TECHNISCHE UNIVERSITÄT WIEN

institute of telecommunications

LEARNING SIGNED GRAPHS FROM DATA

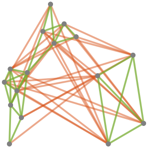
Gerald Matz and Thomas Dittrich

Institute of Telecommunications
TU Wien (Austria)

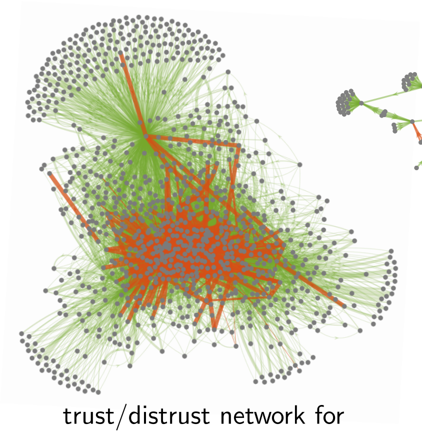
ICASSP 2020 May 5, 2020

Motivation

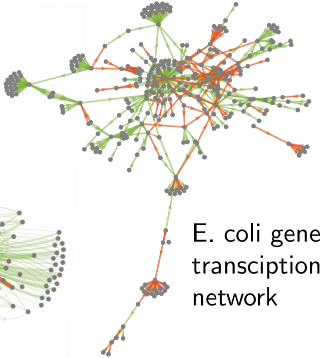
Signed graphs have positive and negative edge weights to model similarity/dissimilarity, like/dislike, trust/distrust, friend/foe, activation/inhibition



friend-foe network for 16 Guinea highland tribes



trust/distrust network for 1000 users on Bitcoin Alpha



E. coli gene transciption network

Problem Statement

Given: N objects, M measurements $\rightarrow N \times M$ data matrix

$$\mathbf{X} = \begin{pmatrix} \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_N^T \end{pmatrix} = \left(\tilde{\mathbf{x}}_1 \ldots \tilde{\mathbf{x}}_M\right) \qquad \begin{cases} N \text{ measurement vectors } \mathbf{x}_i \text{ (rows)} \\ M \text{ graph signals } \tilde{\mathbf{x}}_m \text{ (columns)} \end{cases}$$

Signed graph learning ➡ find edge weight matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$

$$\begin{cases} W_{ij} = 0, & \text{no edge between nodes } i \text{ and } j \\ W_{ij} > 0, & \text{similarity edge between nodes } i \text{ and } j \\ W_{ij} < 0, & \text{dissimilarity edge between nodes } i \text{ and } j \end{cases}$$

Rationale:

- graph signals $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_M$ shall be smooth on the graph
- $+/-$ edge between nodes i and j if $\mathbf{x}_i$ and $\mathbf{x}_j$ are similar/dissimilar

NB: unsigned graph learning ($W_{ij} \geq 0$) has received a lot of attention!

Smoothness/(Dis)similarity Metrics

Signed Laplacian form:

$$\mathbf{x}^T\mathbf{L}\mathbf{x} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}|W_{ij}| \cdot \big(x_i - \mathsf{sign}(W_{ij})x_j\big)^2$$

with signed Laplacian

$$\mathbf{L} = \mathsf{Diag}\{d_1, \ldots, d_N\} - \mathbf{W}, \qquad d_i = \sum_{j=1}^{N}|W_{ij}|$$

Signed total variation:

$$\|\mathbf{x}\|_{\mathsf{TV}} = \sum_{i=1}^{N}\sum_{j=1}^{N}|W_{ij}| \cdot |x_i - \mathsf{sign}(W_{ij})x_j|$$

Both metrics are small if $x_i \approx x_j$ for $W_{ij} > 0$ and $x_i \approx -x_j$ for $W_{ij} < 0$

Optimization Problem

Unified formulation of cost function:

$$f_p(\mathbf{W}) = \sum_{i=1}^{N} \sum_{j=1}^{N} |W_{ij}| \cdot \|\mathbf{x}_i - \mathsf{sign}(W_{ij})\mathbf{x}_j\|_p^p$$

Optimization problem:

$$\min_{\mathbf{W} \in \mathcal{W}} \; f_p(\mathbf{W}) + \frac{\mu}{2}\|\mathbf{W}\|_2^2$$

$$\mathcal{W} = \left\{ \mathbf{W} : \begin{array}{ll} 0 \leq |W_{ij}| \leq C_{ij}, & \ldots \text{ prior knowledge} \\ a_i \leq \sum_{j=1}^{N} |W_{ij}| \leq b_i, & \ldots \text{ control node degrees} \\ \sum_{j=1}^{N} |W_{ij}| = 2N & \ldots \text{ weight normalization} \end{array} \right\}$$

This problem is not convex ...

