Weighted Networks

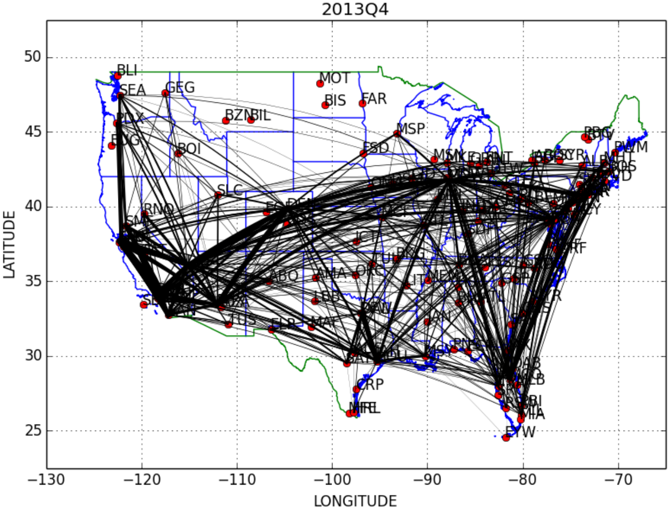
with Application to U.S. Domestic Airlines

Alexandre Dossin Steve Lawford

ENAC

September 22, 2016

Why Are We Interested in Networks?



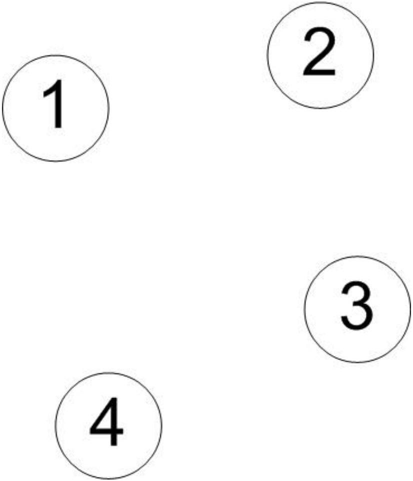
Southwest Airlines' (WN) Route Network, 2013Q4

What Do We Find?

- Develop simple weighted centrality measures
  - Standard unweighted measures can be generalized with only minor modifications (typically to ensure scaling to $[0, 1]$).
  - Can avoid misleading unweighted results e.g. minimum-step or minimum-distance paths.
- Apply to U.S. domestic airline networks
  - Generally, networks have one (or several) dominant hubs.
  - Some changes in rankings for non-dominant airports, when using weighted rather than unweighted centrality.
- Significant centrality (hub) premium
  - One-standard-deviation increase in unweighted airport centrality implies fare increase of about $17 ($8) for most (least) central endpoint on route i.e. 5% (2%) of $350 ticket.[1]
  - Quantitatively similar results for weighted centrality measures.

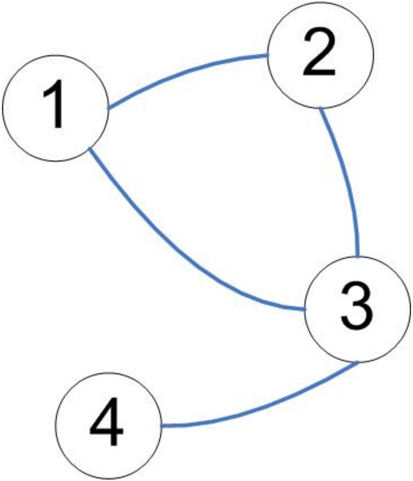
[1] e.g. to achieve this fare increase, the most (least) central endpoint would need to be on 30% (3%) more shortest paths, or have 26% (11%) more routes (of total).

What Is a Graph?



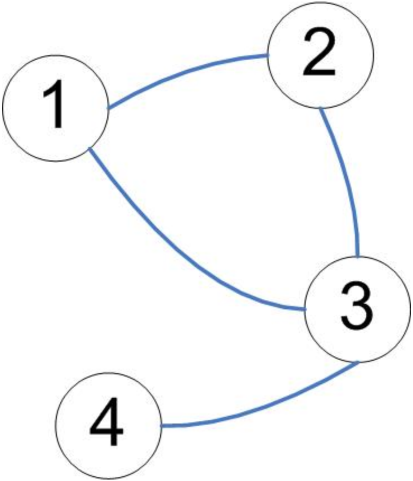
- *nodes* i, j (or $i_1, i_2, \ldots, i_n$)

What Is a Graph?



- *nodes* i, j (or $i_1, i_2, \ldots, i_n$)
- *edges* $g = (g_{ij})$ or $ij \in g$

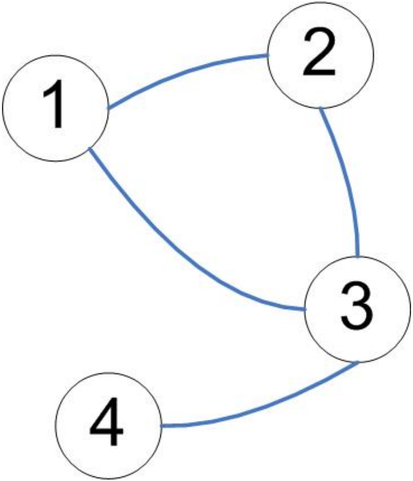
What Is a Graph?



- *nodes* i, j (or $i_1, i_2, \ldots, i_n$)
- *edges* $g = (g_{ij})$ or $ij \in g$
- *adjacency matrix*

$$g = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

What Is a Graph?

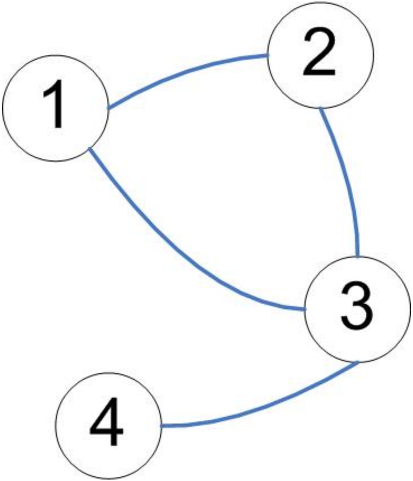


- *nodes* i, j (or $i_1, i_2, \ldots, i_n$)
- *edges* $g = (g_{ij})$ or $ij \in g$
- *adjacency matrix*

$$g = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

- $g_{ij} = 0, 1$ (unweighted)

What Is a Graph?

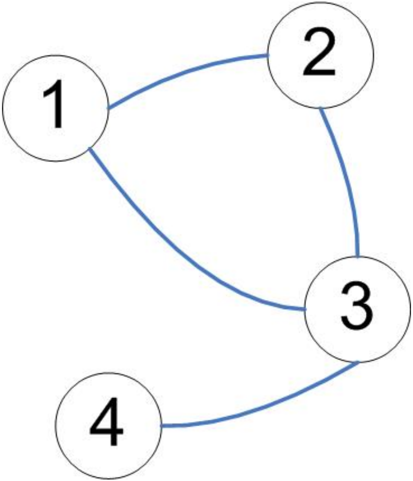


- *nodes* i, j (or $i_1, i_2, \ldots, i_n$)
- *edges* $g = (g_{ij})$ or $ij \in g$
- *adjacency matrix*

$$g = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

- $g_{ij} = 0, 1$ (unweighted)
- $g = g^{\mathsf{T}}$ (undirected)

What Is a Graph?

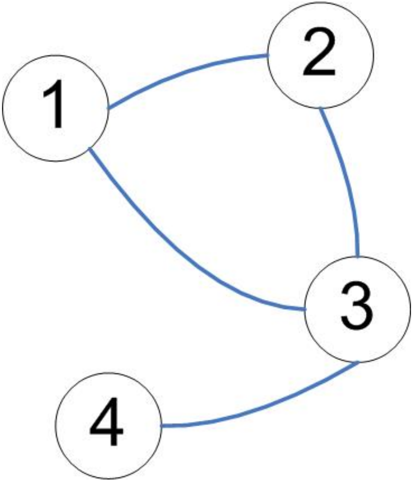


- *nodes* i, j (or $i_1, i_2, \ldots, i_n$)
- *edges* $g = (g_{ij})$ or $ij \in g$
- *adjacency matrix*

$$g = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

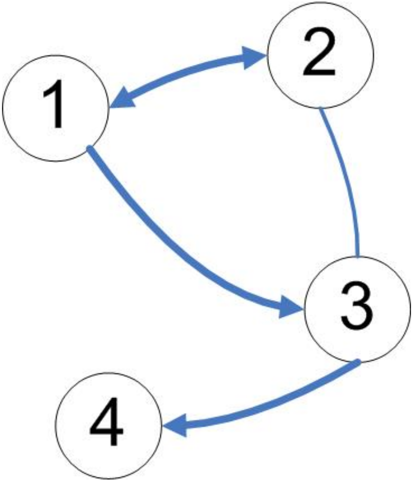
- $g_{ij} = 0, 1$ (unweighted)
- $g = g^{\mathsf{T}}$ (undirected)
- $g_{ii} = 0$ (no self-links)

Walks and Paths



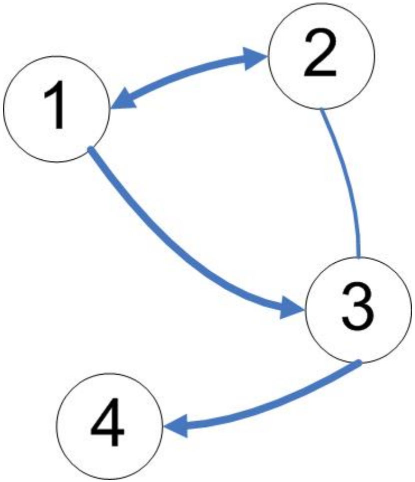
- A *walk* is a sequence of links $i_k i_{k+1} \in g$ with $i_1 = i$ and $i_K = j$, for $k = 1, \ldots, K-1$.

Walks and Paths



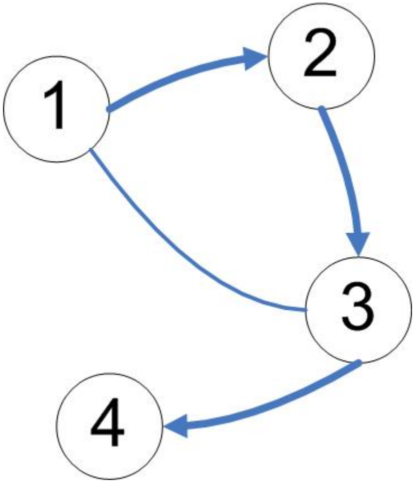
- A *walk* is a sequence of links $i_k i_{k+1} \in g$ with $i_1 = i$ and $i_K = j$, for $k = 1, \ldots, K-1$.
- e.g. walk 1–2–1–3–4

Walks and Paths



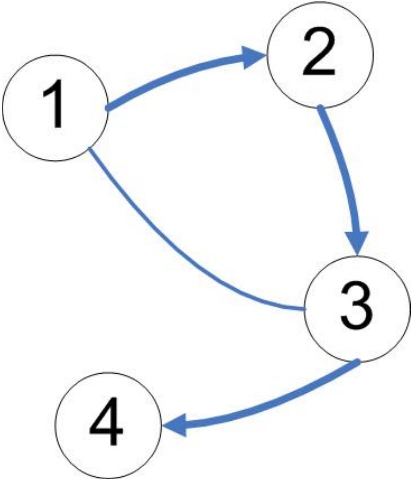
- A *walk* is a sequence of links $i_k i_{k+1} \in g$ with $i_1 = i$ and $i_K = j$, for $k = 1, \ldots, K-1$.
- e.g. walk 1–2–1–3–4
- A *path* is a walk with distinct nodes.

