The Resistance Perturbation Distance: A Metric for the Analysis of Dynamic Networks

françois g. MEYER
University of Colorado at Boulder
Acknowledgments

- Nathan D. Monning


Acknowledgments

- U.S. DOE, grants (SC) DE-FC02-03ER25574
  Lawrence Livermore National Laboratory under contract B600360
- Tom Manteuffel and Geoff Sanders
- NSF DMS 0941476, NSF DMS 1407340
- Leto Peel
Outline

1. Introduction
2. The Resistance Perturbation Distance
3. Fast Computation of the RP-2 Distance
4. Analysis of Dynamic Networks
5. Discussion
1 Introduction
Statement of the Problem

• dynamic graphs: edges appear and disappear, weights along the edges change [2, 15, 20, 21, 27].

• study the coupling between changes in network topology and latent dynamical processes that drive the evolution of the connectivity

• one needs a notion of temporal difference: quantify significant configurational changes between two successive instants

→ pairwise comparison of graphs [30]
Statement of the Problem

• measure the distance between two graphs on the same vertex set; known vertex correspondence ✓

• challenge: not all distances are equivalent, or useful

• axiomatic definition of metrics between graphs:
  → inspired by the pioneering work of Koutra et al. [19, 18]
Metrics Between Graphs: an Axiomatic Definition

Axiom 1. A distance on a space of graphs should meet all the conditions of a distance: non-negativity, identity, symmetry, and subadditivity.

Four additional principles [19, 18]:

Principle 1 (Edge Importance). Changes that create disconnected components should be penalized more than changes that maintain the connectivity properties of the graphs.

Principle 2 (Weight Awareness). In weighted graphs, the larger the weight of the removed edge is, the greater the impact on the distance should be.

Principle 3 (Edge-“Submodularity”). A specific change is more important in a graph with few edges than in a much denser, but equally sized graph.

Principle 4 (Focus Awareness). Random changes in graphs are less important than targeted changes of the same extent.
• $G = (V, E, w)$ undirected weighted graph, connected, no self-loops
• $V = \{1, \ldots, n\}$ = vertex set, $E$ = edge set
• $n \times n$ weighted adjacency matrix

\[ A_{ij} = A_{ji} = \begin{cases} w_{ij} & \text{if the edge } [i, j] \in E, \\ 0 & \text{otherwise.} \end{cases} \]  

• combinatorial Laplacian matrix, $L = D - A$,

• degree matrix $D_{ii} = \sum_{j=1}^{n} A_{ij}$.

• pseudoinverse of $L$, $L^\dagger = (L + \frac{1}{n}J)^{-1} - \frac{1}{n}J$
• \((\varphi_k, \lambda_k)\) eigenvector and eigenvalues of \(L\),
\[ 0 = \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_n, \]

\[
L = \sum_{k=2}^{n} \lambda_k \varphi_k \varphi_k^T, \quad L^\dagger = \sum_{k=2}^{n} \frac{1}{\lambda_k} \varphi_k \varphi_k^T. \tag{2}
\]

• assign an (arbitrary) orientation to each edge \(e\), and define the \(m \times n\) signed edge incidence matrix, \(B\),

\[
B_{ei} = \begin{cases} 
1 & \text{if vertex } i \text{ is at the head of } e, \\
-1 & \text{if vertex } i \text{ is at the tail of } e, \\
0 & \text{otherwise.}
\end{cases} \tag{3}
\]

• diagonal edge weight matrix \(dA \in M_m\) with entries \(dA_{ee} = w_e\)

\[
L = B^T dAB. \tag{4}
\]
Related Work: True Metrics

• edit distance

\[ \| A^{(1)} - A^{(2)} \|_1 = \sum_{i,j} |A^{(1)}_{ij} - A^{(2)}_{ij}|. \]

• cut distance

\[
\max_{S, T \subseteq V} |E_{G^{(1)}}(S, T) - E_{G^{(2)}}(S, T)|, \quad \text{with} \quad E_G(S, T) = \sum_{i \in S, j \in T} \omega_{ij}
\]

computation: \( O(2^{2n}) \)

• difference in path lengths \([10]\) (pseudo-metric)

\( d_{G^{(i)}}(u, v) \): shortest distance from \( u \) to \( v \) in the graph \( G^{(i)} \)

\[
\min_{\Pi} \sum_{u, v \in V} |d_{G^{(1)}}(u, v) - d_{G^{(2)}}(\Pi(u), \Pi(v))|
\]

\( \Pi \): permutation of the vertices
Related Work: True Metrics

- size of the largest edge-, or vertex-induced subgraph common to $G^{(1)}$ and $G^{(2)}$ [5]

  computation = maximum common subgraph: NP-complete problem.
Related Work: Similarities and Graph Kernels

1. compute feature vector = signature of the graph characteristics
2. compare respective feature vectors of the two graphs are compared
   - isomorphism problem need not be solved
   - typically not injective; no triangular inequality
   - example: spectral similarity
     \[ d_\lambda \left( G^{(1)}, G^{(2)} \right) = \sqrt{\sum_{i=1}^{n} \left( \lambda^{(1)}_i - \lambda^{(2)}_i \right)^2}. \]
     \( \lambda_i \): spectrum of the adjacency, Laplacian, or normalized Laplacian matrices [8, 26, 36]
   - other examples: [6, 19, 25, 1, 7, 9, 3, 12, 28, 35]
Related Work: The DeltaCon Similarity

