Script generated by TTT

Title: profile1 (28.05.2013)

Date: Tue May 28 12:04:37 CEST 2013

Duration: 86:47 min

Pages: 67

Distances: Eccentricity

- Example: Facility location problems: Objective function on d(u,v): e.g. minimax (minimize maximal distance (e.g.: hospital emergency)) → can be mapped to social case
- For the moment: G is undirected and unweighted (e.g. "friendship"). Mapping to weighted and / or directed case is possible.
- Eccentricity e(u)=max{d(u,v); v∈V}

Distances: Centroids

- Competitive objective: Given number c … e to open a store (Customers will just choose store b… stance)?
- Social Problem: Example: find "social e… computer scientists try to find a partner at a compu… at social science parties?) ☺
- Formalization: For u, v : $\gamma_u(v)$=number … u than to v; If one salesman selects u and competitor selects v as locations, the first will have

Weiter
Zurück
Zuletzt angesehen
Alle Folien anzeigen
Vergrößern
Zielgruppenorientierte Präsentation
Referentenansicht
Bildschirm
Zeigeroptionen
Hilfe
Anhalten
Präsentation beenden

$$\gamma_u(v)+\frac{1}{2}(|V|-\gamma_u(v)-\gamma_v(u))=\frac{1}{2}|V|+\frac{1}{2}(\gamma_u(v)-\gamma_v(u))$$

customers

Distances: Centroids

- Competitive objective: Given number of competitors: where to open a store (Customers will just choose store based on minimal distance)?
- Social Problem: Example: find "social ecological niche" (do computer scientists try to find a partner at a computer science party or at social science parties?) ☺
- Formalization: For u, v : $\gamma_u(v)$=number of vertices closer to u than to v; If one salesman selects u and competitor selects v as locations, the first will have

$$\gamma_u(v)+\frac{1}{2}(|V|-\gamma_u(v)-\gamma_v(u))=\frac{1}{2}|V|+\frac{1}{2}(\gamma_u(v)-\gamma_v(u))$$

customers

Distances: Centroids

- Competitive objective: Given number of competitors: where to open a store (Customers will just choose store based on minimal distance)?
- Social Problem: Example: find “social ecological niche” (do computer scientists try to find a partner at a computer science party or at social science parties?) ☺
- Formalization: For u, v : $\gamma_u(v)$=number of vertices closer to u than to v; If one salesman selects u and competitor selects v as locations, the first will have

$$\gamma_u(v)+\frac{1}{2}(|V|-\gamma_u(v)-\gamma_v(u))=\frac{1}{2}|V|+\frac{1}{2}(\gamma_u(v)-\gamma_v(u))$$

customers

Distances: Centroids

- Competitive objective: Given number of competitors: where to open a store (Customers will just choose store based on minimal distance)?
- Social Problem: Example: find “social ecological niche” (do computer scientists try to find a partner at a computer science party or at social science parties?) ☺
- Formalization: For u, v : $\gamma_u(v)$=number of vertices closer to u than to v; If one salesman selects u and competitor selects v as locations, the first will have

$$\gamma_u(v)+\frac{1}{2}(|V|-\gamma_u(v)-\gamma_v(u))=\frac{1}{2}|V|+\frac{1}{2}(\gamma_u(v)-\gamma_v(u))$$

customers

Distances: Centroids

- Competitive objective: Given number of competitors: where to open a store (Customers will just choose store based on minimal distance)?
- Social Problem: Example: find “social ecological niche” (do computer scientists try to find a partner at a computer science party or at social science parties?) ☺
- Formalization: For u, v : $\gamma_u(v)$=number of vertices closer to u than to v; If one salesman selects u and competitor selects v as locations, the first will have

$$\gamma_u(v)+\frac{1}{2}(|V|-\gamma_u(v)-\gamma_v(u))=\frac{1}{2}|V|+\frac{1}{2}(\gamma_u(v)-\gamma_v(u))$$

customers

Distances: Centroids

- →Competitor will want to minimize

$$f(u,v)=\gamma_u(v)-\gamma_v(u)$$

- → Possible centrality index: First salesman knows the strategy of the competitor and calculates for each location the worst case:

$$c(u)=\min_v\{f(u,v):v\in V/\{u\}\}$$

- c(u) is called centroid value: measures the advantage of location u compared to other locations: Minimal loss of customers if he choses u and a competitor choses v

Distances: Centroids

- → Competitor will want to minimize

$$f(u,v) = \gamma_u(v) - \gamma_v(u)$$

- → Possible centrality index: First salesman knows the strategy of the competitor and calculates for each location the worst case:

$$c(u) = \min_v \{f(u,v) : v \in V / \{u\}\}$$

- c(u) is called centroid value: measures the advantage of location u compared to other locations: Minimal loss of customers if he choses u and a competitor choses v

Distances: Centroids

- → Competitor will want to minimize

$$f(u,v) = \gamma_u(v) - \gamma_v(u)$$

- → Possible centrality index: First salesman knows the strategy of the competitor and calculates for each location the worst case:

$$c(u) = \min_v \{f(u,v) : v \in V / \{u\}\}$$

- c(u) is called centroid value: measures the advantage of location u compared to other locations: Minimal loss of customers if he choses u and a competitor choses v

Distances: Centroids

- → Competitor will want to minimize

$$f(u,v) = \gamma_u(v) - \gamma_v(u)$$

- → Possible centrality index: First salesman knows the strategy of the competitor and calculates for each location the worst case:

$$c(u) = \min_v \{f(u,v) : v \in V / \{u\}\}$$

- c(u) is called centroid value: measures the advantage of location u compared to other locations: Minimal loss of customers if he choses u and a competitor choses v

Shortest Paths: Stress

- Heuristic: If a vertex is part of many shortest paths → "much information will run through it" if information is routed along shortest paths
- Social analogon: People that are asked to contribute to a workflow more often than others
- → A vertex v is more central the more shortest paths run through it. Let $\sigma_{ab}(v)$ denote the number of shortest paths from node a to node b containing v. $\sigma_{ab}(v)$ can be >1 if there there are several paths with the same minimal length

stress centrality: $$c(v) = \sum_{a \in V; a \neq v} \sum_{b \in V; b \neq v} \sigma_{ab}(v)$$

