Script generated by TTT

Title: groh: profile1 (06.05.2015)

Date: Wed May 06 08:13:57 CEST 2015

Duration: 91:37 min

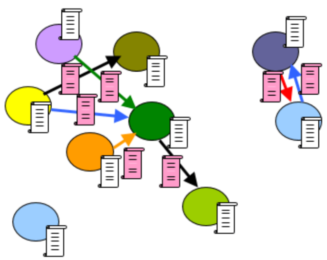
Pages: 55

Long-Term Social Context: Social Networks

Social Network

slightly **refined Social Network Model**: Graph $G=(V,E,P_V,P_E, f_{P_V},f_{P_E})$

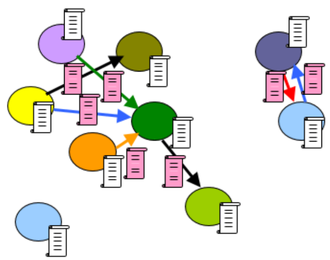
- Nodes $V = \biguplus V_i$: represent humans (actors) of "sorts" (←→ modes) V_i;
- Edges $E \subseteq V \times V$; $E = \biguplus E_i$: represent directed binary social relations (ties) of "sorts" E_i
- P_V: Set of Node Profiles
- P_E: Set of Edge Profiles
- f_{P_V}: $V \rightarrow P_V$
- f_{P_E}: $V \rightarrow P_E$



Social Network

slightly **refined Social Network Model**: Graph $G=(V,E,P_V,P_E, f_{P_V},f_{P_E})$

- Nodes $V = \biguplus V_i$: represent humans (actors) of "sorts" (←→ modes) V_i;
- Edges $E \subseteq V \times V$; $E = \biguplus E_i$: represent directed binary social relations (ties) of "sorts" E_i
- P_V: Set of Node Profiles
- P_E: Set of Edge Profiles
- f_{P_V}: $V \rightarrow P_V$
- f_{P_E}: $V \rightarrow P_E$



6 degrees of separation

av. path length in real world SN is ~ 6

- First occurrence of a claim similar to 6 degrees: G. Marconi (Italian Physicist & Nobel Prize laureate) 1909: Number of radio stations necc. to cover inhabited world → any transmission path needs about 6 stations
- 1920s: Hungarian writer F. Karinthy claims six degrees of separation in Budapest in a short story (prob. inspired by Marconi)
- Most famous: S. Milgram (Inspired by unpublished paper by M. Kochen & I. de Sola Pool claiming ~3 degrees in USA): "Small world experiment" [20]. Randomly chosen people: mail letter to target person; track record → ~ 6 av. Path length

6 degrees of separation

- Popular culture: Erdös number / Kevin-Bacon Number / Erdös-Bacon-Number
- Several newer experiments (see [20], [21]) on degree of separation on the web (Facebook, Email-studies (D. Watts, Columbia U.) etc.) also showed degree of separation ~ 6
- More thorough mathematical investigation → Random Graph Theory
- Watts and Strogatz [22] : Small World graph (informal): high clustering coefficient, small mean av. Path-length

Technical intermezzo: Clustering coefficient

- Undirected Graph: Clustering Coefficient C_i of node v_i : Measures "how close" v_i and its neighbors $\{v_j \in N_i\}$ (where neighborhood N_i is $\{v_j \mid \{v_i, v_j\} \in E;\ E \subseteq \binom{V}{2}\}$) are to a complete subgraph (clique):

$$C_i = \frac{|\{e_{\{kj\}} \mid v_k, v_j \in N_i\}|}{\frac{d_i(d_i-1)}{2}}$$

Degree d_i of node v_i : $d_i=|N_i|$

- Directed Graph: Clustering Coefficient C_i of node v_i : Measures "how close" v_i 's neighbors $\{v_j \in N_i = N_i^{out} \cup N_i^{in}\}$ (where out-neighborhood N^{out}_i is $\{v_j \mid (v_i, v_j) \in E;\ E \subseteq V \times V\}$ and in-neighborhood N^{in}_i is $\{v_j \mid (v_j, v_i) \in E;\ E \subseteq V \times V\}$) are to a complete subgraph (clique):

