Statistical Analysis of Networks

Closeness & Betweenness Centrality

When we say a *node* is "central,"
what do we mean conceptually?

Degree Centrality

- Nodes that have many edges are “central”.
  - We can operationalize this using the degree (for undirected graphs) or the indegree/outdegree (for directed graphs).

Motivating Example

J Youth Adolescence (2014) 43:104–115
DOI 10.1007/s10964-013-9946-0

EMPIRICAL RESEARCH

"Role Magnets"? An Empirical Investigation of Popularity Trajectories for Life-Course Persistent Individuals During Adolescence

Jacob T. N. Young

PNAS

RESEARCH ARTICLE | POLITICAL SCIENCES

OPEN ACCESS

Disrupting hate: The effect of deplatforming hate organizations on their online audience

Daniel Robert Thomas[a,1] and Laila A. Wahedi[a]

Edited by Timothy Wilson, University of Virginia, Charlottesville, VA; received August 17, 2022; accepted January 20, 2023

Conceptualization

- “Everyone agrees, it seems, that centrality is an important structural attribute of networks. All concede that it is related to a high degree to other important group properties and processes. But there consensus ends.” (Freeman, 1978/1979: 217)
- The type of measure we use depends on the substantive question of interest.
  - Various measures of centrality are correlated, but they operationalize different concepts.

What are some other ways a node can be “central”?

Motivating Example

- Problem: Bullying
- Question: Who is likely to be a bully?
  - Typical explanation…
  - Faris & Felmlee's argument…

American Sociological Review
76(1) 48–73

DOI: 10.1177/0003122410396196
http://asr.sagepub.com
SAGE

Status Struggles: Network Centrality and Gender Segregation in Same- and Cross-Gender Aggression

Robert Faris[a] and Diane Felmlee[a]

Motivating Example

- Findings?

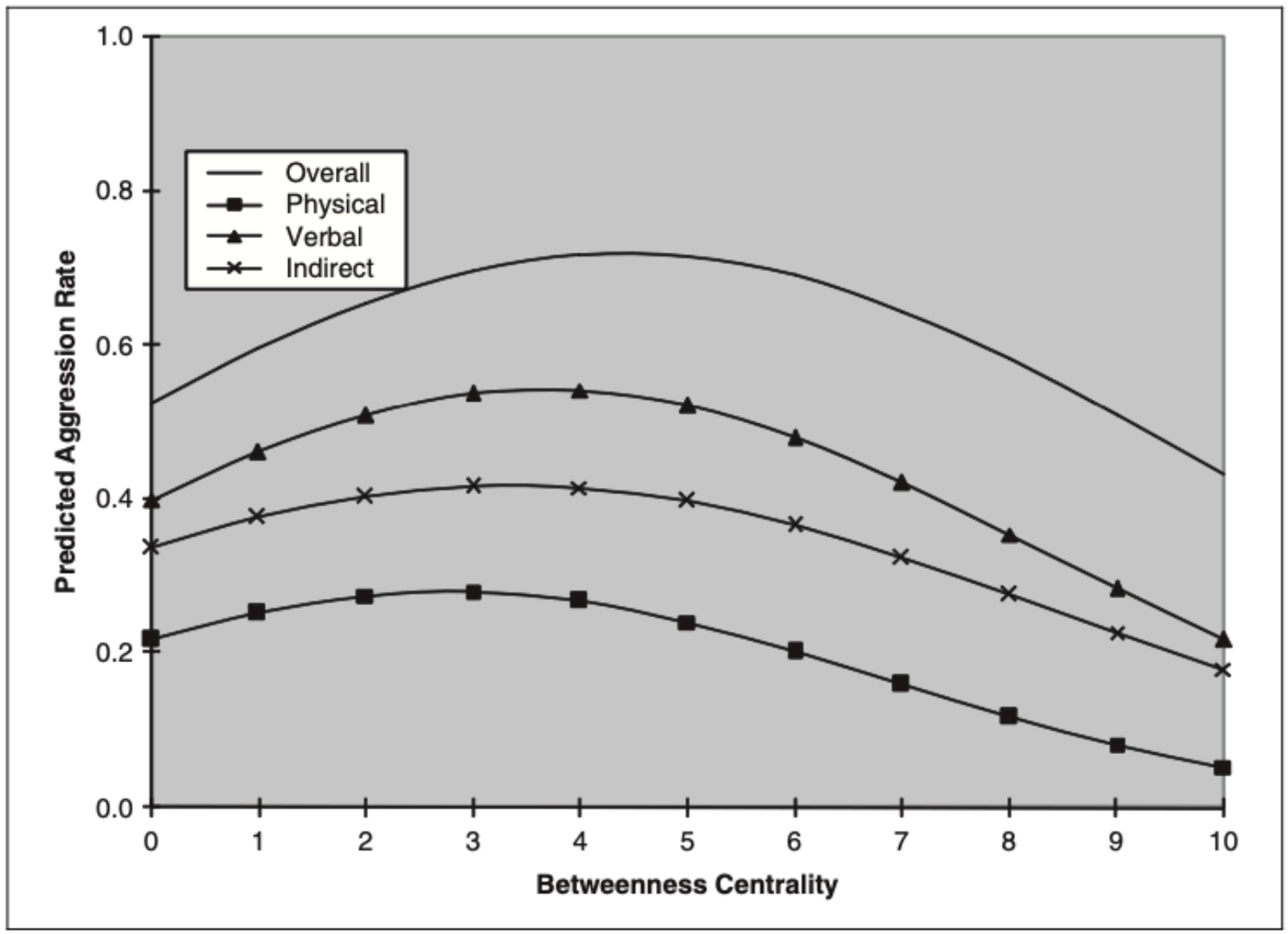


Figure 2. Predicted Aggression Rate by Social Network Centrality

Motivating Example

- Findings?

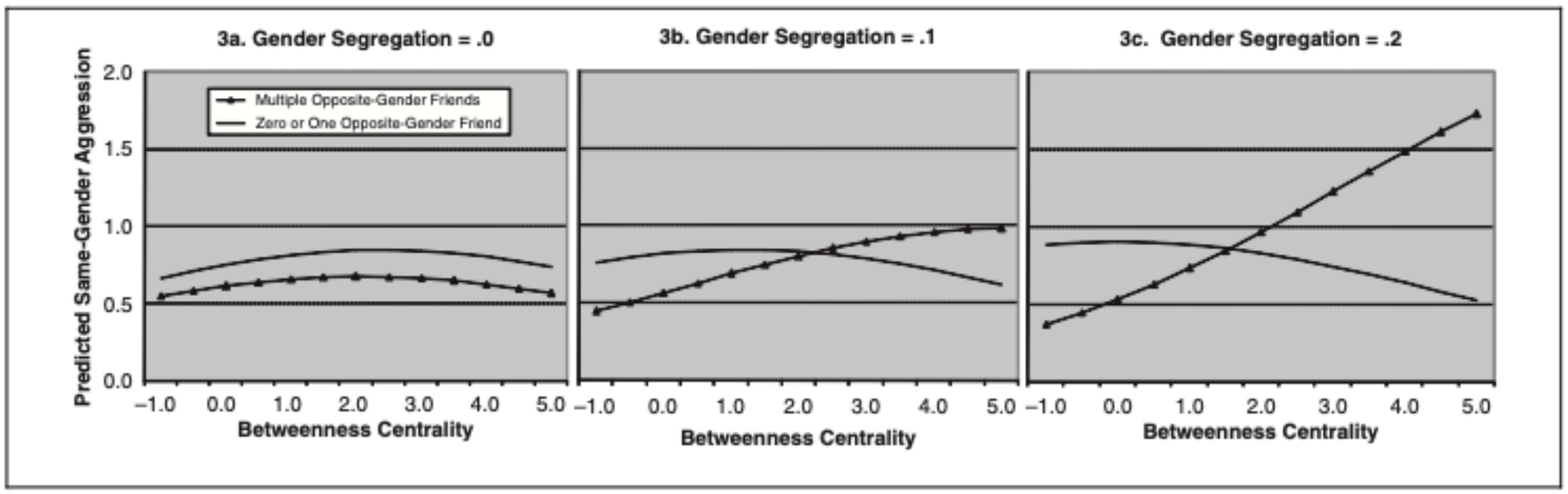


Figure 3. Predicted Rate of Same-Gender Aggression, by Centrality, Gender Segregation, and Cross-Gender Friendships

Learning Goals

- At the end of the lecture, you should be able to answer these questions:
  - What are some different ways we can conceptualize "centrality"?
  - What is *closeness* and *betweenness* centrality?
  - How do we calculate these measures for undirected and directed graphs?
  - What do comparing these measures tell us about the structure of a network?

Closeness Centrality

- How "close" is a node to other nodes?
- In a graph, *closeness centrality* measures how near a node is to the other nodes in the network.
  - Nodes that are central are those that can reach other nodes in short distances.

Closeness Centrality

- Closeness is based on the inverse of the distance of each actor to every other actor.
- Terminology:
  - A **geodesic** is the shortest path between two nodes.
  - The **distance**, $d(n_i,n_j)$, is the length of the path between i and j.

Closeness Centrality: Undirected Binary Graphs

$$g$$

$$C_C(n_i) = [\sum_{j=1} d(n_i, n_j)]^{-1}$$

Closeness Centrality: Undirected Binary Graphs

$$g$$

$$C_C(n_i) = [\sum_{j=1} d(n_i, n_j)]^{-1}$$

How far is *i* from every node?

Closeness Centrality: Undirected Binary Graphs

$$g$$

$$C_C(n_i) = [\sum_{j=1} d(n_i, n_j)]^{-1}$$

How far is *i* from every node?

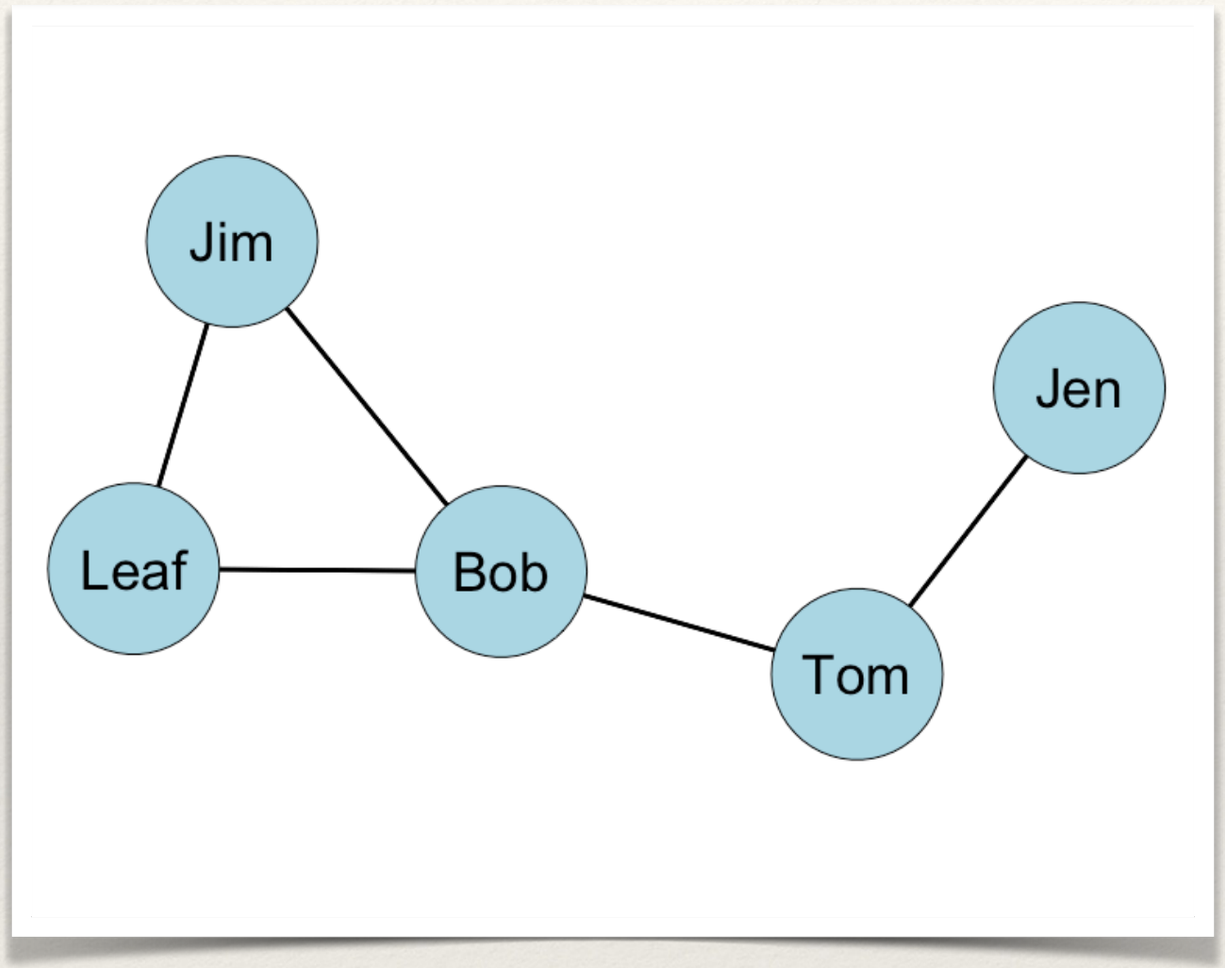
Then take the inverse of the sum.

