Signed graphs: theory and applications

Bruno Ordozgoiti[1]

[1] Dept. of Computer Science - Aalto University

BYMAT 2020

Some theory

Signed graphs

Signed graphs: each edge labeled $+$ or $-$.

Definitions:

- $G = (V, E^+, E^-)$,
- $G = (V, E, \sigma), \sigma : E \to \{-, +\}$.

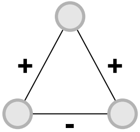
Signed graphs

Signed graphs: each edge labeled $+$ or $-$.

Definitions:

- $G = (V, E^+, E^-)$,
- $G = (V, E, \sigma), \sigma : E \to \{-, +\}$.

Adjacency matrix: $A = A_{E^+} - A_{E^-}$



$$\mapsto \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix}$$

Densest subgraph problem in unsigned graphs:

$$\max_{S \subseteq V} \frac{2e(S)}{|S|} = \max_{x \in \{0,1\}^n} \frac{x^T A x}{x^T x}.$$

Polynomial-time solvable (Goldberg, 1984).

Densest subgraph problem in unsigned graphs:

$$\max_{S \subseteq V} \frac{2e(S)}{|S|} = \max_{x \in \{0,1\}^n} \frac{x^T A x}{x^T x}.$$

Polynomial-time solvable (Goldberg, 1984).

Densest subgraph problem in signed graphs:

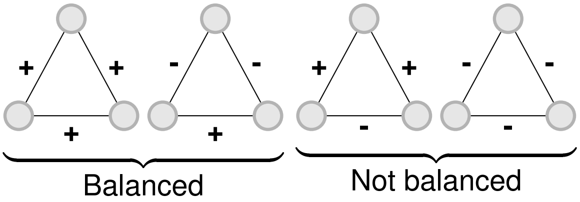
$$\max_{x \in \{-1,0,1\}^n} \frac{x^T A x}{x^T x}.$$

NP-hard ! (Bonchi et al., 2019; Tsourakakis et al., 2019).

Motivation

Motivation: **balance** in social networks (Harary, 1953).

"*The friend of a friend is a friend*" (or "*the enemy of a friend is an enemy*").



The four possible non-isomorphic signed triangles.

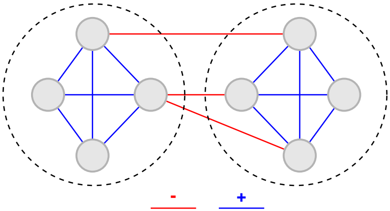
Motivation

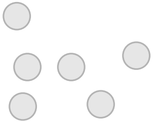
Characterizations of balance

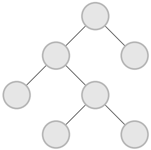
G is balanced iff

- It contains no negative (unbalanced) cycles.

Some balanced graphs







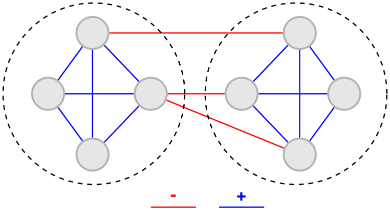
Motivation

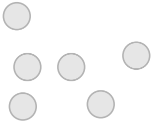
Characterizations of balance

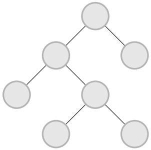
G is balanced iff

- It contains no negative (unbalanced) cycles.
- There exists a sign-compliant partition of G: $V = V_1 \cup V_2$, $V_1 \cap V_2 = \varnothing$, all **+** edges within sets, all **-** edges between sets.

