Graph Signal Processing: Foundational Advances For Learning From Network Data

Gonzalo Mateos

Dept. of ECE and Goergen Institute for Data Science
University of Rochester
gmateosb@ece.rochester.edu
http://www.ece.rochester.edu/~gmateosb/

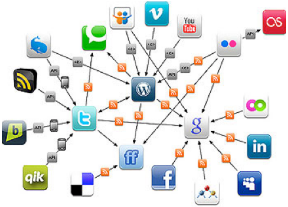
Collaborators: R. Shafipour, S. Segarra, A. G. Marques, and A. Ribeiro

Acknowledgment: NSF Awards CCF-1750428 and ECCS-1809356

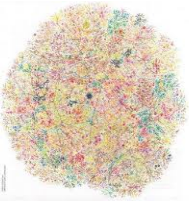
Rochester, NY, October 4, 2019

Network Science analytics

Online social media



Internet



Clean energy and grid analytics

- Network as graph $G = (\mathcal{V}, \mathcal{E})$: encode pairwise relationships
- Desiderata: Process, analyze and learn from network data [Kolaczyk'09]

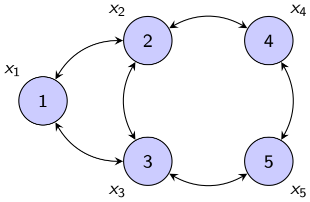
  $\Rightarrow$ Use G to study graph signals, data associated with nodes in $\mathcal{V}$
- Ex: Opinion profile, buffer congestion levels, neural activity, epidemic

Graph signal processing (GSP)

- Graph G with adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$

  $\Rightarrow A_{ij} =$ proximity between i and j
- Define a signal $\mathbf{x} \in \mathbb{R}^{N}$ on top of the graph

  $\Rightarrow x_i =$ signal value at node i

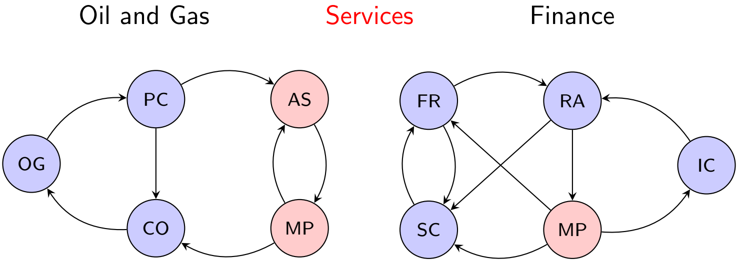


- Graph Signal Processing $\rightarrow$ Exploit structure encoded in $\mathbf{A}$ to process $\mathbf{x}$

  $\Rightarrow$ Our view: GSP well suited to study (network) diffusion processes
- **Q:** Graph signals common and interesting as networks are?
- **Q:** Why do we expect the graph structure to be useful in processing $\mathbf{x}$?

Network of economic sectors of the United States

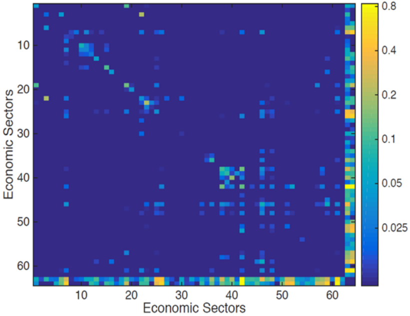
- Bureau of Economic Analysis of the U.S. Department of Commerce
  - A_{ij} = Output of sector i that becomes input to sector j (62 sectors)



- Oil extraction (OG), Petroleum and coal products (PC), Construction (CO)
- Administrative services (AS), Professional services (MP)
- Credit intermediation (FR), Securities (SC), Real state (RA), Insurance (IC)
- Only interactions stronger than a threshold are shown

Network of economic sectors of the United States

- Bureau of Economic Analysis of the U.S. Department of Commerce
  - A_{ij} = Output of sector i that becomes input to sector j (62 sectors)