Walks and Paths



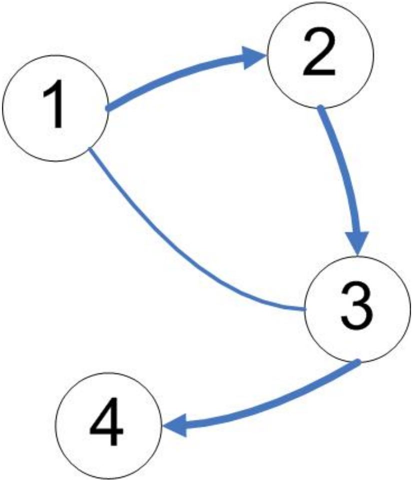
- A *walk* is a sequence of links $i_k i_{k+1} \in g$ with $i_1 = i$ and $i_K = j$, for $k = 1, \ldots, K-1$.
- e.g. walk 1–2–1–3–4
- A *path* is a walk with distinct nodes.
- e.g. path 1–2–3–4

Walks and Paths



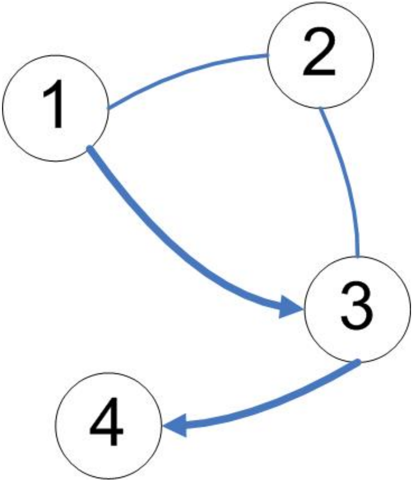
- A *walk* is a sequence of links $i_k i_{k+1} \in g$ with $i_1 = i$ and $i_K = j$, for $k = 1, \ldots, K-1$.
- e.g. walk 1–2–1–3–4
- A *path* is a walk with distinct nodes.
- e.g. path 1–2–3–4
- A *connected* graph has a path between every i and j.

Walks and Paths



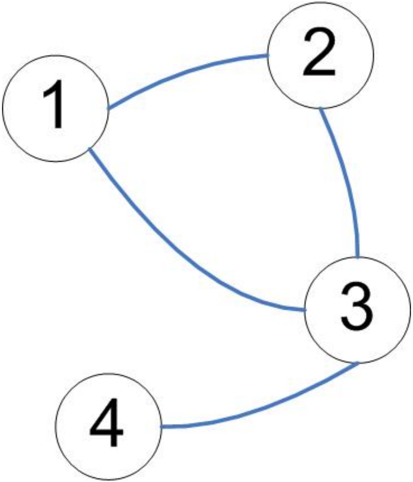
- A *walk* is a sequence of links $i_k i_{k+1} \in g$ with $i_1 = i$ and $i_K = j$, for $k = 1, \ldots, K-1$.
- e.g. walk 1–2–1–3–4
- A *path* is a walk with distinct nodes.
- e.g. path 1–2–3–4
- A *connected* graph has a path between every i and j.
- A *geodesic* is the shortest path between two nodes.

Walks and Paths



- A *walk* is a sequence of links $i_k i_{k+1} \in g$ with $i_1 = i$ and $i_K = j$, for $k = 1, \ldots, K-1$.
- e.g. walk 1–2–1–3–4
- A *path* is a walk with distinct nodes.
- e.g. path 1–2–3–4
- A *connected* graph has a path between every i and j.
- A *geodesic* is the shortest path between two nodes.
- e.g. shortest path 1–3–4

Diameter of a Graph (= Longest Shortest Path)

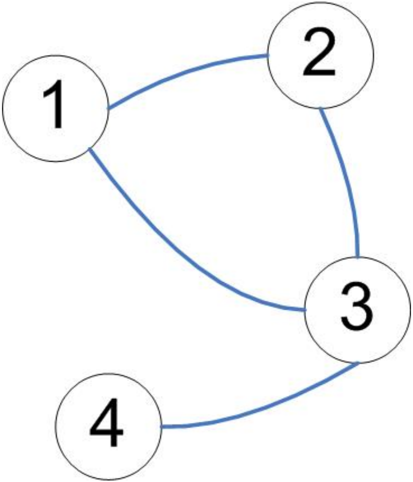


- powers of adjacency matrix

$$g^1 = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 1 \\ 4 & 0 & 0 & 1 & 0 \end{array}$$

$$g^2 = \begin{array}{c|cccc} & 1 & 2 & 3 & 4 \\ \hline 1 & 2 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 & 1 \\ 3 & 1 & 1 & 3 & 0 \\ 4 & 1 & 1 & 0 & 1 \end{array}$$

Diameter of a Graph (= Longest Shortest Path)



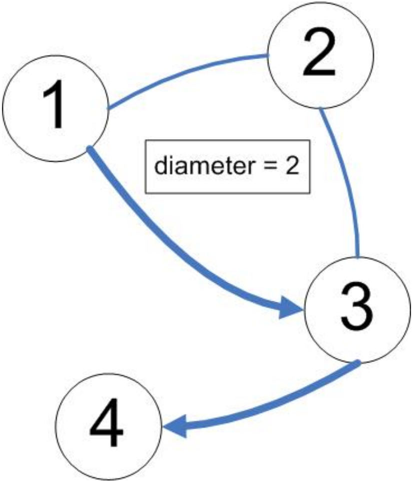
- powers of adjacency matrix

$$g^1 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

$$g^2 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix} \end{array}$$

- $\sum_k g^k$ gives walks (paths) of length $\leq k$ ($k \leq n-1$).

Diameter of a Graph (= Longest Shortest Path)



- powers of adjacency matrix

	1	2	3	4
1	0	1	1	0
2	1	0	1	0
3	1	1	0	1
4	0	0	1	0

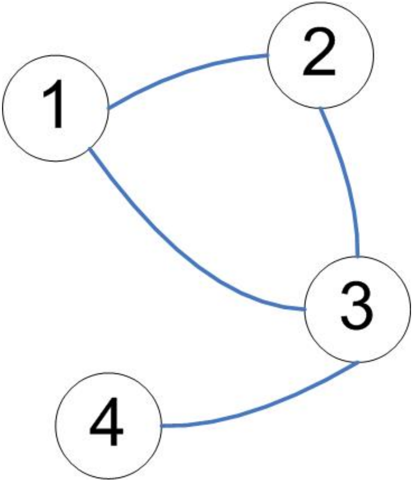
$$g^1 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

	1	2	3	4
1	2	1	1	1
2	1	2	1	1
3	1	1	3	0
4	1	1	0	1

$$g^2 = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

- $\sum_k g^k$ gives walks (paths) of length $\leq k$ ($k \leq n-1$).

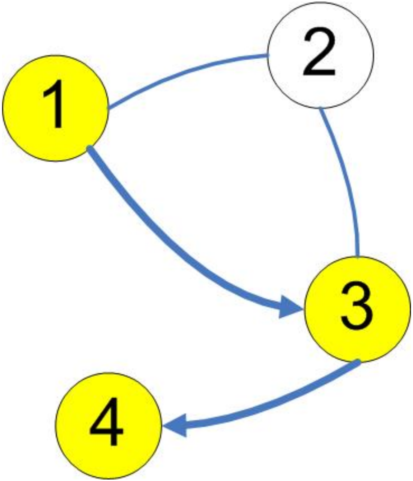
Distance Matrix and Shortest Paths



- *distance matrix* $D = (l_{ij})$ contains geodesic lengths

$$D = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 \\ 2 & 2 & 1 & 0 \end{pmatrix} \end{array}$$

Distance Matrix and Shortest Paths

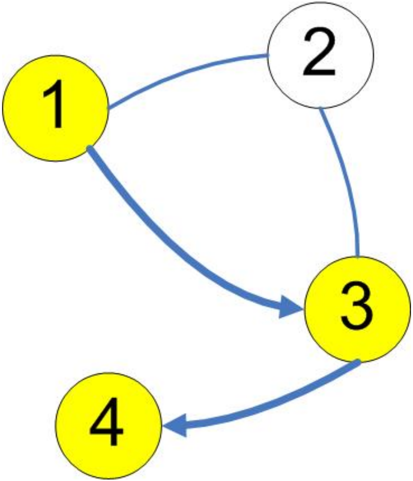


- *distance matrix* $D = (l_{ij})$ contains geodesic lengths

$$D = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 \\ 2 & 2 & 1 & 0 \end{pmatrix} \end{array}$$

- paths reconstructed using variant of Dijkstra (1959) or Floyd-Warshall (1962) algos.

Distance Matrix and Shortest Paths

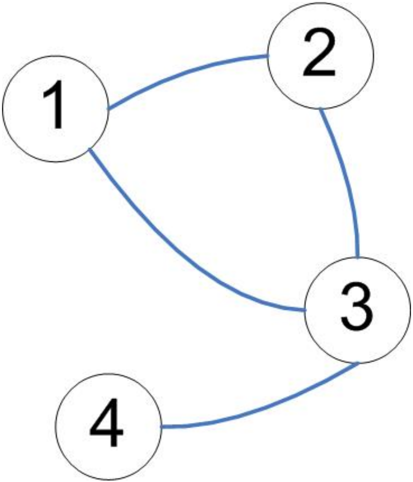


- *distance matrix* $D = (l_{ij})$ contains geodesic lengths

$$D = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 \\ 2 & 2 & 1 & 0 \end{pmatrix} \end{array}$$

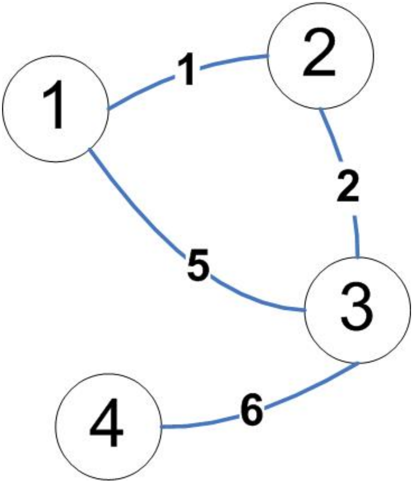
- paths reconstructed using variant of Dijkstra (1959) or Floyd-Warshall (1962) algos.
- we will need **all** shortest paths between i and j . . .