Some balanced graphs







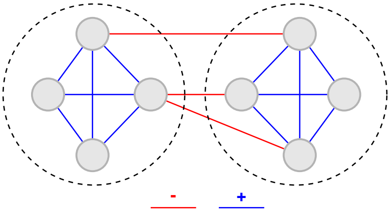
Motivation

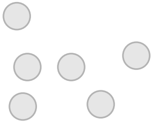
Characterizations of balance

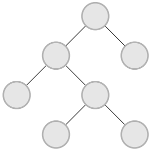
G is balanced iff

- It contains no negative (unbalanced) cycles.
- There exists a sign-compliant partition of G: $V = V_1 \cup V_2$, $V_1 \cap V_2 = \varnothing$, all **+** edges within sets, all **-** edges between sets.
- All paths between any pair u, v have the same sign.

Some balanced graphs



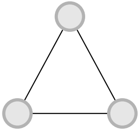




Spectral theory

Review of unsigned spectral theory:

Laplacian: $L = D - A$

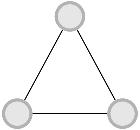


$$L = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

Spectral theory

Review of unsigned spectral theory:

Laplacian: $L = D - A$



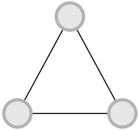
$$Lv_1 = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0$$

- $\lambda_{min}(L) = 0$ (Multiplicity of 0 = n. of connected components)

Spectral theory

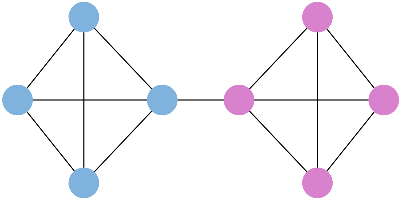
Review of unsigned spectral theory:

Laplacian: $L = D - A$



$$Lv_1 = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0$$

- $\lambda_{min}(L) = 0$ (Multiplicity of 0 = n. of connected components)
- Eigenvector v_2 gives a "good" partition (Cheeger inequality).



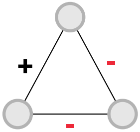
$$v_2 \approx \begin{pmatrix} -0.25 \\ -0.38 \\ -0.38 \\ -0.38 \\ 0.38 \\ 0.38 \\ 0.38 \\ 0.25 \end{pmatrix}, \lambda_2(L) \approx 0.35.$$

Spectral theory

Signed spectral theory:

Laplacian: $L = D - A$

Unsigned	Signed
L is positive semidefinite	
$D_{ii} = \sum_j A_{ij}$	$D_{ii} = \sum_j \lvert A_{ij} \rvert$
$\lambda_{min}(L) = 0$	$\lambda_{min}(L) \geq 0$



$$Lv_1 = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = 0$$

Spectral characterizations of balance

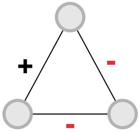
- Connected and $\lambda_{min} = 0$ (or one zero-eigval per connected component).

Spectral theory

Signed spectral theory:

Laplacian: $L = D - A$

Unsigned	Signed
L is positive semidefinite	
$D_{ii} = \sum_j A_{ij}$	$D_{ii} = \sum_j \lvert A_{ij} \rvert$
$\lambda_{min}(L) = 0$	$\lambda_{min}(L) \geq 0$



$$Lv_1 = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = 0$$

Spectral characterizations of balance

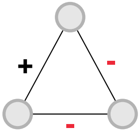
- Connected and $\lambda_{min} = 0$ (or one zero-eigval per connected component).
- Spectrum of A = spectrum of $\bar{A}$ (underlying graph).

Spectral theory

Signed spectral theory:

Laplacian: $L = D - A$

Unsigned	Signed
L is positive semidefinite	
$D_{ii} = \sum_j A_{ij}$	$D_{ii} = \sum_j \lvert A_{ij} \rvert$
$\lambda_{min}(L) = 0$	$\lambda_{min}(L) \geq 0$



$$Lv_1 = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = 0$$

Spectral characterizations of balance

- Connected and $\lambda_{min} = 0$ (or one zero-eigval per connected component).
- Spectrum of A = spectrum of $\bar{A}$ (underlying graph).
- A switches to $\bar{A}$ (switching preserves the spectrum).