- only algorithm that adheres to the set of axioms and principles [19]

\[
\text{sim}_{\text{DC}_0} \left( G^{(1)}, G^{(2)} \right) = \frac{1}{1 + d_{\text{rootED}} \left( S^{(1)}, S^{(2)} \right)},
\]

root Euclidean distance

\[
d_{\text{rootED}} \left( S^{(1)}, S^{(2)} \right) = \left\{ \sum_{i, j=1}^{n} \left( \sqrt{S^{(1)}_{ij}} - \sqrt{S^{(2)}_{ij}} \right)^2 \right\}^{1/2},
\]

(5)

\( S^{(i)} \) is the fast belief propagation matrix

\[
S^{(i)} = \left[ I + \epsilon^2 D^{(i)} - \epsilon A^{(i)} \right]^{-1}
\]

(6)

\( \epsilon = 1/(1 + \max_i D_{ii}) \).
Related Work: The DeltaCon Similarity

- intuition: assume $\varepsilon \ll 1$, and drop the term $\varepsilon^2 D$

$$S \approx (I - \varepsilon A)^{-1} = I + \varepsilon A + \varepsilon^2 A^2 + \varepsilon^3 A^3 + \ldots.$$ (7)

- unweighted case, $A^k_{ij} = \text{count of paths of length } k \text{ between } i \text{ and } j$

- $S$: connectivity between vertices at all scales: longer paths having a reduced impact
Related Work: The DeltaCon Similarity

- DeltaCon$_0$ of a single-edge perturbation: $G \rightarrow G + \Delta w_{kl}$
- analytic computation of $d_{\text{rootED}}(G, G + \Delta w_{kl})$ for simple graphs

- complete graph, $K_n$

$$d_{\text{rootED}}(K_n, K_n + \Delta w_{kl})) = \left| \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2} + \Delta w_{kl}} \right| n + O(1).$$

- star graph, $S_n$

$$d_{\text{rootED}}(S_n, S_n + \Delta w_{kl}) = \frac{\sqrt{2\Delta w_{kl}}}{\sqrt{n}} - \frac{\sqrt{2}}{n} + O(1/n^{3/2}).$$

analysis: Sherman–Morrison–Woodbury theorem [14], inverse of a low-rank perturbation of a non-singular matrix.
DeltaCon\(_0\) of a single-edge perturbation: \(G \rightarrow G + \Delta w_{kl}\)

DeltCon\(_0\) does not meet Principle 3 [19]: “A specific change is more important in a graph with few edges than in a much denser, but equally sized graph.”
Summary of Related Work and Contribution

- many existing distances do not conform with the principles [19]
- many true distances suffer from a prohibitive computational cost
  
→ need for novel distances between graphs

- novel framework for constructing graph distances
- study one specific distance: resistance perturbation distance
- fast algorithms to compute this metric
2 The Resistance Perturbation Distance
A Unified Framework for Graph Distances

Definition 1. Given a matrix-to-matrix function, \( \varphi \),

\[ \varphi : M_n \to M_n, \]

and a distance \( d \) on \( M_n \), we define the pseudo-distance \( d_\varphi \) between two graphs \( G^{(1)} \) and \( G^{(2)} \) as follows,

\[ d_\varphi(G^{(1)}, G^{(2)}) = d(\varphi(A^{(1)}), \varphi(A^{(2)})), \]

where \( A^{(1)} \) and \( A^{(2)} \) are the adjacency matrices representing \( G^{(1)} \) and \( G^{(2)} \), respectively.

If \( \varphi \) is injective, then \( d_\varphi \) defines a distance.
Significance of the definition:

- natural mechanism to construct new distances
- decoupling two aspects of the distance $d_\phi$:
  1. matrix function $\phi$ extracts configurational or geometric properties about each graph
  2. distance $d$ emphasize large or small variations in the matrix function $\phi$
  3. choice of $d$: existence of fast algorithms
A Unified Framework for Graph Distances

Very general structure

- if $\varphi = I$ and $d = \ell_1$, then $d_{\varphi}$ is the edit distance
- if $\varphi = I$ and $d$ is the cut norm, then $d_{\varphi}$ is the cut distance
- if $\varphi(A) = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$ and $d = \| \cdot \|_F$, then $d_{\varphi}$ is the spectral pseudo-distance
- if $\varphi(A) = [I + \epsilon^2 D - \epsilon A]^{-1}$ and $d = \text{root Euclidean distance}$ (5), then $d_{\varphi}$ is DeltaCon$_0$.
- if $\varphi(A)_{ij} =$ shortest distance between nodes $i$ and $j$ and $d = \ell_1$, then $d_{\varphi}$ is the difference in path lengths
- if $\varphi(A)_{ij} =$ effective resistance between nodes $i$ and $j$ and $d = \| \cdot \|_p$, then $d_{\varphi}$ is the resistance perturbation distance
The Effective Resistance

element $[i, j] \mapsto$ resistor with resistance $= 1/w_{ij}$

The effective resistance between two vertices $i$ and $j$ is the voltage between $i$ and $j$ necessary to maintain a unit current through the edge $[i, j]$.

We have [4]

$$R_{ij} = L_{ii}^\dagger + L_{jj}^\dagger - 2L_{ij}^\dagger.$$ 

Intuition:
• $R_{ij}$ is a distance on the graph
• $R_{ij} \leq 1/w_{ij}$,
• $R_{ij} \ll 1/w_{ij}$ if there are many paths from $i$ to $j$.

The Kirchhoff index is defined by

$$\text{KI}(G) = \sum_{i, j \in V} R_{ij}.$$
Resistance Perturbation Distance

\( G^{(1)} = (V, E^{(1)}, w^{(1)}) \) and \( G^{(2)} = (V, E^{(2)}, w^{(2)}) \), connected and undirected graphs on the same vertex set, with effective resistance matrices, \( R^{(1)} \) and \( R^{(2)} \).