Shortest Paths: Stress

- Heuristic: If a vertex is part of many shortest paths → "much information will run through it" if information is routed along shortest paths
- Social analogon: People that are asked to contribute to a workflow more often than others
- → A vertex v is more central the more shortest paths run through it. Let $\sigma_{ab}(v)$ denote the number of shortest paths from node a to node b containing v. $\sigma_{ab}(v)$ can be >1 if there there are several paths with the same minimal length

stress centrality: $$c(v) = \sum_{a \in V; a \neq v} \sum_{b \in V; b \neq v} \sigma_{ab}(v)$$

Shortest Paths: Shortest Path Betweenness

- Again assume that communication (workflows etc.) happen along shortest paths only. Let

$$\delta_{ab}(v) = \frac{\sigma_{ab}(v)}{\sigma_{ab}}$$

with σ_{ab}: total number of shortest paths between nodes a and b.

Interpretation. Probability that v is involved in a communication between a and b

Shortest Paths: Shortest Path Betweenness

- Shortest Path Betweenness (SPB) centrality is then:

$$c(v) = \sum_{a \neq v} \sum_{b \neq v} \delta_{ab}(v)$$

- Interpretation: Control that v exceeds on the communication in the graph
- Also applicable to disonnected graphs
- Algorithm by Ulrik Brandes computes SPB in $O(|V||E| + |V|^2 \log|V|)$ time

Shortest Paths: Shortest Path Betweenness

- Define c_SPB for edges analogously

$$c(e) = \sum_{a \in V} \sum_{b \in V} \delta_{ab}(e)$$

- Possible: Interpret quantity $\delta_{ab}(v)$ as general relative information flow through v ("rush")
- Other variants: Instead of shortest paths between a and b regard
  - the set of all paths
  - the set of the k-shortest paths (interesting for social case; choose small k)
  - the set of the k-shortest node disjoint paths
  - the set of paths not longer than $(1+\varepsilon)d(a,b)$

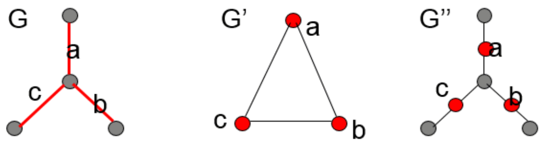
Deriving edge centralities from vertex centralities

- What we have seen so far: Various centrality measures mostly for vertices (based on degree, closeness, betweenness)
- → Formal way to translate a given vertex centrality index to a corresponding edge centrality: Apply the given vertex centrality to a transformed version of G, the edge graph
- Given original G =(V,E) then the edge graph G' = (E,K) is defined by taking original edges as vertices. Two original edges are connected in G' if they are originally incident to the same original node.
- Size of G' may be quadratic (w.r.t. number of nodes) compared to G

Deriving edge centralities from vertex centralities

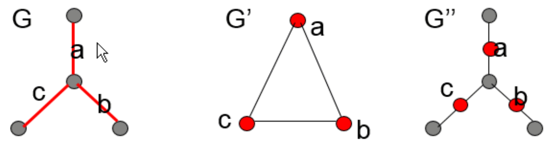
- What we have seen so far: Various centrality measures mostly for vertices (based on degree, closeness, betweenness)
- → Formal way to translate a given vertex centrality index to a corresponding edge centrality: Apply the given vertex centrality to a transformed version of G, the edge graph
- Given original G =(V,E) then the edge graph G' = (E,K) is defined by taking original edges as vertices. Two original edges are connected in G' if they are originally incident to the same original node.
- Size of G' may be quadratic (w.r.t. number of nodes) compared to G

Deriving edge centralities from vertex centralities



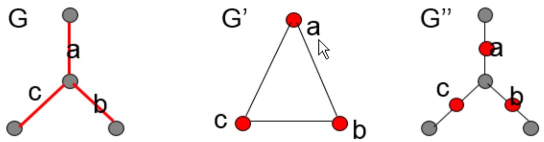
- Remember: Vertex stress centrality for node x: Number of shortest paths that use x; Straightforward version for edge e: Number of shortest paths that use e;
- → Upper Example: G: Stress centrality of edge a would be 3; But in edge graph G' stress centrality of original edge a (now a node) is 0.
- → Formal translations of vertex centrality indices to edge centralities with edge graphs are not well suited for all purposes
- → Introduce incidence graph G'': Each original edge is split by new "edge vertex" that represents the edge → Now vertex indices can be applied, preserving the intuition.

Deriving edge centralities from vertex centralities



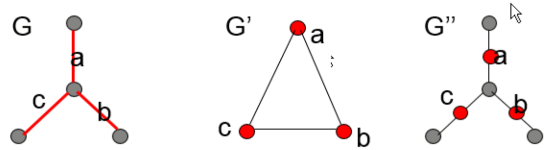
- Remember: Vertex stress centrality for node x: Number of shortest paths that use x; Straightforward version for edge e: Number of shortest paths that use e;
- → Upper Example: G: Stress centrality of edge a would be 3; But in edge graph G' stress centrality of original edge a (now a node) is 0.
- → Formal translations of vertex centrality indices to edge centralities with edge graphs are not well suited for all purposes
- → Introduce incidence graph G'': Each original edge is split by new "edge vertex" that represents the edge → Now vertex indices can be applied, preserving the intuition.