Adding Edge-Weights to the Graph



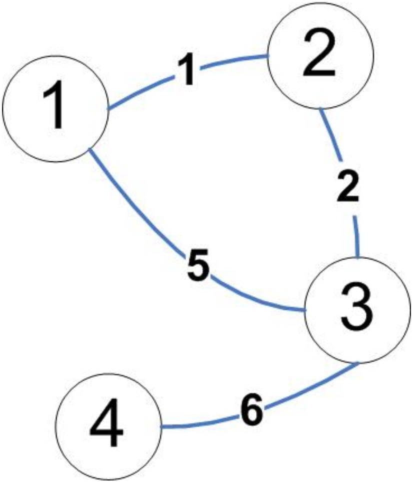
- unweighted graph no info. on strength of links $ij \in g$

Adding Edge-Weights to the Graph



- unweighted graph no info. on strength of links $ij \in g$
- replace $g_{ij} = 0, 1$ with *edge-weights* $g_w = (g_{w,ij})$

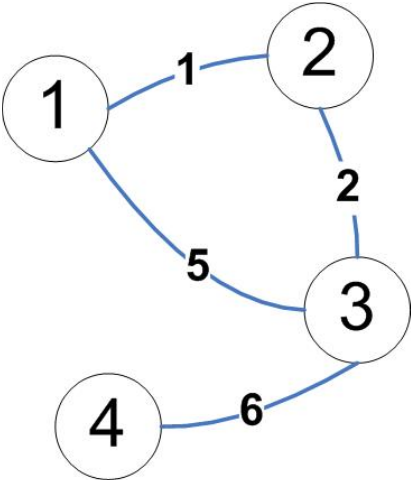
Adding Edge-Weights to the Graph



- unweighted graph no info. on strength of links $ij \in g$
- replace $g_{ij} = 0, 1$ with *edge-weights* $g_w = (g_{w,ij})$
- weighted adjacency matrix

$$g_w = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 5 & 0 \\ 1 & 0 & 2 & 0 \\ 5 & 2 & 0 & 6 \\ 0 & 0 & 6 & 0 \end{pmatrix} \end{array}$$

Adding Edge-Weights to the Graph

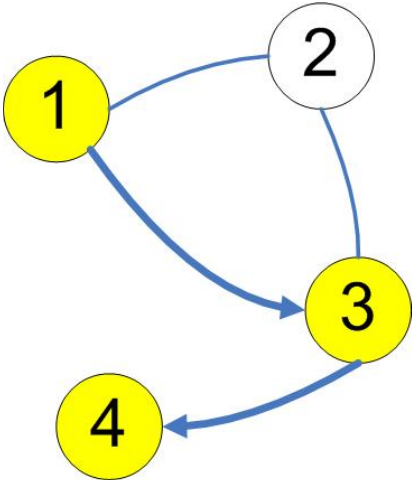


- unweighted graph no info. on strength of links $ij \in g$
- replace $g_{ij} = 0, 1$ with *edge-weights* $g_w = (g_{w,ij})$
- weighted adjacency matrix

$$g_w = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 5 & 0 \\ 1 & 0 & 2 & 0 \\ 5 & 2 & 0 & 6 \\ 0 & 0 & 6 & 0 \end{pmatrix} \end{array}$$

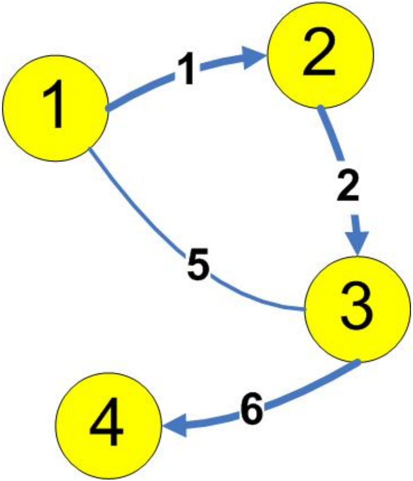
- applications in economics use topology, not weights

Shortest Paths in a Graph with Edge-Weights



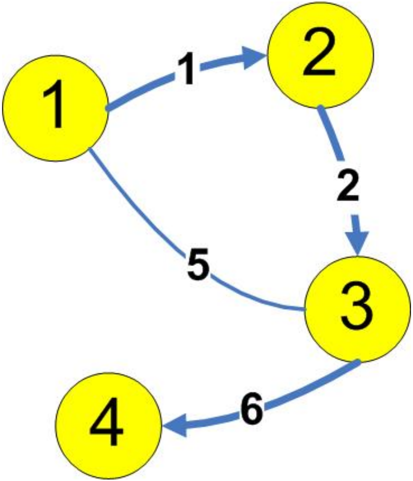
- shortest paths can change when edge-weights are used
- remember from g that $l_{14} = 2$ (path 1–3–4)

Shortest Paths in a Graph with Edge-Weights



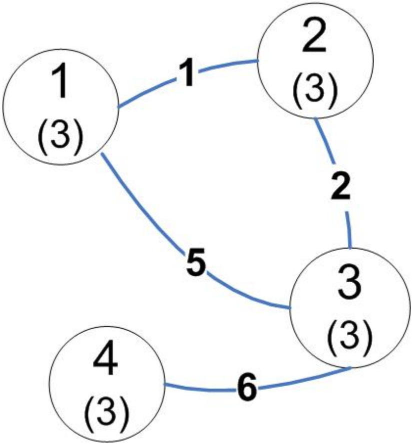
- shortest paths can change when edge-weights are used
- remember from g that $l_{14} = 2$ (path 1–3–4)
- however, g_w gives $l_{w,14} = 9$ (path 1–2–3–4 has lower weight but more steps than shortest topological path)

Shortest Paths in a Graph with Edge-Weights



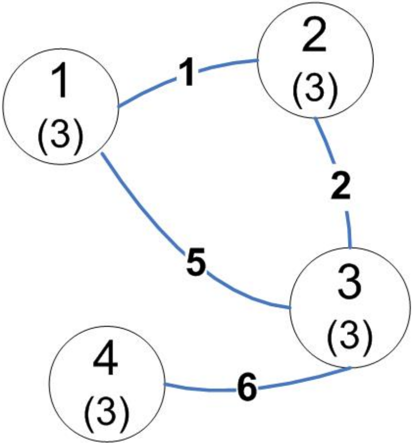
- shortest paths can change when edge-weights are used
- remember from g that $l_{14} = 2$ (path 1–3–4)
- however, g_w gives $l_{w,14} = 9$ (path 1–2–3–4 has lower weight but more steps than shortest topological path)
- measures based on shortest paths (and shortest path lengths) may give different results in weighted networks

Adding Edge-Weights and Node-Weights to the Graph



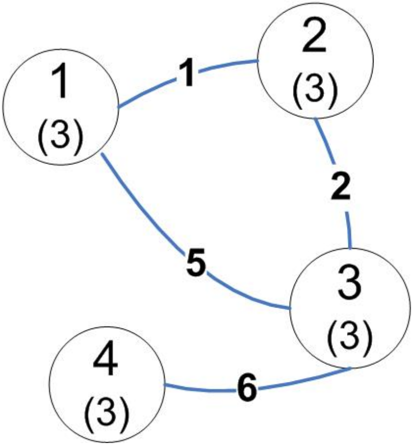
- further, we can impose a *node-weight*, or *penalty*, x_i
- for simplicity, assume that $x_i = x$ for all i (here, $x = 3$)

Adding Edge-Weights and Node-Weights to the Graph



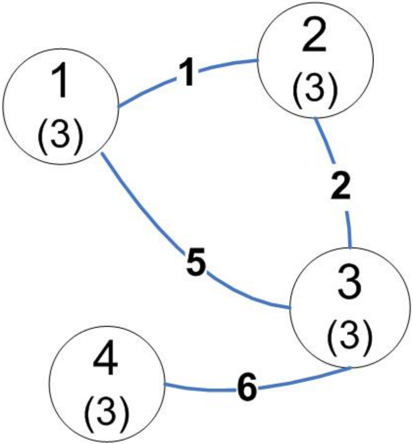
- further, we can impose a *node-weight*, or *penalty*, x_i
- for simplicity, assume that $x_i = x$ for all i (here, $x = 3$)
- e.g. edge-weight physical distance between airports, node-weight $\propto$ expected waiting time at airport

Adding Edge-Weights and Node-Weights to the Graph



- further, we can impose a *node-weight*, or *penalty*, x_i
- for simplicity, assume that $x_i = x$ for all i (here, $x = 3$)
- e.g. edge-weight physical distance between airports, node-weight $\propto$ expected waiting time at airport
- node-weights may influence shortest path computations
- let $g_w(x) = (g_{w,ij}(x))$, with shortest path lengths $l_{w,ij}(x)$

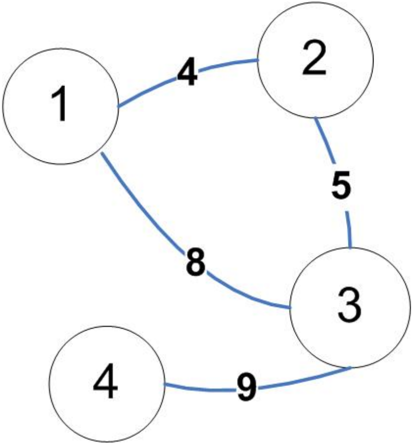
Shortest Paths in a Graph with Edge-/Node-Weights



- transform g_w by $g_{w,ij}(x) =$

$$\begin{cases} g_{w,ij}(0) + x, & g_{w,ij}(0) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

Shortest Paths in a Graph with Edge-/Node-Weights



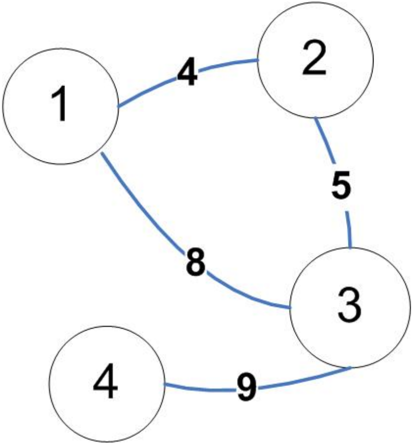
- transform g_w by $g_{w,ij}(x) =$

$$\begin{cases} g_{w,ij}(0) + x, & g_{w,ij}(0) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

- modified adjacency matrix

$$g_w(3) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 4 & 8 & 0 \\ 4 & 0 & 5 & 0 \\ 8 & 5 & 0 & 9 \\ 0 & 0 & 9 & 0 \end{pmatrix} \end{array}$$

Shortest Paths in a Graph with Edge-/Node-Weights



- transform g_w by $g_{w,ij}(x) =$

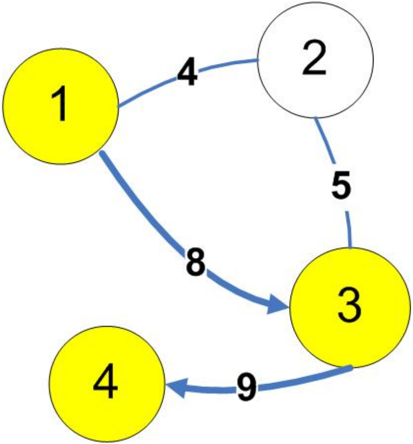
$$\begin{cases} g_{w,ij}(0) + x, & g_{w,ij}(0) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

- modified adjacency matrix

$$g_w(3) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 4 & 8 & 0 \\ 4 & 0 & 5 & 0 \\ 8 & 5 & 0 & 9 \\ 0 & 0 & 9 & 0 \end{pmatrix} \end{array}$$

- remember from g_w that $l_{w,14} = 9$ (path 1–2–3–4)

Shortest Paths in a Graph with Edge-/Node-Weights



- transform g_w by $g_{w,ij}(x) =$

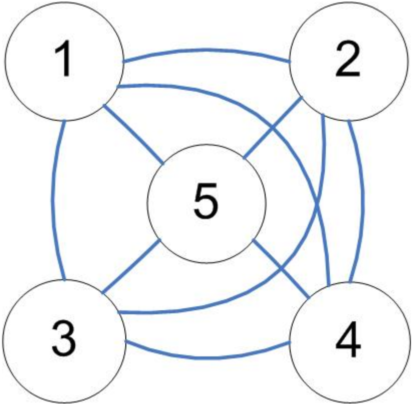
$$\begin{cases} g_{w,ij}(0) + x, & g_{w,ij}(0) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

- modified adjacency matrix

$$g_w(3) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 4 & 8 & 0 \\ 4 & 0 & 5 & 0 \\ 8 & 5 & 0 & 9 \\ 0 & 0 & 9 & 0 \end{pmatrix} \end{array}$$

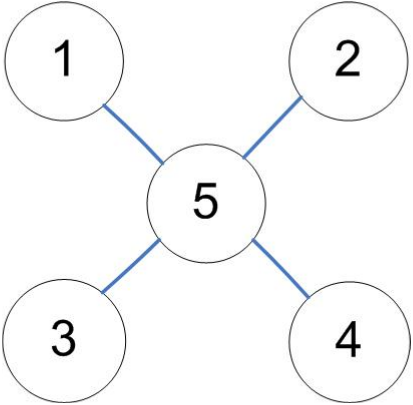
- remember from g_w that $l_{w,14} = 9$ (path 1–2–3–4)
- $g_w(3)$ gives $l_{w,14} = 17$ (path 1–3–4), or $l_{w,14} - x = 14$