**Definition 2.** The RP-\( p \) distance, \( d_{\text{rp}(p)} \), between \( G^{(1)} \) and \( G^{(2)} \) is

\[
d_{\text{rp}(p)}(G^{(1)}, G^{(2)}) = \left\| R^{(1)} - R^{(2)} \right\|_p = \left[ \sum_{i,j \in V} \left| R_{ij}^{(1)} - R_{ij}^{(2)} \right|^p \right]^{1/p},
\]

(9)

**Theorem 1.** For \( 1 \leq p \leq \infty \), the RP-\( p \) distance defines a distance on the space of connected, weighted, undirected graphs with the same vertex set.

Idea of the proof: \( (R^{(1)} = R^{(2)}) \Rightarrow (L^{(1)} = L^{(2)}) \Rightarrow (A^{(1)} = A^{(2)}) \).

**Step 1:**

\[
L = \left( -\frac{1}{2} \left[ R - \frac{1}{n}(RJ + JR) + \frac{1}{n^2}JRJ \right] + \frac{1}{n}J \right)^{-1} - \frac{1}{n}J.
\]
Theorem 2. If $G + \Delta w_{i_0j_0}$ is the graph obtained from $G$ by a perturbation $\Delta w_{i_0j_0}$ to the edge $[i_0, j_0]$, then

$$d_{\text{rp}1}(G, G + \Delta w_{i_0j_0}) = \frac{2n \left| \Delta w_{i_0j_0} \right|}{1 + \Delta w_{i_0j_0} R_{i_0j_0}} \sum_{k=2}^{n} \frac{1}{\lambda_k^2} \left[ \varphi_k(i_0) - \varphi_k(j_0) \right]^2$$

$$= \frac{2n \left| \Delta w_{i_0j_0} \right|}{1 + \Delta w_{i_0j_0} \sum_{k=2}^{n} \frac{1}{\lambda_k} \left[ \varphi_k(i_0) - \varphi_k(j_0) \right]^2} \sum_{k=2}^{n} \frac{1}{\lambda_k^2} \left[ \varphi_k(i_0) - \varphi_k(j_0) \right]^2$$
Size of $\Delta w_{i_0j_0}/(1 + \Delta w_{i_0j_0} R_{i_0j_0})$

1. we always have $R_{i_0j_0} \leq 1/w_{i_0j_0}$ and $\Delta w_{i_0j_0} \geq -w_{i_0j_0}$

2. therefore $1 + \Delta w_{i_0j_0} R_{i_0j_0} \geq 0$.

3. the condition $d_{rp 1}(G, G + \Delta w_{i_0j_0}) = \infty$ implies
   
   a) $1 + \Delta w_{i_0j_0} R_{i_0j_0} = 0$

   b) a targeted removal of the edge $[i_0, j_0]$: $w_{i_0j_0} \rightarrow \Delta w_{i_0j_0} + w_{i_0j_0} = 0$
   and $1/R_{i_0j_0} = w_{i_0j_0}$

   c) $R_{i_0j_0} = 1/w_{i_0j_0} \Rightarrow [i_0, j_0]$ is the only path between $i_0$ and $j_0$

   d) the change disconnects the graph

4. $d_{rp 1}(G, G + \Delta w_{i_0j_0}) = \infty \iff \Delta w_{i_0j_0} + w_{i_0j_0} = 0$ disconnects the graph

Principle 1 is satisfied
Size of $\sum_{k=2}^{n} [\varphi_k(i_0) - \varphi_k(j_0)]^2 / \lambda_k^2$

1. large $k$
   - eigenvectors $\varphi_k$ “oscillate” very quickly on the graph
   - difficult to estimate the contribution of $[\varphi_k(i_0) - \varphi_k(j_0)]^2$
   - weights $1/\lambda_k^2$ are relatively small

2. small $k$
   - weights $1/\lambda_k^2$ are large
   - $\varphi_k(i_0) \approx \varphi_k(j_0)$ unless $i_0$ and $j_0$ belong to different nodal regions
   - example: network formed by densely connected communities, which are weakly connected to one another

$\rightarrow d_{rp1}(G, G + \Delta w_{i_0j_0})$ is maximal if $i_0$ and $j_0$ are in different communities
RP-1 Metric Created by Small Perturbations of Simple Graphs

claim: RP-1 distance can detect edge perturbations that have a profound effect on the functionality of the network

- consider a simple graph $G$
- perturb $G$ along edge $[i_0, j_0]$: $w_{i_0j_0} \rightarrow w_{i_0j_0} + \Delta w_{i_0j_0}$
- edit distance: $d_1(G, G + \Delta w_{i_0j_0}) = |\Delta w_{i_0j_0}|$
- compute $d_{rp1}(G, G + \Delta w_{i_0j_0})$ analytically
- compute a normalized RP-1 distance by dividing by $\|R\|_1$ (=Kirchhoff index)
The RP-1 Metric Created by Small Perturbations of the Complete Graph

$K_n$: complete graph on $n$ vertices

**Theorem 3.** If we perturb the weight of the edge $[i_0, j_0]$ by $\Delta w_{i_0j_0}$, then the RP-1 distance between the original and the perturbed graph is

$$d_{rp1}(K_n, K_n + \Delta w_{i_0j_0}) = \frac{4|\Delta w_{i_0j_0}|}{n + 2\Delta w_{i_0j_0}}.$$ 

The Kirchhoff index for the complete graph is $KI(K_n) = 2(n - 1)$.