Closeness Centrality: Undirected Binary Graphs

$$C_C(n_i) = [\sum_{j=1}^{g} d(n_i, n_j)]^{-1} = \frac{1}{[\sum_{j=1}^{g} d(n_i, n_j)]}$$

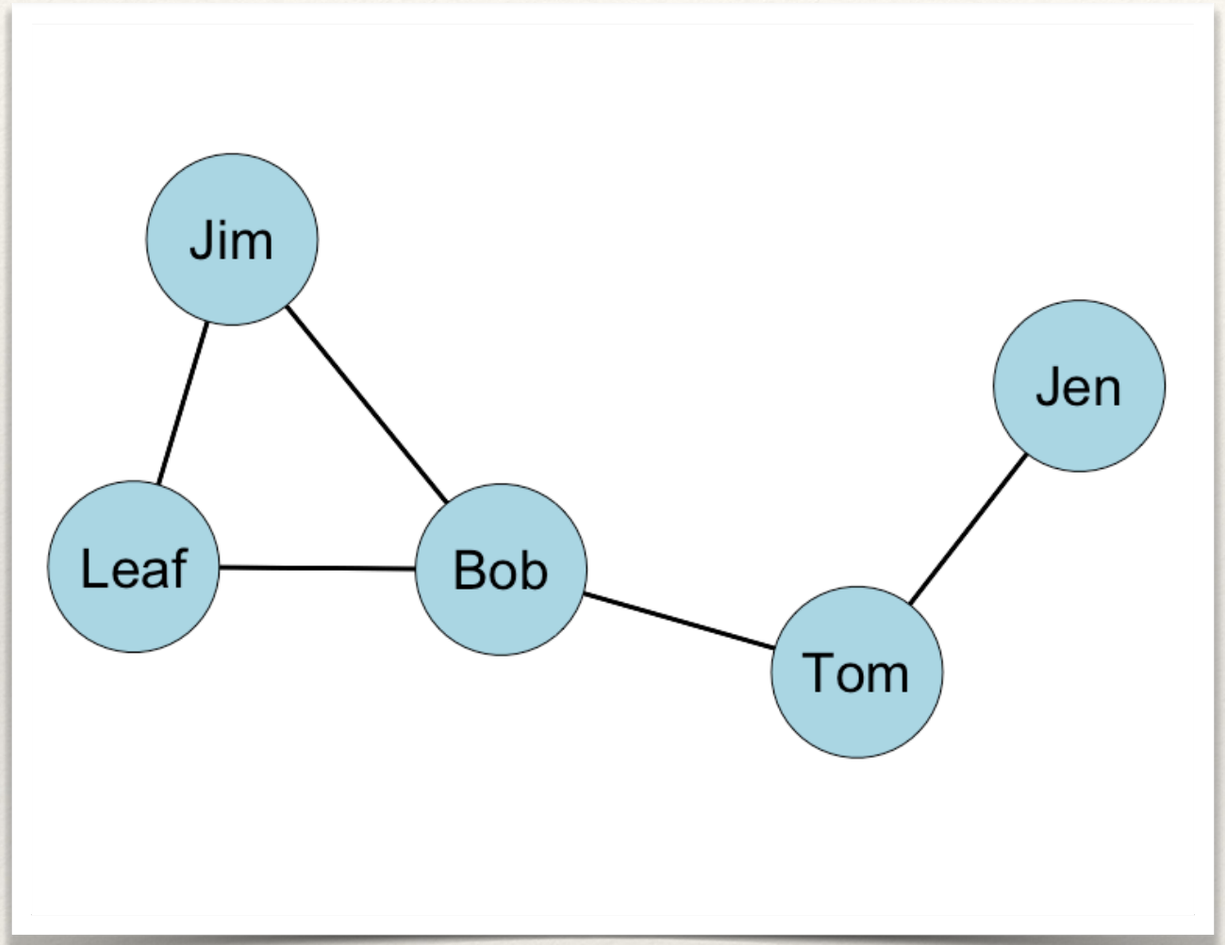
Undirected Networks

Example: Undirected, Binary Network



We want a matrix of the distances between each node.

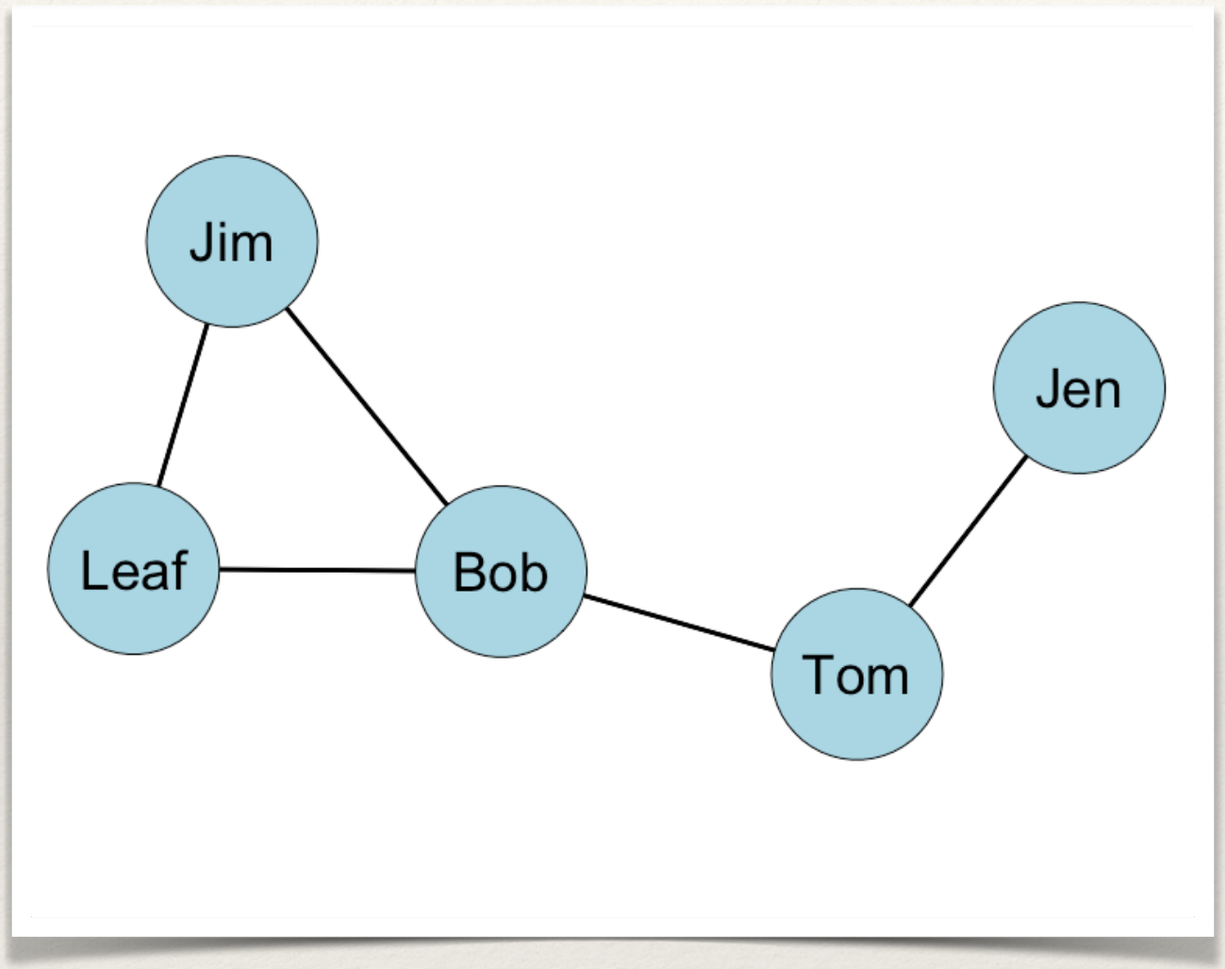
Example: Undirected, Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen		?	?		
Tom					
Bob					
Leaf					
Jim					

How far is Jen from Tom? From Bob?

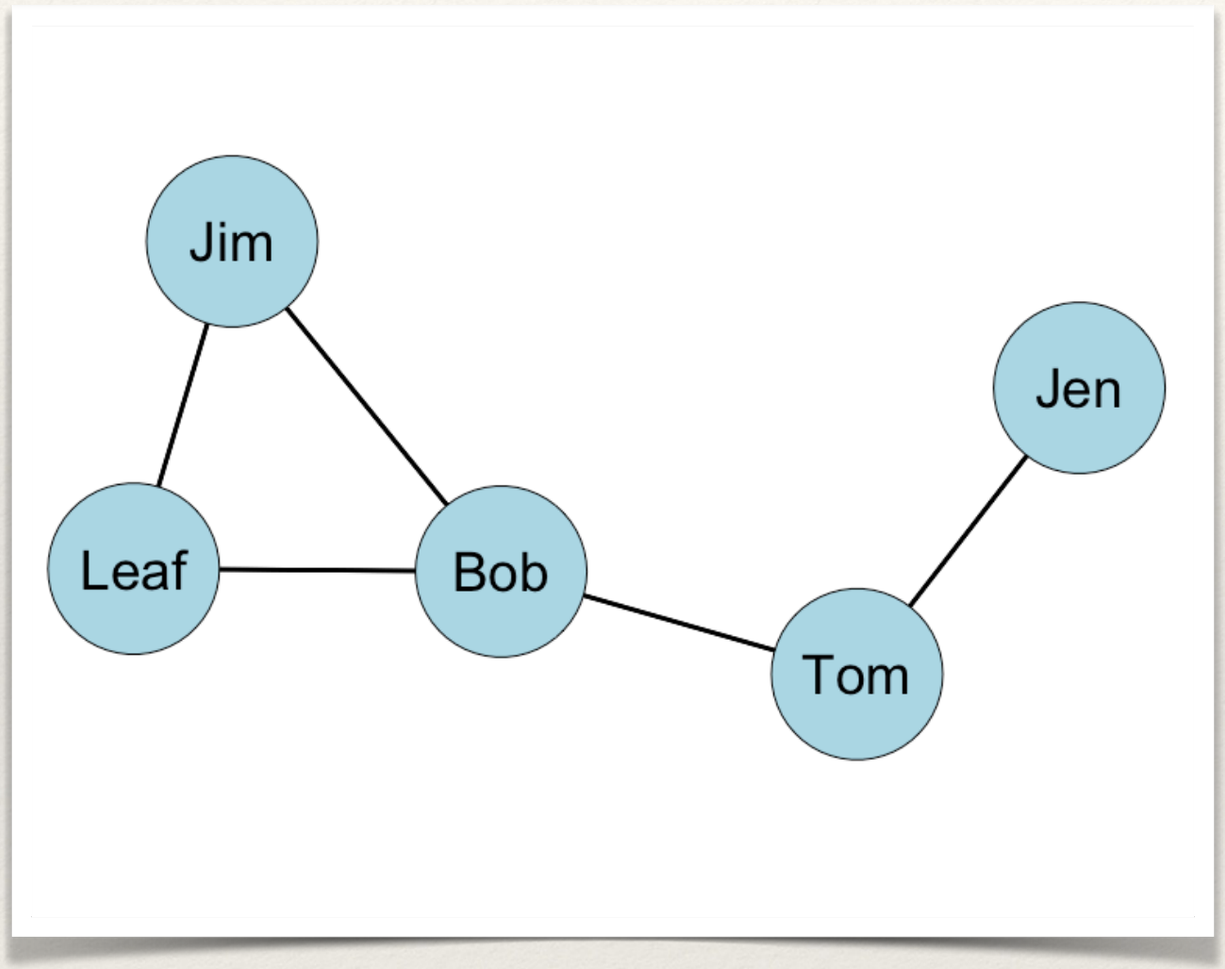
Example: Undirected, Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen		1			
Tom					
Bob					
Leaf					
Jim					

How far is Jen from Tom? From Bob?

Example: Undirected, Binary Network

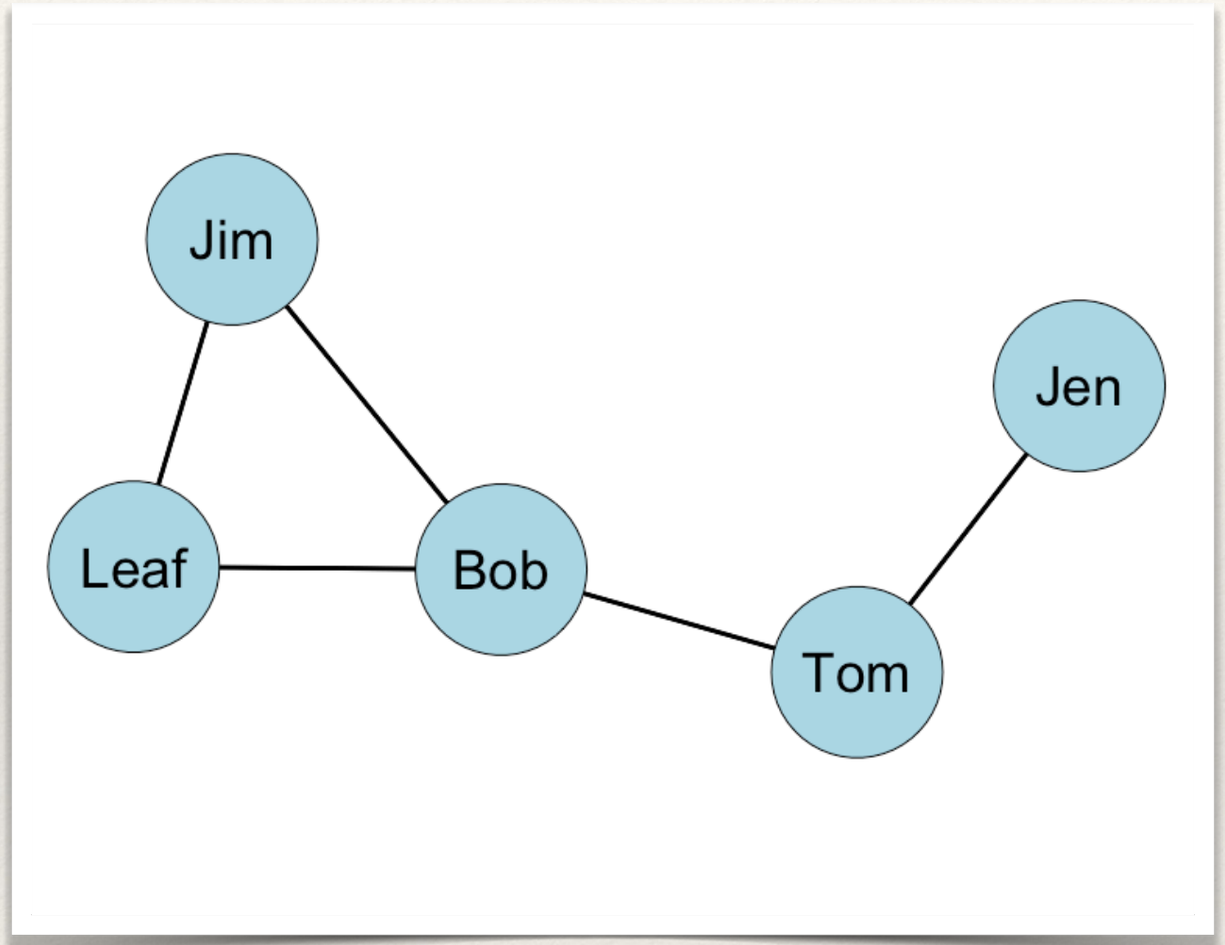


Distance Matrix

	Jen	Tom	Bob	Leaf	Jim
Jen		1	2		
Tom					
Bob					
Leaf					
Jim					

How far is Jen from Tom? From Bob?