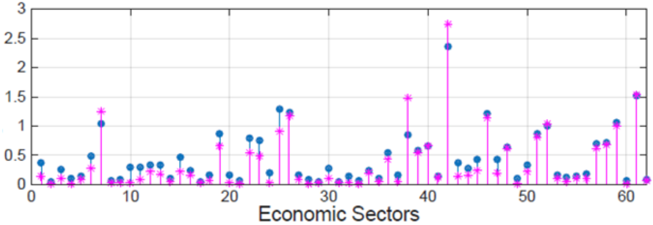


- A few sectors have widespread strong influence (services, finance, energy)
- Some sectors have strong indirect influences (oil)
- The heavy last row is final consumption

- This is an interesting network $\Rightarrow$ Signals on this graph are as well

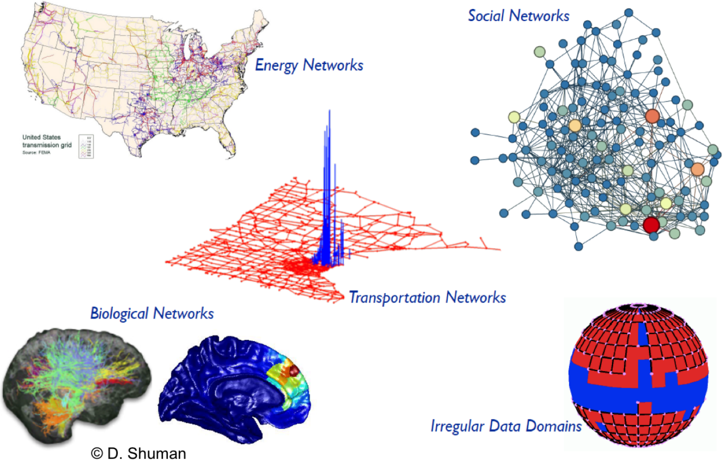
Disaggregated GDP of the United States

- Signal $\mathbf{x}$ = output per sector = disaggregated GDP
  - $\Rightarrow$ Network structure used to, e.g., reduce GDP estimation noise



- Signal is as interesting as the network itself. Arguably more
  - Same is true for brain connectivity and fMRI brain signals, ...
  - Gene regulatory networks and gene expression levels, ...
  - Online social networks and information cascades, ...

Graph signals are ubiquitous

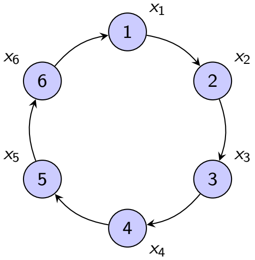


Importance of signal structure in time

- Signal and Information Processing is about exploiting signal structure

- Discrete time described by cyclic graph
  - $\Rightarrow$ Time n follows time $n-1$
  - $\Rightarrow$ Signal value x_n similar to x_{n-1}

- Formalized with the notion of frequency

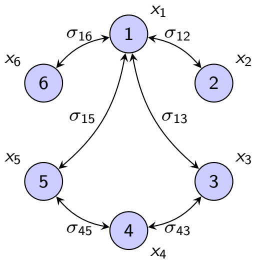


- Cyclic structure $\Rightarrow$ Fourier transform $\Rightarrow \tilde{\mathbf{x}} = \mathbf{F}^H \mathbf{x}$ $\left(F_{kn} = \frac{e^{j2\pi kn/N}}{\sqrt{N}} \right)$

- **Fourier transform** $\Rightarrow$ Projection on eigenvector space of cycle

Covariances and principal components

- Random signal with mean $\mathbb{E}[\mathbf{x}] = 0$ and covariance $\mathbf{C}_x = \mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right]$
  - $\Rightarrow$ Eigenvector decomposition $\mathbf{C}_x = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^H$
- Covariance matrix $\mathbf{A} = \mathbf{C}_x$ is a graph
  - $\Rightarrow$ Not a very good graph, but still
- Precision matrix $\mathbf{C}_x^{-1}$ a common graph too
  - $\Rightarrow$ Conditional dependencies of Gaussian $\mathbf{x}$