$$C_i = \frac{|\{e_{kj} \mid v_k, v_j \in N_i\}|}{d_i(d_i-1)}$$

Degree d_i of node v_i : $d_i=|N_i|$

Technical intermezzo: Clustering coefficient

- Undirected Graph: Clustering Coefficient C_i of node v_i : Measures "how close" v_i and its neighbors $\{v_j \in N_i\}$ (where neighborhood N_i is $\{v_j \mid \{v_i, v_j\} \in E;\ E \subseteq \binom{V}{2}\}$) are to a complete subgraph (clique):

$$C_i = \frac{|\{e_{\{kj\}} \mid v_k, v_j \in N_i\}|}{\frac{d_i(d_i-1)}{2}}$$

Degree d_i of node v_i : $d_i=|N_i|$

- Directed Graph: Clustering Coefficient C_i of node v_i : Measures "how close" v_i 's neighbors $\{v_j \in N_i = N_i^{out} \cup N_i^{in}\}$ (where out-neighborhood N^{out}_i is $\{v_j \mid (v_i, v_j) \in E;\ E \subseteq V \times V\}$ and in-neighborhood N^{in}_i is $\{v_j \mid (v_j, v_i) \in E;\ E \subseteq V \times V\}$) are to a complete subgraph (clique):

$$C_i = \frac{|\{e_{kj} \mid v_k, v_j \in N_i\}|}{d_i(d_i-1)}$$

Degree d_i of node v_i : $d_i=|N_i|$

Technical intermezzo: Clustering coefficient

- Undirected Graph: Clustering Coefficient C_i of node v_i: Measures "how close" v_i and its neighbors $\{v_j \in N_i\}$ (where neighborhood N_i is $\{v_j \mid \{v_i, v_j\} \in E;\ E \subseteq \binom{V}{2}\}$) are to a complete subgraph (clique):

$$C_i = \frac{|\{e_{\{kj\}} \mid v_k, v_j \in N_i\}|}{\frac{d_i(d_i-1)}{2}}$$

Degree d_i of node v_i: $d_i=|N_i|$

- Directed Graph: Clustering Coefficient C_i of node v_i: Measures "how close" v_i 's neighbors $\{v_j \in N_i = N_i^{out} \cup N_i^{in}\}$ (where out-neighborhood N^{out}_i is $\{v_j \mid (v_i, v_j) \in E;\ E \subseteq V \times V\}$ and in-neighborhood N^{in}_i is $\{v_j \mid (v_j, v_i) \in E;\ E \subseteq V \times V\}$) are to a complete subgraph (clique):

$$C_i = \frac{|\{e_{kj} \mid v_k, v_j \in N_i\}|}{d_i(d_i-1)}$$

Degree d_i of node v_i: $d_i=|N_i|$

Technical intermezzo: Clustering coefficient

- Undirected Graph: Clustering Coefficient C_i of node v_i: Measures "how close" v_i and its neighbors $\{v_j \in N_i\}$ (where neighborhood N_i is $\{v_j \mid \{v_i, v_j\} \in E;\ E \subseteq \binom{V}{2}\}$) are to a complete subgraph (clique):

$$C_i = \frac{|\{e_{\{kj\}} \mid v_k, v_j \in N_i\}|}{\frac{d_i(d_i-1)}{2}}$$

Degree d_i of node v_i: $d_i=|N_i|$

- Directed Graph: Clustering Coefficient C_i of node v_i: Measures "how close" v_i 's neighbors $\{v_j \in N_i = N_i^{out} \cup N_i^{in}\}$ (where out-neighborhood N^{out}_i is $\{v_j \mid (v_i, v_j) \in E;\ E \subseteq V \times V\}$ and in-neighborhood N^{in}_i is $\{v_j \mid (v_j, v_i) \in E;\ E \subseteq V \times V\}$) are to a complete subgraph (clique):

$$C_i = \frac{|\{e_{kj} \mid v_k, v_j \in N_i\}|}{d_i(d_i-1)}$$

Degree d_i of node v_i: $d_i=|N_i|$

History of Social Network Analysis, Main Contributors

see e.g. [9]:

- Main contributing fields of science: Sociology (surprisingly ☺), Anthropology, Urban Studies, Mathematics (modeling & evaluation formalisms), Physics (large community (surprisingly)), Computer Science (graph algorithms etc.), Economic Sciences
- 1887: F. Tönnies (German sociologist): 2 basic "sorts" of groups: Gemeinschaft (Family, Friends etc.; supported by "Wesenwille") ←→ Gesellschaft (Goal oriented; (Firm, State etc.); supported by "Kürwille")
- 1911: G. Simmel (German sociologist): Sociability of humans (especially in larger cities): One of the first to impose a "social network" view

