = 0, \ldots, N-1$$

$$u_k^c[\ell] = \cos\left(2\pi k\ell/N\right),\ \ell = 0, \ldots, N-1$$

$$u_k^s[\ell] = \sin\left(2\pi k\ell/N\right),\ \ell = 0, \ldots, N-1$$

Proposition: eigendecomposition of $\mathbf{L}$ and structure of G

The number of connected components c of G is the dimension of the nullspace of $\mathbf{L}$. Furthermore the null space of $\mathbf{L}$ has a basis of indicator vectors of the connected components of G

Indicator of a subset H of V is

$$x \in \mathbf{R}^N \text{ s.t. } \begin{cases} x[i] = 1 & \text{if } \in H \\ x[i] = 0 & \text{otherwise} \end{cases}$$

Normalized Graph Laplacian

Note: we will sometimes need to consider the generalised problem

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

In this case it makes sense to introduce the normalised Laplacian

$$\mathbf{L}_{\text{norm}} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$$

Eigenvectors are closely related

$$\mathbf{L}_{\text{norm}} f = \lambda f \rightarrow u = D^{-1/2} f$$

Normalized Graph Laplacian

Eigenvalues of the normalised Laplacian

$$0 = \lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_N \leqslant 2$$

Algebraic connectivity

IFF bipartite graph!

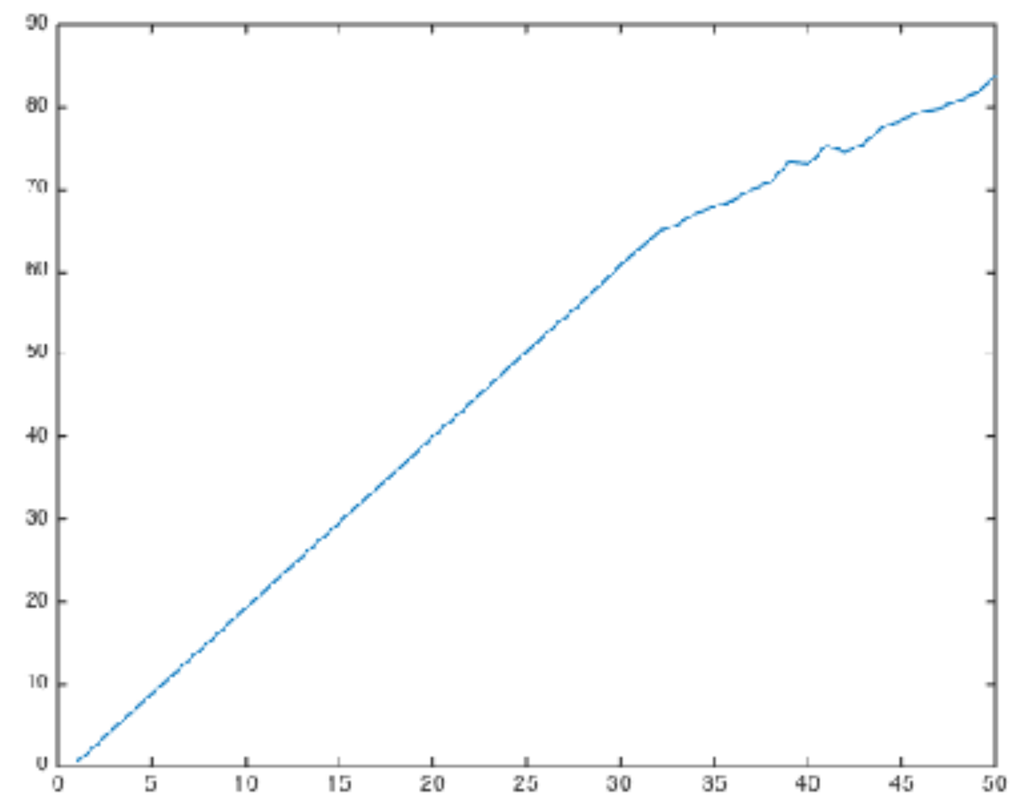
Algebraic Connectivity, Fiedler Vector

Multiplicity of eigenvalue 0 gives connectedness of graph

What if $\lambda_2 > 0$?

