Algebraic Graph-theoretic Measures of Conflict

Jérôme Kunegis

Based on work performed in collaboration with Christian Bauckhage, Andreas Lommatzsch, Stephan Spiegel, Jürgen Lerner, Fariba Karimi and Christoph Carl Kling

Journée Graphes et Systèmes Sociaux (JGSS), March 18, 2016, Avignon

WeST
People and Knowledge Networks
Social Network
Algebraic Graph-theoretic Measures of Conflict

Jérôme Kunegis
The Slashdot Zoo

Slashdot Zoo: Tag users as **friends** and **foes**

Graph has **two** types of edges: friendship and enmity
Signed Directed Social Network

You are the fan of your friends and the freak of your foes.

The resulting graph is sparse, square, asymmetric and has signed edge weights.
Wikipedia Edit Wars
Signed Bipartite Networks
### Other Signed Networks

<table>
<thead>
<tr>
<th>Network</th>
<th>Description</th>
<th>Type</th>
<th>Conflicts</th>
<th>Edges</th>
<th>Conflicts</th>
<th>Edges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mv</td>
<td>Dutch college</td>
<td>HumanSocial</td>
<td>D ±</td>
<td>32</td>
<td>3,062</td>
<td></td>
</tr>
<tr>
<td>EP</td>
<td>Epinions trust</td>
<td>Social</td>
<td>D ±</td>
<td>131,828</td>
<td>841,372</td>
<td></td>
</tr>
<tr>
<td>HT</td>
<td>Highland tribes</td>
<td>HumanSocial</td>
<td>U ±</td>
<td>16</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>SZ</td>
<td>Slashdot Zoo</td>
<td>Social</td>
<td>D ±</td>
<td>79,120</td>
<td>515,397</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>Wikipedia conflict</td>
<td>OnlineContact</td>
<td>U ±</td>
<td>118,100</td>
<td>2,917,785</td>
<td></td>
</tr>
<tr>
<td>EL</td>
<td>Wikipedia elections</td>
<td>OnlineContact</td>
<td>D ±</td>
<td>7,118</td>
<td>103,675</td>
<td></td>
</tr>
<tr>
<td>W2</td>
<td>WikiSigned</td>
<td>OnlineContact</td>
<td>D ±</td>
<td>138,592</td>
<td>740,397</td>
<td></td>
</tr>
<tr>
<td>AR</td>
<td>Amazon ratings</td>
<td>Rating</td>
<td>B ±</td>
<td>3,376,972</td>
<td>5,838,041</td>
<td></td>
</tr>
<tr>
<td>Bx</td>
<td>BookCrossing (ratings)</td>
<td>Rating</td>
<td>B ±</td>
<td>263,757</td>
<td>433,652</td>
<td></td>
</tr>
<tr>
<td>ER</td>
<td>Epinions product ratings</td>
<td>Rating</td>
<td>B ±</td>
<td>876,252</td>
<td>13,668,320</td>
<td></td>
</tr>
<tr>
<td>Fr</td>
<td>Filmitips</td>
<td>Rating</td>
<td>B ±</td>
<td>144,671</td>
<td>19,554,219</td>
<td></td>
</tr>
<tr>
<td>J1</td>
<td>Jester 100</td>
<td>Rating</td>
<td>B ±</td>
<td>73,521</td>
<td>4,136,360</td>
<td></td>
</tr>
<tr>
<td>J2</td>
<td>Jester 150</td>
<td>Rating</td>
<td>B ±</td>
<td>50,832</td>
<td>1,728,847</td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>Libimseti.cz</td>
<td>Social</td>
<td>D ±</td>
<td>220,970</td>
<td>17,359,346</td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>MovieLens 100k</td>
<td>Rating</td>
<td>B ±</td>
<td>2,625</td>
<td>100,000</td>
<td></td>
</tr>
<tr>
<td>M3</td>
<td>MovieLens 10M</td>
<td>Rating</td>
<td>B ±</td>
<td>80,555</td>
<td>10,000,054</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>MovieLens 1M</td>
<td>Rating</td>
<td>B ±</td>
<td>9,746</td>
<td>1,000,209</td>
<td></td>
</tr>
<tr>
<td>NX</td>
<td>Netflix</td>
<td>Rating</td>
<td>B ±</td>
<td>497,959</td>
<td>100,480,507</td>
<td></td>
</tr>
<tr>
<td>MO</td>
<td>Residence hall</td>
<td>HumanSocial</td>
<td>D ±</td>
<td>217</td>
<td>2,672</td>
<td></td>
</tr>
<tr>
<td>Ms</td>
<td>Sampson</td>
<td>HumanSocial</td>
<td>D ±</td>
<td>18</td>
<td>188</td>
<td></td>
</tr>
<tr>
<td>YS</td>
<td>Yahoo songs</td>
<td>Rating</td>
<td>B ±</td>
<td>1,625,951</td>
<td>256,804,235</td>
<td></td>
</tr>
<tr>
<td>SX</td>
<td>Sexual escorts</td>
<td>Rating</td>
<td>B ±</td>
<td>16,730</td>
<td>50,632</td>
<td></td>
</tr>
</tbody>
</table>

konect.uni-koblenz.de/networks
Polarization

- Slashdot: 23.9% of edges are negative

Polarization, No conflict
Dyadic Conflict (Reciprocity of Valences)

Balance

Conflict
Measuring Dyadic Conflict

\[ C_2 = \frac{\#\text{conflictDyads}}{\#\text{totalDyads}} \]

\[ C_2 = 0 \]
Tryadic Conflict (Balance Theory)

Balance

Conflict

(Harary 1953)
Measuring Tryadic Conflict

• Definition:

\[ C_3 = \frac{\#\text{negTriangles}}{\#\text{totalTriangles}} \]

\[ C_3 = 0 \]

(cf. Signed relative clustering coefficient, Kunegis et al. 2009)
Balance on Longer Cycles

- Equivalent definitions: a graph is balanced when

  (a) all cycles contain an even number of negative edges
  (b) its nodes can be partitioned into two groups such that all positive edges are within each group, and all negative edges connect the two groups
Frustration

• Definition: The minimum number of edges $f$ that have to be removed from a signed graph to make the graph balanced.

Example: $f = 1$
Frustration (partitioning view)

• Definition: minimum number of edges that are *frustrated* (i.e., inconsistent with balance) given any partition of the graph's nodes into two groups.

Example: $f = 1$
Frustration: Computation

- Computation of frustration is equivalent to MAX-2-XORSAT

\[
\max (a \text{ XOR } b) + (\neg a \text{ XOR } c) + (b \text{ XOR } \neg d) \ldots
\]

- MAX-2-XORSAT is NP complete

- Solution: Relax the problem

see overview in (Facchetti & al. 2011)
Algebraic Formulation

• Let $G = (V, E, \sigma)$ be a signed graph. $\sigma_{uv} = \pm 1$ is the sign of edge $(uv)$.

• Given a partition $V = S \cup T$, let $x$ be the characteristic node-vector:

$$x_u = \begin{cases} 
+1/2 & \text{when } u \in S \\
-1/2 & \text{when } u \in T 
\end{cases}$$

• Number of frustrated edges:

$$\sum_{uv \in E} (x_u - \sigma_{uv} x_v)^2$$
Frustration as Minimization

• $f$ is given by the solution to:

$$f^* = \min_{x} \sum_{uv \in E} (x_u - \sigma_{uv} x_v)^2$$

s.t.  $x \in \{\pm 1/2\}^V$

$$\sum_{u} x_u^2 = |V| / 4$$

$\Leftrightarrow ||x|| = \sqrt{|V|} / 2$
Using Matrices

• The quadratic form can be expressed using matrices

\[
\sum_{uv \in E} (x_u - \sigma_{uv} x_v)^2 = \frac{1}{2} x^T L x
\]

where \( L \in \mathbb{R}^{V \times V} \) is the matrix given by

\( L_{uv} = -\sigma_{uv} \) when \((uv)\) is an edge

\( L_{uu} = d(u) \) is the degree of node \( u \)

• \( L = D - A \) is the \textbf{signed} graph Laplacian
Minimizing Quadratic Forms

\[ f^* = \min_x \frac{1}{2} x^T L x \]

s.t. \( \|x\| = \sqrt{|V|} / 2 \)

\[ \frac{8}{|V|} f^* = \min_x \frac{x^T L x}{x^T x} \]

\[ \frac{8}{|V|} f^* = \lambda_{\min}[L] \]

\[ f^* = \frac{|V|}{8} \lambda_{\min}[L] \]
Relative Relaxed Frustration

- Definition: Proportion of edges that have to be removed to make the graph balanced

\[ F^* = \frac{f^*}{|E|} \]

\[ F^* = \frac{|V|}{8|E|} \lambda_{\min}[L] \]

\[ 0 \leq F^* \leq \frac{f}{|E|} \leq 1 \]

\[ \lambda_{\min}[L] \leq \frac{8|E|}{|V|} \]
Properties of $L$ (Unsigned Graphs)

- $L$ is positive-semidefinite (all $\lambda[L] \geq 0$)
- Multiplicity of $\lambda = 0$ equals number of connected components
- Smallest eigenvalue measures conflict
- Second-smallest eigenvalue measures connectivity ("algebraic connectivity")
Properties of $L$ (Signed Graphs)

\[ L = \sum_{uv \in E} L^{(uv)} \]

\[ L^{(uv)} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \text{when } \sigma_{uv} = +1 \]

\[ L^{(uv)} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{when } \sigma_{uv} = -1 \]
Minimal Eigenvalue of $L$ (Signed Graphs)

- $L$ is positive-semidefinite (all eigenvalues $\geq 0$)
- $L$ is positive-definite (all eigenvalues $> 0$) iff all connected components are unlabanced
  - Proof “$\Rightarrow$”: by equivalence by all-positive graph
  - Proof “$\Leftarrow$”: by contradiction (eigenvector would be all-zero)
What Unsigned L can Do
What Signed L Can Do
Computing $\lambda_{\min}[L]$

- Sparse LU decomposition + inverse power iteration: $O(|V|^2)$ memory, but then very fast

% Matlab pseudocode

```matlab
[X Y] = sparse_lu(L);
[U D] = eigs(@(x)(Y \ X \ x), k, 'sm');
```
Temporal Analysis of $C_2$

- **Epinions ratings**
  - Dyadic conflict ($\eta$) vs. Volume (m) [edges]
  - $10^3 \times$ scale

- **Wikipedia elections**
  - Dyadic conflict ($\eta$) vs. Volume (m) [edges]
  - $10^4 \times$ scale
Temporal Analysis of $C_3$

Epinions ratings

Wikipedia elections
F* over Time

Wikipedia elections

MovieLens 100k
Cross-Dataset Comparison of F*

For larger networks, a smaller proportion of edges must be removed to make it balanced.
We want more datasets with negative edges, and timestamps.

- In particular: with changing and/or disappearing edges.