The normalized $d_{rp1}$ distance is

$$\frac{d_{rp1}(K_n, K_n + \Delta w_{i_0j_0})}{KI(K_n)} = O \left( \frac{1}{n^2} \right).$$
Small Perturbations of a dense Erdős-Rényi Graph

- $G(n, p)$, when $p > \log(n)/n$
- Commute time $[24, 29]$,

$$n(1 + o(1)) \leq \mathbb{E}[\kappa_{i,j}] \leq n(2 + o(1)).$$

- expected number of edges $\mathbb{E}[m] \sim pn^2/2$
- estimate of the effective resistance, $R_{ij} = \kappa_{i,j}/2m$

$$\mathbb{E}[R_{ij}] \sim \frac{1}{np}.$$

- estimate of the RP-1 distance

$$\mathbb{E}\left[ d_{rp 1}(G, G + \Delta 1_{i_0j_0}) \right] \sim \frac{1}{np}.$$
The RP-1 Metric Created by Small Perturbations of the Star Graph

\( S_n \): star graph, every leaf 2, \ldots, \( n \) is connected to the root 1.

**Theorem 4.** *If we perturb a branch connecting the root 1 to a leaf \( i_0 \), then*

\[
d_{\text{rp}1}(S_n, S_n + \Delta w_{1i_0}) = \frac{2(n - 1)|\Delta w_{1i_0}|}{1 + \Delta w_{1i_0}}.
\]

*If we add an edge with weight \( \Delta w_{i_0j_0} \geq 0 \) between two leaves \( i_0 \) and \( j_0 \), then*

\[
d_{\text{rp}1}(S_n, S_n + \Delta w_{i_0j_0}) = \frac{4n\Delta w_{i_0j_0}}{1 + 2\Delta w_{i_0j_0}}.
\]

The Kirchhoff index for the star graph is \( \text{KI}(S_n) = 2(n - 1)^2 \)

The normalized \( d_{\text{rp}1} \) distance is

\[
\frac{d_{\text{rp}1}(S_n, S_n + \Delta w_{i_0j_0})}{\text{KI}(S_n)} = O\left(\frac{1}{n}\right).
\]
The RP-1 Metric Created by Small Perturbations of the Path Graph

$P_n$: path graph on $n$ vertices.

**Theorem 5.** If we add an edge with weight $\Delta w_{i_0j_0} \geq 0$ between $i_0$ and $j_0 > i_0$, then

$$d_{rp1}(P_n, P_n + \Delta w_{i_0j_0}) = |\Delta w_{i_0j_0}|(j_0 - i_0) \frac{2n [1 + (j_0 - i_0)(2j_0 + 4i_0 - 3)] - 3(j_0 - i_0)(i_0 + j_0 - 1)^2}{6 (\Delta w_{i_0j_0}(j_0 - i_0) + 1)}.$$ 

The Kirchhoff index for the path graph is $\text{KI}(P_n) = \frac{1}{3}(n - 1)n(n + 1)$.

If we assume that $i_0 = O(1) \leq j_0 \leq O(n)$, then

$$d_{rp1}(P_n, P_n + \Delta w_{i_0j_0}) \frac{1}{\text{KI}(P_n)} = O \left( \left[ \frac{j_0}{n} \right]^2 \right).$$
The RP-1 Metric Created by Small Perturbations of the Path Graph

• two regimes:
  1. if $j_0 = \mathcal{O}(1)$, then $i_0 \approx j_0$
     • $[i_0, j_0]$ has little impact on the graph
  2. if $j_0 = \mathcal{O}(n)$, then
     • $[i_0, j_0]$ = short circuit between the head and the end of the path
     • $d_{rp1}(P_n, P_n + \Delta w_{i_0,j_0})$ grows at the same rate as $KI(P_n)$

• behavior very different from that of the star graph:
  • both graphs are trees...
  • star: all the nodes are well connected
    $d_{rp1}(P_n, P_n + \Delta w_{i_0,j_0})$ independent of the location of $i_0$ and $j_0$
  • path: head and the tail at a distance $n$
    $d_{rp1}(P_n, P_n + \Delta w_{i_0,j_0})$ very dependent of the location of $i_0$ and $j_0$
The RP-1 Metric Created by Small Perturbations of the Path Graph

- maximum of $d_{rp1}$ is not when $i_0 = 1$ and $j_0 = n$
- faster diffusion: $i_0 = n/8$ and $j_0 = 7n/8 + 1$
- path $\rightarrow$ a cycle of perimeter $3n/4$ with two small tails of length $n/8$
The RP-1 Metric Created by Small Perturbations of the Cycle

\( C_n \): cycle on \( n \) vertices.

**Theorem 6.** If we add an edge with weight \( \Delta w_{i_0j_0} \geq 0 \) between \( i_0 \) and \( j_0 > i_0 \), then

\[
d_{\text{rp}1}(C_n, C_n + \Delta w_{i_0j_0}) = \frac{1}{6} \Delta w_{i_0j_0} [j_0 \ominus i_0] n \frac{[j_0 \ominus i_0]^3 - 2n [(j_0 \ominus i_0)^2 - 1] + [j_0 \ominus i_0] (n^2 - 2)}{n^2 + \Delta w_{i_0j_0} n [j_0 \ominus i_0] [n - (j_0 \ominus i_0)]},
\]

with \( j_0 \ominus i_0 = j_0 - i_0 \) (mod \( n \)).

The Kirchhoff index for the cycle graph is \( \text{KI}(C_n) = \frac{1}{6} (n - 1)n(n + 1) \).