Experiment:

Gradually increase connections between two Erdos-Renyi subgraphs



$$\lambda_2 \geqslant \frac{1}{\mathrm{vol}(G)d(G)}$$ where $d(G)$ is the diameter of the graph

The Cheeger Constant

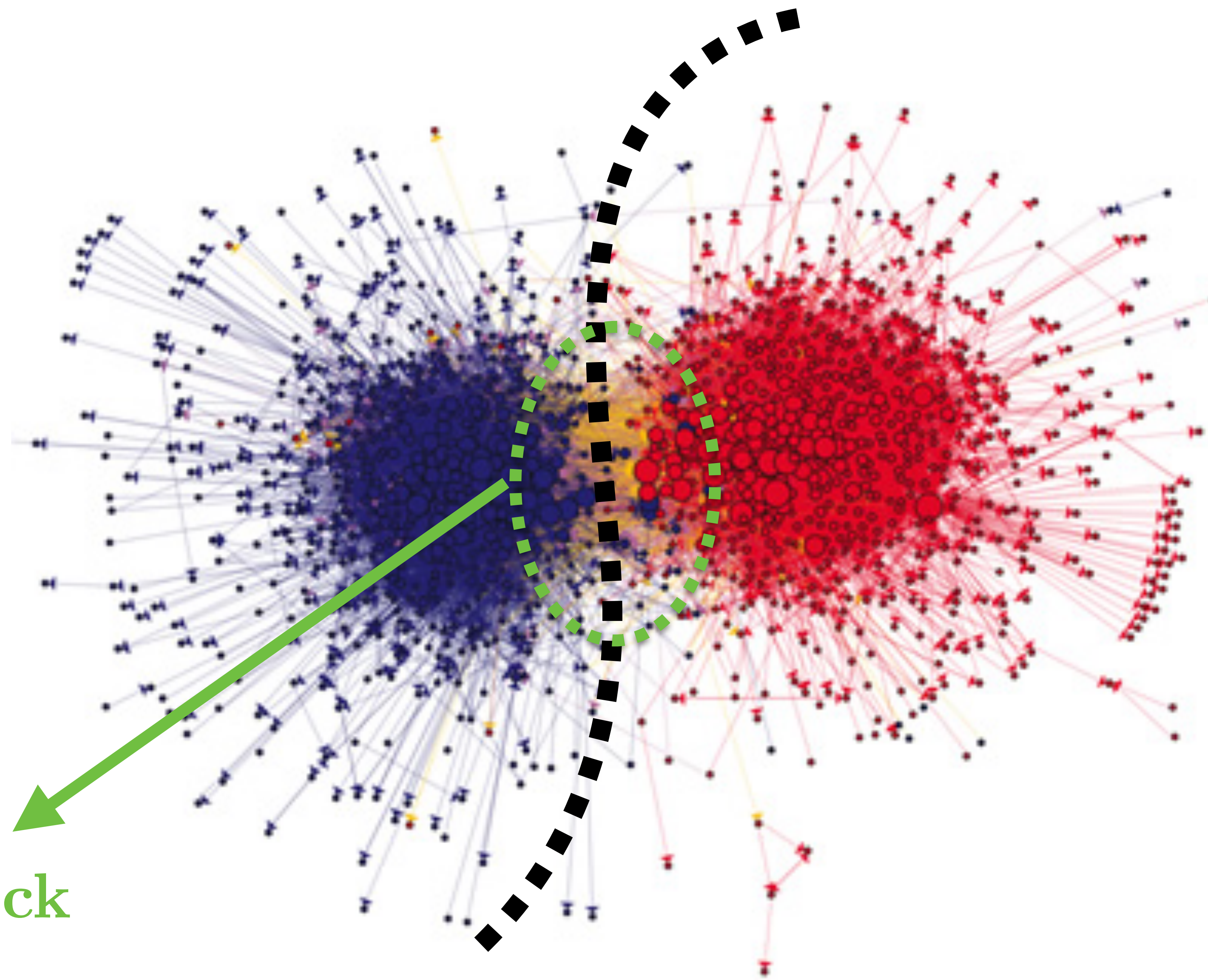
Cheeger constant measures presence of a "bottleneck"

