An edge centrality measure based on the Kemeny constant

D. Altafini$^1$, D. A. Bini$^2$, V. Cutini$^1$, B. Meini$^2$, F. Poloni$^3$

$^1$University of Pisa, Department of Energy Engineering, Systems, Land and Buildings
$^2$ University of Pisa, Department of Mathematics
$^3$ University of Pisa, Department of Computer Science

24$^{th}$ ILAS Conference
Galway, June 2022
The Kemeny constant

$A \in \mathbb{R}^{n \times n}$ adjacency matrix of an undirected, connected, weighted network; $P = D^{-1}A \in \mathbb{R}_{\geq 0}^{n \times n}$ transition matrix of the random walk on it (discrete-time Markov chain). $\text{eig}(P) = \{\lambda_1 = 1, \lambda_2, \ldots, \lambda_n\}$.

Kemeny constant [Kemeny,Snell ’60]

\[
K(P) = \sum_{i=2}^{n} \frac{1}{1 - \lambda_i}.
\]

Probabilistic definition: the mean first passage time from a fixed state $i$ to a state $j$ drawn according the invariant distribution.

Car-based interpretation: Car 1 runs for a long time on a road network and then breaks down. How many steps does car 2 take (on average) to get to the same spot as $A$ with a random walk?

$K(P)$ small $\iff$ $A$ well-connected as a network.
Centralities

We study a centrality measure for roads (edges) based on the Kemeny constant: a road is important if its removal causes a large increase in $K(P)$:

$$c(e) = K(\hat{P}) - K(P).$$

Many other centrality measures are available in literature. [Estrada, book ’13]

Main inspirations for us:

- [Estrada, D.Higham, Hatano ’09]: communicability betweenness centrality: variation in communicability centrality caused by the removal of an edge.

- [Crisostomi, Kirkland, Shorten ’11]: Kemeny constant variation in a Markov chain model of road circulation. Main difference: we do not want to rely on external traffic data, just on the map.
Collaboration with our civil engineering department; research question: is industry location driven by well-connected outskirts?

Large scale maps, e.g., continental Tuscany: 1.56M edges; no traffic data.
Weak ties

Goal: highlight weak ties [Granovetter, '73], i.e., crucial edges that separate (strongly-connected) sections of the map. Example: bridges.
Challenges

- Deal with negative centralities;
- Deal with cut-edges;
- Make it fast enough for $1.5M$ road elements.
Negative centralities

Sometimes, the Kemeny constant **decreases** when removing an edge!

**Example** $K(\text{left}) \approx 2.54$, $K(\text{right}) = 2.5$.

Not ideal: intuition of “connectedness” says more roads are always better.

This phenomenon is known as **Braess paradox** [Braess ’68, Kirkland, Zeng ’16].
Kemeny constant

\[ K(P) = \sum_{i=2}^{n} \frac{1}{1 - \lambda_i}. \]

\[ \{\lambda_1 = 1, \ldots, \lambda_n\} = \text{eig}(D^{-1}A) = \text{eig}(D^{-1/2}AD^{-1/2}) \]

\[ := W, \text{ symmetrized adjacency matrix} \]

The edge removal changes \( W \) in a non-trivial way.

\[
\begin{bmatrix}
0 & 1/2 & 6^{-1/2} & 0 \\
1/2 & 0 & 6^{-1/2} & 1 \\
6^{-1/2} & 6^{-1/2} & 0 & 3^{-1/2} \\
0 & 0 & 3^{-1/2} & 0
\end{bmatrix} \rightarrow \begin{bmatrix}
0 & 0 & 3^{-1/2} & 0 \\
0 & 0 & 3^{-1/2} & 0 \\
3^{-1/2} & 3^{-1/2} & 0 & 3^{-1/2} \\
0 & 0 & 3^{-1/2} & 0
\end{bmatrix}
\]
Solution

Idea Replace the removed edge with two loop edges, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

This changes the model in an easier-to-predict way:

$\begin{bmatrix} 0 & 1/2 & 6^{-1/2} & 0 \\ 1/2 & 0 & 6^{-1/2} & 1 \\ 6^{-1/2} & 0 & 3^{-1/2} & 0 \\ 0 & 0 & 3^{-1/2} & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1/2 & 0 & 6^{-1/2} & 0 \\ 0 & 1/2 & 6^{-1/2} & 1 \\ 6^{-1/2} & 6^{-1/2} & 0 & 3^{-1/2} \\ 0 & 0 & 3^{-1/2} & 0 \end{bmatrix}$

$W \mapsto \hat{W} := W + \frac{1}{\sqrt{d_i d_j}} (e_i - e_j)(e_i - e_j)^T$.

Theorem

With this definition, $c(e) = k(\hat{P}) - k(P) \geq 0$ after each edge removal.

Proof Standard eigenvalue inequalities for symmetric matrices:

$\hat{W} \succeq W \implies \hat{\lambda}_i \geq \lambda_i \implies \sum \frac{1}{1-\hat{\lambda}_i} \geq \frac{1}{1-\lambda_i}$.
Cut-edges

(Color scheme: blue edge = higher = important.)

Problem If the removed edge is a cut-edge, \( \hat{G} \) is disconnected, \( \hat{\lambda}_2 = 1 \), and \( K(\hat{\mathbf{P}}) = +\infty \).

On a road network, cut-edges are often unimportant dead ends, but sometimes they are crucial for connectivity and cannot be ignored/dismissed.
Solution

First idea Change the definition to

\[ K_r(P) = \sum_{i=2}^{n} \frac{1}{1 + r - \lambda_i}. \]

for a small regularization factor \( r > 0 \), e.g., \( r = 10^{-6} \).