Some Special Graphs (K_n)



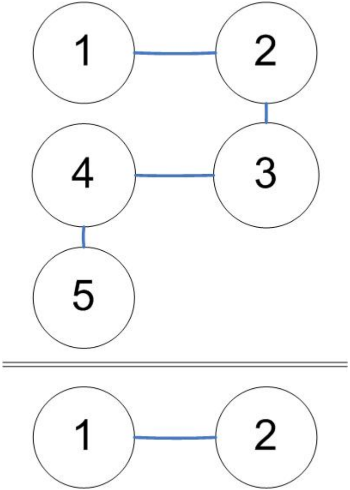
- *complete* graph K_n
- $\{g \mid ij \in g \, \forall \, i, j \mid i \neq j\}$
- here, we see K_5
- in economic theory, K_n is sometimes called a (perfect) *point-to-point* network
- e.g. Hendricks, Piccione & Tan (1995, Review of Economic Studies), and Hendricks, Piccione & Tan (1999, Econometrica)

Some Special Graphs ($S_{1,n-1}$)



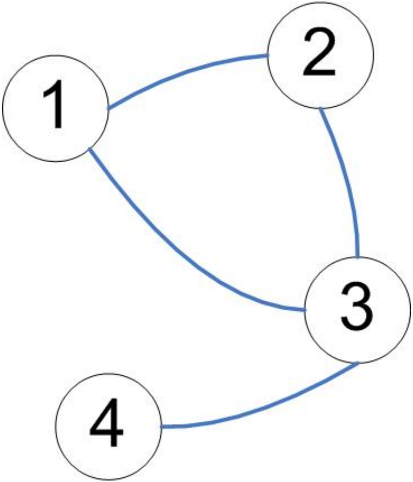
- *star* graph $S_{1,n-1}$, with *center* i_1
- $\{g \mid i_1 i_k \in g \mid k = 2, \ldots, n\}$
- we see $S_{1,4}$, center $i_1 = 5$
- in economic theory, $S_{1,n-1}$ is sometimes called a (perfect) *hub-and-spoke* network
- e.g. Hendricks, Piccione & Tan (1995, Review of Economic Studies), and Hendricks, Piccione & Tan (1999, Econometrica)

Some Special Graphs (P_n)



- *path graph* P_n
- $\{g \mid i_k i_{k+1} \in g \mid k = 1, \ldots, n-1\}$
- here, we see P_5 and P_2
- P_2 is called a *dyad*

Unweighted Centrality Measures (Degree)

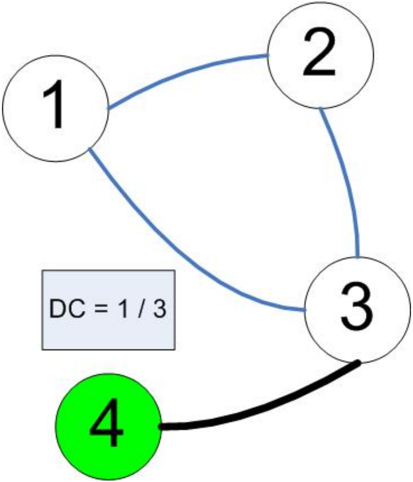


- *degree centrality*

$$DC_i(g) = \frac{d_i}{n-1}$$

- where $d_i = \sum_j g_{ij}$ is the *degree* of node i

Unweighted Centrality Measures (Degree)



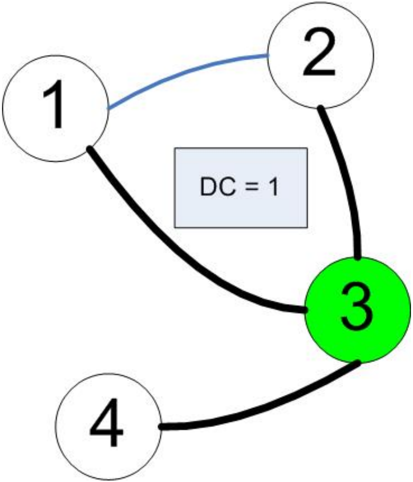
- *degree centrality*

$$DC_i(g) = \frac{d_i}{n-1}$$

- where $d_i = \sum_j g_{ij}$ is the *degree* of node i
- limits on DC

$$DC \in \left[\frac{1}{n-1}, 1\right]$$

Unweighted Centrality Measures (Degree)



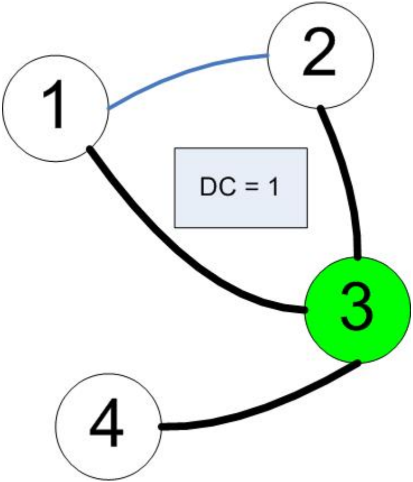
- *degree centrality*

$$DC_i(g) = \frac{d_i}{n-1}$$

- where $d_i = \sum_j g_{ij}$ is the *degree* of node i
- limits on DC

$$DC \in \left[\frac{1}{n-1}, 1\right]$$

Unweighted Centrality Measures (Degree)



- *degree centrality*

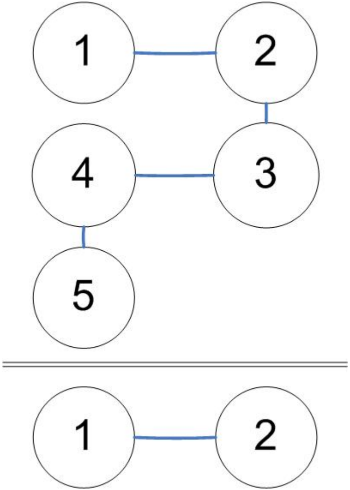
$$DC_i(g) = \frac{d_i}{n-1}$$

- where $d_i = \sum_j g_{ij}$ is the *degree* of node i
- limits on DC

$$DC \in \left[\frac{1}{n-1}, 1\right]$$

- this is the simplest measure of node centrality
- algorithm: counting!

Unweighted Centrality Measures (Closeness)

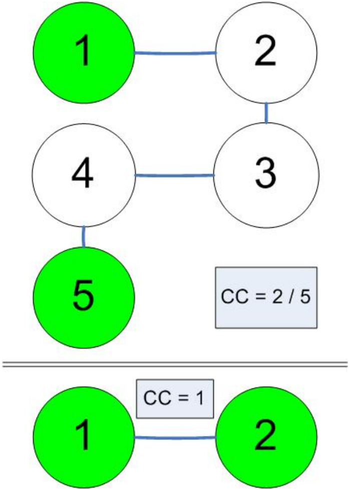


- *closeness centrality*

$$CC_i(g) = \frac{n-1}{\sum_j l_{ij}}$$

- where l_{ij} is the length of the geodesic between i and j

Unweighted Centrality Measures (Closeness)



- *closeness centrality*

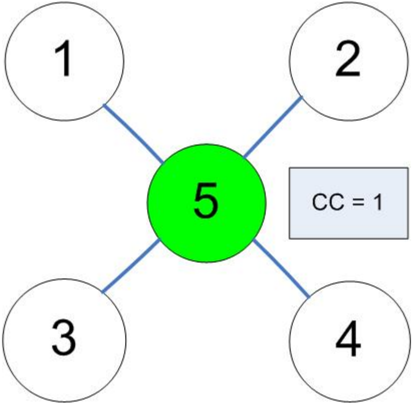
$$CC_i(g) = \frac{n-1}{\sum_j l_{ij}}$$

- where l_{ij} is the length of the geodesic between i and j
- limits on CC

$$CC \in \left[\frac{2}{n}, 1\right]$$

- note $\sum_j l_{ij} = n(n-1)/2$

Unweighted Centrality Measures (Closeness)



- *closeness centrality*

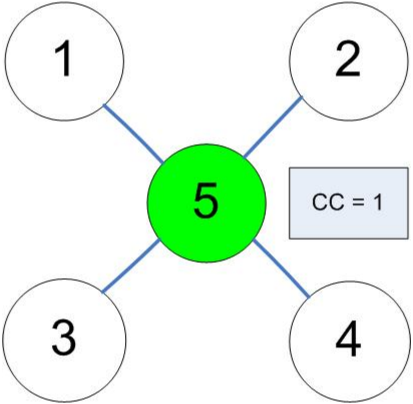
$$CC_i(g) = \frac{n-1}{\sum_j l_{ij}}$$

- where l_{ij} is the length of the geodesic between i and j
- limits on CC

$$CC \in \left[\frac{2}{n}, 1\right]$$

- note $\sum_j l_{ij} = n(n-1)/2$

Unweighted Centrality Measures (Closeness)



- *closeness centrality*

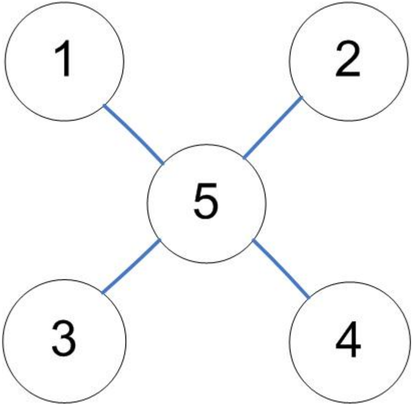
$$CC_i(g) = \frac{n-1}{\sum_j l_{ij}}$$

- where l_{ij} is the length of the geodesic between i and j
- limits on CC

$$CC \in \left[\frac{2}{n}, 1\right]$$

- note $\sum_j l_{ij} = n(n-1)/2$
- algorithm: distance $D = (l_{ij})$

Unweighted Centrality Measures (Betweenness)

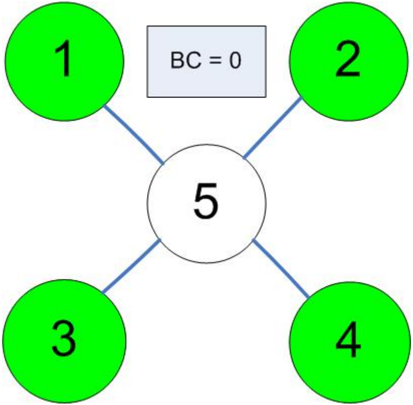


- *betweenness centrality* $BC_i(g) =$

$$\frac{2}{(n-1)(n-2)} \sum_{k<j \,|\, i \neq j \neq k} \frac{P_i(k,j)}{P(k,j)}$$

- where $P(k,j)$ is the **number** of geodesics between nodes k and j, and $P_i(k,j)$ is the **number** that include node i

Unweighted Centrality Measures (Betweenness)



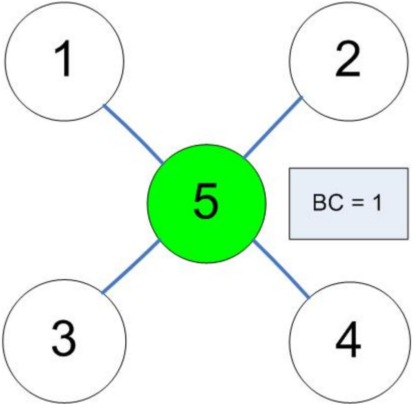
- *betweenness centrality* $BC_i(g) =$

$$\frac{2}{(n-1)(n-2)} \sum_{k<j \,|\, i \neq j \neq k} \frac{P_i(k,j)}{P(k,j)}$$

- where $P(k,j)$ is the **number** of geodesics between nodes k and j, and $P_i(k,j)$ is the **number** that include node i
- limits on BC

$$BC \in [0,1]$$

Unweighted Centrality Measures (Betweenness)