Some recent results

Signed densest subgraph

(Bonchi, Galimberti, Gionis, Ordozgoiti, and Ruffo, 2019)

$\max_{x \in \{-1,0,1\}^n} \frac{x^T A x}{x^T x}$. (**NP**-hard)

Signed densest subgraph

(Bonchi, Galimberti, Gionis, Ordozgoiti, and Ruffo, 2019)

$\max_{x \in \{-1,0,1\}^n} \frac{x^T A x}{x^T x}$. (**NP**-hard)

Randomized algorithm:

Input: $n \times n$ adjacency matrix A

1: Compute v, leading eigenvector of A.
2: Set $x_i = sgn(v_i)$ with probability $|v_i|$, $x_i = 0$ w.p. $1 - |v_i|$.
3: Output x.

Signed densest subgraph

(Bonchi, Galimberti, Gionis, Ordozgoiti, and Ruffo, 2019)

$\max_{x \in \{-1,0,1\}^n} \frac{x^T A x}{x^T x}$. (**NP**-hard)

Randomized algorithm:

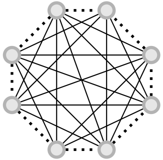
Input: $n \times n$ adjacency matrix A

1: Compute v, leading eigenvector of A.
2: Set $x_i = sgn(v_i)$ with probability $|v_i|$, $x_i = 0$ w.p. $1 - |v_i|$.
3: Output x.

Theorem

Our algorithm gives a $\sqrt{n}$-approximation.

$$\mathcal{O}(\sqrt{n})\mathbb{E}\left[\frac{x^T A x}{x^T x}\right] \geq \lambda_{\max} \geq OPT.$$

Tight analysis: 

Signed densest subgraph

(Bonchi, Galimberti, Gionis, Ordozgoiti, and Ruffo, 2019)

$\max_{x\in\{-1,0,1\}^n} \frac{x^T A x}{x^T x}$. (**NP**-hard)

Randomized algorithm:

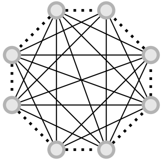
Input: $n \times n$ adjacency matrix A

1: Compute v, leading eigenvector of A.
2: Set $x_i = sgn(v_i)$ with probability $|v_i|$, $x_i = 0$ w.p. $1 - |v_i|$.
3: Output x.

Theorem

Our algorithm gives a $\sqrt{n}$-approximation.

$$\mathcal{O}(\sqrt{n})\mathbb{E}\left[\frac{x^T A x}{x^T x}\right] \geq \lambda_{\max} \geq OPT.$$

Tight analysis: 

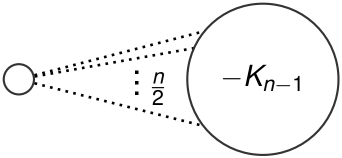
(Bhaskara, Charikar, Manokaran, and Vijayaraghavan, 2012) give an SDP-based $n^{1/3}$-approximation.

Signed densest subgraph

$\mathcal{O}(\sqrt{n})$ is the best possible approximation of $\lambda_{\max}$:

Signed densest subgraph

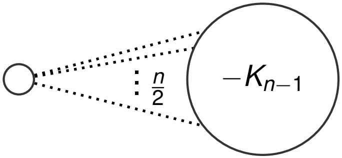
$\mathcal{O}(\sqrt{n})$ is the best possible approximation of $\lambda_{\max}$:



$$OPT = \frac{2e(S^*)}{|S^*|} = \frac{2(n-1)}{n} = \mathcal{O}(1).$$

Signed densest subgraph

$\mathcal{O}(\sqrt{n})$ is the best possible approximation of $\lambda_{\max}$:



$$OPT = \frac{2e(S^*)}{|S^*|} = \frac{2(n-1)}{n} = \mathcal{O}(1).$$

$$\text{Consider } z = \left(\sqrt{\tfrac{n+1}{2n}}, \underbrace{\frac{1}{\sqrt{2n}}, \ldots, \frac{1}{\sqrt{2n}}}_{n/2-1}, \underbrace{\frac{-1}{\sqrt{2n}}, \ldots, \frac{-1}{\sqrt{2n}}}_{n/2} \right).$$

$z^T z = 1$ and $z^T A z = \frac{\sqrt{n+1}+1}{2} - \frac{1}{n} = \Omega(\sqrt{n})$.