If we assume that \( O(1) \leq j_0 \ominus i_0 \leq O(n) \), then

\[
\frac{d_{\text{rp}1}(C_n, C_n + \Delta w_{i_0j_0})}{\text{KI}(C_n)} = O\left(\left[\frac{j_0 \ominus i_0}{n}\right]^2\right).
\]
The RP-1 Metric Created by Small Perturbations of the Cycle

- maximum of $d_{rp_1}$ is $j_0 - i_0 = n/2$
- short circuit in the middle of the cycle
- “small world” model
DeltaCon₀ and non-normalized \( d_{rp1} \) for \( G \rightarrow G + \Delta w_{kl} \)
Adherence to Axioms and Principles

1. **Axiom 1. ✓**
   All the RP-p distances are proper distances.

1. **Edge Importance. ✓**
   \[ d_{rp} \left( G, G + \Delta w_{i_0,j_0} \right) \to \infty \] if and only if removing \([i_0, j_0]\) disconnects the graph.

2. **Weight Awareness. ✓**
   As \( w_{i_0,j_0} \to \infty \), then \( R_{i_0,j_0} \approx \frac{1}{w_{i_0,j_0}} \)
   if \([i_0, j_0]\) is removed \((\Delta w_{i_0,j_0} = -w_{i_0,j_0})\), then \( d_{rp} \left( G, G + \Delta w_{i_0,j_0} \right) \to \infty \).
3. **Edge-“Submodularity”**. No formal proof.
   \( d_{rp_1} \downarrow \) as density of edges \( \uparrow \)
   - complete graph: \( O(n^2) \) edges, \( d_{rp_1}(K_n, K_n + \Delta w_{i_0j_0}) = O(1/n) \)
   - star graph: \( O(n) \) edges, and \( d_{rp_1}(S_n, S_n + \Delta w_{i_0j_0}) = O(n) \)

4. **Focus Awareness**. No formal proof.
   - densely connected communities, weakly connected to each other
   - \( d_{rp_1}(G, G + \Delta w_{i_0j_0}) \) is maximal if \( i_0 \) and \( j_0 \) are in different communities
   \( \rightarrow \) targeted change
3 Fast Computation of the RP-2 Distance
Naive evaluation: cost of at least $O(n^2)$

$$d_{rp}(p)(G^{(1)}, G^{(2)}) = \left\| \text{diag}(L^{(1)\dagger} - L^{(2)\dagger})1^T + 1 \text{diag}(L^{(1)\dagger} - L^{(2)\dagger})^T - 2(L^{(1)\dagger} - L^{(2)\dagger}) \right\|_p$$

We propose a fast (linear in the number of edges) randomized algorithm to compute an approximation to the RP-2 distance

Key ingredient: algorithm of Spielman and Srivastava [33, 34] for approximating pairwise effective resistances:

\[ \forall \varepsilon > 0, \exists \tilde{O}(m \log \frac{w}{\varepsilon^2}) \text{ time algorithm [33] that computes a} \]
\[ (24 \log \frac{n}{\varepsilon^2}) \times n \text{ matrix } \tilde{Z} \text{ such that with probability at least } 1 - 1/n, \]
\[ (1 - \varepsilon)R_{ij} \leq \left\| \tilde{Z}(e_i - e_j) \right\|_2^2 \leq (1 + \varepsilon)R_{ij}, \quad \forall i, j \in V, \]

\[ w_{\min} \text{ and } w_{\max} \text{ are the minimum and maximum edge weights,} \]
\[ \overline{w} = w_{\max}/w_{\min}, \text{ and } e_i: i^{th} \text{ vector of the canonical basis in } \mathbb{R}^n. \]
Fast Approximation of Pairwise Resistances

Recall crucial ideas of Spielman and Srivastava [33, 34]

1. embed vertices in $\mathbb{R}^m$: Euclidean distance $\rightarrow$ effective resistance

   $R_{ij} = (e_i - e_j)^T L^\dagger (e_i - e_j) = \left\| dA^{1/2} BL^\dagger (e_i - e_j) \right\|_2^2 = \left\| Z(e_i - e_j) \right\|_2^2.$

   ... but $Z = dA^{1/2} B$ is too large: $m \times n$

2. replace $dA^{1/2} B$ with $Y = Q dA^{1/2} B$ of size $s \times n$,

   $Q \in \mathbb{R}^{s \times m}$: random entries $\pm 1/\sqrt{s}$, $s = 24 \log n/\varepsilon^2$

3. compute $\tilde{Z}^T = L^\dagger Y^T$, $\tilde{z}_i = i^{th}$ column of $\tilde{Z}^T$, is the solution of

   $L \tilde{z}_i = y_i, i = 1, \ldots, s,$

   where $y_i$ is the $i^{th}$ column of $Y^T$
Algorithm 1 Compute the matrix $\tilde{Z}$ [34]

1: Generate a realization $Q \in M_{s \times m}$, with random entries $\pm 1/\sqrt{s}$, with $s = 24 \log n/\varepsilon^2$.

2: $Y \leftarrow QdA^{1/2}B$

// $\delta$ controls the relative error, $\|x - L^\dagger y\|_L \leq \delta \|L^\dagger y\|_L$, and $\|y\|_L = \sqrt{y^T Ly}$

3: $\delta \leftarrow \frac{\varepsilon}{3} \sqrt{\frac{2}{n^3}} \left( \frac{1 - \varepsilon}{1 + \varepsilon} \right) \frac{w_{\text{min}}}{w_{\text{max}}}$.  