$$A \subset V \qquad \partial A = \left\{(u, v) \in E \text{ s.t. } u \in A,\, v \in \overline{A}\right\}$$

$$\mathrm{vol}(A) = \sum_{u \in A} d(u)$$

$$h(G) = \min_{A \subset V} \left\{ \frac{|\partial A|}{\min\left(\mathrm{vol}(A), \mathrm{vol}(\overline{A})\right)} \text{ s.t. } 0 < |A| < \frac{1}{2}|V| \right\}$$

The Cheeger Constant



Bottleneck

A Cheeger Inequality

The Cheeger constant and algebraic connectivity are related by Cheeger inequalities. A simple example:

Theorem: Cheeger Inequality[Polya, Szego]

For a general graph G,

$$2h(G) \geqslant \lambda_2 \geqslant \frac{h^2(G)}{2}$$

Remark: the eigenvector associated to the algebraic connectivity is called the Fiedler vector

Algebraic Connectivity, Fiedler Vector

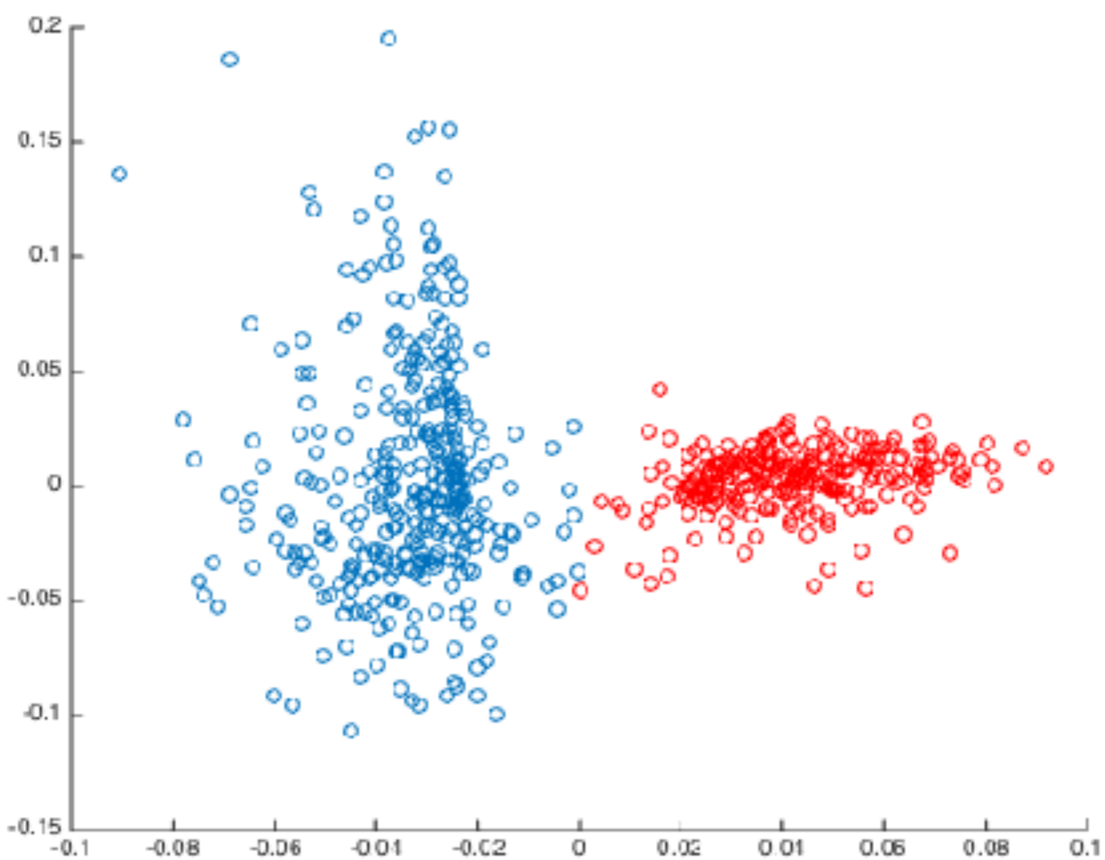
Set of 1490 US political blogs, labelled “Dem” or “Rep”