Deriving edge centralities from vertex centralities



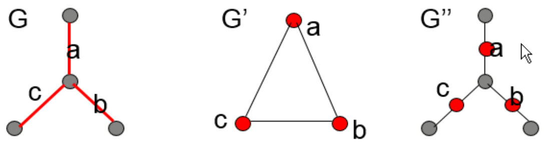
- Remember: Vertex stress centrality for node x: Number of shortest paths that use x; Straightforward version for edge e: Number of shortest paths that use e;
- → Upper Example: G: Stress centrality of edge a would be 3; But in edge graph G’ stress centrality of original edge a (now a node) is 0.
- → Formal translations of vertex centrality indices to edge centralities with edge graphs are not well suited for all purposes
- → Introduce incidence graph G’’: Each original edge is split by new “edge vertex” that represents the edge → Now vertex indices can be applied, preserving the intuition.

Deriving edge centralities from vertex centralities



- Remember: Vertex stress centrality for node x: Number of shortest paths that use x; Straightforward version for edge e: Number of shortest paths that use e;
- → Upper Example: G: Stress centrality of edge a would be 3; But in edge graph G’ stress centrality of original edge a (now a node) is 0.
- → Formal translations of vertex centrality indices to edge centralities with edge graphs are not well suited for all purposes
- → Introduce incidence graph G’’: Each original edge is split by new “edge vertex” that represents the edge → Now vertex indices can be applied, preserving the intuition.

Deriving edge centralities from vertex centralities



- Remember: Vertex stress centrality for node x: Number of shortest paths that use x; Straightforward version for edge e: Number of shortest paths that use e;
- → Upper Example: G: Stress centrality of edge a would be 3; But in edge graph G’ stress centrality of original edge a (now a node) is 0.
- → Formal translations of vertex centrality indices to edge centralities with edge graphs are not well suited for all purposes
- → Introduce incidence graph G’’: Each original edge is split by new “edge vertex” that represents the edge → Now vertex indices can be applied, preserving the intuition.

Vitality

- Intuition: Measure importance of vertex (or edge) by the difference of a given quality measure q on G with or without the vertex (edge):
  - → Vitality v(x) of graph element x : v(x) = q(G) - q(G\{x})
- Example 1 for quality measure q: Flow:
  - Given directed graph G with positive edge weights w modeling capacities. The flow $f(s,t)$ from node s (source) to node t (sink) is defined as:

$$f(s,t) = \sum_{e \in \{Out-Edges\ of\ s\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ t\}} \tilde{f}(e)$$

where the local flows $\tilde{f}$ respect capacity contraints: $0 \le \tilde{f}(e) \le w(e)$ and balance conditions:

$$\forall v \in V \setminus \{s,t\}: \sum_{e \in \{Out-Edges\ of\ v\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ v\}} \tilde{f}(e)$$

Vitality

- Intuition: Measure importance of vertex (or edge) by the difference of a given quality measure q on G with or without the vertex (edge):
  - → Vitality v(x) of graph element x : v(x) = q(G) - q(G\{x})

- Example 1 for quality measure q: Flow:
  - Given directed graph G with positive edge weights w modeling capacities. The flow $f(s,t)$ from node s (source) to node t (sink) is defined as:

$$f(s,t) = \sum_{e \in \{Out-Edges\ of\ s\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ t\}} \tilde{f}(e)$$

where the local flows $\tilde{f}$ respect capacity contraints: $0 \le \tilde{f}(e) \le w(e)$ and balance conditions:

$$\forall v \in V \setminus \{s,t\}: \sum_{e \in \{Out-Edges\ of\ v\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ v\}} \tilde{f}(e)$$

Vitality

- Intuition: Measure importance of vertex (or edge) by the difference of a given quality measure q on G with or without the vertex (edge):
  - → Vitality v(x) of graph element x : v(x) = q(G) - q(G\{x})

- Example 1 for quality measure q: Flow:
  - Given directed graph G with positive edge weights w modeling capacities. The flow $f(s,t)$ from node s (source) to node t (sink) is defined as:

$$f(s,t) = \sum_{e \in \{Out-Edges\ of\ s\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ t\}} \tilde{f}(e)$$

where the local flows $\tilde{f}$ respect capacity contraints: $0 \le \tilde{f}(e) \le w(e)$ and balance conditions:

$$\forall v \in V \setminus \{s,t\}: \sum_{e \in \{Out-Edges\ of\ v\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ v\}} \tilde{f}(e)$$

Vitality

- Intuition: Measure importance of vertex (or edge) by the difference of a given quality measure q on G with or without the vertex (edge):
  - → Vitality v(x) of graph element x : v(x) = q(G) - q(G\{x})

- Example 1 for quality measure q: Flow:
  - Given directed graph G with positive edge weights w modeling capacities. The flow $f(s,t)$ from node s (source) to node t (sink) is defined as:

$$f(s,t) = \sum_{e \in \{Out-Edges\ of\ s\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ t\}} \tilde{f}(e)$$

where the local flows $\tilde{f}$ respect capacity contraints: $0 \le \tilde{f}(e) \le w(e)$ and balance conditions:

$$\forall v \in V \setminus \{s,t\}: \sum_{e \in \{Out-Edges\ of\ v\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ v\}} \tilde{f}(e)$$

Vitality

- Intuition: Measure importance of vertex (or edge) by the difference of a given quality measure q on G with or without the vertex (edge):
  - → Vitality v(x) of graph element x : v(x) = q(G) - q(G\{x})

- Example 1 for quality measure q: Flow:
  - Given directed graph G with positive edge weights w modeling capacities. The flow $f(s,t)$ from node s (source) to node t (sink) is defined as:

$$f(s,t) = \sum_{e \in \{Out-Edges\ of\ s\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ t\}} \tilde{f}(e)$$

where the local flows $\tilde{f}$ respect capacity contraints: $0 \le \tilde{f}(e) \le w(e)$ and balance conditions:

$$\forall v \in V \setminus \{s,t\}: \sum_{e \in \{Out-Edges\ of\ v\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ v\}} \tilde{f}(e)$$

Vitality

- Intuition: Measure importance of vertex (or edge) by the difference of a given quality measure q on G with or without the vertex (edge):
  - → Vitality v(x) of graph element x : v(x) = q(G) - q(G\{x})