// Use Laplacian solver STSolve of Spielman and Teng [31, 32]

4: Compute: $\tilde{z}_i \leftarrow \text{STSolve}(L, y_i, \delta), \quad \forall i = 1, \ldots, s.$

approximate effective resistance matrix: $R \approx \tilde{R}$

$$\tilde{R} = \text{diag}(\tilde{Z}^T \tilde{Z}) 1^T + 1 \text{ diag}(\tilde{Z}^T \tilde{Z})^T - 2\tilde{Z}^T \tilde{Z}.$$
Last Ingredient: Fast Frobenius Norm

Frobenius norm = $O(n^2) \rightarrow$ nearly linear time in $n$.

**Theorem 7.** $\|\tilde{R}^{(1)} - \tilde{R}^{(2)}\|_F$ can be computed in $\tilde{O}(n) = O(n \log^2 n)$ time.

**Proof.** Let $d = \text{diag} \left( [\tilde{Z}^{(1)}]^T \tilde{Z}^{(1)} - [\tilde{Z}^{(2)}]^T \tilde{Z}^{(2)} \right) \in \mathbb{R}^n$. Using the invariance of the trace under cyclic permutations, we can show that

$$\|\tilde{R}^{(1)} - \tilde{R}^{(2)}\|_F^2 = 2 \left\{ [1^T d]^2 + n \|d\|_2^2 + 4 \left( 1^T [\tilde{Z}^{(2)}]^T \tilde{Z}^{(2)} d - 1^T [\tilde{Z}^{(1)}]^T \tilde{Z}^{(1)} d \right) \right. \\
+ 2 \left( \|\tilde{Z}^{(1)}[\tilde{Z}^{(1)}]^T\|_F^2 \|\tilde{Z}^{(2)}[\tilde{Z}^{(2)}]^T\|_F^2 - 2 \|\tilde{Z}^{(2)}[\tilde{Z}^{(1)}]^T\|_F^2 \right) \right\},$$

which can be computed in $\tilde{O}(n) = O(n \log^2 n)$ time. \qed

A Nearly Linear-time Algorithm for the RP-2 Distance

- two graphs: $G^{(1)} = (V, E^{(1)}, w^{(1)})$, $G^{(2)} = (V, E^{(2)}, w^{(2)})$
- number of edges: $m^{(1)} = |E^{(1)}|$, $m^{(2)} = |E^{(2)}|$
- weight factors: $\overline{w}^{(1)} = w_{max}^{(1)}/w_{min}^{(1)}$, $\overline{w}^{(2)} = w_{max}^{(2)}/w_{min}^{(2)}$
- $m \log \overline{w} = \max(m^{(1)} \log \overline{w}^{(1)}, m^{(2)} \log \overline{w}^{(2)})$.

**Theorem 8.** There is an algorithm with complexity $\tilde{O}\left(n + \frac{m \log \overline{w}}{\varepsilon^2}\right)$ that computes $\tilde{d}_{rp2}(G^{(1)}, G^{(2)})$, such that

$$d_{rp2}(G^{(1)}, G^{(2)}) - \varepsilon \left\| R^{(1)} + R^{(2)} \right\|_F \leq \tilde{d}_{rp2}(G^{(1)}, G^{(2)}) \leq d_{rp2}(G^{(1)}, G^{(2)}) + \varepsilon \left\| R^{(1)} + R^{(2)} \right\|_F.$$  

with probability at least $1 - 2/n$. 
Fast RP-2 Distance

Algorithm 2 Compute $\tilde{d}_{rp2}(G^{(1)}, G^{(2)})$

1: Input: $E^{(1)}, w^{(1)}, E^{(2)}, w^{(2)}$, tolerance $\varepsilon > 0$.
2: Compute $\tilde{Z}^{(1)}, \tilde{Z}^{(2)} \in M_{s \times n}$ using Algorithm 1.
3: $d \leftarrow \text{diag}([\tilde{Z}^{(1)}]^T \tilde{Z}^{(1)} - [\tilde{Z}^{(2)}]^T \tilde{Z}^{(2)})$.
4: $\tilde{d}_{rp2} \leftarrow \sqrt{2 \left\{ \left[ 1^T d \right]^2 + n \|d\|^2 + 4 \left( 1^T [\tilde{Z}^{(2)}]^T \tilde{Z}^{(2)} d - 1^T [\tilde{Z}^{(1)}]^T \tilde{Z}^{(1)} d \right) 
+ 2 \left( \|\tilde{Z}^{(1)}[\tilde{Z}^{(1)}]^T \|_F^2 + \|\tilde{Z}^{(2)}[\tilde{Z}^{(2)}]^T \|_F^2 - 2 \|\tilde{Z}^{(2)}[\tilde{Z}^{(1)}]^T \|_F^2 \right) \right\}^{1/2}}$
5: return $\tilde{d}_{rp2}$.

- Algorithm 1: Spielman and Srivastava’s algorithm written by Richard Garcia-Lebron [13]
- combinatorial multigrid solver [17] written by Ioannis Koutis, and Gary Miller
Computation Time of $\tilde{d}_{rp2}$ for Sparse Random Graphs

![Graph showing computation time vs number of edges]

- **Computation Time (seconds)**
- **Number of Edges**
- **approx. RP$_2$ distance**
- **linear time**

Graph legend:
- Black dots represent the computation time for approximate RP$_2$ distance.
- The straight line represents linear time.

### Table

<table>
<thead>
<tr>
<th>Number of Edges</th>
<th>Computation Time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^3$</td>
<td>approx. RP$_2$ distance</td>
</tr>
<tr>
<td></td>
<td>linear time</td>
</tr>
<tr>
<td>$10^4$</td>
<td></td>
</tr>
<tr>
<td>$10^5$</td>
<td></td>
</tr>
<tr>
<td>$10^6$</td>
<td></td>
</tr>
<tr>
<td>$10^7$</td>
<td></td>
</tr>
</tbody>
</table>
4 Analysis of Dynamic Networks
Detection of Anomalies in Dynamic Networks

Goal: detect configurational changes in dynamic graphs that are triggered by modifications of the hidden variables controlling the graph dynamics.