History of Social Network Analysis, Main Contributors

- 1930s-1950s: J. Moreno (American Psychiatrist & Sociologist): → Sociometry (quantitative method for measuring social relationships) [11]
- 1930s-1960s: Further contributors: W. Warner (Harvard U., Anthropologist) [12] :→ Native American social structures, E. Mayo (Harvard U., Sociologist) [13]: Hawthorne Studies; A. Radcliffe-Brown (Oxford U., Social Anthropologist): Structural Functionalism (←→ primitive civilizations); M. Gluckman (Manchester U., anthropologist): Urban studies; etc.
- 1960s-1970s-present: H. White (Columbia U. Mathematical Sociologist): Extremely influential contributor to formal SNA [14]; students: M. Granovetter, B. Wellman
- 2000s-present: A. Barabasi, D. Watts, M. Newman, J. Kleinberg: („Physicists take over"), A. Pentland (Reality Mining) etc.

History of Social Network Analysis, Main Contributors

- 1930s-1950s: J. Moreno (American Psychiatrist & Sociologist): → Sociometry (quantitative method for measuring social relationships) [11]
- 1930s-1960s: Further contributors: W. Warner (Harvard U., Anthropologist) [12] :→ Native American social structures, E. Mayo (Harvard U., Sociologist) [13]: Hawthorne Studies; A. Radcliffe-Brown (Oxford U., Social Anthropologist): Structural Functionalism (←→ primitive civilizations); M. Gluckman (Manchester U., anthropologist): Urban studies; etc.
- 1960s-1970s-present: H. White (Columbia U. Mathematical Sociologist): Extremely influential contributor to formal SNA [14]; students: M. Granovetter, B. Wellman
- 2000s-present: A. Barabasi, D. Watts, M. Newman, J. Kleinberg: („Physicists take over"), A. Pentland (Reality Mining) etc.

History of Social Network Analysis, Main Contributors

- 1930s-1950s: J. Moreno (American Psychiatrist & Sociologist): → Sociometry (quantitative method for measuring social relationships) [11]
- 1930s-1960s: Further contributors: W. Warner (Harvard U., Anthropologist) [12] :→ Native American social structures, E. Mayo (Harvard U., Sociologist) [13]: Hawthorne Studies; A. Radcliffe-Brown (Oxford U., Social Anthropologist): Structural Functionalism (←→ primitive civilizations); M. Gluckman (Manchester U., anthropologist): Urban studies; etc.
- 1960s-1970s-present: H. White (Columbia U. Mathematical Sociologist): Extremely influential contributor to formal SNA [14]; students: M. Granovetter, B. Wellman
- 2000s-present: A. Barabasi, D. Watts, M. Newman, J. Kleinberg: („Physicists take over"), A. Pentland (Reality Mining) etc.

History of Social Network Analysis, Main Contributors

- 1930s-1950s: J. Moreno (American Psychiatrist & Sociologist): → Sociometry (quantitative method for measuring social relationships) [11]
- 1930s-1960s: Further contributors: W. Warner (Harvard U., Anthropologist) [12] :→ Native American social structures, E. Mayo (Harvard U., Sociologist) [13]: Hawthorne Studies; A. Radcliffe-Brown (Oxford U., Social Anthropologist): Structural Functionalism (←→ primitive civilizations); M. Gluckman (Manchester U., anthropologist): Urban studies; etc.
- 1960s-1970s-present: H. White (Columbia U. Mathematical Sociologist): Extremely influential contributor to formal SNA [14]; students: M. Granovetter, B. Wellman
- 2000s-present: A. Barabasi, D. Watts, M. Newman, J. Kleinberg: („Physicists take over"), A. Pentland (Reality Mining) etc.

History of Social Network Analysis, Main Contributors

- 1930s-1950s: J. Moreno (American Psychiatrist & Sociologist): → Sociometry (quantitative method for measuring social relationships) [11]
- 1930s-1960s: Further contributors: W. Warner (Harvard U., Anthropologist) [12] :→ Native American social structures, E. Mayo (Harvard U., Sociologist) [13]: Hawthorne Studies; A. Radcliffe-Brown (Oxford U., Social Anthropologist): Structural Functionalism (←→ primitive civilizations); M. Gluckman (Manchester U., anthropologist): Urban studies; etc.
- 1960s-1970s-present: H. White (Columbia U. Mathematical Sociologist): Extremely influential contributor to formal SNA [14]; students: M. Granovetter, B. Wellman
- 2000s-present: A. Barabasi, D. Watts, M. Newman, J. Kleinberg: („Physicists take over"), A. Pentland (Reality Mining) etc.