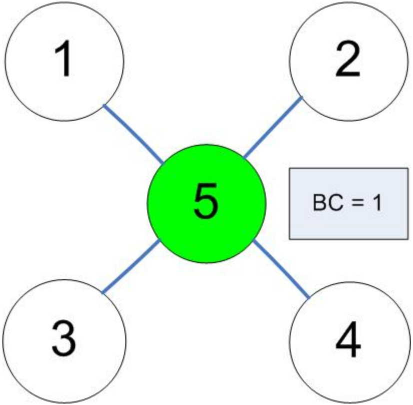
- *betweenness centrality*
  $BC_i(g) =$

$$\frac{2}{(n-1)(n-2)} \sum_{k<j \mid i \neq j \neq k} \frac{P_i(k,j)}{P(k,j)}$$

- where $P(k,j)$ is the **number** of geodesics between nodes k and j, and $P_i(k,j)$ is the **number** that include node i
- limits on BC

$$BC \in [0,1]$$

Unweighted Centrality Measures (Betweenness)



- *betweenness centrality* $BC_i(g) =$

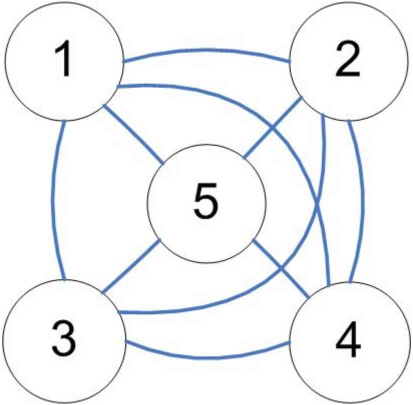
$$\frac{2}{(n-1)(n-2)} \sum_{k<j \,|\, i \neq j \neq k} \frac{P_i(k,j)}{P(k,j)}$$

- where $P(k,j)$ is the **number** of geodesics between nodes k and j, and $P_i(k,j)$ is the **number** that include node i
- limits on BC

$$BC \in [0,1]$$

- algorithm: reconstruct all shortest paths from $D = (l_{ij})$

Unweighted Centrality Measures (Eigenvector)

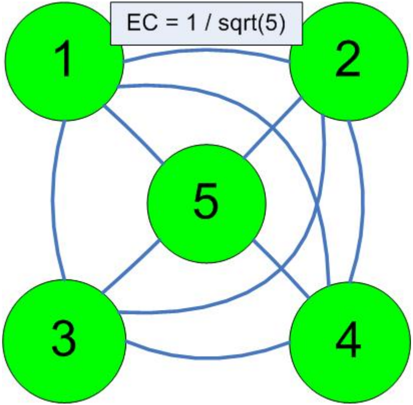


- *eigenvector centrality*

$$\lambda\, EC(g) = g\, EC(g) \iff$$

$$\mathrm{EC}_i(g) = \frac{1}{\lambda} \sum_j g_{ij}\, EC_j(g)$$

Unweighted Centrality Measures (Eigenvector)



- *eigenvector centrality*

$$\lambda\, EC(g) = g\, EC(g) \iff$$

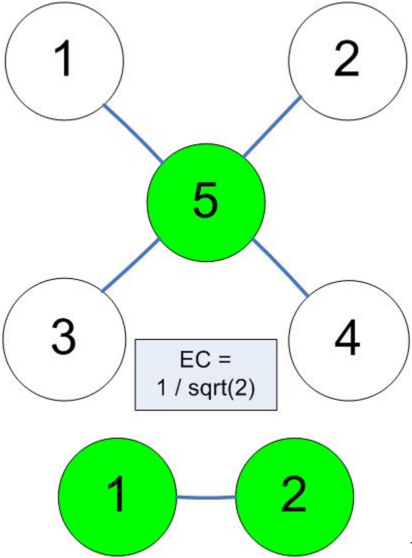
$$\mathrm{EC}_i(g) = \tfrac{1}{\lambda} \textstyle\sum_j g_{ij}\, EC_j(g)$$

- limits on EC (use $\sqrt{2}\,EC$)

$$EC \in \left[0, \frac{1}{\sqrt{2}}\right]$$

- no connected graph attains minimum (K_n is $O(1/\sqrt{n})$)

Unweighted Centrality Measures (Eigenvector)



- *eigenvector centrality*

$$\lambda\, EC(g) = g\, EC(g) \iff$$

$$\mathrm{EC}_i(g) = \tfrac{1}{\lambda} \textstyle\sum_j g_{ij}\, EC_j(g)$$

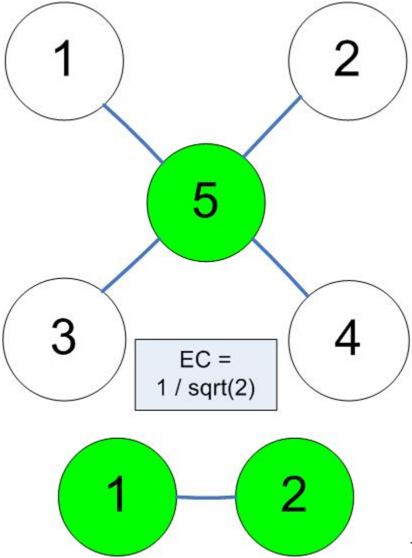
- limits on EC (use $\sqrt{2}\, EC$)

$$EC \in \left[0, \frac{1}{\sqrt{2}}\right]$$

- no connected graph attains minimum (K_n is $O(1/\sqrt{n})$)
- maximum $\sim$ eigenvector normalization ($p = 2$)[a]

[a] Papendieck & Recht (2000, LAA).

Unweighted Centrality Measures (Eigenvector)



- *eigenvector centrality*

$$\lambda\, EC(g) = g\, EC(g) \iff$$

$$\mathrm{EC}_i(g) = \tfrac{1}{\lambda} \sum_j g_{ij}\, EC_j(g)$$

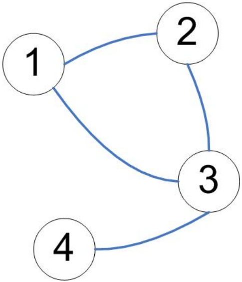
- limits on EC (use $\sqrt{2}\, EC$)

$$EC \in \left[0, \frac{1}{\sqrt{2}}\right]$$

- no connected graph attains minimum (K_n is $O(1/\sqrt{n})$)
- maximum $\sim$ eigenvector normalization ($p = 2$)[a]
- algorithm: eigenvector

[a] Papendieck & Recht (2000, LAA).

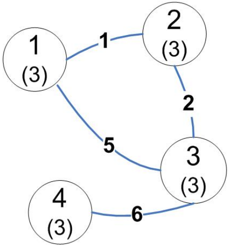
Variation across Measures



- measures capture different aspects of node centrality
- e.g. node 2 on no shortest paths, close to other nodes
- node 3 is always "dominant"

node	BC	CC	DC	EC
1	0	0.75	0.67	0.74
2	0	0.75	0.67	0.74
3	0.67	1	1	0.86
4	0	0.60	0.33	0.40

Weighted Network Measures[2]



$$g_w(0) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 5 & 0 \\ 1 & 0 & 2 & 0 \\ 5 & 2 & 0 & 6 \\ 0 & 0 & 6 & 0 \end{pmatrix} \end{array}$$

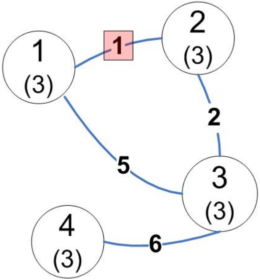
- Weighted Degree

$$DC_{w,i}(g_w(x)) = \frac{d_{w,i}(0)}{(n-1)g_w^{\vee}},$$

where $d_{w,i}(0) = \sum_j g_{w,ij}(0)$ and $g_w^{\vee} = \max_{\mathcal{A}} g_{w,ij}(0)$, with $\mathcal{A}$ the set of non-zero elements of $g_w(0)$; not a function of x.

[2] Clustering / centrality: **Barrat, Barthélemy, Pastor-Satorras & Vespignani (2004, PNAS)**, Opsahl & Panzarasa (2009, Social Networks); Centrality: Newman (2001, Physical Review E), Brandes (2008, Social Networks), Wang, Hernandez & Van Mieghem (2008, Physical Review E), Opsahl, Agneessens & Skvoretz (2010, Social Networks), Wei, Pfeffer, Reminga & Carley (2011, Carnegie Mellon tech. report).

Weighted Network Measures



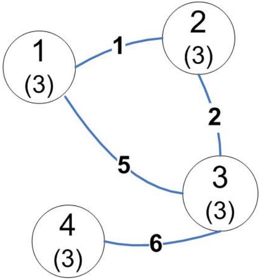
$$D_w(3) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 5 & 14 \\ 1 & 0 & 2 & 11 \\ 5 & 2 & 0 & 6 \\ 14 & 11 & 6 & 0 \end{pmatrix} \end{array}$$

- Weighted Closeness

$$CC_{w,i}(g_w(x)) = \frac{(n-1)g_w^{\wedge}}{\sum_j l_{w,ij}(x)},$$

where $g_w^{\wedge} = \min_{\mathcal{A}} g_{w,ij}(0)$, with $\mathcal{A}$ non-zero elements of $g_w(0)$.

Weighted Network Measures



$$D_w(3) = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 5 & 14 \\ 1 & 0 & 2 & 11 \\ 5 & 2 & 0 & 6 \\ 14 & 11 & 6 & 0 \end{pmatrix} \end{array}$$

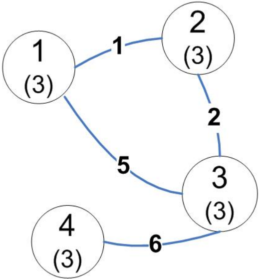
unique shortest paths ($x = 3$):
1–2, 1–3, 1–3–4, 2–3, 2–3–4, 3–4

- Weighted Betweenness

$$BC_{w,i}(g_w(x)) = \frac{2}{(n-1)(n-2)} \sum_{k<j \,|\, i \neq j \neq k} \frac{P_{w(x),i}(k,j)}{P_{w(x)}(k,j)}$$

- where $P_{w(x)}(k,j)$ is the **number** of weighted shortest paths between k and j, and $P_{w(x),i}(k,j)$ number that include node i

Weighted Network Measures



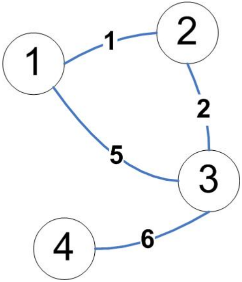
$$g_w(0) = \begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} & \begin{pmatrix} 0 & 1 & 5 & 0 \\ 1 & 0 & 2 & 0 \\ 5 & 2 & 0 & 6 \\ 0 & 0 & 6 & 0 \end{pmatrix} \end{array}$$

- Weighted Eigenvector

$$\lambda\, EC_w(g_w(x)) = g_w(0)\, EC_w(g_w(x)).$$

- EC_w is not a function of x. As above, we report $\sqrt{2}\, EC_w$.
- In some applications (if higher weight is "bad") may need to invert the non-zero elements of $g_w(0)$.