We need to check that:

- the RP-p distance is indifferent to normal random fluctuation
- the RP-p distance tracks the hidden variables influencing significant structural changes

on

- synthetic random dynamic networks
- real dynamic networks
Random Dynamic Networks

• generate random realizations of graphs sampled from ensembles

• three different families:

  1. random graph with a latent space,

  2. two-communities block stochastic model (planted partition),

  3. small world (Watts and Strogatz) model

• each model depends on a single scalar that characterizes the structure of the graph

• compute \( d_{rp(p)}(G^{(1)}, G^{(2)}) \) and compare to the (unobserved) changes in the latent parameter that controls the organization of the graph
Experimental procedure

1. a baseline graph $G^{(1)}$ is randomly selected using the baseline value of the parameter of the corresponding model

2. generate 50 random realizations of a second graph $G^{(2)}$, using the same value of the parameter

3. compute: $d_{rp1}(G^{(1)}, G^{(2)})$, $d_{rp2}(G^{(1)}, G^{(2)})$, and $d_1(G^{(1)}, G^{(2)})$
   → quantify the sensitivity of the three distances to normal random fluctuations in the baseline (no changes) condition

4. repeat 10 times:
   a) increase the parameter that controls $G^{(2)}$
   b) generate random realizations of $G^{(2)}$
   c) compute: $d_{rp1}(G^{(1)}, G^{(2)})$, $d_{rp2}(G^{(1)}, G^{(2)})$, and $d_1(G^{(1)}, G^{(2)})$
   d) display: distance(parameter) normalized by maximum distance
Unit Circle Latent Space Model

• Sample 2,000 points on the unit circle in $\mathbb{R}^2$,

$$x_i^{(1)} = \begin{bmatrix} \cos(\theta_i^{(1)}) \\ \sin(\theta_i^{(1)}) \end{bmatrix}, \quad \theta_i^{(1)} \sim U[0, 2\pi], \quad i = 1, \ldots, 2,000.$$

• Randomly connect each pair of vertices $\{i, j\}$ with an edge $[i, j]$

$$P([i, j] \in E) = \frac{20}{\sqrt{\pi}} \exp\left(-400\|x_i - x_j\|^2\right), \quad \text{for } i \neq j.$$

• $G^{(2)}$ same as $G^{(1)}$ but $\{x_i^{(2)}\}$ is a random perturbation of $\{x_i^{(1)}\}$,

$$x_i^{(2)} = \begin{bmatrix} \cos(\theta_i^{(2)}) \\ \sin(\theta_i^{(2)}) \end{bmatrix}, \quad \theta_i^{(2)} = \theta_i^{(1)} + \gamma_i, \quad \gamma_i \sim \mathcal{N}(0, \sigma^2), \quad i = 1, \ldots, 2,000.$$

• Generate 50 random realizations of $G^{(2)}$ for $\sigma \in [0, 1]$
Unit Circle Latent Space Model

\[ d_{RP}(G^{(1)}, G^{(2)}) \text{ (normalized by max)} \]

error bars = [min, max] computed over 50 random realizations
Two Communities Stochastic Block Model

- $n = 2,000$ nodes are divided into two communities of size $n/2$
- connect vertices according to

\[
\text{Prob} \left( \{i, j\} \in E \right) = \begin{cases} 
    p_{\text{in}} = 0.9 & \text{if } i \text{ and } j \text{ are in the same community}, \\
    p_{\text{out}} & \text{otherwise}
\end{cases}
\]

- $p_{\text{out}}^{(1)} = 0.005$ for $G^{(1)}$
- $p_{\text{out}}^{(2)} \in [0.005, 0.01]$ for $G^{(2)}$
- compute $d_{\text{rp}1}(G^{(1)}, G^{(2)}), d_{\text{rp}2}(G^{(1)}, G^{(2)}),$ and $d_1(G^{(1)}, G^{(2)})$
  as a function of $\Delta p_{\text{out}} = p_{\text{out}}^{(2)} - p_{\text{out}}^{(1)}$. 
Two Communities Stochastic Block Model

\[ \Delta p_{\text{out}} \times 10^{-3} \]

\[ d_{\text{RP}}(G^{(1)}, G^{(2)}) \text{ (normalized by max)} \]

error bars = [min, max] computed over 50 random realizations
Small World (Watts and Strogatz) Model

- randomly re-wire a regular ring lattice of constant degree $k = 40$
- random rewiring with probability $\beta_1 = 0.01$ for $G^{(1)}$
- random rewiring with probability $\beta_2 \in [0.01, 0.03]$ for $G^{(2)}$
- compute $d_{rp1}(G^{(1)}, G^{(2)})$, $d_{rp2}(G^{(1)}, G^{(2)})$, and $d_1(G^{(1)}, G^{(2)})$
  as a function of $\Delta \beta = \beta_2 - \beta_1$
Small World (Watts and Strogatz) Model

\[ d_{RP}(G^{(1)}, G^{(2)}) \text{ (normalized by max)} \]

\[ \Delta \]

\[ \beta \]

\[ 0 \quad 0.002 \quad 0.004 \quad 0.006 \quad 0.008 \quad 0.01 \]

error bars = [min, max] computed over 50 random realizations
Real Dynamic Network