Centrality

- Centrality indices formalize intuitive feeling that some nodes (or edges) are more central (important, meaningful etc.) than others.
- Interpretations of “centrality”: “influence”, “prestige”, “control”, “heavily required for information flow”
- Example: n persons vote for a leader; (u,v) ∈ E if *u voted for v*; Winner (most central node): node with most incoming edges (highest in-degree).
  → Degree Centrality
  Other variant: (u,v) ∈ E if *u has convinced v to vote for u’s favorite candidate*. (Influence network) → node with large out-degree is central
- Other Example: If graph can be split up into groups X and Y and if node u has many edges to X and many edges to Y → u mediates most information between groups → u is central
  → Betweenness centrality

Centrality

- Centrality indices formalize intuitive feeling that some nodes (or edges) are more central (important, meaningful etc.) than others.
- Interpretations of “centrality”: “influence”, “prestige”, “control”, “heavily required for information flow”
- Example: n persons vote for a leader; (u,v) ∈ E if *u voted for v*; Winner (most central node): node with most incoming edges (highest in-degree).
  → Degree Centrality
  Other variant: (u,v) ∈ E if *u has convinced v to vote for u’s favorite candidate*. (Influence network) → node with large out-degree is central
- Other Example: If graph can be split up into groups X and Y and if node u has many edges to X and many edges to Y → u mediates most information between groups → u is central
  → Betweenness centrality

General “Definition”: Structural Index

- “Importance” has many aspects but minimal def. for centrality: Only depends on structure of graph:
- Structural Index: Let G = (V,E,w) be a weighted directed or undirected multigraph. A function s: V → $\mathbb{R}$ (or s: E → $\mathbb{R}$) is a structural index iff

$$\forall x : G \simeq H \rightarrow s_G(x) = s_H(\phi(x))$$

(Recall: Two graphs G and H are isomorphic (G≃H) iff exists a bijective mapping Φ: G → H so that (u,v) ∈ G iff (Φ(u),Φ(v)) ∈H)

- structural index induces order (≤) on nodes/edges
- → centrality can usually only be viewed as measured on an ordinal scale only (not interval or ratio scale)

Distance- and Neighborhood-based Centralities

- Centrality-measures defined on the basis of distances or neighbourhoods in the graph:

  Centrality of vertex ←→ “reachability” of a vertex

Neighborhoods: Degree Centrality

- Most basic: Degree centrality: c(u) = deg(u) (or c(u)=in-deg(u) or
  c(u) = out-deg(u)) → local measure
- Applicable: If edges have “direct vote” semantics

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}

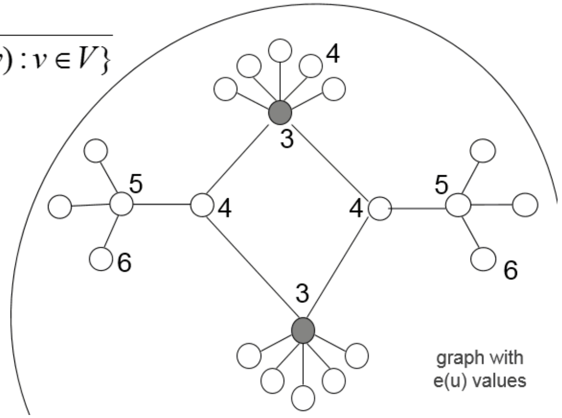
Distances: Eccentricity

- Eccentricity e(u)=max{d(u,v); v∈V}

- Center of a graph: Set of all nodes with minimum eccentricity
- Eccentricity based centrality measure:

$$c(u) = \frac{1}{e(u)} = \frac{1}{\max\{d(u,v) : v \in V\}}$$

- → nodes in the center of the graph have maximal centrality ☺

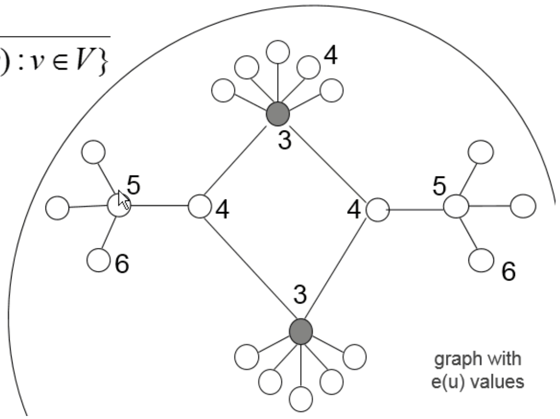


Distances: Eccentricity

- Eccentricity e(u)=max{d(u,v); v∈V}
- Center of a graph: Set of all nodes with minimum eccentricity
- Eccentricity based centrality measure:

$$c(u) = \frac{1}{e(u)} = \frac{1}{\max\{d(u,v) : v \in V\}}$$

- → nodes in the center of the graph have maximal centrality ☺

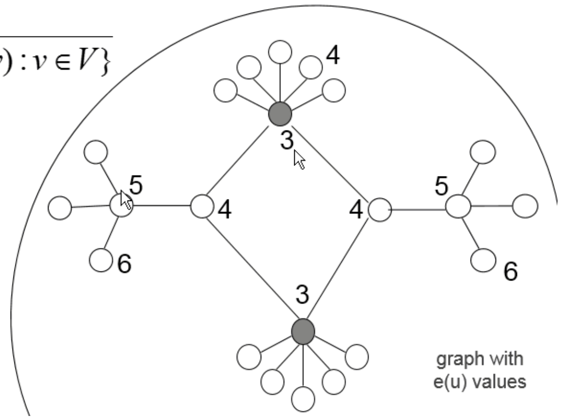


Distances: Eccentricity

- Eccentricity e(u)=max{d(u,v); v∈V}
- Center of a graph: Set of all nodes with minimum eccentricity
- Eccentricity based centrality measure:

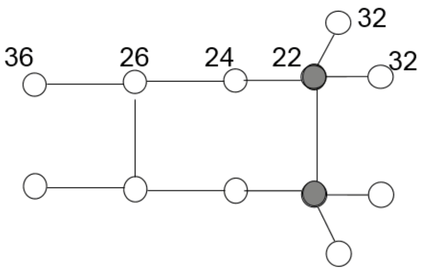
$$c(u) = \frac{1}{e(u)} = \frac{1}{\max\{d(u,v) : v \in V\}}$$

- → nodes in the center of the graph have maximal centrality ☺



Distances: Closeness

- Minisum problem: find nodes whose sum of distances to other nodes is minimal (→ service facility location problem): For all u minimize total sum of minimal distances $\sum_{v \in V} d(u,v)$
- Social analog: Determine central figure for coordination tasks
- Example:



graph with $\sum_{v \in V} d(u,v)$ values

Distances: Closeness

- Possible resulting centrality index: closeness:

$$c(u) = \frac{1}{\sum_{v \in V} d(u,v)}$$

Only applicable to connected graphs; disconnected graph: all nodes will get the same centrality 1/∞

- Other possibility

$$c(u) = \frac{\sum_{v \in V} (\Delta_G + 1 - d(u,v))}{|V| - 1}$$

Δ_G is the diameter of the graph

  - if computed on directed graph: (set d(u,u) = 0 and set d(u,v) = 0 if u,v are unreachable via directed path → problematic !): using in-distances: „integration", using out-distances „radiality" (see [6])

Distances: Closeness

- Possible resulting centrality index: closeness:

$$c(u) = \frac{1}{\sum_{v \in V} d(u,v)}$$

Only applicable to connected graphs; disconnected graph: all nodes will get the same centrality 1/∞

- Other possibility

$$c(u) = \frac{\sum_{v \in V} (\Delta_G + 1 - d(u,v))}{|V| - 1}$$

Δ_G is the diameter of the graph

  - if computed on directed graph: (set d(u,u) = 0 and set d(u,v) = 0 if u,v are unreachable via directed path → problematic !): using in-distances: „integration", using out-distances „radiality" (see [6])

Distances: Closeness

- Possible resulting centrality index: closeness:

$$c(u) = \frac{1}{\sum_{v \in V} d(u,v)}$$

Only applicable to connected graphs; disconnected graph: all nodes will get the same centrality 1/∞