- Example 1 for quality measure q: Flow:
  - Given directed graph G with positive edge weights w modeling capacities. The flow $f(s,t)$ from node s (source) to node t (sink) is defined as:

$$f(s,t) = \sum_{e \in \{Out-Edges\ of\ s\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ t\}} \tilde{f}(e)$$

where the local flows $\tilde{f}$ respect capacity contraints: $0 \leq \tilde{f}(e) \leq w(e)$
and balance conditions:

$$\forall v \in V \setminus \{s,t\}: \sum_{e \in \{Out-Edges\ of\ v\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ v\}} \tilde{f}(e)$$

Vitality

- Intuition: Measure importance of vertex (or edge) by the difference of a given quality measure q on G with or without the vertex (edge):
  - → Vitality v(x) of graph element x : v(x) = q(G) - q(G\{x})

- Example 1 for quality measure q: Flow:
  - Given directed graph G with positive edge weights w modeling capacities. The flow $f(s,t)$ from node s (source) to node t (sink) is defined as:

$$f(s,t) = \sum_{e \in \{Out-Edges\ of\ s\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ t\}} \tilde{f}(e)$$

where the local flows $\tilde{f}$ respect capacity contraints: $0 \leq \tilde{f}(e) \leq w(e)$
and balance conditions:

$$\forall v \in V \setminus \{s,t\}: \sum_{e \in \{Out-Edges\ of\ v\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ v\}} \tilde{f}(e)$$

Vitality

- Intuition: Measure importance of vertex (or edge) by the difference of a given quality measure q on G with or without the vertex (edge):
  - → Vitality v(x) of graph element x : v(x) = q(G) - q(G\{x})

- Example 1 for quality measure q: Flow:
  - Given directed graph G with positive edge weights w modeling capacities. The flow $f(s,t)$ from node s (source) to node t (sink) is defined as:

$$f(s,t) = \sum_{e \in \{Out-Edges\ of\ s\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ t\}} \tilde{f}(e)$$

where the local flows $\tilde{f}$ respect capacity contraints: $0 \leq \tilde{f}(e) \leq w(e)$
and balance conditions:

$$\forall v \in V \setminus \{s,t\}: \sum_{e \in \{Out-Edges\ of\ v\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ v\}} \tilde{f}(e)$$

Vitality

- Intuition: Measure importance of vertex (or edge) by the difference of a given quality measure q on G with or without the vertex (edge):
  - → Vitality v(x) of graph element x : v(x) = q(G) - q(G\{x})

- Example 1 for quality measure q: Flow:
  - Given directed graph G with positive edge weights w modeling capacities. The flow $f(s,t)$ from node s (source) to node t (sink) is defined as:

$$f(s,t) = \sum_{e \in \{Out-Edges\ of\ s\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ t\}} \tilde{f}(e)$$

where the local flows $\tilde{f}$ respect capacity contraints: $0 \leq \tilde{f}(e) \leq w(e)$
and balance conditions:

$$\forall v \in V \setminus \{s,t\}: \sum_{e \in \{Out-Edges\ of\ v\}} \tilde{f}(e) = \sum_{e \in \{In-Edges\ of\ v\}} \tilde{f}(e)$$

Vitality

- Computing a flow f: E → ℝ of maximum value (tweaking the local flows): O(|V| |E| log(|V|²/|E|)) (Algorithm by Goldberg & Tarjan (see [2]))
- Now define quality measure by e.g.:

$$q(G) = \sum_{s,t \in V} \max f(s,t)$$

- Social analog of flow: Workflow, Information-flow, “Doing favors flow” etc.

Vitality

- Besides vitality-based centrality c(x) = v(x) = q(G) - q(G\{x}) we may also define a centrality as max-flow betweenness:
denote: $f_{st}(G) = \max_G f(s,t)$
we may then define:

$$c(u) = \sum_{s,t \in V; u \neq s,t} \frac{f_{st}(G) - f_{st}(G \setminus \{u\})}{f_{st}(G)}$$

- The numerator denotes the amount of flow that must go through node u

Vitality

- Example 2: Mobile (Peer to Peer) communication-network: Each node should be connected to each other node by as few intermediaries as possible. → quality measure: Wiener Index

$$q(G) = \sum_{v \in V} \sum_{w \in V} d(v,w)$$

- Possible: write Wiener Index with the help of closeness centrality $c_c(v)$

$$q(G) = \sum_{v \in V} \frac{1}{c_c(v)}$$

- Define centrality „closeness vitality“ of graph element x as vitality:

$$c(x) = q(G) - q(G \setminus \{x\})$$

Vitality

- Example 2: Mobile (Peer to Peer) communication-network: Each node should be connected to each other node by as few intermediaries as possible. → quality measure: Wiener Index

$$q(G) = \sum_{v \in V} \sum_{w \in V} d(v,w)$$

- Possible: write Wiener Index with the help of closeness centrality $c_c(v)$

$$q(G) = \sum_{v \in V} \frac{1}{c_c(v)}$$

- Define centrality „closeness vitality“ of graph element x as vitality:

$$c(x) = q(G) - q(G \setminus \{x\})$$

Vitality

- Example 2: Mobile (Peer to Peer) communication-network: Each node should be connected to each other node by as few intermediaries as possible. → quality measure: Wiener Index

$$q(G) = \sum_{v \in V} \sum_{w \in V} d(v,w)$$

- Possible: write Wiener Index with the help of closeness centrality $c_c(v)$

$$q(G) = \sum_{v \in V} \frac{1}{c_c(v)}$$

- Define centrality „closeness vitality" of graph element x as vitality:

$$c(x) = q(G) - q(G \setminus \{x\})$$

Stress Centrality as Vitality

- We had: stress centrality of v or e is equal to number of shortest paths through v or e

$$c_{stress}(v) = \sum_{a \in V; a \neq v} \sum_{b \in V; b \neq v} \sigma_{ab}(v) \qquad c_{stress}(e) = \sum_{a \in V} \sum_{b \in V} \sigma_{ab}(e)$$

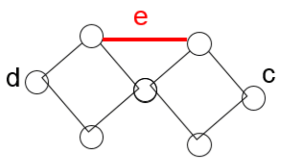
- Intuition: $c_{stress}(v)$ seems to measure the number of shortest paths that would be lost if v wasn't avaliable any more

- Why can't we directly use c_{stress} as a graph quality index to construct a vitality index ?