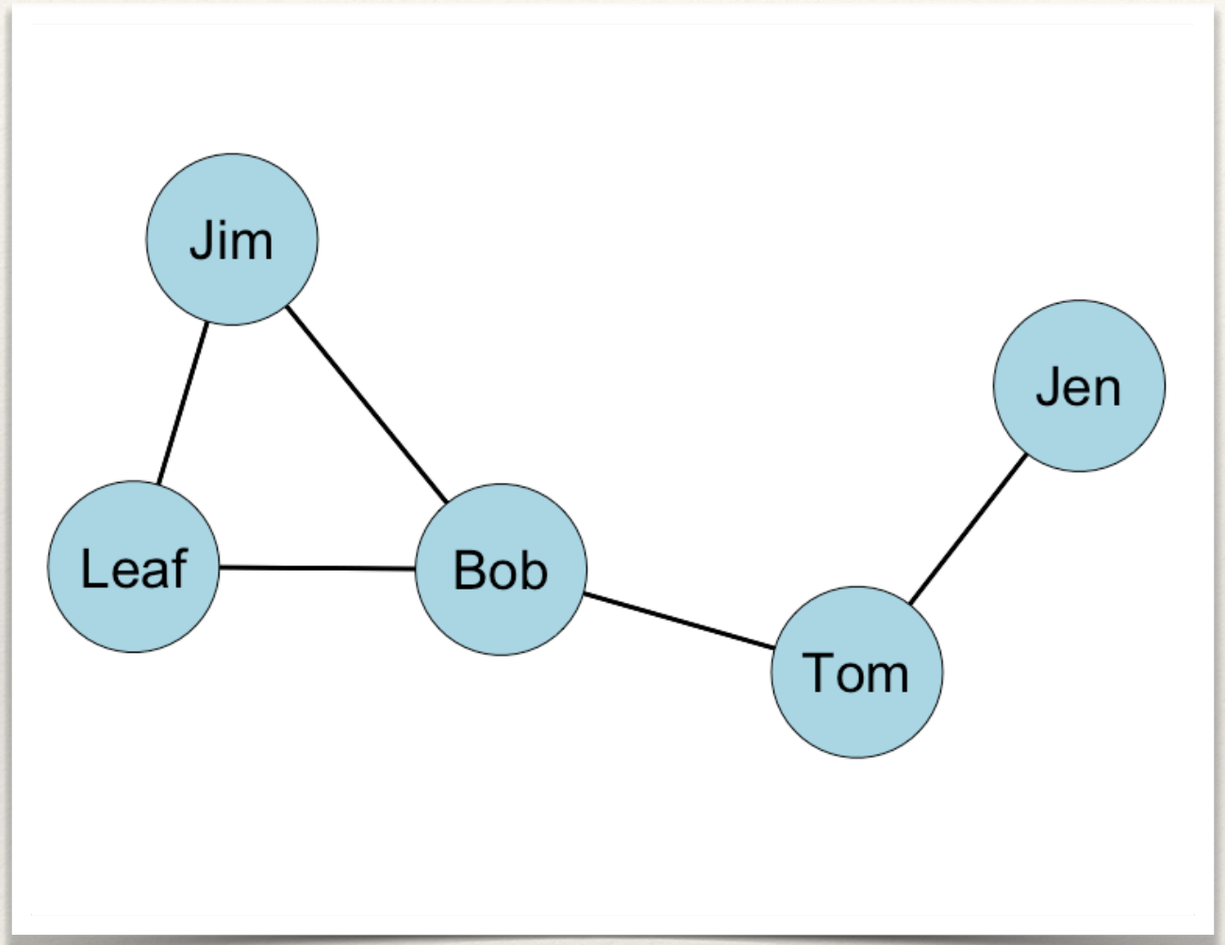
Example: Undirected, Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen		1	2	?	?
Tom			?	?	?
Bob				?	?
Leaf					?
Jim					

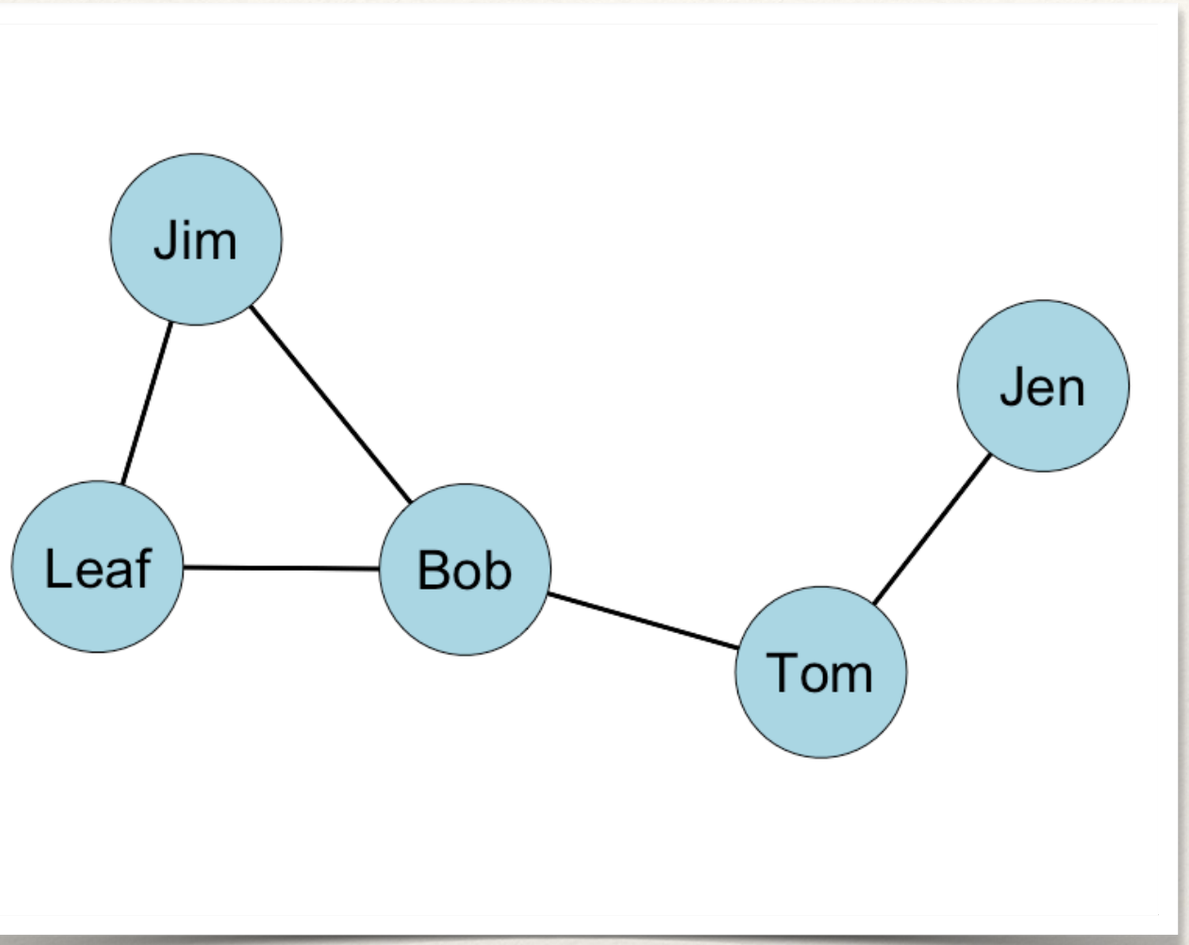
Now, fill in the rest…

Example: Undirected, Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen		1	2	3	3
Tom			1	2	2
Bob				1	1
Leaf					1
Jim					

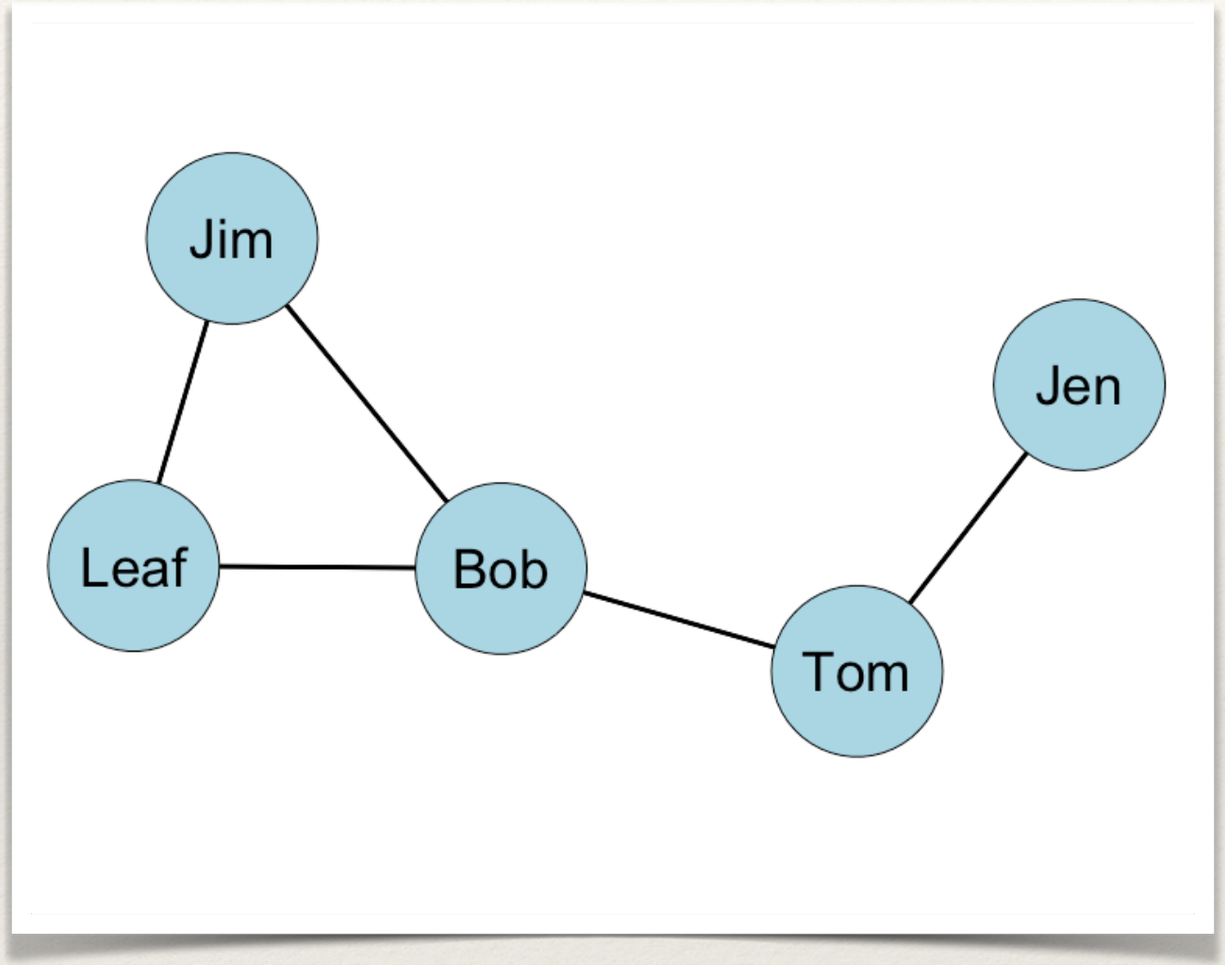
Example: Undirected, Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen		1	2	3	3
Tom	1		1	2	2
Bob	2	1		1	1
Leaf	3	2	1		1
Jim	3	2	1	1	

Since the graph is undirected, the distance matrix is symmetric about the diagonal.