- Covariance matrix structure $\Rightarrow$ Principal components (PCA) $\Rightarrow \tilde{\mathbf{x}} = \mathbf{V}^H\mathbf{x}$
- **PCA transform** $\Rightarrow$ Projection on eigenvector space of (inverse) covariance
- **Q:** Can we extend these principles to general graphs and signals?

Graph Fourier Transform

- Adjacency $\mathbf{A}$, Laplacian $\mathbf{L}$, or, generically graph shift $\mathbf{S} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$
  $\Rightarrow$ $S_{ij} = 0$ for $i \neq j$ and $(i,j) \notin \mathcal{E}$ (captures local structure in G)
- The Graph Fourier Transform (GFT) of $\mathbf{x}$ is defined as

$$\tilde{\mathbf{x}} = \mathbf{V}^{-1}\mathbf{x}$$

- While the inverse GFT (iGFT) of $\tilde{\mathbf{x}}$ is defined as

$$\mathbf{x} = \mathbf{V}\tilde{\mathbf{x}}$$

  $\Rightarrow$ Eigenvectors $\mathbf{V} = [\mathbf{v}_1, ..., \mathbf{v}_N]$ are the frequency basis (atoms)

- Additional structure
  $\Rightarrow$ If $\mathbf{S}$ is normal, then $\mathbf{V}^{-1} = \mathbf{V}^H$ and $\tilde{x}_k = \mathbf{v}_k^H\mathbf{x} = <\mathbf{v}_k, \mathbf{x}>$
  $\Rightarrow$ Parseval holds, $\|\mathbf{x}\|^2 = \|\tilde{\mathbf{x}}\|^2$
- **GFT** $\Rightarrow$ Projection on eigenvector space of graph shift operator $\mathbf{S}$

Frequency modes of the Laplacian

- Total variation of signal $\mathbf{x}$ with respect to $\mathbf{L}$

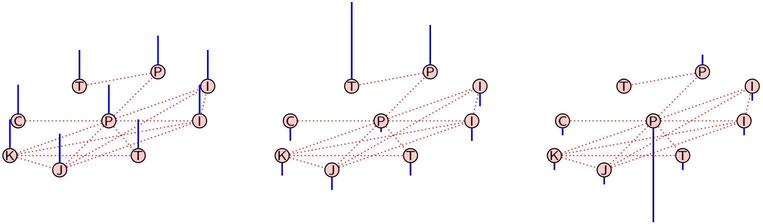
$$\text{TV}(\mathbf{x}) = \mathbf{x}^T\mathbf{L}\mathbf{x} = \sum_{i,j=1,j>i}^{N} A_{ij}(x_i - x_j)^2$$

$\Rightarrow$ Smoothness measure on the graph G (Dirichlet energy)

- For Laplacian eigenvectors $\mathbf{V} = [\mathbf{v}_1, \cdots, \mathbf{v}_N]$ $\Rightarrow$ $\text{TV}(\mathbf{v}_k) = \lambda_k$

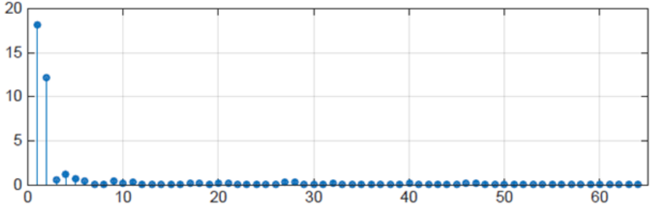
$\Rightarrow$ Can view $0 = \lambda_1 < \cdots \leq \lambda_N$ as frequencies

- Ex: gene network, $N=10$, $k=1$, $k=2$, $k=9$



Is this a reasonable transform?