Solution

Sign-magnitude representation: $W_{ij} = |W_{ij}| \operatorname{sign}(S_{ij}) = \widetilde{W}_{ij} S_{ij}$

$$\Rightarrow \min_{\substack{\widetilde{W}_{ij} \in \mathcal{W} \\ S_{ij} \in \{-1,1\}}} \sum_{i=1}^{N} \sum_{j=1}^{N} \widetilde{W}_{ij} \|\mathbf{x}_i - S_{ij}\mathbf{x}_j\|_p^p + \frac{\mu}{2} \widetilde{W}_{ij}^2$$

Sign optimization has explicit solution:

$$\min_{S_{ij} \in \{-1,1\}} \|\mathbf{x}_i - S_{ij}\mathbf{x}_j\|_p^p = D_{ij}^{(p)} \triangleq \min \Big\{ \underbrace{\|\mathbf{x}_i - \mathbf{x}_j\|_p^p}_{S_{ij}=1}, \underbrace{\|\mathbf{x}_i + \mathbf{x}_j\|_p^p}_{S_{ij}=-1} \Big\}$$

Magnitude optimization amounts to ordinary unsigned graph learning:

$$\min_{\widetilde{\mathbf{W}} \in \mathcal{W}} \operatorname{tr}\{\widetilde{\mathbf{W}} \mathbf{D}^{(p)}\} + \frac{\mu}{2} \|\widetilde{\mathbf{W}}\|_2^2$$

. . . quadratic problem, can be solved efficiently e.g. using ADMM

Illustrative Example

$N = 3$, $\mathbf{x}^T = (1, -2, 4)$, $p = 1$, no weight and degree constraints

$|x_1 - x_2| = 3$

$|x_1 + x_2| = 1$

$S_{12} = -1,\ D_{12}^{(1)} = 1$

$|x_1 - x_3| = 3$

$|x_1 + x_3| = 5$

$S_{13} = 1,\ D_{13}^{(1)} = 3$

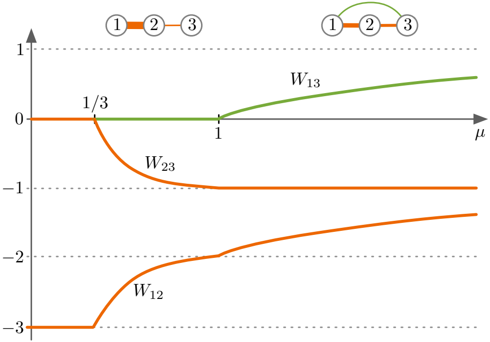
$|x_2 - x_3| = 6$

$|x_2 + x_3| = 2$

$S_{23} = -1,\ D_{23}^{(1)} = 2$

Waterfilling solution:

$$W_{ij} = \frac{S_{ij}}{\mu} \max\left\{0, \omega - D_{ij}^{(1)}\right\}$$



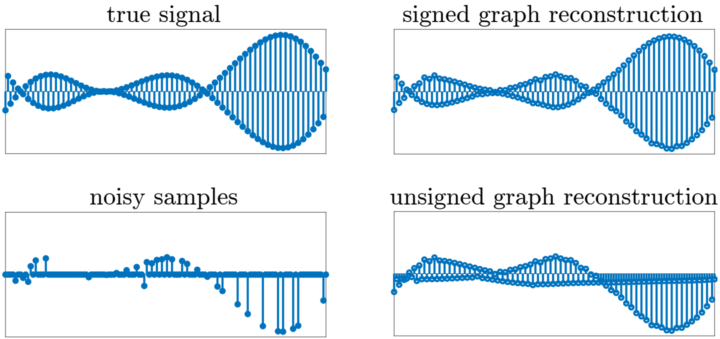
Application 1: Signal Reconstruction

Given: K noisy samples $y_i = s_i + z_i$, $i \in \mathcal{S}$

Goal: reconstruct bandpass signal $\mathbf{s}$ of length $N = 128$

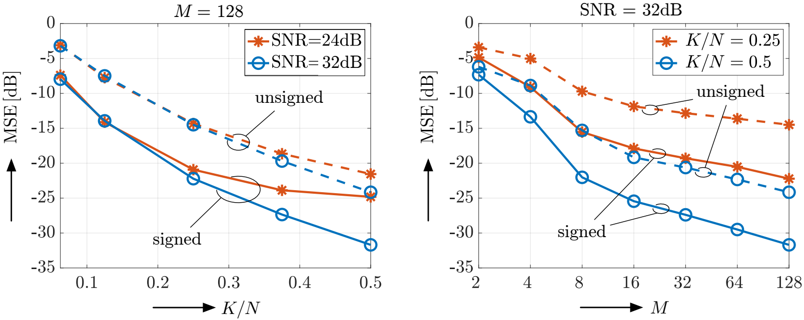
Laplacian regression: $\mathbf{s}^{\text{opt}} = \arg\min_{\mathbf{s}} \mathbf{s}^T \mathbf{L} \mathbf{s} + \nu \sum_{i \in \mathcal{S}} (y_i - s_i)^2$

Signed/unsigned graph learned from M examples $\tilde{\mathbf{x}}_m = \mathbf{s}_m + \mathbf{z}_m$ ($p = 2$)



Reconstruction Performance

Reconstruction MSE obtained via 1000 Monte Carlo trials



- signed outperforms unsigned by typically 7 dB
- unsigned requires twice as many samples as signed

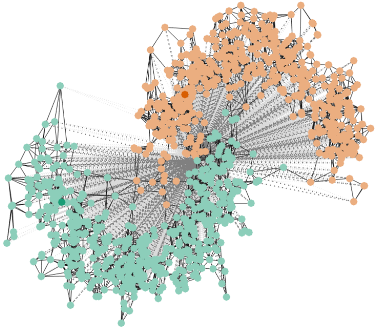
Application 2: Clustering

Given: $N = 1000$ points $\mathbf{x}_i$ of the two-moon model ($M = 2$)

Goal: Find cluster labels $\{-1, 1\}^N$

Total variation clustering: $\mathbf{y}^{\mathsf{opt}} = \underset{\mathbf{y}\in[-1,1]^N}{\arg\min} \|\mathbf{y}\|_{\mathsf{TV}}$

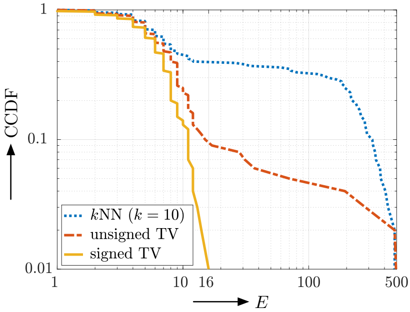
Signed/unsigned graph learned from $\mathbf{x}_i$ with $p = 1$



signed clustering unsigned clustering

Clustering Performance

Misclassification rate E obtained from 100 Monte Carlo runs



- signed clustering: average $E = 6.42 \pm 3.01$, worst-case $E = 16$
- unsigned clustering: average $E = 22.64 \pm 74.92$, worst-case $E = 500$
- kNN clustering: average $E = 93.75 \pm 133.4$, worst-case $E = 500$

Conclusion

- Signed graphs offer increased modeling flexibility
- Proposed learning strategy for signed graphs
  - based on notions of (dis)similarity and smoothness
  - non-convex problem, can be reduced to unsigned learning
  - efficient implementation
- Application examples
  - signal sampling and reconstruction
  - data clustering
- at same graph sparsity, signed outperforms unsigned
- Future work: distinct penalty and constraints for $+/-$ edges

THANKS FOR WATCHING!

Get in touch with me!

Email: gerald.matz@tuwien.ac.at

Skype: gerald_matz

#staysafe