- Other possibility

$$c(u) = \frac{\sum_{v \in V} (\Delta_G + 1 - d(u,v))}{|V| - 1}$$

Δ_G is the diameter of the graph

  - if computed on directed graph: (set d(u,u) = 0 and set d(u,v) = 0 if u,v are unreachable via directed path → problematic !): using in-distances: „integration", using out-distances „radiality" (see [6])

Distances: Closeness

- Possible resulting centrality index: closeness:

$$c(u) = \frac{1}{\sum_{v \in V} d(u,v)}$$

Only applicable to connected graphs; disconnected graph: all nodes will get the same centrality 1/∞

- Other possibility

$$c(u) = \frac{\sum_{v \in V} (\Delta_G + 1 - d(u,v))}{|V| - 1}$$

Δ_G is the diameter of the graph

  - if computed on directed graph: (set d(u,u) = 0 and set d(u,v) = 0 if u,v are unreachable via directed path → problematic !): using in-distances: „integration“, using out-distances „radiality“ (see [6])

Distances: Closeness

- Possible resulting centrality index: closeness:

$$c(u) = \frac{1}{\sum_{v \in V} d(u,v)}$$

Only applicable to connected graphs; disconnected graph: all nodes will get the same centrality 1/∞

- Other possibility

$$c(u) = \frac{\sum_{v \in V} (\Delta_G + 1 - d(u,v))}{|V| - 1}$$

Δ_G is the diameter of the graph

  - if computed on directed graph: (set d(u,u) = 0 and set d(u,v) = 0 if u,v are unreachable via directed path → problematic !): using in-distances: „integration“, using out-distances „radiality“ (see [6])

Distances: Closeness

- Possible resulting centrality index: closeness:

$$c(u) = \frac{1}{\sum_{v \in V} d(u,v)}$$

Only applicable to connected graphs; disconnected graph: all nodes will get the same centrality 1/∞

- Other possibility

$$c(u) = \frac{\sum_{v \in V} (\Delta_G + 1 - d(u,v))}{|V| - 1}$$

Δ_G is the diameter of the graph

  - if computed on directed graph: (set d(u,u) = 0 and set d(u,v) = 0 if u,v are unreachable via directed path → problematic !): using in-distances: „integration“, using out-distances „radiality“ (see [6])

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”

- 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"

- 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"

- 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

- Indices of this section can be applied to weighted, unweighted, directed, undirected and multigraphs and to edges and vertices (“graph elements” x).
- Assume that set of all shortest paths APSP is known (e.g. by application of Floyd Warshall algorithm in $O(|V|^3)$ worst case time)
- Reminder:
  - BFS: SSSP; O(|V|+|E|) worst case time complexity, edge-weights==1
  - Djikstra: SSSP; O(|V| log|V| +|E|) with Fibonacci heap; edge-weights ≥ 0
  - Floyd Warshall: APSP, $O(|V|^3)$ worst case time, arbitrary weights, no negative cycles allowed (but can be detected via the alg.), dynamic programming;
  - Bellman Ford: SSSP; O(|V| |E|), arbitrary weights, no negative cycles allowed (but can be detected via the alg.)

Shortest Paths

- Indices of this section can be applied to weighted, unweighted, directed, undirected and multigraphs and to edges and vertices (“graph elements” x).
- Assume that set of all shortest paths APSP is known (e.g. by application of Floyd Warshall algorithm in $O(|V|^3)$ worst case time)
- Reminder:
  - BFS: SSSP; O(|V|+|E|) worst case time complexity, edge-weights==1
  - Djikstra: SSSP; O(|V| log|V| +|E|) with Fibonacci heap; edge-weights ≥ 0
  - Floyd Warshall: APSP, $O(|V|^3)$ worst case time, arbitrary weights, no negative cycles allowed (but can be detected via the alg.), dynamic programming;
  - Bellman Ford: SSSP; O(|V| |E|), arbitrary weights, no negative cycles allowed (but can be detected via the alg.)

Shortest Paths

- Indices of this section can be applied to weighted, unweighted, directed, undirected and multigraphs and to edges and vertices (“graph elements” x).
- Assume that set of all shortest paths APSP is known (e.g. by application of Floyd Warshall algorithm in $O(|V|^3)$ worst case time)
- Reminder:
  - BFS: SSSP; O(|V|+|E|) worst case time complexity, edge-weights==1
  - Djikstra: SSSP; O(|V| log|V| +|E|) with Fibonacci heap; edge-weights ≥ 0
  - Floyd Warshall: APSP, $O(|V|^3)$ worst case time, arbitrary weights, no negative cycles allowed (but can be detected via the alg.), dynamic programming;
  - Bellman Ford: SSSP; O(|V| |E|), arbitrary weights, no negative cycles allowed (but can be detected via the alg.)