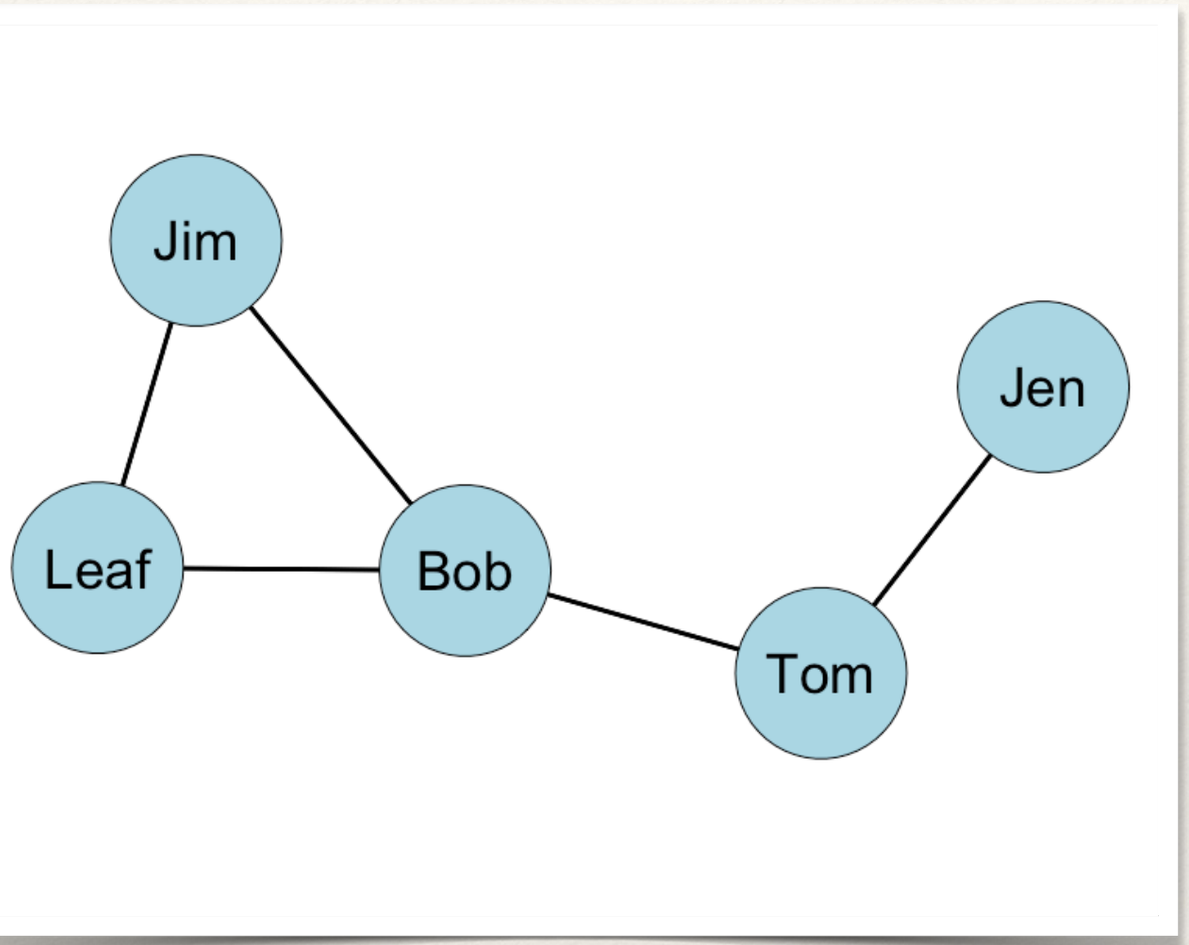
Example: Undirected, Binary Network



Distance Matrix						
	Jen	Tom	Bob	Leaf	Jim	Sum
Jen		1	2	3	3	9
Tom	1		1	2	2	6
Bob	2	1		1	1	5
Leaf	3	2	1		1	7
Jim	3	2	1	1		7

The row sum of these is the distances.

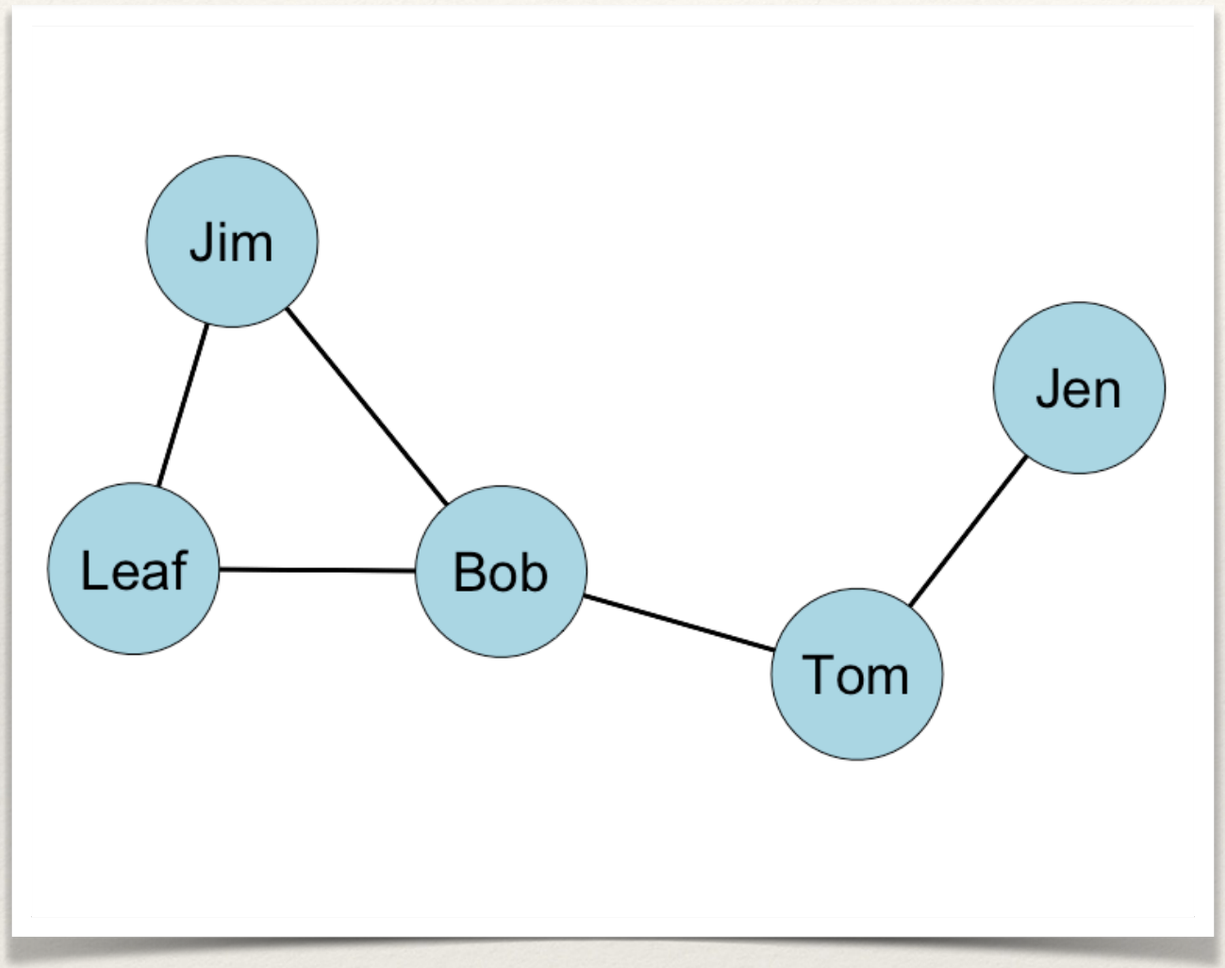
Example: Undirected, Binary Network



Distance Matrix						
	Jen	Tom	Bob	Leaf	Jim	Sum
Jen		1	2	3	3	9
Tom	1		1	2	2	6
Bob	2	1		1	1	5
Leaf	3	2	1		1	7
Jim	3	2	1	1		7
Sum	9	6	5	7	7	

Since the graph is undirected, the column sum reports the same information.

Example: Undirected, Binary Network



By taking the reciprocal, we get the closeness centrality score.

Closeness Centrality

Jen = 1/9 = 0.111

Tom = 1/6 = 0.167

Bob = 1/5 = 0.200

Leaf = 1/7 = 0.143

Jim = 1/7 = 0.143

$$C_C(n_i) = \frac{1}{[\sum_{j=1}^{g} d(n_i, n_j)]}$$

Closeness Centrality: Undirected Binary Graphs

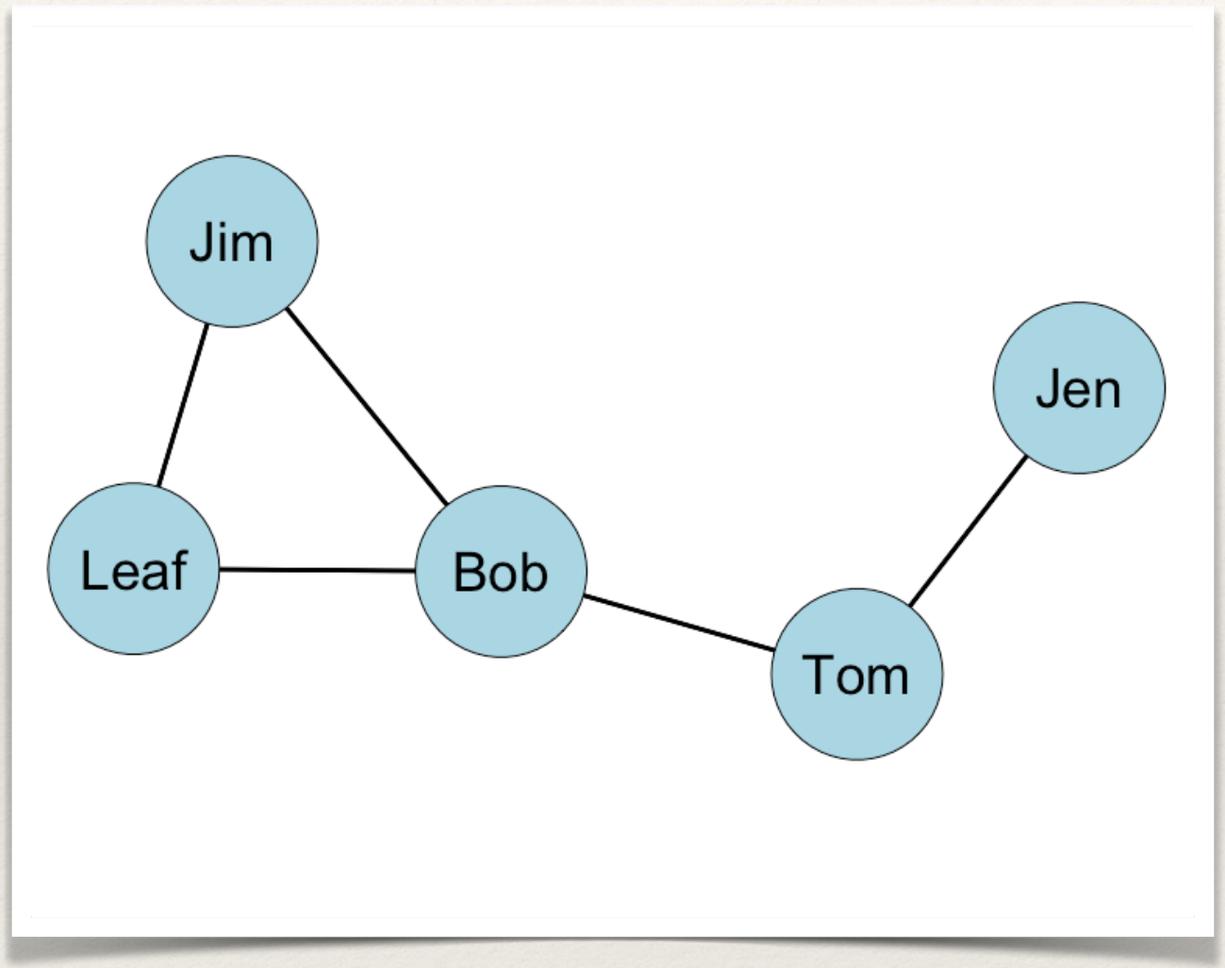
- Actor closeness centrality not only reflects each node's connectivity to other nodes but also depends on the size of the network, g.
- Summing over more nodes will push scores closer to zero.
  - *Solution?*

Standardized Closeness Centrality

$$C'_C(n_i) = [\sum_{j=1}^{g} d(n_i, n_j)]^{-1} \times [g - 1]$$

Multiply by g -1!