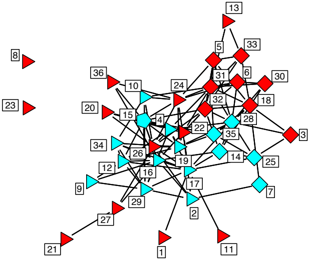
- Particularized to cyclic graphs $\Rightarrow$ GFT $\equiv$ Fourier transform
- Also for covariance graphs $\Rightarrow$ GFT $\equiv$ PCA transform
- But really, this is an empirical question. GFT of disaggregated GDP



- Spectral domain representation characterized by a few coefficients
  - $\Rightarrow$ Notion of bandlimitedness: $\mathbf{x} = \sum_{k=1}^{K} \tilde{x}_k \mathbf{v}_k$
  - $\Rightarrow$ Sampling, compression, filtering, pattern recognition

Predicting law practice

- Working relationships among lawyers [Lazega'01]
  - Graph: 36 partners, edges indicate partners worked together



- Signal: various node-level attributes $\mathbf{x} = \{x_i\}_{i \in \mathcal{V}}$ including

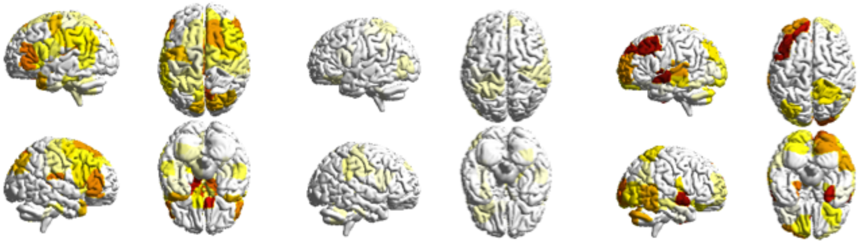
  $\Rightarrow$ Type of practice, i.e., litigation (red) and corporate (cyan)

- Suspect lawyers collaborate more with peers in same legal practice

  $\Rightarrow$ Knowledge of collaboration useful in predicting type of practice

Graph frequency analysis of brain signals

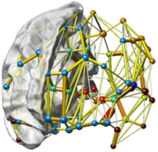
- GFT of brain signals during a visual-motor learning task [Huang et al'16]
  - $\Rightarrow$ Decomposed into low, medium and high frequency components



- Brain: Complex system where regularity coexists with disorder [Sporns'11]
  - $\Rightarrow$ Signal energy mostly in the low and high frequencies
  - $\Rightarrow$ In brain regions akin to the visual and sensorimotor cortices

Learning graphs from data

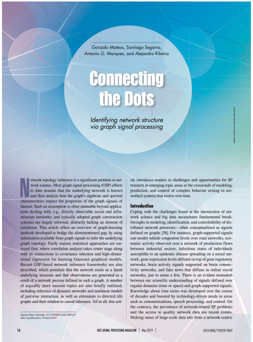
- Learning graphs from nodal observations
- Key in neuroscience
  ⇒ Functional network from fMRI signals

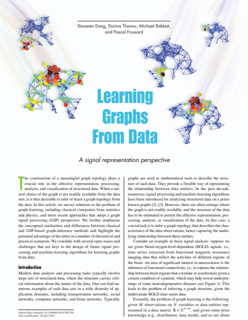


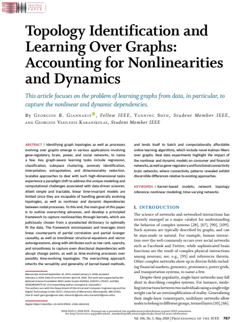
- Most GSP works: how known graph **S** affects signals and filters
- Here, reverse path: how to use GSP to infer the graph topology?
  - Gaussian graphical models [Egilmez et al'16], [Rabbat'17], ...
  - Smooth signals [Dong et al'15], [Kalofolias'16], [Sardellitti et al'17], ...
  - Graph filtering models [Shafipour et al'17], [Thanou et al'17], ...
  - Stationary signals [Pasdeloup et al'15], [Segarra et al'16], ...
  - Directed graphs [Mei-Moura'15], [Shen et al'16], ...