Variation across Measures



- node 2 on no unweighted shortest paths, but on 2/3 weighted $x = 0$ sh. paths (1–2–3, 1–2–3–4, not 3–4)
- nodes 2 and 3 same $CC_w(0)$
- node 2 higher EC_w than node 3 (inverted weights)

Node	BC	$BC_w(0)$	$BC_w(3)$	CC	$CC_w(0)$	$CC_w(3)$	DC	DC_w	EC	EC_w
1	0	0	0	0.75	0.23	0.15	0.67	0.33	0.74	0.88
2	0	0.67	0	0.75	0.27	0.21	0.67	0.17	0.74	0.96
3	0.67	0.67	0.67	1	0.27	0.23	1	0.72	0.86	0.55
4	0	0	0	0.60	0.13	0.10	0.33	0.33	0.40	0.08

Data: Origin & Destination

Source:

- U.S. Department of Transportation DB1B and T-100 databases, covering 1999Q1 to 2013Q4 (domestic tickets)[3]

Details:

- Nonstop round-trip coach-class tickets, continental U.S.
- Aggregated to nondirectional route-carrier-quarter level
- 102,526 route-carrier-quarters, e.g., DEN_PHX_WN_2013_4 (Denver to Phoenix Sky Harbor, with Southwest Airlines)
- Raw database 150GB, parsed database 189MB
- 37 carriers, serving 231 airports (from ABE to YNG)
- **In most of this work, we focus on 2013Q4: 1,623 route-carriers, 12 carriers, 1,134 routes, 135 airports**

---

[3] e.g. Goolsbee & Syverson (2008, Quarterly Journal of Economics), Ciliberto & Tamer (2009, Econometrica), Aguirregabiria & Ho (2012, Journal of Econometrics), Dai, Liu & Serfes (2014, Review of Economics and Statistics).

Data: Carriers of Interest (1999Q1 – 2013Q4)[4]

Legacy:

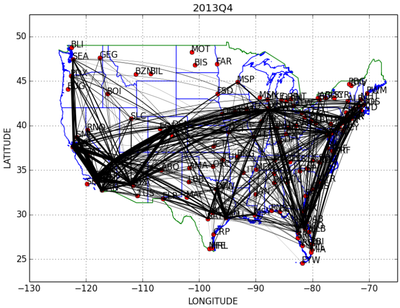
- AA (American Airlines)
- AS (Alaska Airlines)
- DL (Delta Air Lines)
- UA (United Airlines)
- US (US Airways)

Low-Cost:

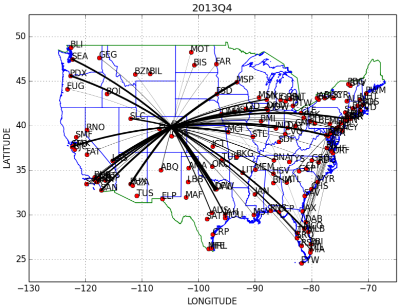
- B6 (JetBlue Airways)
- F9 (Frontier Airlines)
- FL (AirTran Airways)
- NK (Spirit Airlines)
- SY (Sun Country Airlines)
- VX (Virgin America)
- WN (Southwest Airlines)

[4] Incomplete samples for B6 (2000Q2–), SY (1999Q3–), VX.

Data: Representative Networks



Southwest Airlines (WN)

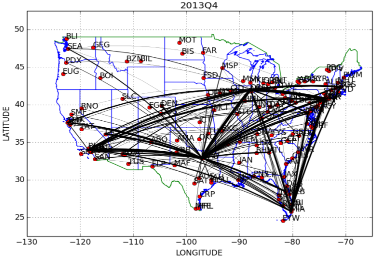


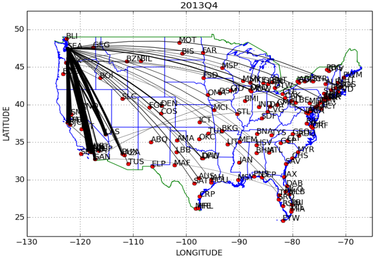
Frontier Airlines (F9)

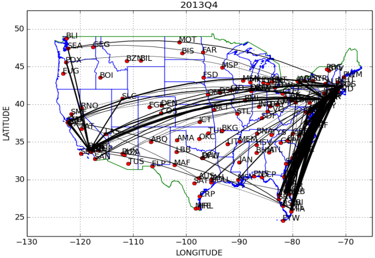
					Denver International (DEN)			
	# nodes	# edges	diameter	density[5]	BC	CC	DC	EC
WN	88	522	3	0.14	0.12	0.73	0.62	0.27
F9	58	70	4	0.04	0.96	0.92	0.95	0.68

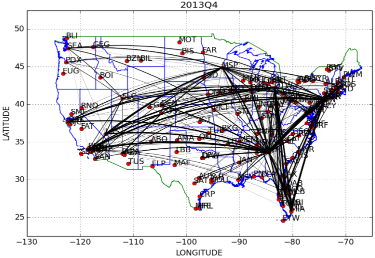
[5] Density is $(\sum_{i,j} g_{ij})/n(n-1)$.

Networks: American, Alaska, JetBlue, Delta[6]



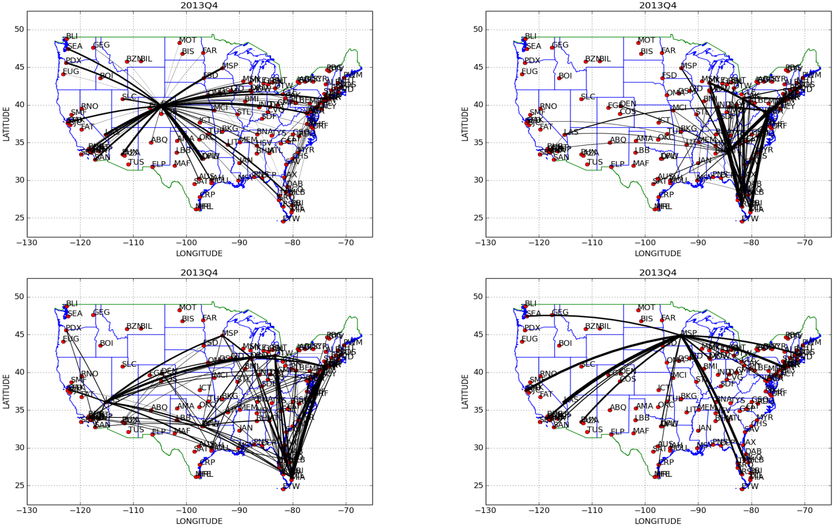






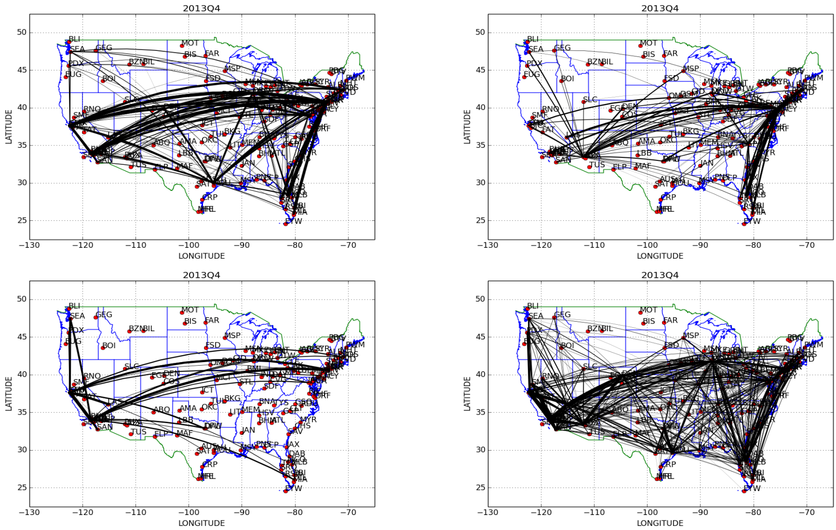
[6] American Airlines (AA), Alaska Airlines (AS), JetBlue Airways (B6), Delta Air Lines (DL).

Networks: Frontier, AirTran, Spirit, Sun Country[7]



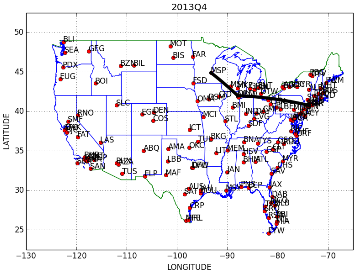
[7] Frontier Airlines (F9), AirTran Airways (FL), Spirit Airlines (NK), Sun Country Airlines (SY).

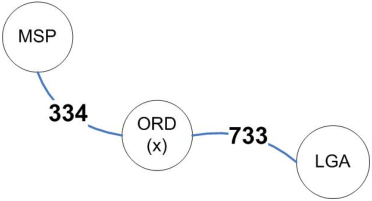
Networks: United, US, Virgin America, Southwest[8]



[8] United Airlines (UA), US Airways (US), Virgin America (VX), Southwest Airlines (WN).

Choice of Node Penalty





- Penalty x corresponds to distance-equivalent waiting time at stop on route
- Airbus 320/321, Boeing 737 Next Generation planes: cruise 0.78 Mach at 35,000 feet (approx. 500 miles/h)
- Waiting time 0h, 1h, 2h $\Longrightarrow\ x = 0, 500, 1000$.
- e.g. American Airlines route MSP–ORD–LGA: 1,067 or 1,567 or 2,067 (miles).

Airport Rankings by Centrality: Frontier Airlines

Rank	BC	$BC_w(0)$	$BC_w(1000)$	CC	$CC_w(0)$	$CC_w(1000)$	DC	DC_w	EC	EC_w
1	DEN (0.964)	DEN (0.954)	DEN (0.958)	DEN (0.919)	DEN (0.357)	DEN (0.327)	DEN (0.947)	DEN (0.517)	DEN (0.967)	DEN (0.987)
2	TTN (0.076)	MDW (0.088)	MDW (0.085)	MCO (0.528)	ABQ (0.264)	ABQ (0.146)	TTN (0.158)	TTN (0.068)	MCO (0.241)	ABQ (0.283)
3	MCO (0.020)	TTN (0.083)	TTN (0.079)	ILG (0.514)	SLC (0.256)	OMA (0.144)	MCO (0.105)	MCO (0.065)	ILG (0.228)	SLC (0.252)
4	MDW (0.012)	MCO (0.009)	MCO (0.008)	MDW (0.514)	OMA (0.244)	SLC (0.143)	ILG (0.088)	ILG (0.051)	MDW (0.178)	OMA (0.220)
5	RSW (0.012)	ILG (0.006)	ILG (0.004)	RSW (0.514)	FSD (0.240)	FSD (0.138)	MDW (0.053)	RSW (0.037)	RSW (0.178)	FSD (0.204)
6	TPA (0.012)	ABQ (0.000)	ABQ (0.000)	TPA (0.514)	OKC (0.238)	OKC (0.138)	RSW (0.053)	TPA (0.035)	TPA (0.178)	OKC (0.199)
7	ATL (0.012)	ATL (0.000)	ATL (0.000)	ATL (0.509)	BIS (0.235)	BIS (0.137)	TPA (0.053)	FLL (0.028)	TTN (0.166)	BIS (0.191)
8	DTW (0.012)	AUS (0.000)	AUS (0.000)	DTW (0.509)	BZN (0.234)	BZN (0.136)	ATL (0.035)	MDT (0.024)	BMI (0.158)	BZN (0.188)
9	FLL (0.012)	BIS (0.000)	BIS (0.000)	FLL (0.509)	MCI (0.232)	MCI (0.136)	BMI (0.035)	MDW (0.023)	MDT (0.158)	MCI (0.185)
10	ILG (0.001)	BKG (0.000)	BKG (0.000)	BMI (0.500)	DSM (0.224)	MDW (0.135)	DTW (0.035)	ATL (0.020)	OMA (0.158)	DSM (0.168)

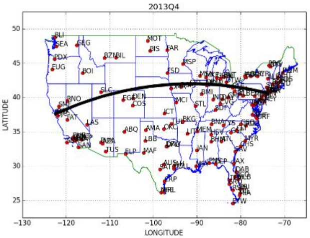
- Top-ranked airport generally same for every measure, for a given carrier (here, Denver): we call this the "dominant hub".
- Importance of dominant hub relative to lower-ranked airports, depends on carrier and measure; notable for BC and $BC_w(x)$.
  - e.g. Denver $BC_w(1000) = 0.958$ (rank 1)
  - e.g. Chicago Midway $BC_w(1000) = 0.085$ (rank 2)