Shortest Paths

- Indices of this section can be applied to weighted, unweighted, directed, undirected and multigraphs and to edges and vertices (“graph elements” x).
- Assume that set of all shortest paths APSP is known (e.g. by application of Floyd Warshall algorithm in $O(|V|^3)$ worst case time)
- Reminder:
  - BFS: SSSP; O(|V|+|E|) worst case time complexity, edge-weights==1
  - Djikstra: SSSP; O(|V| log|V| +|E|) with Fibonacci heap; edge-weights ≥ 0
  - Floyd Warshall: APSP, $O(|V|^3)$ worst case time, arbitrary weights, no negative cycles allowed (but can be detected via the alg.), dynamic programming;
  - Bellman Ford: SSSP; O(|V| |E|), arbitrary weights, no negative cycles allowed (but can be detected via the alg.)

Shortest Paths

- Indices of this section can be applied to weighted, unweighted, directed, undirected and multigraphs and to edges and vertices (“graph elements” x).
- Assume that set of all shortest paths APSP is known (e.g. by application of Floyd Warshall algorithm in $O(|V|^3)$ worst case time)
- Reminder:
  - BFS: SSSP; O(|V|+|E|) worst case time complexity, edge-weights==1
  - Djikstra: SSSP; O(|V| log|V| +|E|) with Fibonacci heap; edge-weights ≥ 0
  - Floyd Warshall: APSP, $O(|V|^3)$ worst case time, arbitrary weights, no negative cycles allowed (but can be detected via the alg.), dynamic programming;
  - Bellman Ford: SSSP; O(|V| |E|), arbitrary weights, no negative cycles allowed (but can be detected via the alg.)

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 analog: 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

- Variant for edges:

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

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

- 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

- 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: Stress

- Variant for edges:

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

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 exerts 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

- Shortest Path Betweenness (SPB) centrality is then:

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

- Interpretation: Control that v exerts 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

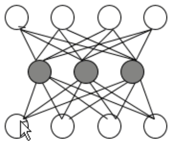
- Shortest Path Betweenness (SPB) centrality is then:

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

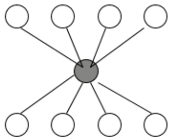
- Interpretation: Control that v exerts on the communication in the graph
- Also applicable to disonnected graphs
- Algorithm by Ulrik Brandes computes SPB in $O(|V||E| + |V|^2 \log|V|)$ time

Shortest Paths: Shortest Path Betweenness

- Example why shortest path betweenness centrality (now denoted as c_SPB) might be more interesting than the basic stress centrality (now denoted as c_S):



each ● *has*
c_S=$\binom{8}{2}$=28
c_SPB=1/3 * 28



● *has*
c_S=28
c_SPB=28

Shortest Paths: Shortest Path Betweenness

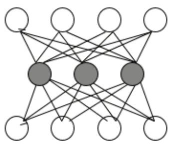
- Shortest Path Betweenness (SPB) centrality is then:

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

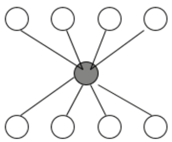
- Interpretation: Control that v exerts on the communication in the graph
- Also applicable to disonnected graphs
- Algorithm by Ulrik Brandes computes SPB in O(|V||E| + $|V|^2$log|V|) time

Shortest Paths: Shortest Path Betweenness

- Example why shortest path betweenness centrality (now denoted as c_SPB) might be more interesting than the basic stress centrality (now denoted as c_S):



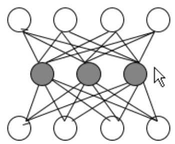
each ● *has*
c_S=$\binom{8}{2}$=28
c_SPB=1/3 * 28



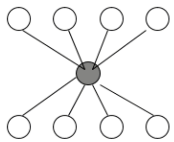
● *has*
c_S=28
c_SPB=28

Shortest Paths: Shortest Path Betweenness

- Example why shortest path betweenness centrality (now denoted as c_SPB) might be more interesting than the basic stress centrality (now denoted as c_S):



each ● *has*
c_S=$\binom{8}{2}$=28
c_SPB=1/3 * 28



● *has*
c_S=28
c_SPB=28