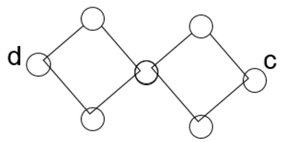
- →Because actual number of shortest paths can INCREASE if e.g. edge is taken away

Stress Centrality as Vitality

- Example



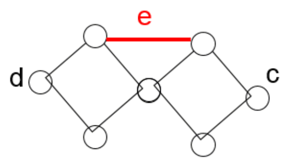
- Number of shortest paths: 54
- Number of shortest paths containing e: 8
- σ_{cd}=1 (length 3)



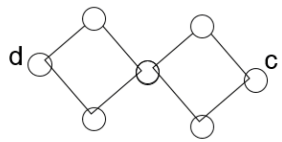
- Number of shortest paths: 64
  (18 of them have increased in length)
- σ_{cd}=4 (length 4)

Stress Centrality as Vitality

- Example



- Number of shortest paths: 54
- Number of shortest paths containing e: 8
- σ_{cd}=1 (length 3)



- Number of shortest paths: 64
  (18 of them have increased in length)
- σ_{cd}=4 (length 4)

Stress Centrality as Vitality

- → In order to define a vitality-like version of stress: Consider only those shortest paths that haven't changed their length:

$$c_{vitality}(v,G) = c_{stress}(v,G) - c_{stress}(v,G \setminus \{v\})$$

with

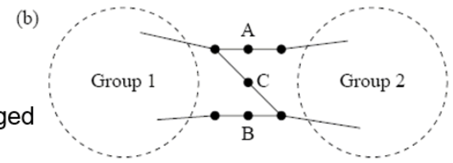
$$c_{stress}(v,G \setminus \{v\}) = \sum_{a \in V; a \neq v} \sum_{b \in V; b \neq v} \sigma_{ab}[d_G(a,b) = d_{G\setminus\{v\}}(a,b)]$$

(Iverson notation)

Critique on Betweenness Based Centralities

- major critique: Max-Flow betweenness centrality (suggested to counteract this drawback) may exhibit similar problems

- here: special Max-Flow betweenness centrality mfb:
  - -- limit edge capacity to one
  - -- mfb(i) := maximum possible flow through i over all possible solutions to the s-t-maximum flow problem, averaged over all s and t.



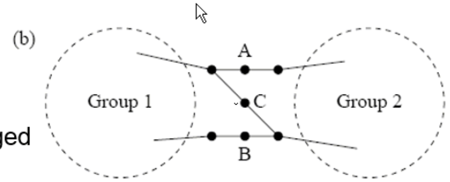
(b) In calculations of flow betweenness, vertices A and B in this configuration will get high scores while vertex C will not.

Source: [5]

Critique on Betweenness Based Centralities

- major critique: Max-Flow betweenness centrality (suggested to counteract this drawback) may exhibit similar problems

- here: special Max-Flow betweenness centrality mfb:
  - -- limit edge capacity to one
  - -- mfb(i) := maximum possible flow through i over all possible solutions to the s-t-maximum flow problem, averaged over all s and t.



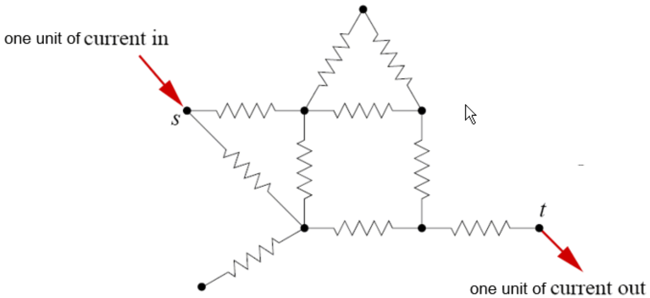
(b) In calculations of flow betweenness, vertices A and B in this configuration will get high scores while vertex C will not.

Source: [5]

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- flow of electric current in a resistor network; V_i = voltage (potential) at vertex i

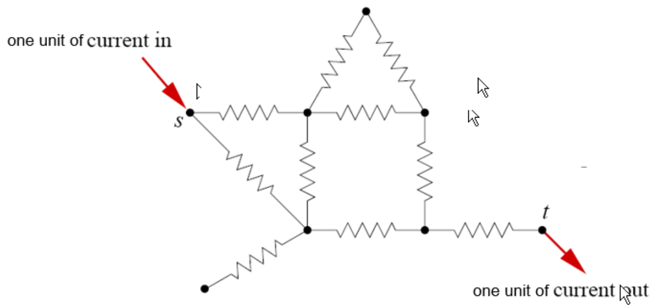
- ↔ Current Flow betweenness cfb centrality : cfb(i) := amount of current that flows through i in this setup, averaged over all s and t.



Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- flow of electric current in a resistor network;
V_i = voltage (potential) at vertex i

- ↔ Current Flow betweenness cfb centrality : cfb(i) := amount of current that flows through i in this setup, averaged over all s and t.