Connecting the dots

- Recent tutorials on learning graphs from data
  - IEEE Signal Processing Magazine and Proceedings of the IEEE







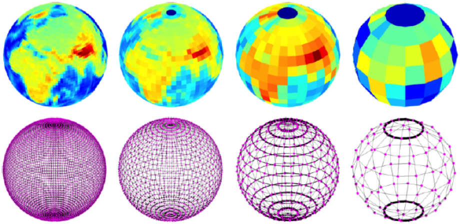
- IEEE Trans. on Signal and Information Processing over Networks
  - Forthcoming issue on Network Topology Inference (Jan. 2020)

Concluding remarks

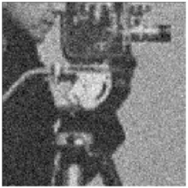
- Network science and big data pose new challenges
  - ⇒ GSP can contribute to address some of those challenges
  - ⇒ Well suited for network (diffusion) processes
- GSP pillars: graph-shift operator, filters and Fourier transform
- GSP tools can be applied to solve practical problems
  - ⇒ Signal representation and compression
  - ⇒ Sampling, interpolation (network control)
  - ⇒ Source localization on graphs (fake news, epileptic seizures)
  - ⇒ Network topology inference
  - ⇒ Geometric deep learning and graph CNNs

Application domains

Visualization / Compression

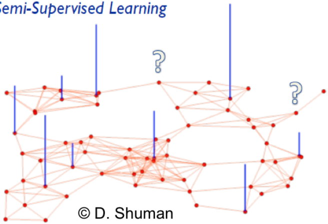


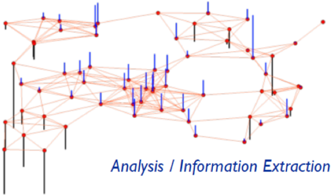
Earth data source: Frederik Simons



Denoising

Semi-Supervised Learning





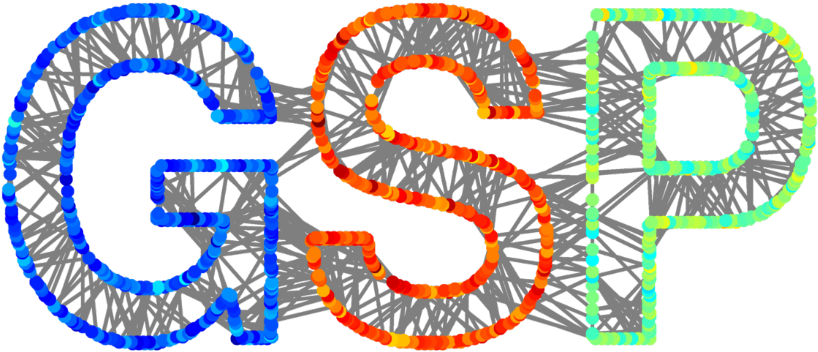
Analysis / Information Extraction

Envisioned application domains

- Gene regulatory and protein interaction networks
- Online social media
- Smart infrastructure networks, IoT
- Economics, finance, social sciences
- Neuroimaging data analysis
  - $\Rightarrow$ Extensive literature on brain network analysis
  - $\Rightarrow$ Classifying neural disorders, predicting learning ability
  - $\Rightarrow$ Analyzed networks because they could not study signals
  - $\Rightarrow$ GSP: Integration of structural and functional perspectives

See also arXiv:1710.01135v3 [eess.IV]

PyGSP: Graph Signal Processing in Python



- PyGSP is a Python package to ease SP on graphs. Free software

Available from https://github.com/epfl-lts2/pygsp