Extension to arbitrary number of conflicting groups

(Tzeng, Ordozgoiti, and Gionis, 2020).

$$\max_{S_1,\ldots,S_k} \frac{\sum_{h\in[k]}(|E_+(S_h)| - |E_-(S_h)|) + \frac{1}{k-1}\sum_{h\neq l\in[k]}(|E_-(S_h,S_\ell)| - |E_+(S_h,S_\ell)|)}{|\cup_{h\in[k]} S_h|}.$$

Extension to arbitrary number of conflicting groups

(Tzeng, Ordozgoiti, and Gionis, 2020).

$$\max_{S_1,\ldots,S_k} \frac{\sum_{h\in[k]}(|E_+(S_h)| - |E_-(S_h)|) + \frac{1}{k-1}\sum_{h\neq l\in[k]}(|E_-(S_h,S_\ell)| - |E_+(S_h,S_\ell)|)}{|\cup_{h\in[k]} S_h|}.$$

$= \dfrac{\langle A, XL_kX^T\rangle_F}{k-1}$, where $L_k = kI_k - J_k$ and $X \in \{0,1\}^{n\times k}$ is a node-group indicator matrix.

Extension to arbitrary number of conflicting groups

$\frac{\langle A, XL_k X^T \rangle_F}{k-1}$.

L_k has a $(k-1)$-dimensional invariant subspace. Let $L_k = UDU^T$, $Y = XU$. We choose U to be

$$\begin{array}{lll} (U_{:,1})^T = 1/\sqrt{k}\,[1,\ldots,1], & & (U_{:,2})^T = c_1\,[k-1,-1,\ldots,-1], \\ (U_{:,3})^T = c_2\,[0,k-2,-1,\ldots,-1], & \ldots & (U_{:,k})^T = c_{k-1}\,[0,\ldots,0,1,-1], \end{array} \tag{1}$$

Extension to arbitrary number of conflicting groups

$\frac{\langle A, XL_k X^T \rangle_F}{k-1}$.

L_k has a $(k-1)$-dimensional invariant subspace. Let $L_k = UDU^T$, $Y = XU$. We choose U to be

$$\begin{aligned} (U_{:,1})^T &= 1/\sqrt{k}\,[1,\ldots,1], & (U_{:,2})^T &= c_1\,[k-1,-1,\ldots,-1], \\ (U_{:,3})^T &= c_2\,[0,k-2,-1,\ldots,-1], \quad \ldots & (U_{:,k})^T &= c_{k-1}\,[0,\ldots,0,1,-1], \end{aligned} \tag{1}$$

Equivalent formulation:

$$\max_{Y \in \mathbb{R}^{n\times(k-1)} \setminus \{0\}} \frac{\mathrm{Tr}(Y^T A Y)}{\mathrm{Tr}(Y^T Y)} \quad \text{subject to} \quad Y_{i,j} = \begin{cases} c_j(k-j), & \text{if } i \in S_j \\ 0, & \text{if } i \in \cup_{h=1}^{j-1} S_h \text{ or } i \notin \cup_{h\in[k]} S_h \\ -c_j, & \text{if } i \in \cup_{h=j+1}^{k} S_h \end{cases}.$$

Extension to arbitrary number of conflicting groups

$\frac{\langle A, X L_k X^T \rangle_F}{k-1}$.

L_k has a $(k-1)$-dimensional invariant subspace. Let $L_k = UDU^T$, $Y = XU$. We choose U to be

$$\begin{array}{lll} (U_{:,1})^T = 1/\sqrt{k}\,[1,\ldots,1], & & (U_{:,2})^T = c_1\,[k-1,-1,\ldots,-1], \\ (U_{:,3})^T = c_2\,[0,k-2,-1,\ldots,-1], & \ldots & (U_{:,k})^T = c_{k-1}\,[0,\ldots,0,1,-1], \end{array} \quad (1)$$

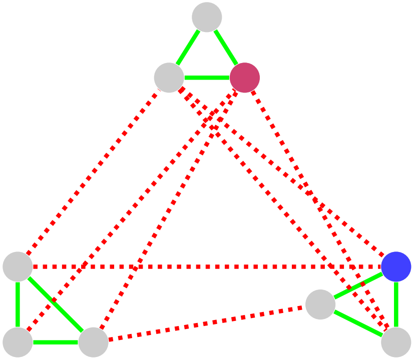
Equivalent formulation:

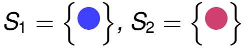
$$\max_{Y \in \mathbb{R}^{n\times(k-1)}\setminus\{0\}} \frac{\mathrm{Tr}(Y^T A Y)}{\mathrm{Tr}(Y^T Y)} \quad \text{subject to} \quad Y_{i,j} = \begin{cases} c_j(k-j), & \text{if } i \in S_j \\ 0, & \text{if } i \in \cup_{h=1}^{j-1} S_h \text{ or } i \notin \cup_{h\in[k]} S_h \\ -c_j, & \text{if } i \in \cup_{h=j+1}^{k} S_h \end{cases}.$$

We give a $\mathcal{O}(k\sqrt{n})$-approximation.

Conflicting groups with queries

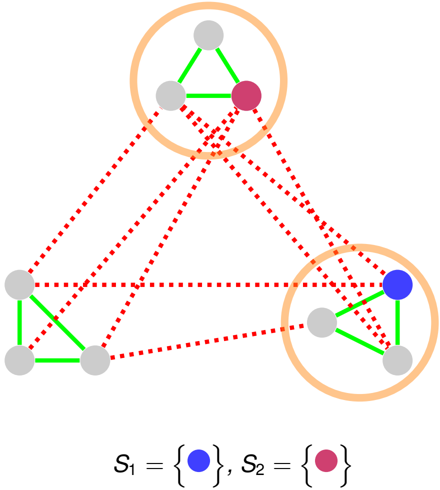
(Xiao, Ordozgoiti, and Gionis, 2020)





Conflicting groups with queries

(Xiao, Ordozgoiti, and Gionis, 2020)



Conflicting groups with queries

For $C_1, C_2 \subseteq V$,

$$\beta(C_1, C_2) = \frac{|E(C_1 \cup C_2, V \setminus (C_1 \cup C_2))| + |E^+(C_1, C_2)| + |E^-(C_1)| + |E^-(C_2)|}{vol(C_1 \cup C_2)}.$$

Conflicting groups with queries

For $C_1, C_2 \subseteq V$,

$$\beta(C_1, C_2) = \frac{|E(C_1 \cup C_2, V \setminus (C_1 \cup C_2))| + |E^+(C_1, C_2)| + |E^-(C_1)| + |E^-(C_2)|}{vol(C_1 \cup C_2)}.$$

minimize $\quad \beta(C_1, C_2)$

s.t $\quad S_1 \subseteq C_1$ and $S_2 \subseteq C_2 \quad$ relevance to seed sets S_1, S_2

$\quad vol(C_1 \cup C_2)/vol(S_1 \cup S_2) \leq k \quad$ control over solution size

Conflicting groups with queries

We rely on the next result:

$$\beta(C_1, C_2) \leq \frac{x^T L x}{x^T D x} \leq 4\beta(C_1, C_2).$$

Combinatorial formulation:

$$\begin{aligned} \min \quad & \beta(C_1, C_2) \\ \text{s.t.} \quad & S_1 \subseteq C_1 \\ & S_2 \subseteq C_2 \\ & vol(C_1 \cup C_2)/vol(S_1 \cup S_2) \leq k \end{aligned}$$