- two real-world dynamic networks
- qualitatively compare the resistance perturbation distance to known events
- compared the metrics $d_{rp1}$ and $d_{rp2}$ to the edit distance
- $d_{rp1}$ and $d_{rp2}$ could be used to infer changes in the hidden variables that govern the evolution of the dynamic graphs
- edit distance was saturated by the random noise in the networks, and failed to resolve the hidden changes occurring in the latent space.
Enron email network

- Enron email corpus [16]
- period = 3 years leading up to the collapse of the company
- emails were aggregated on a weekly basis
- nodes = 150 high-level executives
  (the Enron “core” = individuals most closely involved in the scandal)
- weight for week $t$ = number of emails between individuals $i$ and $j$
- emails with greater than three recipients were excluded
- 25 known significant events

$d_{rp1}$, $d_{rp2}$ can predict catastrophic events (e.g., the collapse of the company)
Enron email network: RP-1 and RP-2 distances

EBS Launch
EBS & Blockbuster partner
Stock reaches all-time high
Attorney discusses Belden's strategies
Enron exonerated by FERC
Memo on Belden's strategies
Announcement that Skilling will become CEO
Memo on Belden's strategies
$53m "profit" from EBS
CA Blackouts
Conference call scheduled to boost stock
Schwarzenegger, Lay, Milken meet
Quarterly conference call
Skilling meets investors/analysts
Watkins questions accounting
Dynegy agrees to buy Enron
Restatement of Q3 earnings
Shares fall below $1
Dynegy pulls offer
Enron files for bankruptcy
Cooper becomes CEO
Lay cancels on Senate
Fastow and Kopper invoke the 5th
Watkins testifies against Skilling & Lay
Anderson indicted

Date
01/10/00 02/07/00 03/06/00 04/03/00 05/01/00 05/29/00 06/28/00 07/24/00 08/21/00 09/18/00 10/16/00 12/1/00 01/09/01 02/05/01 03/05/01 04/02/01 04/30/01 05/29/01 06/25/01 07/23/01 08/20/01 09/17/01 10/15/01 11/12/01 12/10/01 01/07/02 02/04/02 03/04/02 04/01/02 04/29/02 05/27/02 06/24/02

meyer | A Metric for the Analysis of Dynamic Networks

60/75
Enron email network: edit distance

Date

01/10/00 02/07/00 03/08/00 04/03/00 05/01/00 06/08/00 07/24/00 08/21/00 09/18/00 10/16/00 11/13/00 12/06/00 01/08/01 02/05/01 03/05/01 04/02/01 04/30/01 05/28/01 06/26/01 07/24/01 08/21/01 09/18/01 10/16/01 11/13/01 12/06/01 01/03/02 02/08/02 03/04/02 04/29/02 05/27/02 06/24/02

EBS Launch
EBS & Blockbuster partner
Stock reaches all-time high
Attorney discusses Belden's strategies
Enron exonerated by FERC
Memo on Belden's strategies
Announcement that Skilling will become CEO
$53m "profit" from EBS
CA Blackouts
Conference call scheduled to boost stock
Dynegy agrees to buy Enron
Restatement of Q3 earnings
Shares fall below $1
Dynegy pulls offer
Enron files for bankruptcy
Cooper becomes CEO
Lay cancels on Senate
Fastow and Kopper invoke the 5th
Watkins testifies against Skilling & Lay
Anderson indicted

edit distance

IA_{t-1} / \|A_{t-1} \|_1 (normalized by maximum)
MIT reality mining dataset\cite{11}

- collocation information between students and faculty at MIT
- weekly-aggregated proximity data (using bluetooth on cellphones)
- 97 nodes = faculty + graduate students
- period = 35 weeks
- weights for week $t \propto$ amount of time the pair of cellphones were in close physical proximity
- 16 known external events
MIT reality mining dataset: RP-1 and RP-2 distances

Date

08/23/04 08/30/04 09/06/04 09/13/04 09/20/04 09/27/04 10/04/04 10/11/04 10/18/04 10/25/04 11/01/04 11/08/04 11/15/04 11/22/04 12/06/04 12/13/04 12/20/04 12/27/04 01/03/05 01/10/05 01/17/05 01/24/05 01/31/05 02/07/05 02/14/05 02/21/05 02/28/05 03/07/05 03/14/05 03/21/05 03/28/05 04/04/05 04/11/05 04/18/05 04/25/05 05/02/05

meyer | A Metric for the Analysis of Dynamic Networks
MIT reality mining dataset[11]

- significant changes in the network:
  1. between the first and second week of classes at the beginning of the fall semester
  2. week after finals (the beginning of winter break) and the beginning of the independent activities period
  3. beginning and end of spring break, as students depart from and return to their campus routine
MIT reality mining dataset: edit distance

A Metric for the Analysis of Dynamic Networks
5 Discussion
The Resistance Perturbation Distance

• novel graph distance to quantify configurational changes

• perturbations of several prototypical graphs where $d_{rp}(p)$ can be computed analytically

• fast computational algorithms to evaluate the metric $d_{rp^2}$

• detect the changes in the hidden latent variables that control the network dynamics
Open Research Questions

• extension to disconnected graphs
• volume-normalized version of the distance
• statistics to separate baseline noise from significant structural changes: topology, connectivity, etc.
• attribution: localization about the anomalies
• fast $d_{r,p_2}$ algorithm: bulk of the work= $\tilde{O}(\log n)$ Laplacian linear solver (each of size $n$)
• replace combinatorial multigrid solver of Koutis et al. [17]: (constant in the $\tilde{O}$ is significant)
• Lean Algebraic Multigrid (LAMG) [22, 23]
• splitting the $O(\log n)$ independent calls to the Laplacian solver onto independent processors/cores
References