Example: Undirected, Binary Network



Closeness Centrality

Jen = 1/9 = 0.111

Tom = 1/6 = 0.167

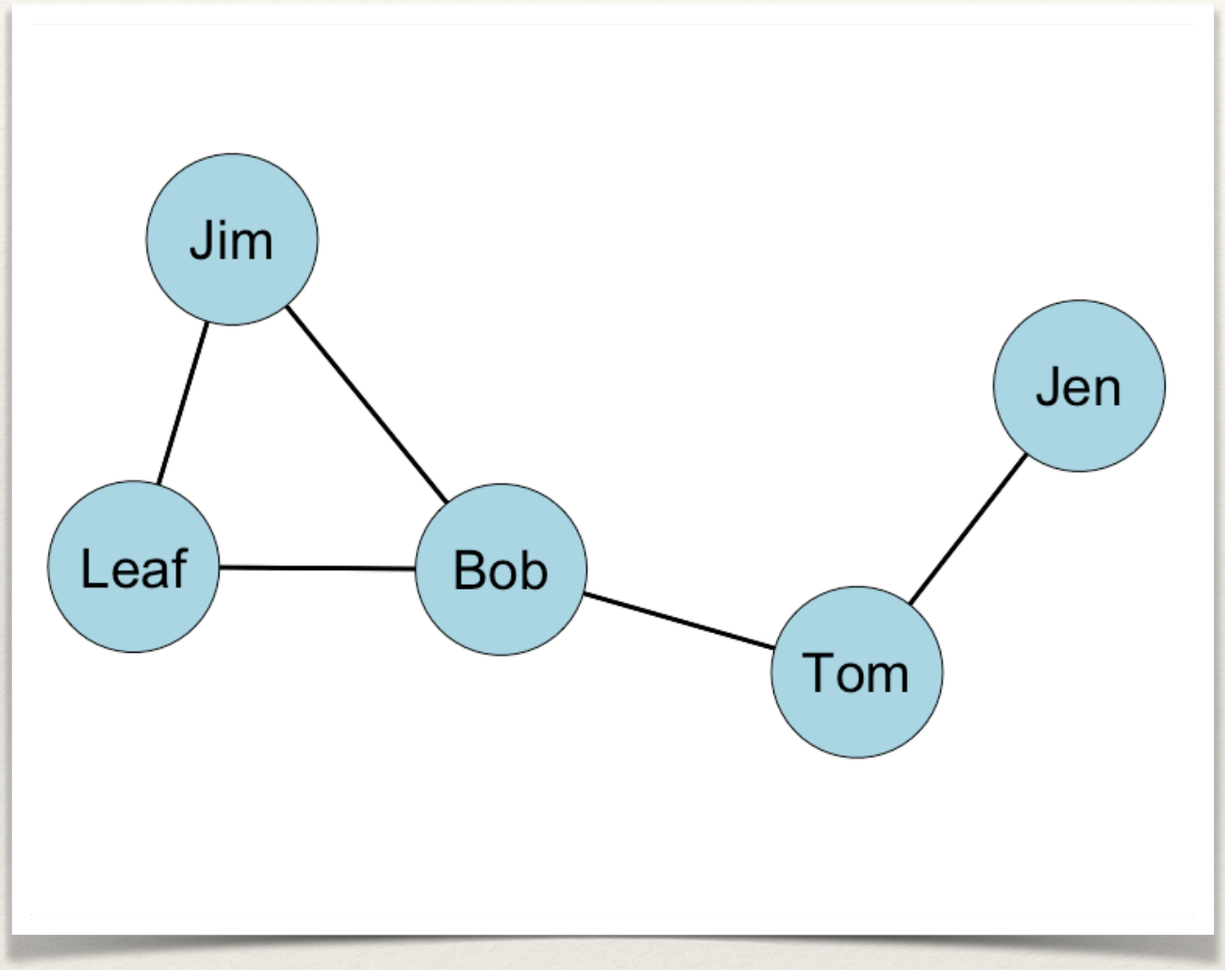
Bob = 1/5 = 0.200

Leaf = 1/7 = 0.143

Jim = 1/7 = 0.143

Multiplying by g-1 gives the standardized value

Example: Undirected, Binary Network



Standardized Closeness Centrality

Jen = 0.111*4 = 0.444

Tom = 0.167*4 = 0.668

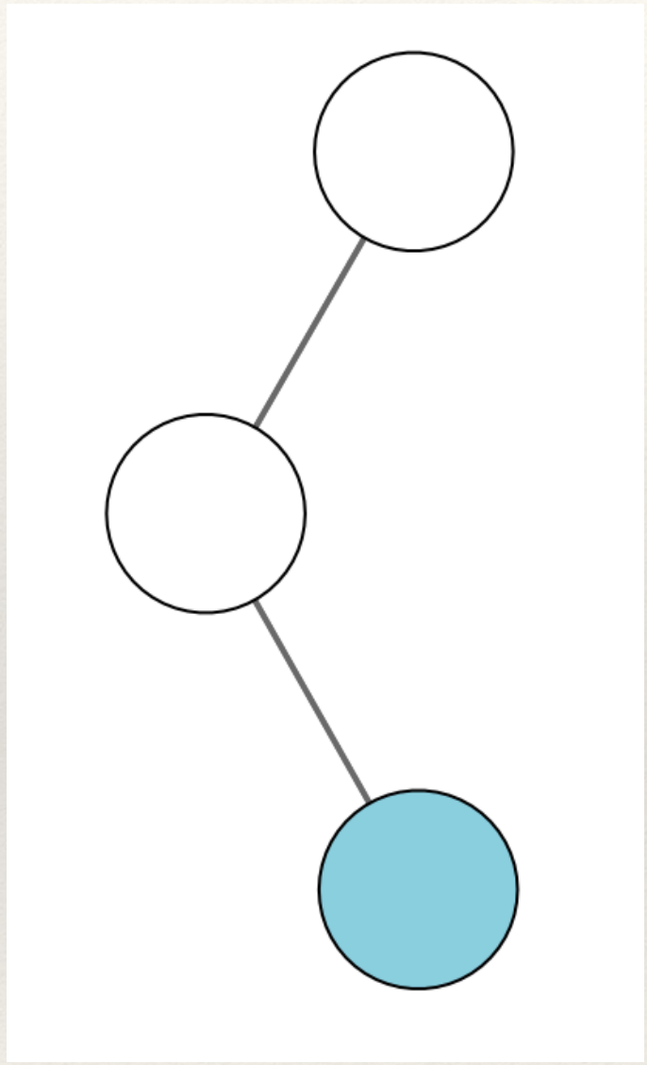
Bob = 0.200*4 = 0.800

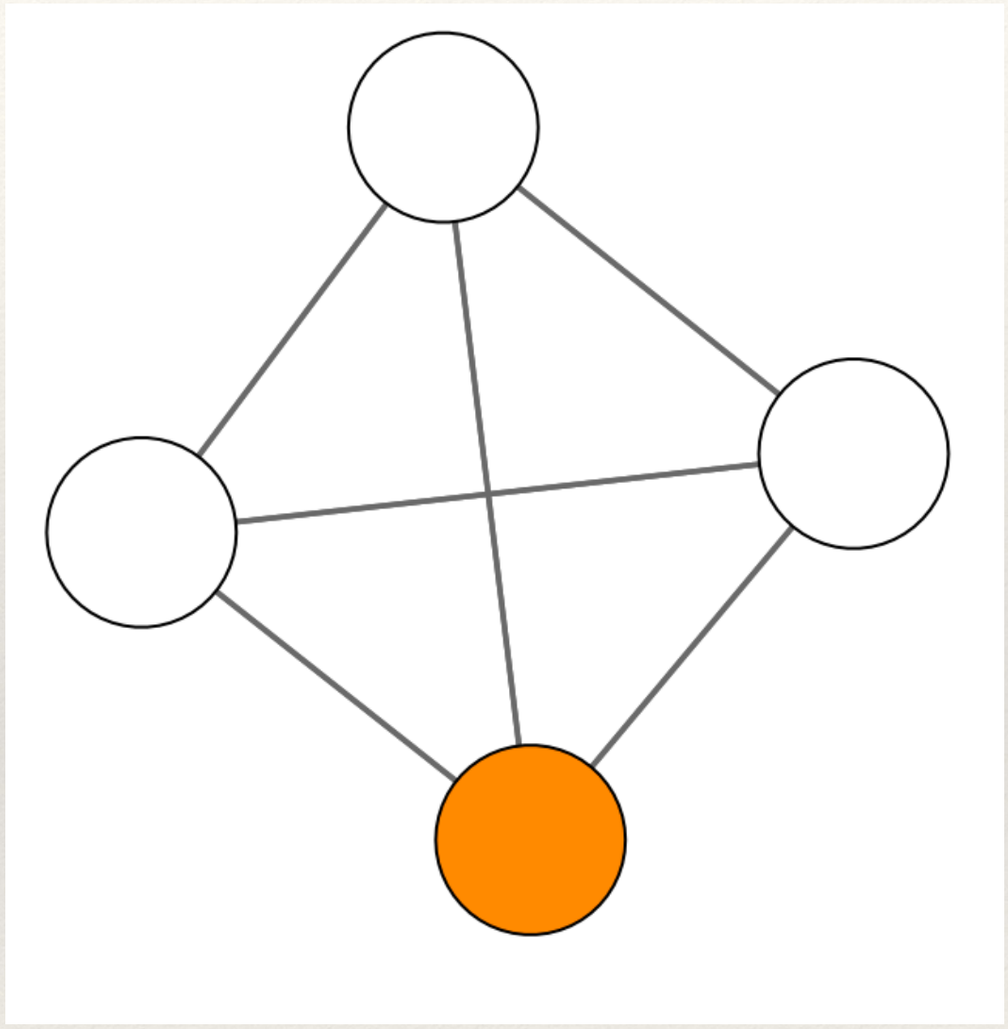
Leaf = 0.143*4 = 0.572

Jim = 0.143*4 = 0.572

Multiplying by g-1 gives the standardized value

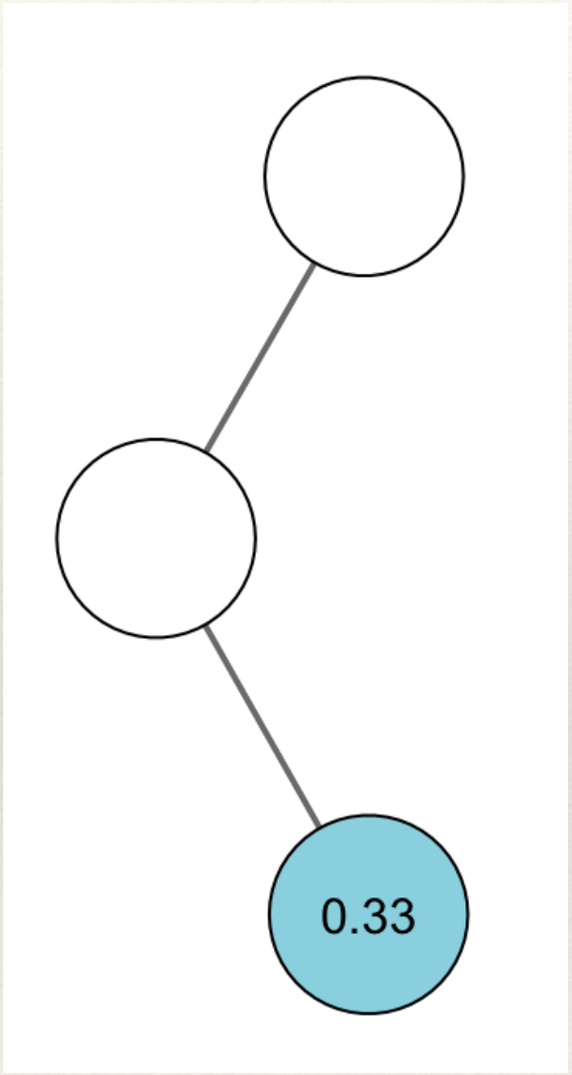
Why standardization matters



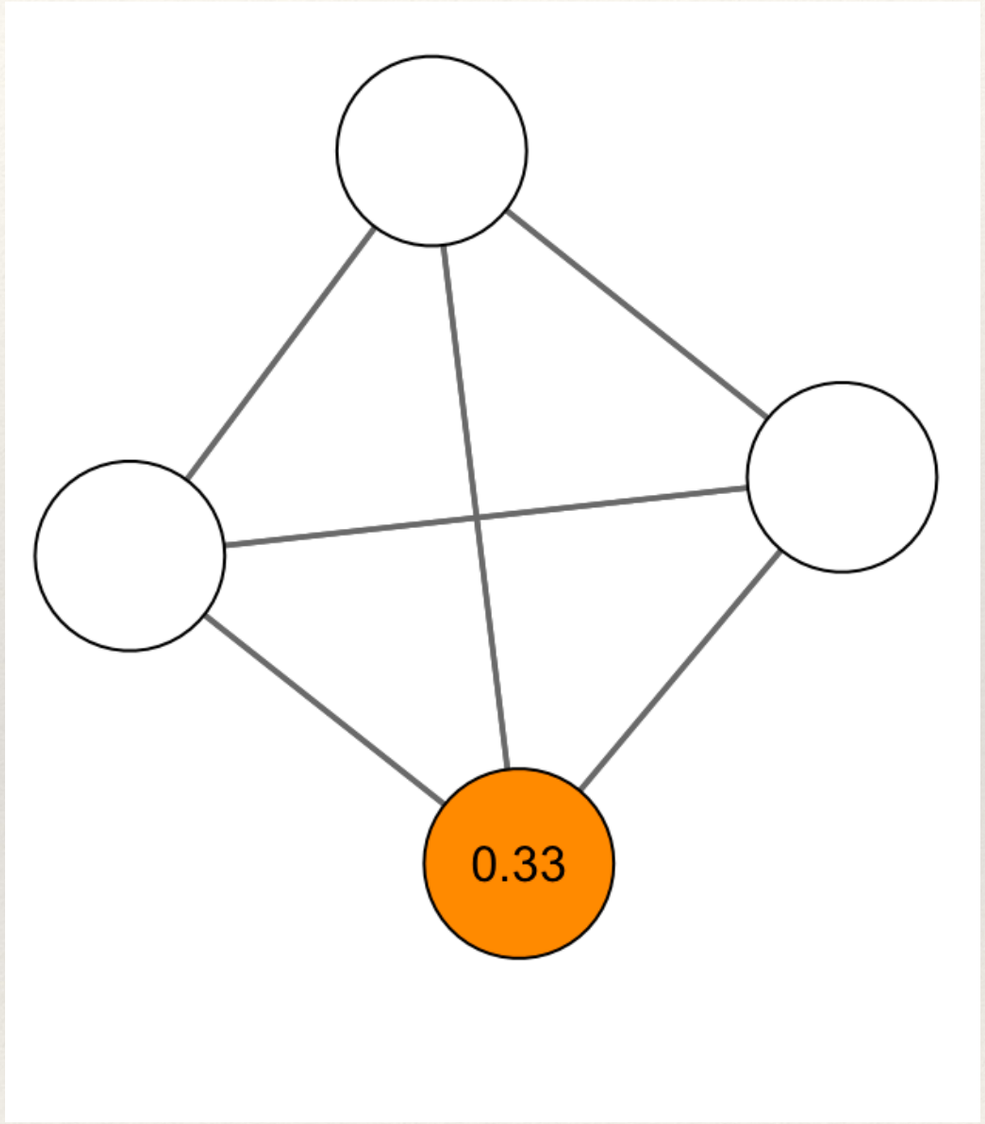


Which node has a higher closeness score?