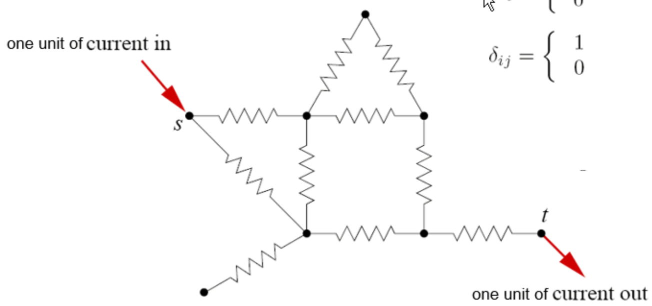
Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- Kirchhoffs point law (current conservation): total current flow in / out of node is zero:

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it},$$

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between } i \text{ and } j, \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$



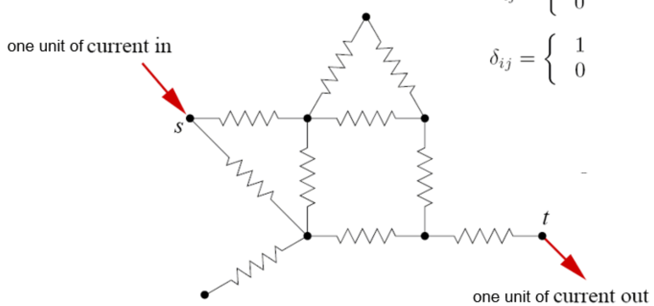
Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- Kirchhoffs point law (current conservation): total current flow in / out of node is zero:

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it},$$

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between } i \text{ and } j, \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$



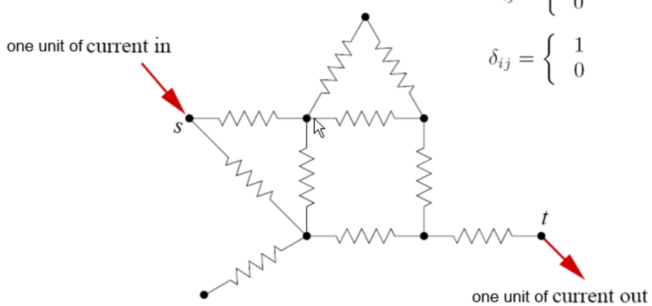
Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- Kirchhoffs point law (current conservation): total current flow in / out of node is zero:

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it},$$

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between } i \text{ and } j, \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$



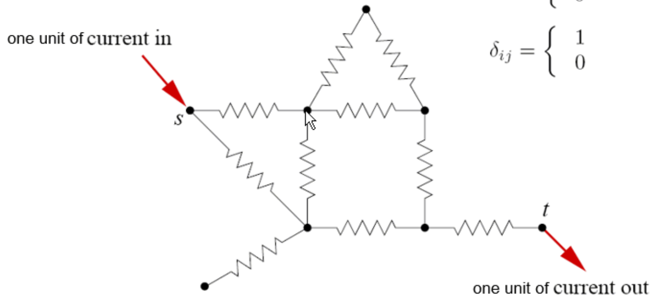
Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- Kirchhoffs point law (current conservation): total current flow in / out of node is zero:

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it},$$

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between } i \text{ and } j, \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$



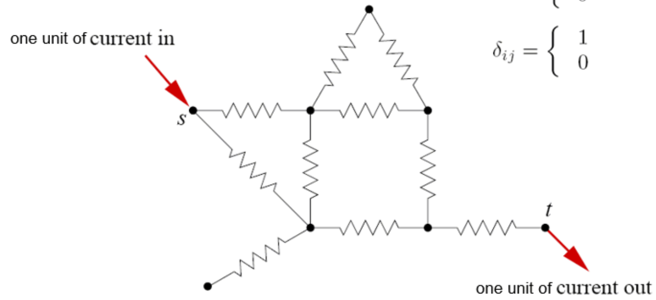
Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- Kirchhoffs point law (current conservation): total current flow in / out of node is zero:

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it},$$

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between } i \text{ and } j, \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$



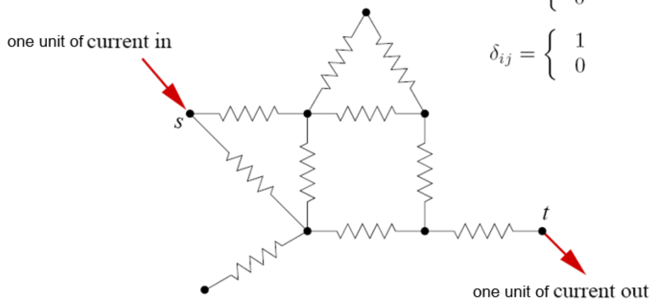
Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- Kirchhoffs point law (current conservation): total current flow in / out of node is zero:

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it},$$

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between } i \text{ and } j, \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$



Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$\sum_j A_{ij} = k_i$, the degree of vertex i

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it} \longrightarrow \underbrace{(\mathbf{D} - \mathbf{A})}_{\text{“Graph Laplacian”}} \cdot \mathbf{V} = \mathbf{s}$$

$\mathbf{D}$ is the diagonal matrix with elements $D_{ii} = k_i$

$$\text{source vector } \mathbf{s} \qquad s_i = \begin{cases} +1 & \text{for } i = s, \\ -1 & \text{for } i = t, \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbf{V} = (\mathbf{D}_v - \mathbf{A}_v)^{-1} \cdot \mathbf{s}$$

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$\sum_j A_{ij} = k_i$, the degree of vertex i

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it} \longrightarrow \underbrace{(\mathbf{D} - \mathbf{A})}_{\text{"Graph Laplacian"}} \cdot \mathbf{V} = \mathbf{s}$$

$\mathbf{D}$ is the diagonal matrix with elements $D_{ii} = k_i$

source vector $\mathbf{s}$ $\quad s_i = \begin{cases} +1 & \text{for } i = s, \\ -1 & \text{for } i = t, \\ 0 & \text{otherwise.} \end{cases}$

$$\mathbf{V} = (\mathbf{D}_v - \mathbf{A}_v)^{-1} \cdot \mathbf{s}$$

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$\sum_j A_{ij} = k_i$, the degree of vertex i

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it} \longrightarrow \underbrace{(\mathbf{D} - \mathbf{A})}_{\text{"Graph Laplacian"}} \cdot \mathbf{V} = \mathbf{s}$$

$\mathbf{D}$ is the diagonal matrix with elements $D_{ii} = k_i$

source vector $\mathbf{s}$ $\quad s_i = \begin{cases} +1 & \text{for } i = s, \\ -1 & \text{for } i = t, \\ 0 & \text{otherwise.} \end{cases}$

$$\mathbf{V} = (\mathbf{D}_v - \mathbf{A}_v)^{-1} \cdot \mathbf{s}$$

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$\sum_j A_{ij} = k_i$, the degree of vertex i