Airport Rankings by Centrality: Southwest Airlines

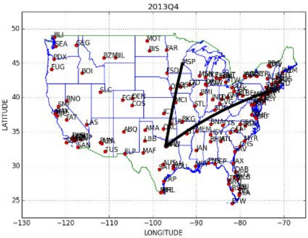
Rank	BC	$BC_w(0)$	$BC_w(1000)$	CC	$CC_w(0)$	$CC_w(1000)$	DC	DC_w	EC	EC_w
1	MDW (0.186)	MDW (0.209)	MDW (0.219)	MDW (0.777)	STL (0.159)	MDW (0.118)	MDW (0.713)	LAS (0.318)	MDW (0.400)	MDW (0.399)
2	LAS (0.134)	BWI (0.169)	BWI (0.170)	LAS (0.737)	MDW (0.155)	DEN (0.095)	LAS (0.644)	DEN (0.266)	LAS (0.387)	HOU (0.380)
3	BWI (0.122)	DEN (0.119)	DEN (0.119)	DEN (0.725)	MCI (0.151)	STL (0.095)	DEN (0.621)	MDW (0.263)	DEN (0.382)	LAS (0.352)
4	DEN (0.117)	STL (0.102)	HOU (0.112)	BWI (0.702)	BNA (0.149)	HOU (0.095)	BWI (0.575)	PHX (0.247)	PHX (0.355)	BWI (0.349)
5	HOU (0.116)	HOU (0.095)	LAS (0.090)	PHX (0.680)	TUL (0.145)	BWI (0.093)	PHX (0.529)	BWI (0.198)	BWI (0.347)	STL (0.344)
6	MCO (0.085)	DAL (0.088)	STL (0.068)	HOU (0.674)	DAL (0.143)	BNA (0.090)	HOU (0.517)	HOU (0.178)	HOU (0.321)	PHX (0.302)
7	PHX (0.058)	LAS (0.085)	BNA (0.057)	MCO (0.644)	SDF (0.142)	MCI (0.086)	MCO (0.448)	MCO (0.175)	STL (0.285)	BNA (0.292)
8	BNA (0.030)	BNA (0.075)	MCO (0.053)	STL (0.617)	OKC (0.142)	LAS (0.082)	STL (0.379)	TPA (0.145)	MCO (0.278)	DEN (0.287)
9	STL (0.029)	MCO (0.052)	DAL (0.038)	BNA (0.608)	CMH (0.138)	MCO (0.080)	BNA (0.356)	LAX (0.117)	BNA (0.272)	DAL (0.270)
10	TPA (0.024)	MCI (0.030)	PHX (0.025)	TPA (0.604)	HOU (0.138)	AUS (0.079)	TPA (0.356)	STL (0.115)	TPA (0.259)	AUS (0.256)

- Similar values for top five or six airports (multiple hubs).
- Dominant hub is Chicago Midway
  - e.g. $BC_w(1000) = 0.219$ and $DC = 0.713$
- Some variation in rankings for non-dominant hubs
  - e.g. Las Vegas McCarran $BC = 0.134$ (rank 2)
  - e.g. Las Vegas McCarran $BC_w(1000) = 0.090$ (rank 5)

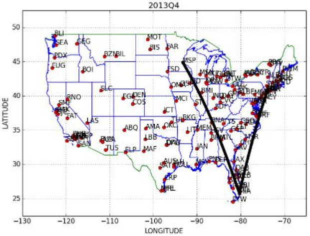
Multiple Shortest Paths: Use BC_w Instead of BC?



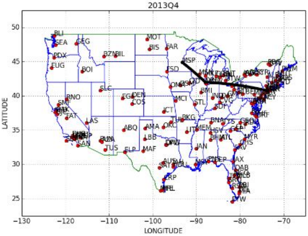
(a) *Minimum-distance JFK–SFO and JFK–ORD–SFO.*



(b) *Minimum-step LGA–DFW–MSP.*



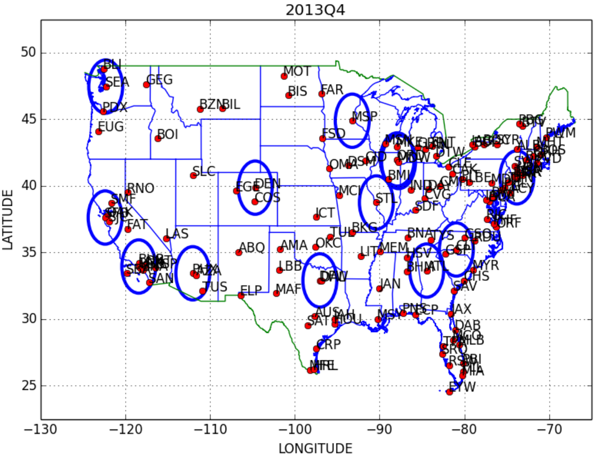
(c) *Minimum-step LGA–MIA–MSP.*



(d) *Minimum-step LGA–ORD–MSP.*

A variety of (un)weighted shortest paths, American Airlines, 2013Q4.

Spatial Distribution of Dominant Hubs



- With two exceptions, dominant hub different for each carrier.
- Dominant hubs spread quite evenly across U.S.

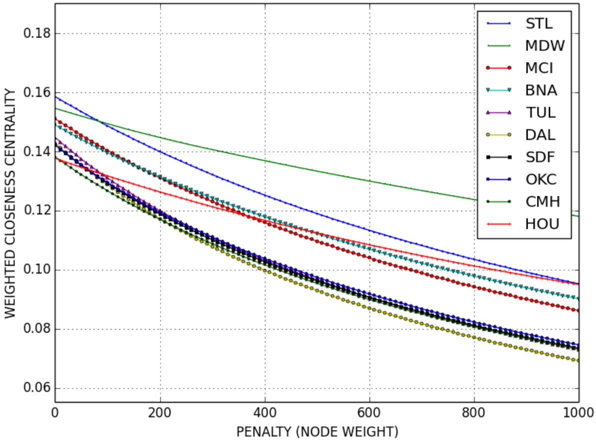
Correlations between Centrality Measures

	$BC_w(0)$	$BC_w(500)$	$BC_w(1000)$	CC	$CC_w(0)$	$CC_w(500)$	$CC_w(1000)$	DC	DC_w	EC	EC_w
BC	0.90	0.95	0.95	0.84	0.26	0.58	0.73	0.91	0.88	0.80	0.77
$BC_w(0)$		0.98	0.98	0.76	0.42	0.68	0.78	0.84	0.75	0.75	0.80
$BC_w(500)$			1.00	0.78	0.35	0.64	0.76	0.85	0.77	0.74	0.77
$BC_w(1000)$				0.78	0.35	0.64	0.76	0.85	0.77	0.74	0.77
CC					0.37	0.72	0.86	0.95	0.94	0.97	0.89
$CC_w(0)$						0.90	0.77	0.28	0.19	0.32	0.49
$CC_w(500)$							0.97	0.64	0.57	0.68	0.77
$CC_w(1000)$								0.79	0.73	0.82	0.86
DC									0.98	0.97	0.91
DC_w										0.95	0.84
EC											0.92

Correlations, Southwest Airlines, 2013Q4

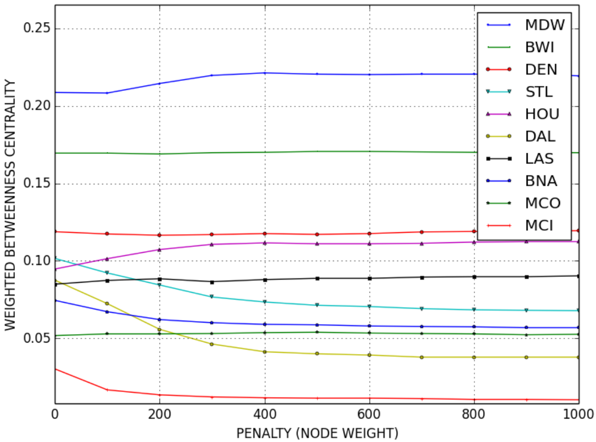
- Bloch, Jackson & Tebaldi (2016, arXiv:1608.05845v1)
  - standard centrality measures characterized by same axioms
  - simulated data: most correlations 0.8 – 1 (Erdős–Rényi, homophily)
- Valente, Coronges, Lakon & Costenbader (2008, Connections), Bolland (1988, Social Networks)
  - real data: average correlation 0.4 – 0.9 (58 datasets, Valente et al.)
  - real data: correlations 0.5 – 0.9 (1 dataset, Bolland)

Robustness of $CC_w(x)$ to Node-Weight: Southwest



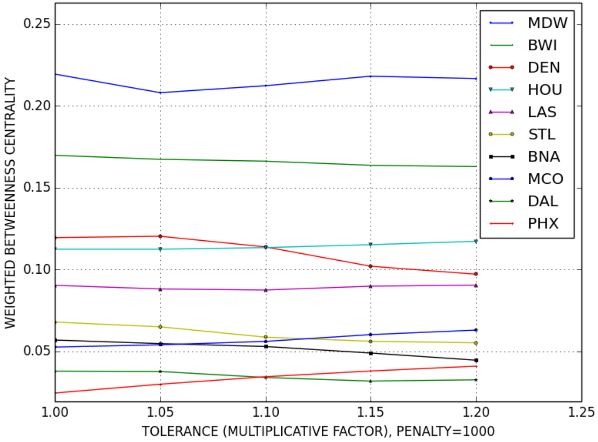
- Note change in rankings of Chicago Midway (MDW) and William P. Hobby Houston (HOU).

Robustness of $BC_w(x)$ to Node-Weight: Southwest



- Note change in rankings of Lambert–St. Louis (STL) and Dallas Love Field (DAL).

Robustness of $BC_w(x)$ to Approximate Shortest Paths



- Robustness of $BC_w(1000)$ to inclusion of paths up to 20% longer than true minimum-distance path.

Illustrative Regression: Econometric Model[9]

- Cross sectional model with carrier fixed effects

$$p_{ij} = \alpha + b_i + x_{ij}'\beta_{network} + w_{ij}'\beta_{controls} + u_{ij}$$

with

- p_{ij} = mean real fare for carrier i, route j
- b_i = carrier fixed effect
- x_{ij} = network variables
- w_{ij} = control variables
- u_{ij} = error term

- Weighted least squares: $p_{ij} = N_{ij}^{-1} \sum_{k=1}^{N_{ij}} p_{ijk}$, where k is an individual ticket and N_{ij} is the carrier-route pax; weight $N_{ij}^{1/2}$
- Data from 2013Q4 used for illustrations

[9] Cochrane (2005, Chicago working paper): "Never use the words "illustrative test" or "illustrative empirical work." Never do illustrative work. Do real empirical work or don't do any at all. Illustrating technique with empirical work you don't believe in is a waste of space. Even if you do it, there is no faster way to get readers to fall asleep than to tell them that what you're doing doesn't really matter."