Why standardization matters

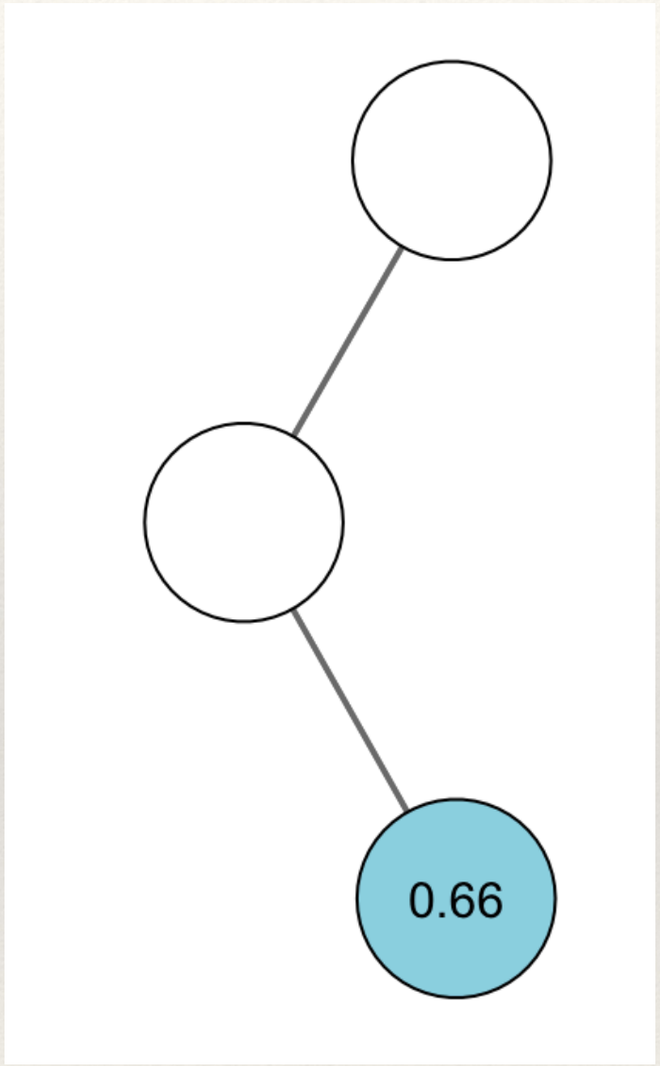


Network A: Raw

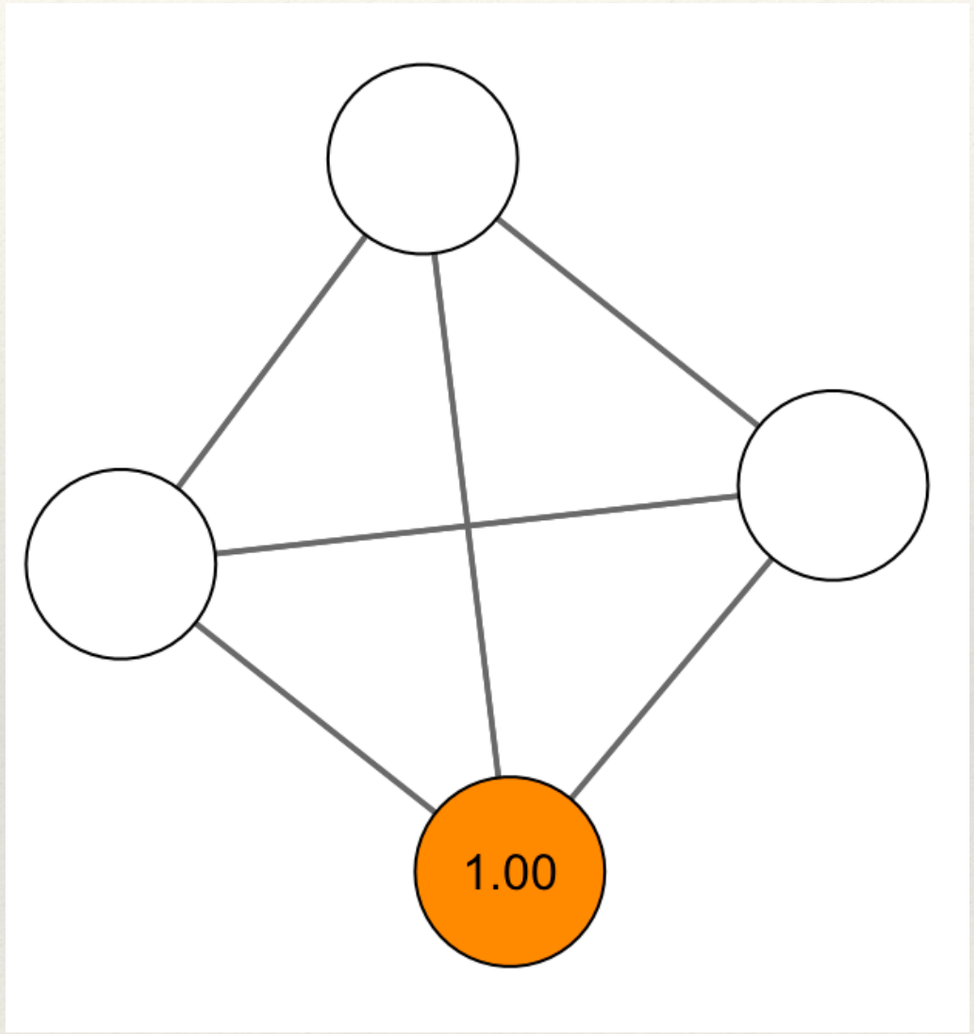


Network B: Raw

Why standardization matters



Network A: Standardized



Network B: Standardized

Centralization

- Recall that we can describe the dispersion in the centrality scores.
  - This is referred to as *centralization*.

?

Group Closeness Centralization: Undirected Binary Graphs

- Group *closeness* centralization describes the dispersion in the closeness scores between the nodes.
- The values are between 0 and 1 (as before).
  - Where scores closer to 0 indicate a decentralized structure and scores closer to 1 indicate a hierarchical structure.

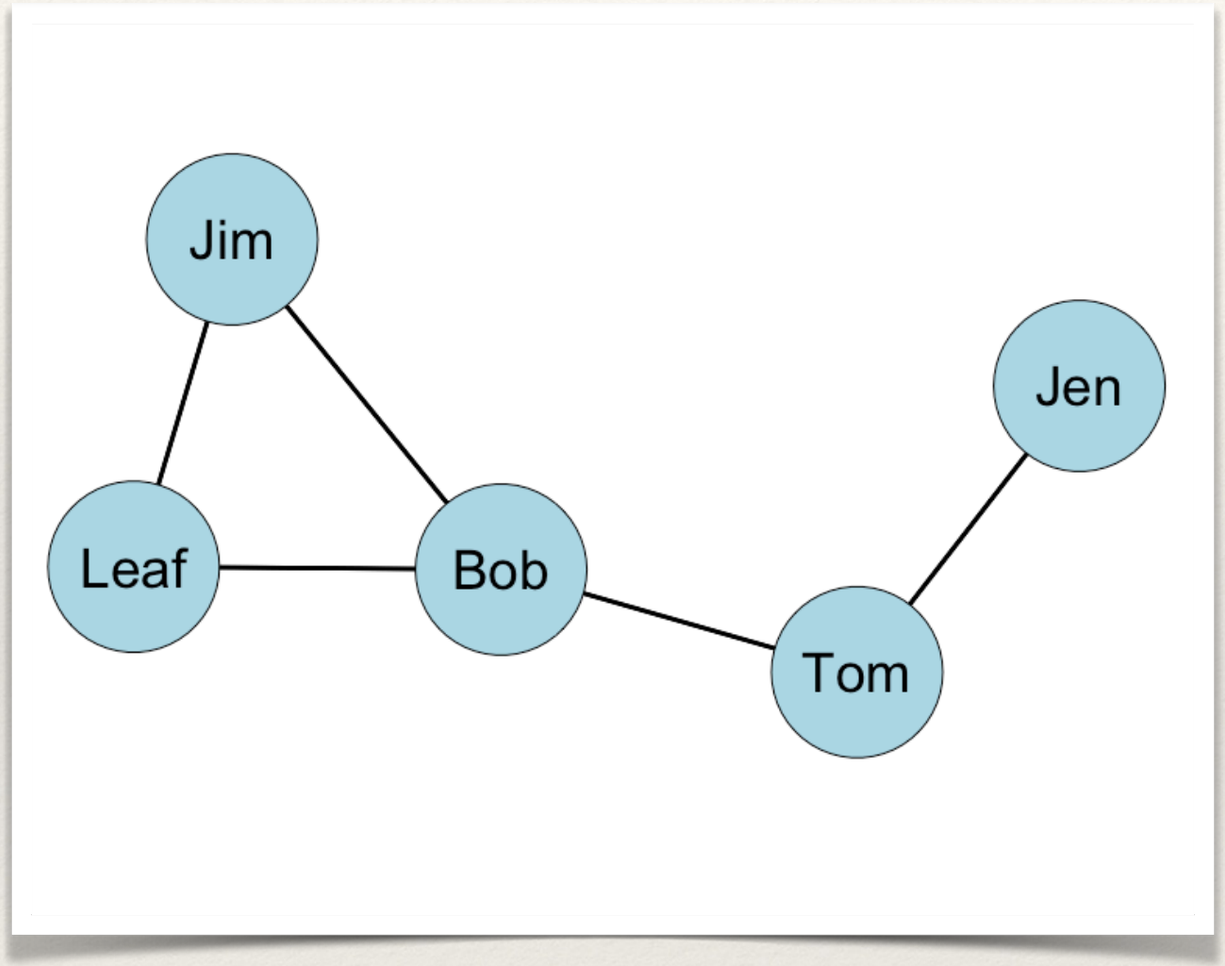
Group Closeness Centralization: Undirected Binary Graphs

Largest standardized closeness score

Standardized closeness score for actor i

$$C_C = \frac{\sum_{i=1}^{g} [C'_C(n^*) - C'_C(n_i)]}{[(g-2)(g-1)]/(2g-3)}$$

Example: Undirected, Binary Network



What is the group closeness centralization score for this graph?