\( \leftrightarrow \) replacing the Laplacian \( L = D - A \) with \( (1 + r)D - A \).

Problem Centrality scores \( c_r(e) = K_r(\hat{P}) - K_r(P) \) of cut-edges become \( \approx \frac{1}{r} \), artificially high.

Solution

Filtered Kemeny-based centrality

\[ \tilde{c}_r(e) = \begin{cases} 
\frac{1}{r} - c_r(e) & \text{e is a cut-edge}, \\
\ c_r(e) & \text{otherwise}. 
\end{cases} \]
Unfiltered vs. filtered

Unfiltered

Filtered

F. Poloni (U Pisa)
Sign reversal

Why $\frac{1}{r} - c_r(e)$ and not the more natural $c_r(e) - \frac{1}{r}$?

Theorem

If $e$ is a cut-edge, $\frac{1}{r} - c_r(e) \geq 0$.

Proof Interlacing inequalities: since $\hat{W} - W \succeq 0$ is rank-1 positive semidefinite,

$$\frac{1}{r} = \hat{\lambda}_2 \geq \lambda_2 \geq \hat{\lambda}_3 \geq \lambda_3 \geq \cdots \geq \hat{\lambda}_n \geq \lambda_n.$$

Hence

$$\frac{1}{r} - c_r(e) = \underbrace{\frac{1}{1+r-\lambda_2} - \frac{1}{1+r-\hat{\lambda}_3}}_{\geq 0} + \underbrace{\frac{1}{1+r-\lambda_3} - \frac{1}{1+r-\hat{\lambda}_4}}_{\geq 0} + \cdots + \underbrace{\frac{1}{1+r-\lambda_n}}_{\geq 0}.$$
Open problem

Filtered Kemeny-based centrality

\[ \tilde{c}_r(e) = \begin{cases} \frac{1}{r} - c_r(e) & e \text{ is a cut-edge}, \\ c_r(e) & \text{otherwise}. \end{cases} \]

Empirical observation

With this definition, centrality scores of cut-edges have centrality scores comparable with non-cut-edges, and they are sorted correctly in order of importance.

We still do not have a good explanation for this observation!
Getting it done

Problem How to reduce the $O(n^4)$ cost and make it fast enough for large graphs?

Theorem [Kemeny ’81, Kirkland ’10, Wang-Dubbeldam-Van Mieghem ’17]

Let $w \in \mathbb{R}^n$ be any vector such that $w^T 1 = 1$. Then,

$$K(P) = \text{Trace}(S^{-1}) - 1, \quad S = I - P + 1w^T.$$  

Since $\hat{P} - P$ and $\hat{S} - S$ is a rank-1 update, we can use the Sherman–Morrison formula

$$(S + uv^T)^{-1} - S^{-1} = \frac{-1}{1 + v^T S^{-1} u} S^{-1} uv^T S^{-1}$$

$$c(e) = K(\hat{P}) - K(P) = \text{Trace} \left( \frac{-1}{1 + v^T S^{-1} u} S^{-1} uv^T S^{-1} \right) = \frac{-u^T S^{-2} v}{1 + v^T S^{-1} u}.$$
Final formula

Some more routine manipulations:

- Introduce regularization parameter $r$;
- Use again Sherman–Morrison to invert $S_r = (1 + r)I - P + 1w^T$;
- Express it in terms of “regularized Laplacian” $L_r = (1 + r)D - A$;
- Choose $w$ to make the problem symmetric.

Final formula

$$c(\{i,j\}) = \frac{A_{ij} d^T (x_i^2)}{1 - A_{ij}(x_i - x_j)}, \quad y = L_r^{-1}(e_i - e_j), \quad x = y - \frac{d^T y}{\gamma} z.$$

where $d = \text{diag}(D)$, $z = L_r^{-1}d$, $\gamma = d^T z + d^T 1$. 
Practical cost

**Final formula**

\[
c(\{i,j\}) = \frac{A_{ij}d^T(x_i^2)}{1 - A_{ij}(x_i - x_j)}, \quad y = L_r^{-1}(e_i - e_j), \quad x = y - \frac{d^Ty}{\gamma}z.
\]

where \(d = \text{diag}(D), z = L_r^{-1}d, \gamma = d^Tz + d^T1\).

1. Precompute Cholesky factorization of \(L_r = (1 + r)D - A\), and \(d, z, \gamma\).
2. To compute \(c(e)\) for each edge (possibly in parallel), solve one linear system with \(L_r\) (using the precomputed factorization) and \(O(n)\) extra operations.

On road networks, often \(n \approx m \approx \text{nnz(chol}(L_r))\), so all these operations are somewhat cheap — but the cost is still \(O(n^2)\) to compute all centralities.
Experiment: a large-scale network

Mainland Tuscany map: \( n = 1.22M, m = 1.56M, \text{nnz}(\text{chol}(L_r)) = 3.36M. \)

1. Precomputation and \texttt{chol}: < 1s.
2. \texttt{parfor} centrality computation: 18 hours.

On a machine with 12 3.4GHz Xeon physical cores.
Experiment: the bridges of Pisa

Kemeny-based centrality $r=1e-8$, filtered

F. Poloni (U Pisa)
Conclusions

- The Kemeny constant variation works well to highlight bottlenecks and weak ties.
- Connectivity/positivity issues can be solved.
- Computationally feasible even in large scale.
- Interesting results for our collaborators in civ-eng.

Altafini, Bini, Cutini, Meini, Poloni. *An edge centrality measure based on the Kemeny constant*. Arxiv:2203.06459.

Thanks for your attention!