Illustrative Regression: Related Literature

- Unweighted centrality measures to characterize network (Airline)
  - Shaw (1993, Journal of Transport Geography)
- Simple hub measures as explanatory variables (Airline)
  - Borenstein (1989, RAND Journal of Economics)
  - Reiss & Spiller (1989, Journal of Law and Economics)
  - Borenstein (1990, American Economics Review)
  - Brueckner, Dyer & Spiller (1992, RAND Journal of Economics)
  - Kahn (1993, Review of Industrial Organization)
- Unweighted centralities as explanatory variables (Sociology)
  - Faris & Felmlee (2011, American Sociological Review)
- Unweighted centralities as explanatory variables (Finance)
  - Robinson & Stuart (J. of Law, Economics & Organization)
  - Hochberg, Ljunqvist & Lu (2007, Journal of Finance)
  - Cohen-Cole, Kirilenko & Patacchini (2014, J. Fin. Economics)
  - El-Khatib, Fogel & Jandik (2015, J. of Financial Economics)

Illustrative Regression: Results (Unweighted Centrality)

$meanrealfare_{ij}$	(1)[10]	(2)	(3)	(4)	(5)
$constant$	142.70^{***}	125.99^{***}	23.22	154.30^{***}	116.51^{***}
$mindegree_{ij} * 10$	-	7.99^{***}	-	-	-
$maxdegree_{ij} * 10$	-	5.85^{***}	-	-	-
$mincloseness_{ij} * 10$	-	-	14.33^{***}	-	-
$maxcloseness_{ij} * 10$	-	-	9.71^{***}	-	-
$minbetweenness_{ij} * 10$	-	-	-	19.90^{***}	-
$maxbetweenness_{ij} * 10$	-	-	-	4.84^{***}	-
$mineigenvector_{ij} * 10$	-	-	-	-	15.54^{***}
$maxeigenvector_{ij} * 10$	-	-	-	-	16.62^{***}
$distance_j/100$	22.08^{***}	19.15^{***}	19.65^{***}	20.36^{***}	19.06^{***}
$(distance_j/100)^2$	-0.40^{***}	-0.30^{***}	-0.32^{***}	-0.34^{***}	-0.31^{***}
$abstempdiff_j$	-0.91^{**}	-0.88^{***}	-0.85^{**}	-0.84^{***}	-0.87^{***}
$meangdppercap_j$	0.42^{*}	0.46^{*}	0.42^{**}	0.39^{*}	0.44^{*}
$t100seats_{ij}/100000$	6.14^{**}	$-3.99^{\cdot}$	-1.71	-0.67	-3.97
$monopoly_j$	32.10^{***}	32.39^{***}	33.11^{***}	31.76^{***}	33.63^{***}
$competitive_j$	-24.80^{***}	-23.93^{***}	-23.85^{***}	-22.92^{***}	-24.61^{***}
carrier dummies	yes	yes	yes	yes	yes
carrier $\times$ $southwest_j$	yes	yes	yes	yes	yes
adjusted R^2	0.786	0.804	0.801	0.798	0.806

[10] Significance: *** 99.9%, ** 99%, * 95%, · 90%; White's s.e.'s.; WN omitted

Illustrative Regression: Results (Weighted Centrality)

$meanrealfare_{ij}$	(1)[11]	(2)	(3)	(4)	(5)
$constant$	142.70^{***}	128.45^{***}	85.94	153.04^{***}	125.72^{***}
$mindegree\_w_{ij} * 10$	-	11.80^{**}	-	-	-
$maxdegree\_w_{ij} * 10$	-	11.13^{***}	-	-	-
$mincloseness\_w(1000)_{ij} * 10$	-	-	61.77^{***}	-	-
$maxcloseness\_w(1000)_{ij} * 10$	-	-	33.19^{***}	-	-
$minbetweenness\_w(1000)_{ij} * 10$	-	-	-	15.47^{**}	-
$maxbetweenness\_w(1000)_{ij} * 10$	-	-	-	4.51^{***}	-
$mineigenvector\_w_{ij} * 10$	-	-	-	-	7.90^{***}
$maxeigenvector\_w_{ij} * 10$	-	-	-	-	7.26^{***}
$distance_j/100$	22.08^{***}	20.42^{***}	20.60^{***}	20.56^{***}	21.11^{***}
$(distance_j/100)^2$	-0.40^{***}	-0.36^{***}	-0.33^{***}	-0.34^{***}	-0.36^{***}
$abstempdiff_j$	-0.91^{**}	-0.90^{***}	-0.90^{**}	-0.81^{**}	-0.89^{***}
$meangdppercap_j$	0.42^{*}	0.46^{*}	0.44^{*}	$0.39^{\cdot}$	0.41^{*}
$t100seats_{ij}/100000$	6.14^{**}	-0.56	-0.39	-0.36	-2.53
$monopoly_j$	32.10^{***}	33.80^{***}	32.49^{***}	31.34^{***}	30.73^{***}
$competitive_j$	-24.80^{***}	-25.44^{***}	-21.887^{***}	-23.58^{***}	-21.22^{***}
carrier dummies	yes	yes	yes	yes	yes
carrier $\times$ $southwest_j$	yes	yes	yes	yes	yes
adjusted R^2	0.786	0.796	0.802	0.797	0.804

[11] Significance: *** 99.9%, ** 99%, * 95%, · 90%; White's s.e's.; WN omitted

Why do (Un)weighted Measures Give Similar Results(?)

- (Possible answer 1) Airline networks are very special, with few nodes, a regulated initial state, dominant hubs, a (local) star-type structure, and are (globally) stable. The centrality measures would differ in other types of network.

Why do (Un)weighted Measures Give Similar Results(?)

- (Possible answer 1) Airline networks are very special, with few nodes, a regulated initial state, dominant hubs, a (local) star-type structure, and are (globally) stable. The centrality measures would differ in other types of network.
- (Possible answer 2) Distance is not a good choice of weight, and the topological and weighted "worlds" (networks) contain very similar information, and are determined endogenously.[12]

[12] Can we just ignore spatial information when analyzing airline networks?!

Why do (Un)weighted Measures Give Similar Results(?)

- (Possible answer 1) Airline networks are very special, with few nodes, a regulated initial state, dominant hubs, a (local) star-type structure, and are (globally) stable. The centrality measures would differ in other types of network.
- (Possible answer 2) Distance is not a good choice of weight, and the topological and weighted "worlds" (networks) contain very similar information, and are determined endogenously.[12]
- (Possible answer 3) Measures give similar results in rankings and regression, but would differ in another application.

[12] Can we just ignore spatial information when analyzing airline networks?!

Why do (Un)weighted Measures Give Similar Results(?)

- (Possible answer 1) Airline networks are very special, with few nodes, a regulated initial state, dominant hubs, a (local) star-type structure, and are (globally) stable. The centrality measures would differ in other types of network.
- (Possible answer 2) Distance is not a good choice of weight, and the topological and weighted "worlds" (networks) contain very similar information, and are determined endogenously.[12]
- (Possible answer 3) Measures give similar results in rankings and regression, but would differ in another application.
- (Possible answer 4) Standard (un)weighted centrality measures are highly correlated, and will always give similar results.

[12] Can we just ignore spatial information when analyzing airline networks?!

Why do (Un)weighted Measures Give Similar Results(?)

- (Possible answer 1) Airline networks are very special, with few nodes, a regulated initial state, dominant hubs, a (local) star-type structure, and are (globally) stable. The centrality measures would differ in other types of network.
- (Possible answer 2) Distance is not a good choice of weight, and the topological and weighted "worlds" (networks) contain very similar information, and are determined endogenously.[12]
- (Possible answer 3) Measures give similar results in rankings and regression, but would differ in another application.
- (Possible answer 4) Standard (un)weighted centrality measures are highly correlated, and will always give similar results.
- (Possible answer 5) Alexandre made a mistake in his code.

[12] Can we just ignore spatial information when analyzing airline networks?!

Why do (Un)weighted Measures Give Similar Results(?)

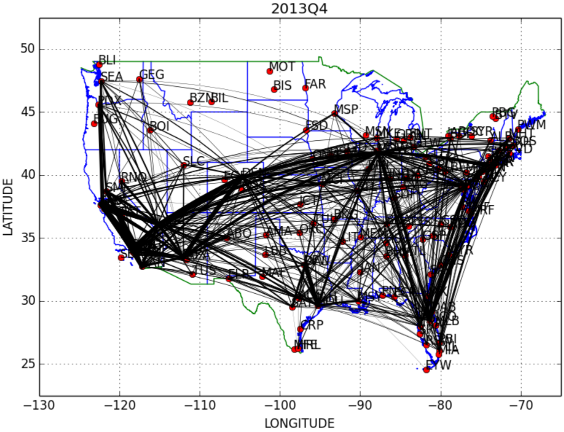
- (Possible answer 1) Airline networks are very special, with few nodes, a regulated initial state, dominant hubs, a (local) star-type structure, and are (globally) stable. The centrality measures would differ in other types of network.
- (Possible answer 2) Distance is not a good choice of weight, and the topological and weighted “worlds” (networks) contain very similar information, and are determined endogenously.[14]
- (Possible answer 3) Measures give similar results in rankings and regression, but would differ in another application.
- (Possible answer 4) Standard (un)weighted centrality measures are highly correlated, and will always give similar results.
- ~~(Possible answer 5) Alexandre made a mistake in his code.~~
- (Possible answer 6) Steve made a mistake in his code.

[14] Can we just ignore spatial information when analyzing airline networks?!

What Next?

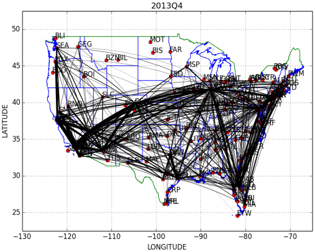
- Different edge-weights (not distance, e.g. pax)
- Different node-weights (not constant)
- Different networks (not airline data)
- Directed networks (g (or g_w) not symmetric)
- Allow self-loops (aggregation)
- More data (include connecting / codesharing tickets)
- New centrality measures (local – ? – global)
- Better regression models (panel, quantile, instruments)
- Advanced econometric models (structural? game theory?)
- Network dynamics (network evolution, centrality evolution)
- Study diffusion across networks (local? global?)

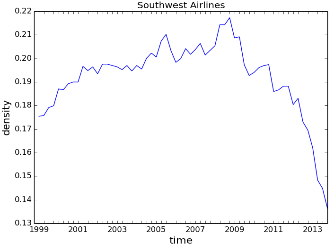
What Next? — Network Evolution

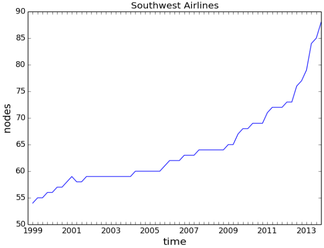


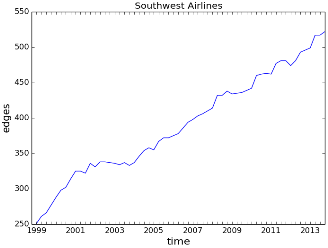
Southwest Airlines, 1999Q1 – 2013Q4 (60 quarters)

Network Evolution — Global[15]



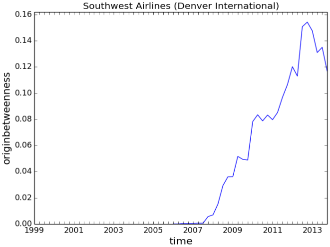


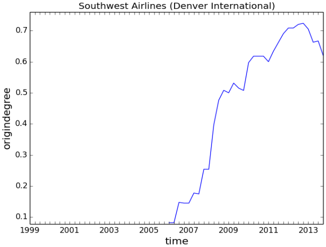


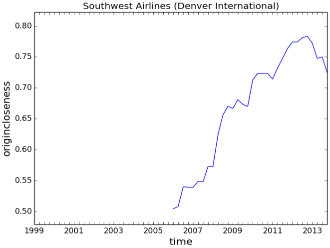


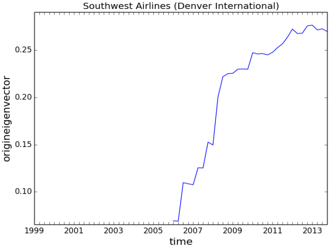
[15] Southwest network plot (2013Q4), density, number of nodes, number of edges

Centrality Measure Evolution — Local[16]









[16] Betweenness, closeness, degree, eigenvector centralities.