<u>Standardized Closeness Centrality</u>

Jen = 0.111*4 = 0.444

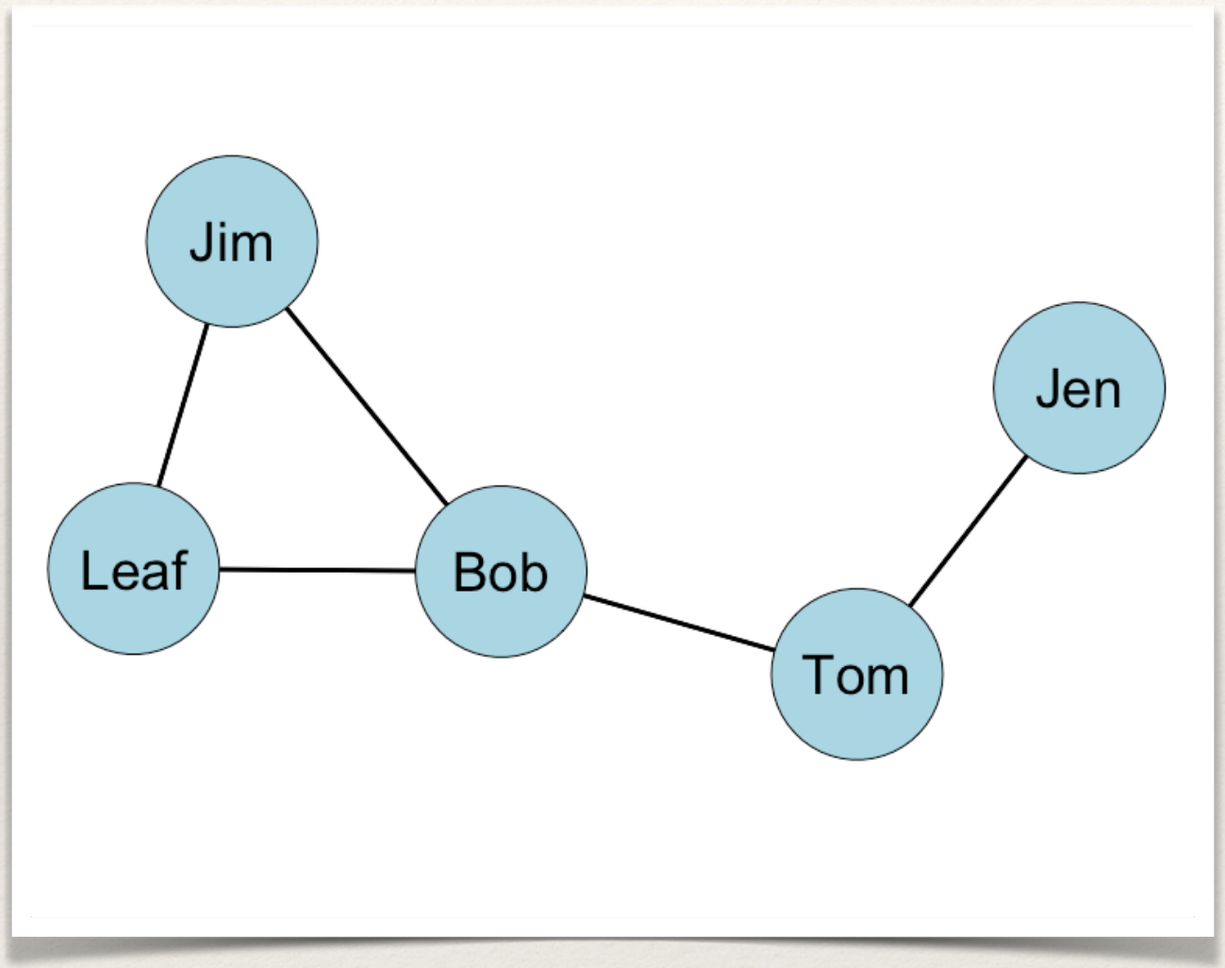
Tom = 0.167*4 = 0.668

Bob = 0.200*4 = 0.800

Leaf = 0.143*4 = 0.572

Jim = 0.143*4 = 0.572

Example: Undirected, Binary Network



Deviations of Standardized Closeness Centrality Scores

Jen = 0.800 - 0.444 = 0.356

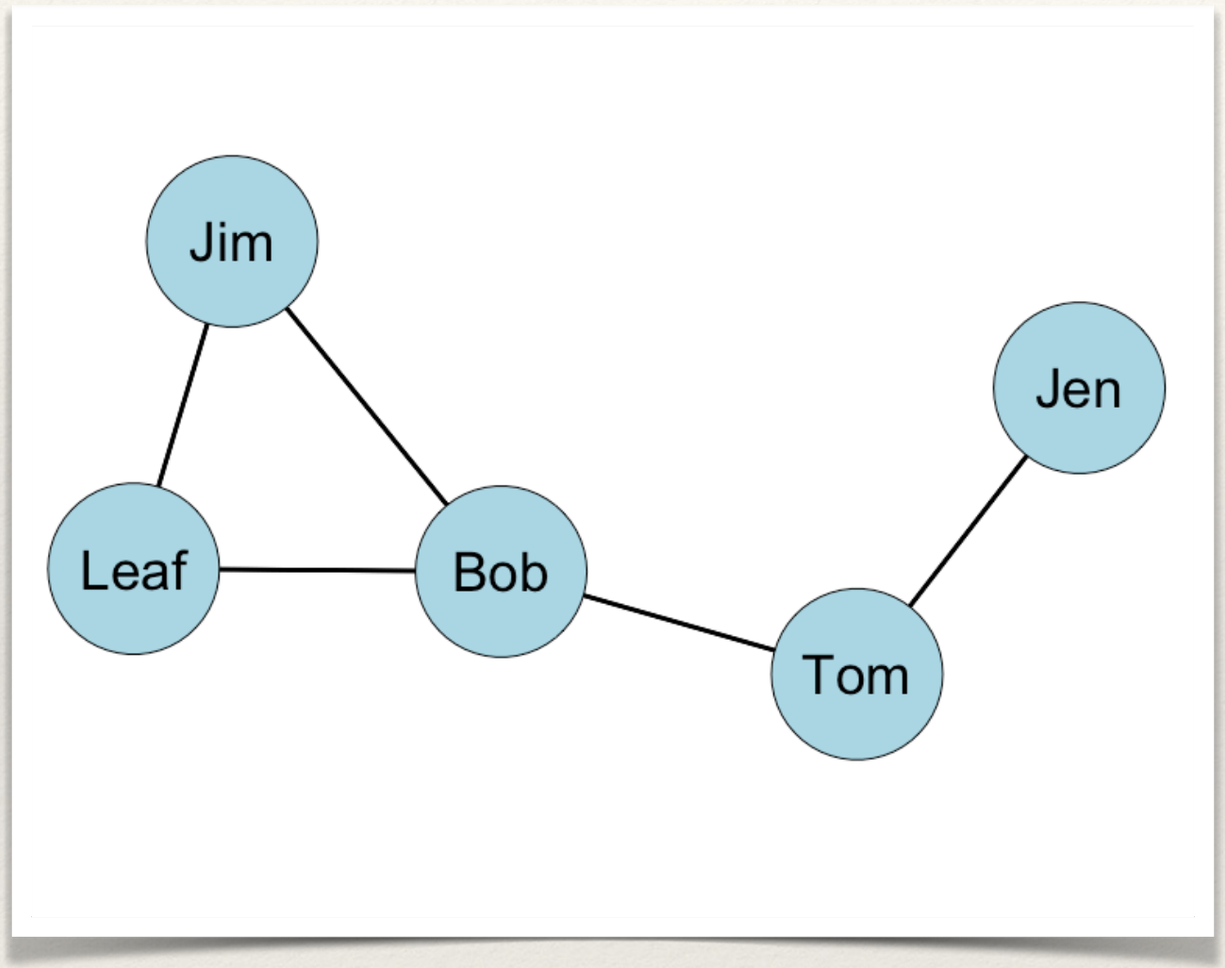
Tom = 0.800 - 0.668 = 0.132

Bob = 0.800 - 0.800 = 0.000

Leaf = 0.800 - 0.572 = 0.228

Jim = 0.800 - 0.572 = 0.228

Example: Undirected, Binary Network



Deviations of Standardized Closeness Centrality Scores

Jen = 0.800 - 0.444 = 0.356

Tom = 0.800 - 0.668 = 0.132

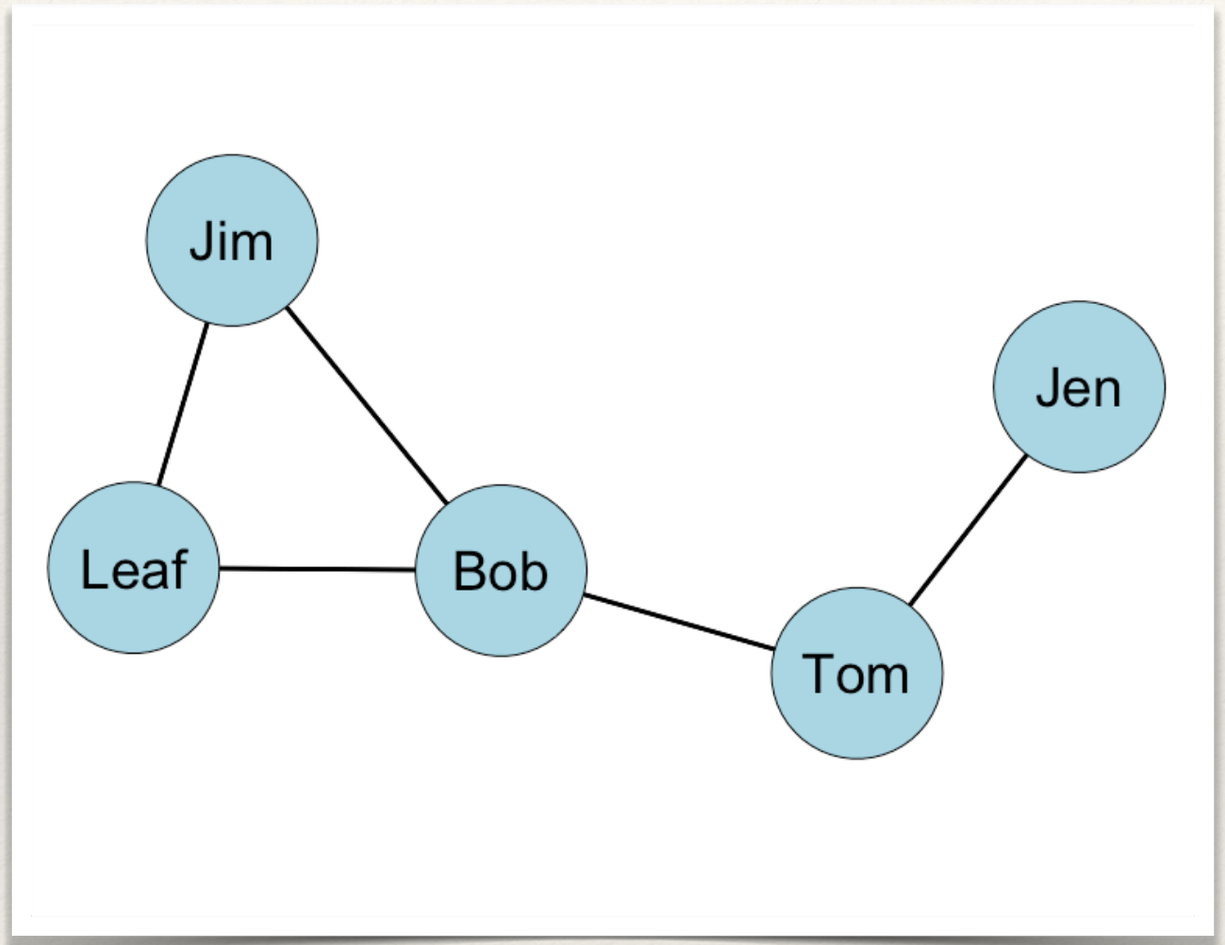
Bob = 0.800 - 0.800 = 0.000

Leaf = 0.800 - 0.572 = 0.228

Jim = 0.800 - 0.572 = 0.228

Sum = 0.944

Example: Undirected, Binary Network



$$C_C = \frac{\sum_{i=1}^{g} [C'_C(n^*) - C'_C(n_i)]}{[(g-2)(g-1)]/(2g-3)}$$

Deviations of Standardized Closeness Centrality Scores

Jen = 0.800 - 0.444 = 0.356

Tom = 0.800 - 0.668 = 0.132

Bob = 0.800 - 0.800 = 0.000

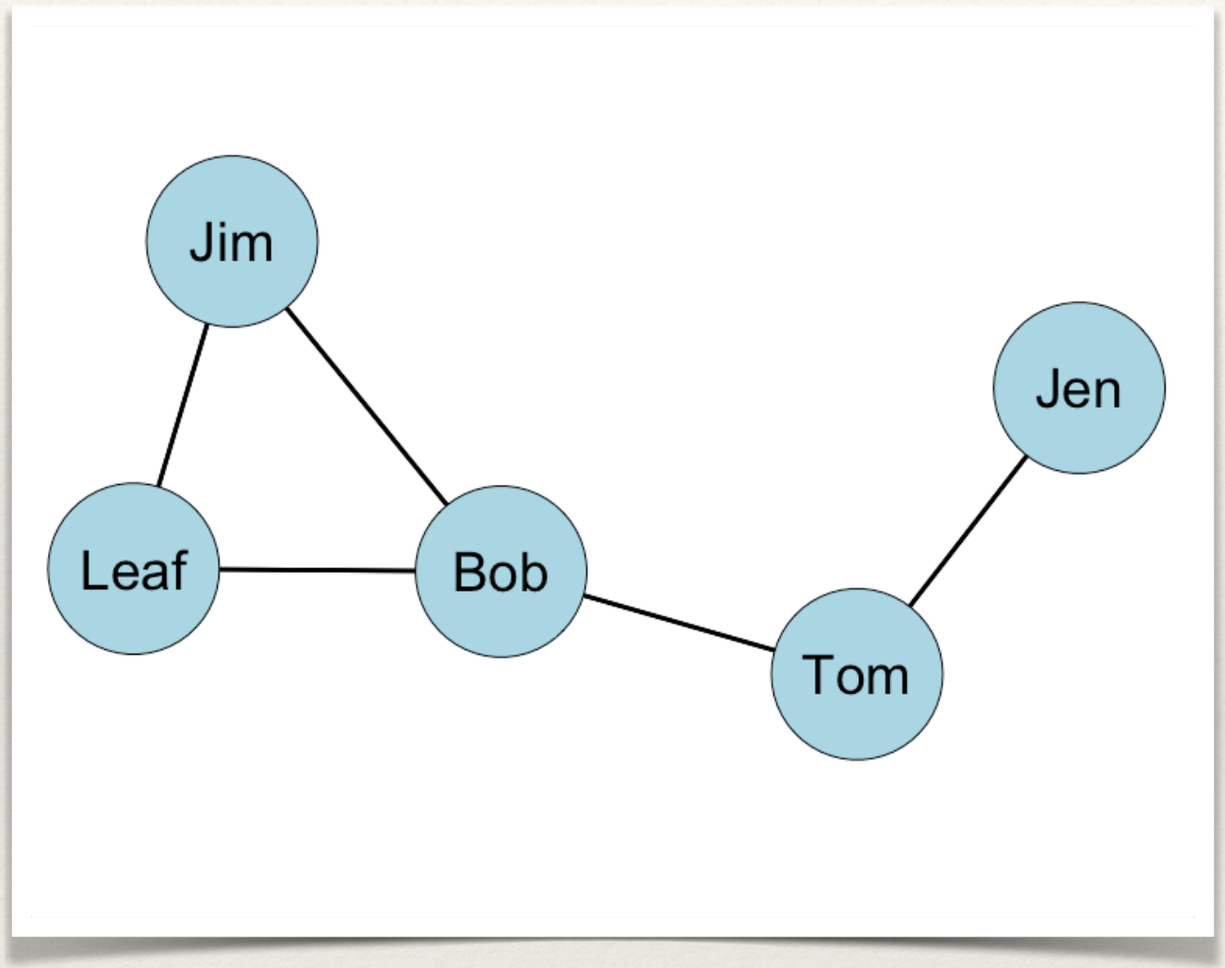
Leaf = 0.800 - 0.572 = 0.228

Jim = 0.800 - 0.572 = 0.228

Sum = 0.944

Denominator = [(5-2)(5-1)/(2*5-3)] = 1.714

Example: Undirected, Binary Network



Approximately 0.551

Deviations of Standardized Closeness Centrality Scores

Jen = 0.800 - 0.444 = 0.356

Tom = 0.800 - 0.668 = 0.132

Bob = 0.800 - 0.800 = 0.000

Leaf = 0.800 - 0.572 = 0.228

Jim = 0.800 - 0.572 = 0.228

Sum = 0.944

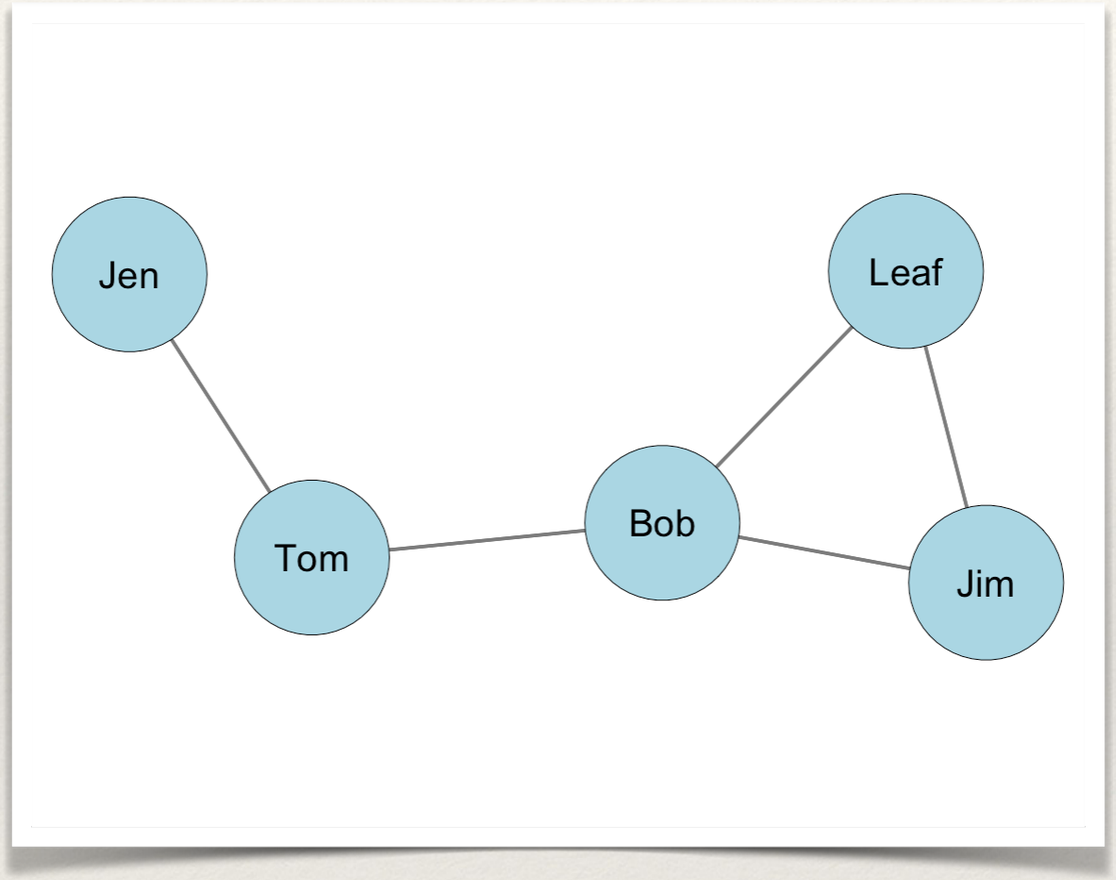
Denominator = [(5-2)(5-1)/(2*5-3)] = 1.714