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it} \longrightarrow \underbrace{(\mathbf{D} - \mathbf{A})}_{\text{"Graph Laplacian"}} \cdot \mathbf{V} = \mathbf{s}$$

$\mathbf{D}$ is the diagonal matrix with elements $D_{ii} = k_i$

source vector $\mathbf{s}$ $\quad s_i = \begin{cases} +1 & \text{for } i = s, \\ -1 & \text{for } i = t, \\ 0 & \text{otherwise.} \end{cases}$

$$\mathbf{V} = (\mathbf{D}_v - \mathbf{A}_v)^{-1} \cdot \mathbf{s}$$

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$\sum_j A_{ij} = k_i$, the degree of vertex i

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it} \longrightarrow \underbrace{(\mathbf{D} - \mathbf{A})}_{\text{"Graph Laplacian"}} \cdot \mathbf{V} = \mathbf{s}$$

$\mathbf{D}$ is the diagonal matrix with elements $D_{ii} = k_i$

source vector $\mathbf{s}$ $\quad s_i = \begin{cases} +1 & \text{for } i = s, \\ -1 & \text{for } i = t, \\ 0 & \text{otherwise.} \end{cases}$

$$\mathbf{V} = (\mathbf{D}_v - \mathbf{A}_v)^{-1} \cdot \mathbf{s}$$

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$$\underbrace{(\mathbf{D}-\mathbf{A})}\cdot\mathbf{V}=\mathbf{s}$$

Laplacian is not invertible, det = 0, because system of eq. is overdetermined → set one V_v=0 and measure voltages relative to v. → remove the v-th row and column (since V_v=0) → now invertible

$$\mathbf{V}=(\mathbf{D}_v-\mathbf{A}_v)^{-1}\cdot\mathbf{s}$$ (matrix inversion: $O(n^3)$)

let us now add a vth row and column back into $(\mathbf{D}_v-\mathbf{A}_v)^{-1}$ with values all equal to zero.
The resulting matrix we will denote $\mathbf{T}$.

$\longrightarrow\ V_i^{(st)}=T_{is}-T_{it}$

$\longrightarrow$ current flow at node i: $I_i^{(st)}=\frac{1}{2}\sum_j A_{ij}|V_i^{(st)}-V_j^{(st)}|$

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$$\underbrace{(\mathbf{D}-\mathbf{A})}\cdot\mathbf{V}=\mathbf{s}$$

Laplacian is not invertible, det = 0, because system of eq. is overdetermined → set one V_v=0 and measure voltages relative to v. → remove the v-th row and column (since V_v=0) → now invertible

$$\mathbf{V}=(\mathbf{D}_v-\mathbf{A}_v)^{-1}\cdot\mathbf{s}$$ (matrix inversion: $O(n^3)$)

let us now add a vth row and column back into $(\mathbf{D}_v-\mathbf{A}_v)^{-1}$ with values all equal to zero.
The resulting matrix we will denote $\mathbf{T}$.

$\longrightarrow\ V_i^{(st)}=T_{is}-T_{it}$

$\longrightarrow$ current flow at node i: $I_i^{(st)}=\frac{1}{2}\sum_j A_{ij}|V_i^{(st)}-V_j^{(st)}|$

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$$\underbrace{(\mathbf{D}-\mathbf{A})}\cdot\mathbf{V}=\mathbf{s}$$

Laplacian is not invertible, det = 0, because system of eq. is overdetermined → set one V_v=0 and measure voltages relative to v. → remove the v-th row and column (since V_v=0) → now invertible

$$\mathbf{V}=(\mathbf{D}_v-\mathbf{A}_v)^{-1}\cdot\mathbf{s}$$ (matrix inversion: $O(n^3)$)

let us now add a vth row and column back into $(\mathbf{D}_v-\mathbf{A}_v)^{-1}$ with values all equal to zero.
The resulting matrix we will denote $\mathbf{T}$.

$\longrightarrow\ V_i^{(st)}=T_{is}-T_{it}$

$\longrightarrow$ current flow at node i: $I_i^{(st)}=\frac{1}{2}\sum_j A_{ij}|V_i^{(st)}-V_j^{(st)}|$

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$$\underbrace{(\mathbf{D}-\mathbf{A})}\cdot\mathbf{V}=\mathbf{s}$$

Laplacian is not invertible, det = 0, because system of eq. is overdetermined → set one V_v=0 and measure voltages relative to v. → remove the v-th row and column (since V_v=0) → now invertible

$$\mathbf{V}=(\mathbf{D}_v-\mathbf{A}_v)^{-1}\cdot\mathbf{s}$$ (matrix inversion: $O(n^3)$)

let us now add a vth row and column back into $(\mathbf{D}_v-\mathbf{A}_v)^{-1}$ with values all equal to zero.
The resulting matrix we will denote $\mathbf{T}$.

$\longrightarrow\ V_i^{(st)}=T_{is}-T_{it}$

$\longrightarrow$ current flow at node i: $I_i^{(st)}=\frac{1}{2}\sum_j A_{ij}|V_i^{(st)}-V_j^{(st)}|$

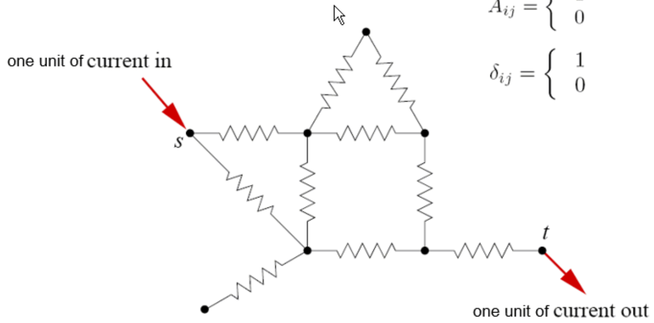
Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- Kirchhoffs point law (current conservation): total current flow in / out of node is zero:

$$\sum_j A_{ij}(V_i - V_j) = \delta_{is} - \delta_{it},$$

$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between } i \text{ and } j, \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$



Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$$\underbrace{(\mathbf{D} - \mathbf{A})}\cdot \mathbf{V} = \mathrm{s}$$

Laplacian is not invertible, det = 0, because system of eq. is overdetermined → set one V_v=0 and measure voltages relative to v. → remove the v-th row and column (since V_v=0) → now invertible

$$\mathbf{V} = (\mathbf{D}_v - \mathbf{A}_v)^{-1} \cdot \mathrm{s}$$ (matrix inversion: O(n^3))

let us now add a vth row and column back into $(\mathbf{D}_v - \mathbf{A}_v)^{-1}$ with values all equal to zero.
The resulting matrix we will denote $\mathbf{T}$.

$$\longrightarrow \quad V_i^{(st)} = T_{is} - T_{it}$$

$\longrightarrow$ current flow at node i: $I_i^{(st)} = \frac{1}{2}\sum_j A_{ij}|V_i^{(st)} - V_j^{(st)}|$

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

$$\underbrace{(\mathbf{D} - \mathbf{A})}\cdot \mathbf{V} = \mathrm{s}$$

Laplacian is not invertible, det = 0, because system of eq. is overdetermined → set one V_v=0 and measure voltages relative to v. → remove the v-th row and column (since V_v=0) → now invertible

$$\mathbf{V} = (\mathbf{D}_v - \mathbf{A}_v)^{-1} \cdot \mathrm{s}$$ (matrix inversion: O(n^3))

let us now add a vth row and column back into $(\mathbf{D}_v - \mathbf{A}_v)^{-1}$ with values all equal to zero.
The resulting matrix we will denote $\mathbf{T}$.

$$\longrightarrow \quad V_i^{(st)} = T_{is} - T_{it}$$

$\longrightarrow$ current flow at node i: $I_i^{(st)} = \frac{1}{2}\sum_j A_{ij}|V_i^{(st)} - V_j^{(st)}|$

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- current flow at node i:

$$I_i^{(st)} = \frac{1}{2}\sum_j A_{ij}|V_i^{(st)} - V_j^{(st)}|$$
$$= \frac{1}{2}\sum_j A_{ij}|T_{is} - T_{it} - T_{js} + T_{jt}|, \quad \text{for } i \neq s, t.$$

- unit current flow at nodes s and t:

$$I_s^{(st)} = 1, \qquad I_t^{(st)} = 1.$$

- cfb(i) (denoted as b_i) is then:

$$b_i = \frac{\sum_{s<t} I_i^{(st)}}{\frac{1}{2}n(n-1)}.$$

(takes O(m n^2) for all i) →
(plus matrix inversion:)
O((m+n) n^2) for everything

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- current flow at node i:

$$\begin{aligned} I_i^{(st)} &= \tfrac{1}{2}\sum_j A_{ij}|V_i^{(st)} - V_j^{(st)}| \\ &= \tfrac{1}{2}\sum_j A_{ij}|T_{is} - T_{it} - T_{js} + T_{jt}|, \quad \text{for } i \neq s,t. \end{aligned}$$

- unit current flow at nodes s and t:

$$I_s^{(st)} = 1, \qquad I_t^{(st)} = 1.$$

- cfb(i) (denoted as b_i) is then:

$$b_i = \frac{\sum_{s<t} I_i^{(st)}}{\frac{1}{2}n(n-1)}.$$

(takes O(m n^2) for all i) →
(plus matrix inversion:)
O((m+n) n^2) for everything

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- current flow at node i:

$$\begin{aligned} I_i^{(st)} &= \tfrac{1}{2}\sum_j A_{ij}|V_i^{(st)} - V_j^{(st)}| \\ &= \tfrac{1}{2}\sum_j A_{ij}|T_{is} - T_{it} - T_{js} + T_{jt}|, \quad \text{for } i \neq s,t. \end{aligned}$$

- unit current flow at nodes s and t:

$$I_s^{(st)} = 1, \qquad I_t^{(st)} = 1.$$

- cfb(i) (denoted as b_i) is then:

$$b_i = \frac{\sum_{s<t} I_i^{(st)}}{\frac{1}{2}n(n-1)}.$$

(takes O(m n^2) for all i) →
(plus matrix inversion:)
O((m+n) n^2) for everything

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- current flow at node i:

$$\begin{aligned} I_i^{(st)} &= \tfrac{1}{2}\sum_j A_{ij}|V_i^{(st)} - V_j^{(st)}| \\ &= \tfrac{1}{2}\sum_j A_{ij}|T_{is} - T_{it} - T_{js} + T_{jt}|, \quad \text{for } i \neq s,t. \end{aligned}$$

- unit current flow at nodes s and t:

$$I_s^{(st)} = 1, \qquad I_t^{(st)} = 1.$$

- cfb(i) (denoted as b_i) is then:

$$b_i = \frac{\sum_{s<t} I_i^{(st)}}{\frac{1}{2}n(n-1)}.$$

(takes O(m n^2) for all i) →
(plus matrix inversion:)
O((m+n) n^2) for everything

Random Walk Centrality == Current Flow Btw. Centrality (see [5])

- cfb == random walk betweenness centrality (rwb):

- rwb(i): move around „messages“: start (absorbing) random walk at s, end at t:

  rwb(i):= net number of times that a message passes through i on its journey (averaged over a large number of trials and averaged over s, t)

  („net“ number of times: „cancel back and fourth passes“)

- if in i, probability that in next step j:

$$M_{ij} = \frac{A_{ij}}{k_j}, \qquad \text{for } j \neq t,$$

$$\mathbf{M} = \mathbf{A} \cdot \mathbf{D}^{-1}$$

with D = diag(k_i)

$$D_{ii} = k_i$$