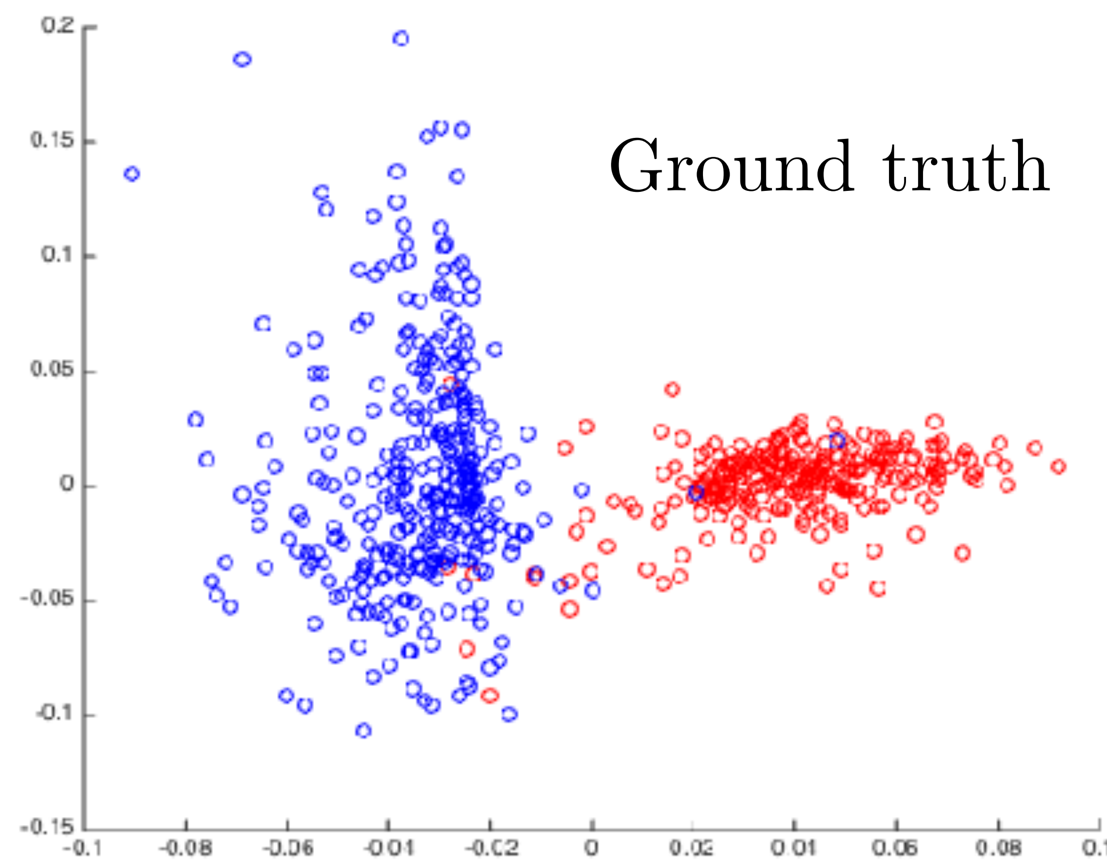
Hyperlinks among blogs

Removed small degrees (<12), keep N = 622 vertices

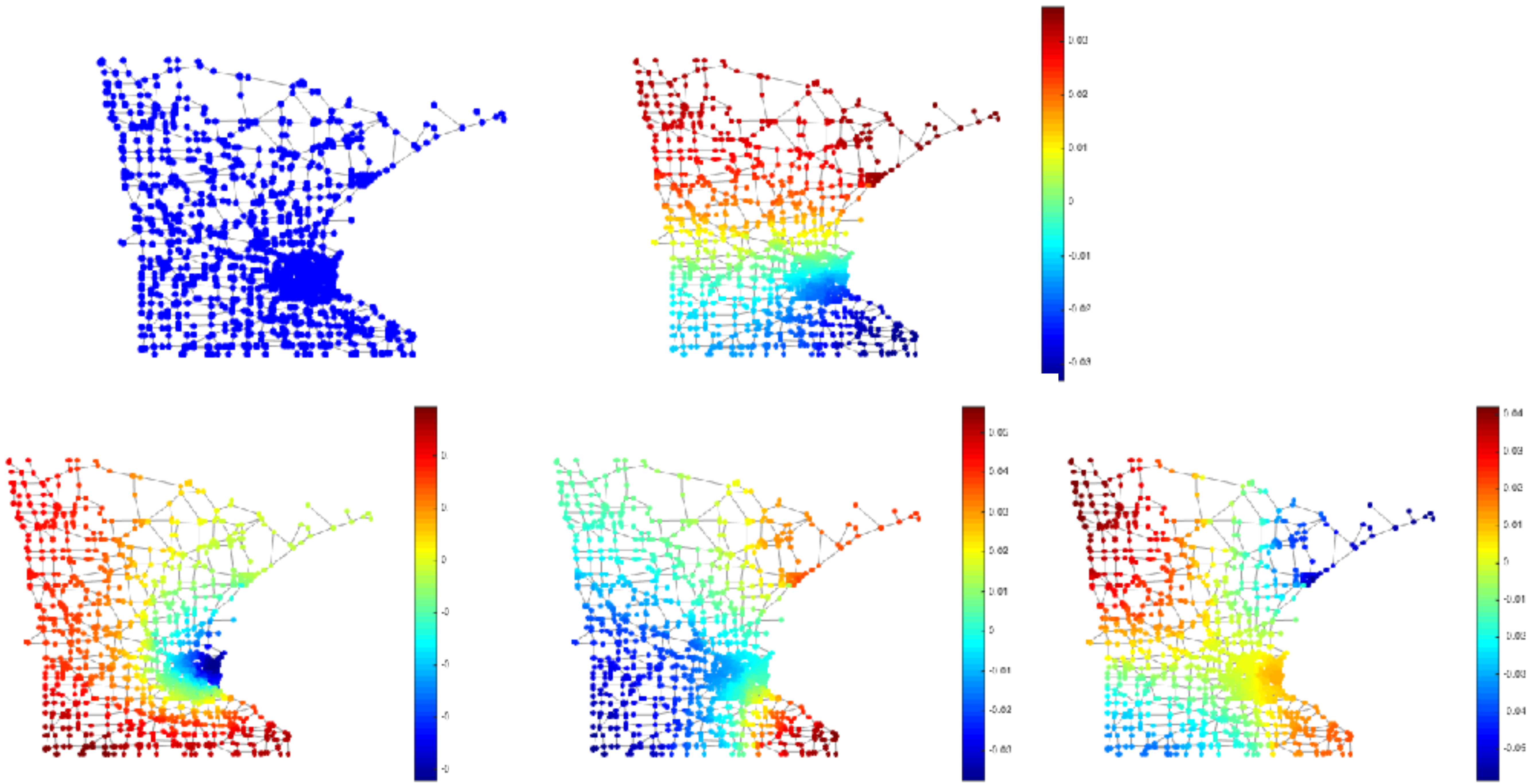
Compute normalised Laplacian, Fiedler vector

Assign attributes +1, -1 by sign of Fiedler vector





A Few Laplacian Eigenvectors



LTS EPFL

LTS2

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE