Random-walk based interpolations between centrality measures on complex networks

Per Arne Rikvold
PoreLab, University of Oslo
and Florida State University

with

Aleks J. Gurfinkel
Florida State University

Supported in part by the Research Council of Norway through the Center of Excellence funding scheme, Project No. 262644, and US NSF Grant No. DMR-1104829
Centrality?

• Centrality measures who (which vertex in a network) is most “important”?

• Importance and centrality can be many things.
  - “Richest” (Jeff Bezos or Scrooge McDuck)
  - Most “friends” (Facebook) or neighbors (Degree)
  - Most relevant to particular Web search (PageRank)
  - Lies on paths joining many vertex pairs (Betweenness)
  - Lies “close” to many other vertices (Closeness)
Electric Circuits and Random Walks

\[ A = \begin{pmatrix}
0 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
\end{pmatrix} \]

Adjacency (conductance) matrix

\[ L = \begin{pmatrix}
4 & -1 & -1 & -1 & -1 \\
-1 & 2 & -1 & 0 & 0 \\
-1 & -1 & 2 & 0 & 0 \\
-1 & 0 & 0 & 1 & 0 \\
-1 & 0 & 0 & 0 & 1 \\
\end{pmatrix} \]

Graph Laplacian. \( |I\rangle = L |\varphi\rangle \)

\[ W = \begin{pmatrix}
0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
\end{pmatrix} \]

Transition matrix. \( <P_{t+1}| = <P_t|W \)
Electric Current vs Geodesic Path

Electric current  Geodesic (Shortest) path
Resistance Distance

- The *effective electrical resistance* $R_{ij}$ between two vertices, $i$ and $j$, is easily found from the pseudo-inverse of the Graph Laplacian, $L^{(-1)}$, as:

$$R_{ij} = L^{(-1)}_{ii} + L^{(-1)}_{jj} - 2L^{(-1)}_{ij}$$

- $R_{ij}$ can be used as a *distance measure*.
- Since electric currents follow many parallel paths, this distance is less than the weighted shortest-path distance, $d_{ij}$:

$$R_{ij} \leq d_{ij}$$
Interpolating between current-based and geodesic-based centralities

Current $i,j$: $I_{ij} = \frac{V}{R_{ij}}$

Conditional current $i,j$: Portion of current injected at $i$ that reaches $j$

Death param.: $\Pi_D = 0$

Death param.: $\Pi_D > 0$

Resistance-Closeness and Current Betweenness Centralities

Walker-Flow Centralities
Electric Current vs Geodesic Path

Conductance-weighted Florida Power Grid

Distance: $R_{ij}$

Resistance Closeness:

$V/R_{ij} = I_{ij}$

Electric current

Distance: $d_{ij}$

Modified Closeness:

$V/d_{ij} = \mathcal{I}_{ij}$

Geodesic (Shortest) path
## Interpolation with $\Pi_D$

<table>
<thead>
<tr>
<th>Resistance Closeness</th>
<th>( \leftarrow \text{Interpolation} \rightarrow )</th>
<th>Modified Closeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Modified Information)</td>
<td>Conditional Resistance Closeness</td>
<td>(physical current)</td>
</tr>
<tr>
<td>( M_{ij}^{RCC} = \frac{1}{R_{ij}^{\text{eff}}} = I_{ij} )</td>
<td>( M_{ij}^{RCC}(\Pi_D) = \frac{1}{R_{ij}^{\text{eff, min}}(\Pi_D)} )</td>
<td>( M_{ij}^{MCL} = \frac{1}{d_{ij}} )</td>
</tr>
<tr>
<td>( \lim_{\Pi_D \to 0} I_{ij} )</td>
<td>( \Pi_D &gt; 0 )</td>
<td>( \lim_{\Pi_D \to \infty} )</td>
</tr>
<tr>
<td>( J_{ij}(\Pi_D) )</td>
<td>( J_{ij}(\Pi_D) )</td>
<td>( J ) only flows on geodesic paths from ( i ) to ( j ).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Current Betweenness</th>
<th>( \leftarrow \text{Interpolation} \rightarrow )</th>
<th>Betweenness</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Random Walk )</td>
<td>Conditional Current Betweenness</td>
<td>(physical current)</td>
</tr>
<tr>
<td>( M_{ij}^{CBT} = \sum_s I_{sij}/I_{sj} )</td>
<td>( M_{ij}^{CBT}(\Pi_D) = \sum_s J_{sij}(\Pi_D)/J_{sj}(\Pi_D) )</td>
<td>(conditional current)</td>
</tr>
<tr>
<td>( \lim_{\Pi_D \to 0} I_{ij} )</td>
<td>( \Pi_D &gt; 0 )</td>
<td>( \lim_{\Pi_D \to \infty} )</td>
</tr>
<tr>
<td>( J_{ij}(\Pi_D) )</td>
<td>( J_{ij}(\Pi_D) )</td>
<td>( J ) only flows on geodesic paths:</td>
</tr>
<tr>
<td>( \sum_s n_{sij}/g_{sj} )</td>
<td>( J_{sj} \propto g_{sj} ) and ( J_{sij} \propto n_{sij} ).</td>
<td></td>
</tr>
</tbody>
</table>
Absorbing Markov Chain

Absorbing transition matrix for network with $N$ vertices:

$$
W = \begin{pmatrix}
    (\text{Abs to Abs})_{2 \times 2} & (\text{Abs to Trn})_{2 \times (N-1)} \\
    (\text{Trn to Abs})_{(N-1) \times 2} & (\text{Trn to Trn})_{(N-1) \times (N-1)}
\end{pmatrix} = \begin{pmatrix}
    \mathbb{I} & \mathbb{O} \\
    (|\text{sink}\rangle |\text{target}\rangle) & T
\end{pmatrix}
$$

$T$ is the transient matrix with elements:

$$
T_{mn} = [1 - p_D(m)]A_{mn}/k_m
$$

The final result for the conditional current along the edge $(a,b)$ is:

$$
\frac{J_{i,b}}{J_{i,j}} = \mathbb{E}(\text{# walker crosses from } a \text{ to } b \mid j) - \mathbb{E}(\text{# walker crosses from } b \text{ to } a \mid j)
$$

$$
= F_{i,a} T_{a,b} F_{b,j} / F_{i,j} - F_{i,b} T_{b,a} F_{a,j} / F_{i,j}
$$

Where $F$ is the Fundamental Matrix:

$$
F = (\mathbb{I}_{(N-1) \times (N-1)} - T)^{-1}
$$
Conditional currents in Kangaroo Social Network

\[ \Pi_D = 10^{-8} \]
\[ \Pi_D = 22 \]
\[ \Pi_D = 66 \]
\[ \Pi_D = 200 \]
\[ \Pi_D = 601 \]
\[ \Pi_D = 1808 \]
Conditional Current–Betweenness Centralities in the Kangaroo Network

Conditional Resistance–Closeness Centralities in the Kangaroo Network

All paths

Only geodesic paths
Summary

- Centralities measure the “importance” of individual vertices.
- Many kinds of importance lead to many different centrality measures.
- Here we concentrate on Betweenness and Closeness measures.
- In both cases we use an Absorbing Markov Chain to interpolate between one limit, in which all paths are explored, and one, in which only geodesics count.
- The interpolation parameter is a Death Parameter, $\Pi_D$.
- Numerical results obtained by a linear-algebra method.
• “Absorbing Random Walks Interpolating Between Centrality Measures on Complex Networks.”
  A.J. Gurfinkel and P.A. Rikvold.

THANK YOU!