Matrix formulation:

$$\begin{aligned} \min \quad & \frac{x^T L x}{x^T D x} \\ \text{s.t.} \quad & s^T x \geq \kappa \\ & x \in \{-1, 0, 1\}^n \end{aligned}$$

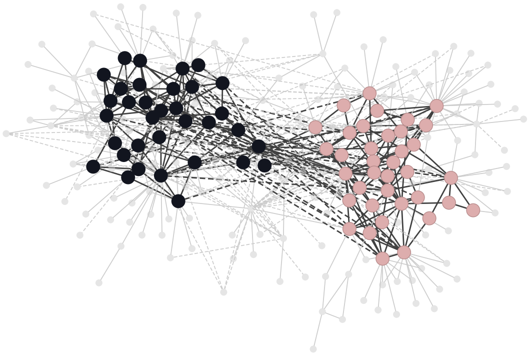
s is an indicator of the seed set.

We give an algorithm guaranteeing $\beta(C_1, C_2) = \mathcal{O}\left(\sqrt{\beta(C_1^*, C_2^*)}\right)$.

Some applications

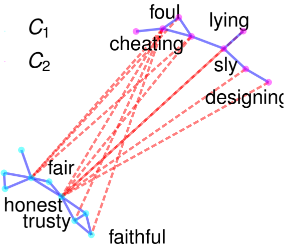
Applications

- Vertices: US Congresspeople.
- Edges: Mentions, favourable and unfavourable.

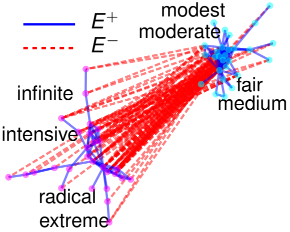


Applications

- Vertices: English words
- Edges: synonym (“happy” vs “joyful”) and antonym (“happy” vs “sad”) relationships.



(a) *fair as without cheating*
$S_1 = \{\textit{fair, honest}\}$
$S_2 = \{\textit{cheating}\}$



(b) *fair as not excessive*
$S_1 = \{\textit{fair, modest}\}$
$S_2 = \{\textit{extreme}\}$

Open problems

- When is $\max_{x \in \{-1,0,1\}^n} \frac{x^T A x}{x^T x}$ easy?
- When is $\max_{x \in \{-1,0,1\}^n} \frac{x^T A x}{x^T x} = \frac{\lambda_{\text{max}}}{\mathcal{O}(1)}$?

Thanks!

References I

Aditya Bhaskara, Moses Charikar, Rajsekar Manokaran, and Aravindan Vijayaraghavan. On quadratic programming with a ratio objective. In *International Colloquium on Automata, Languages, and Programming*, pages 109–120. Springer, 2012.

Francesco Bonchi, Edoardo Galimberti, Aristides Gionis, Bruno Ordozgoiti, and Giancarlo Ruffo. Discovering polarized communities in signed networks. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pages 961–970, 2019.

Andrew V Goldberg. *Finding a maximum density subgraph*. University of California Berkeley, 1984.

P Hansen. Shortest paths in signed graphs. In *North-Holland mathematics studies*, volume 95, pages 201–214. Elsevier, 1984.

References II

Frank Harary. On the notion of balance of a signed graph. *The Michigan Mathematical Journal*, 2(2):143–146, 1953.

Charalampos E Tsourakakis, Tianyi Chen, Naonori Kakimura, and Jakub Pachocki. Novel dense subgraph discovery primitives: Risk aversion and exclusion queries. *arXiv preprint arXiv:1904.08178*, 2019.

Ruo-Chun Tzeng, Bruno Ordozgoiti, and Aristides Gionis. Discovering conflicting groups in signed networks. *Advances in Neural Information Processing Systems*, 33, 2020.

Han Xiao, Bruno Ordozgoiti, and Aristides Gionis. Searching for polarization in signed graphs: a local spectral approach. *arXiv preprint arXiv:2001.09410*, 2020.