Example: Undirected, Binary Network

$$C_C = \frac{\sum_{i=1}^{g} [C'_C(n^*) - C'_C(n_i)]}{[(g-2)(g-1)]/(2g-3)}$$

$$= \frac{0.356 + 0.132 + 0.000 + 0.228 + 0.228}{[(5-2)(5-1)]/(2*5-3)} = \frac{0.944}{1.714} = 0.551$$

Example: Undirected, Binary Network



Compare the centralization scores:

Degree = 0.416
Closeness = 0.551

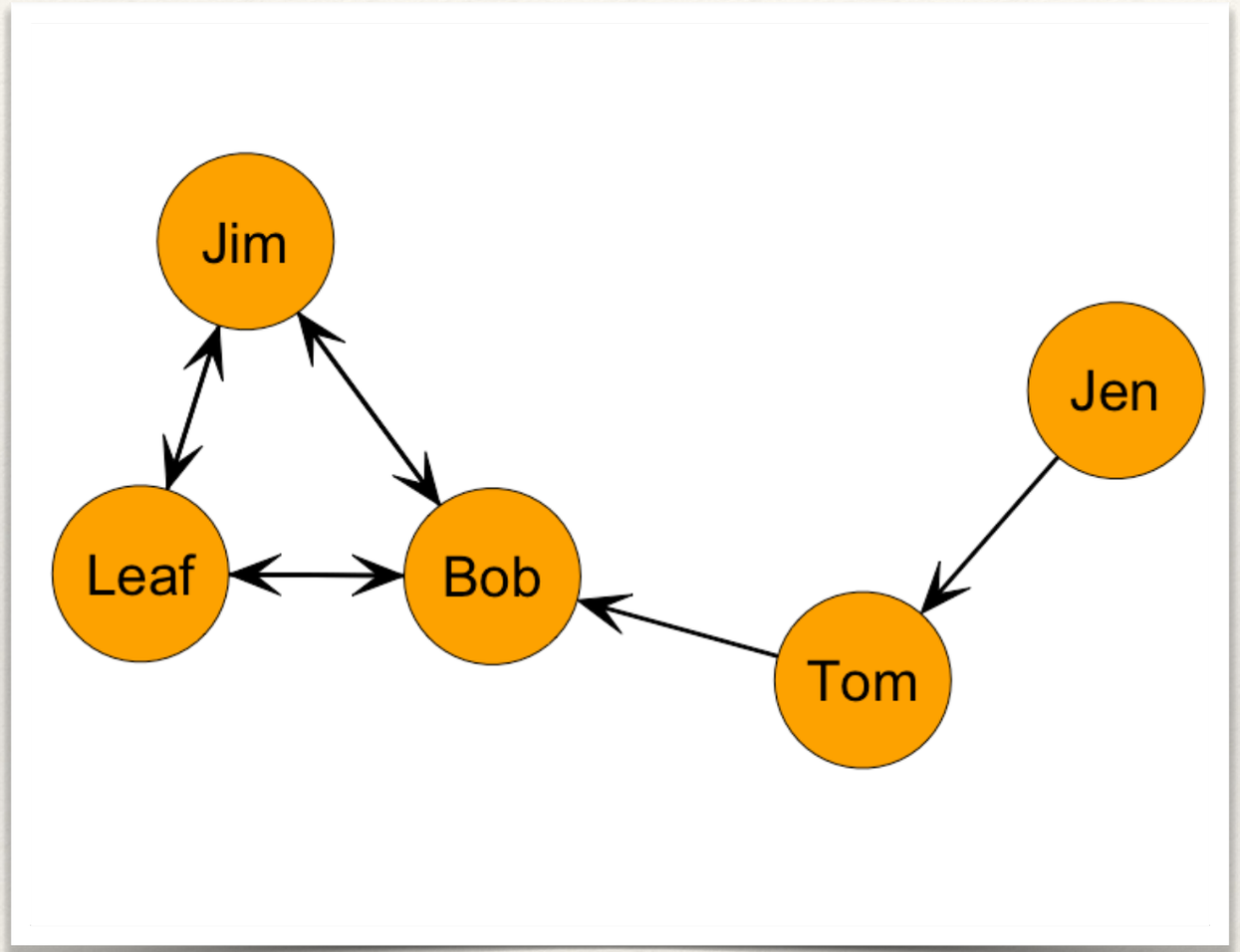
What can we say about the differences in the centralization scores for each type of centrality?

Directed Networks

Closeness Centrality: Directed Binary Graphs

- Recall that in a directed network, the directionality matters.
- As a result, we have to consider how this might influence our measures.

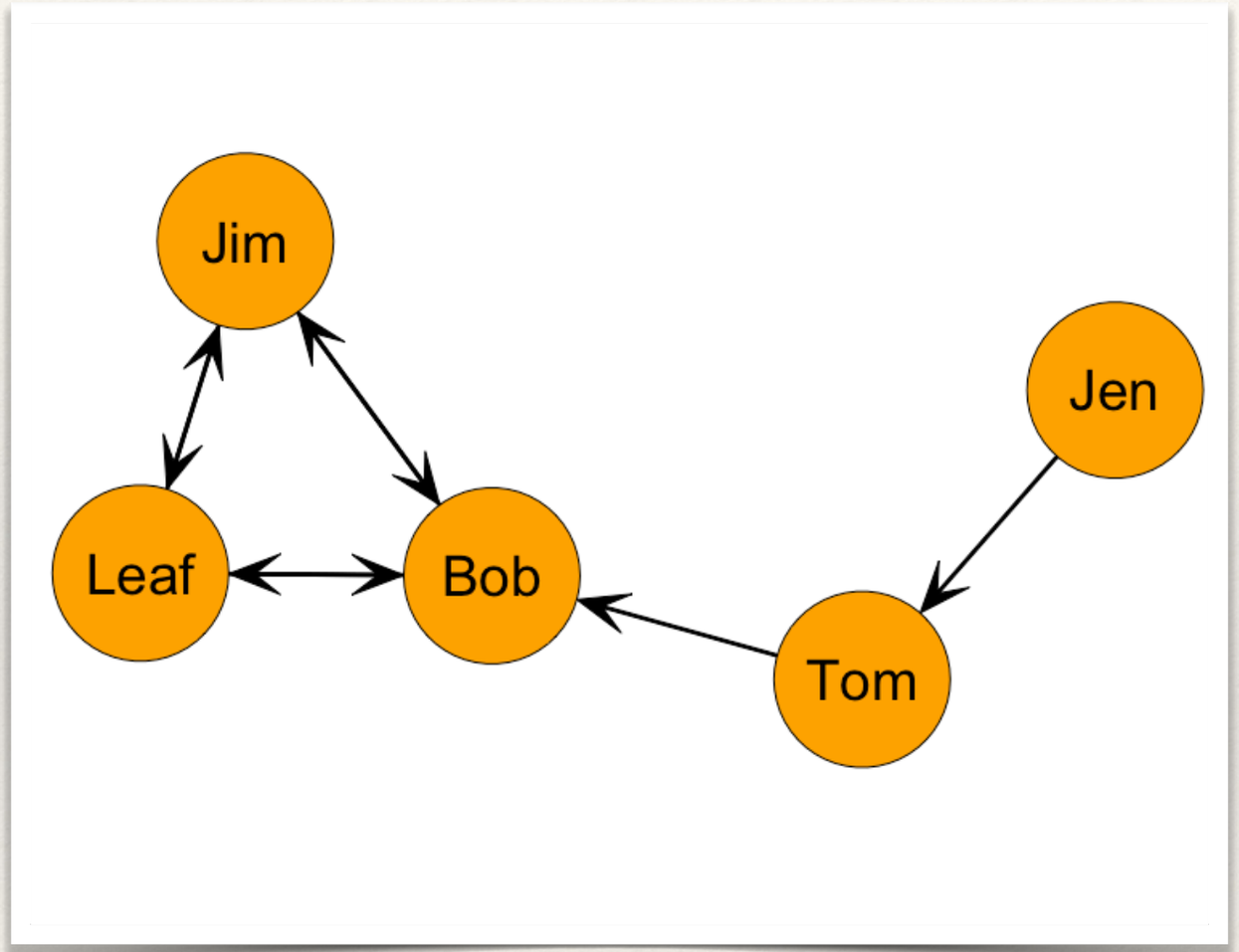
Example: Closeness Centrality for Directed Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen					
Tom					
Bob					
Leaf					
Jim					

*How far is **Jen** from Tom? From Bob?*

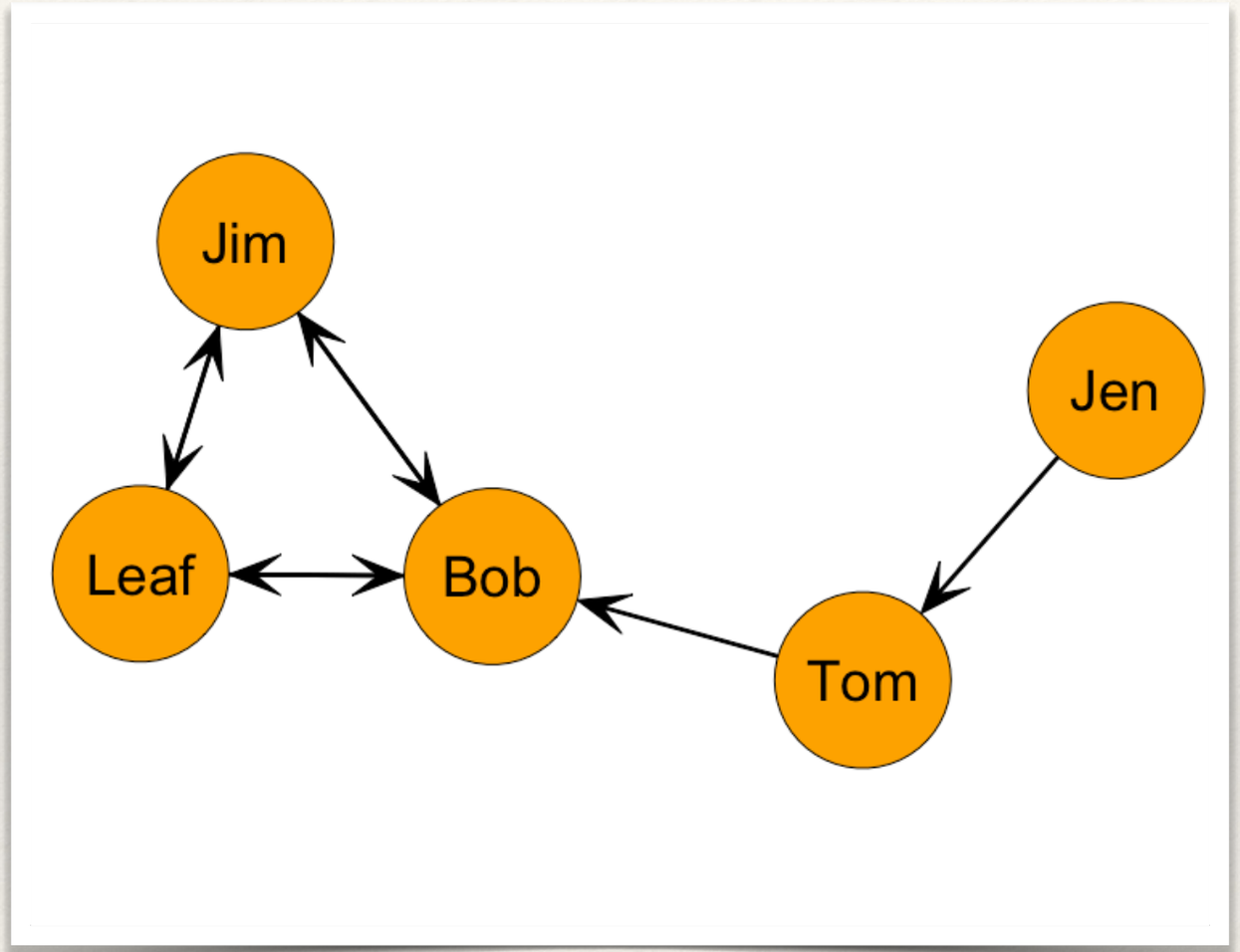
Example: Closeness Centrality for Directed Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen					
Tom					
Bob					
Leaf					
Jim					

*How far is **Jen** from Tom? From Bob? Depends on direction.*

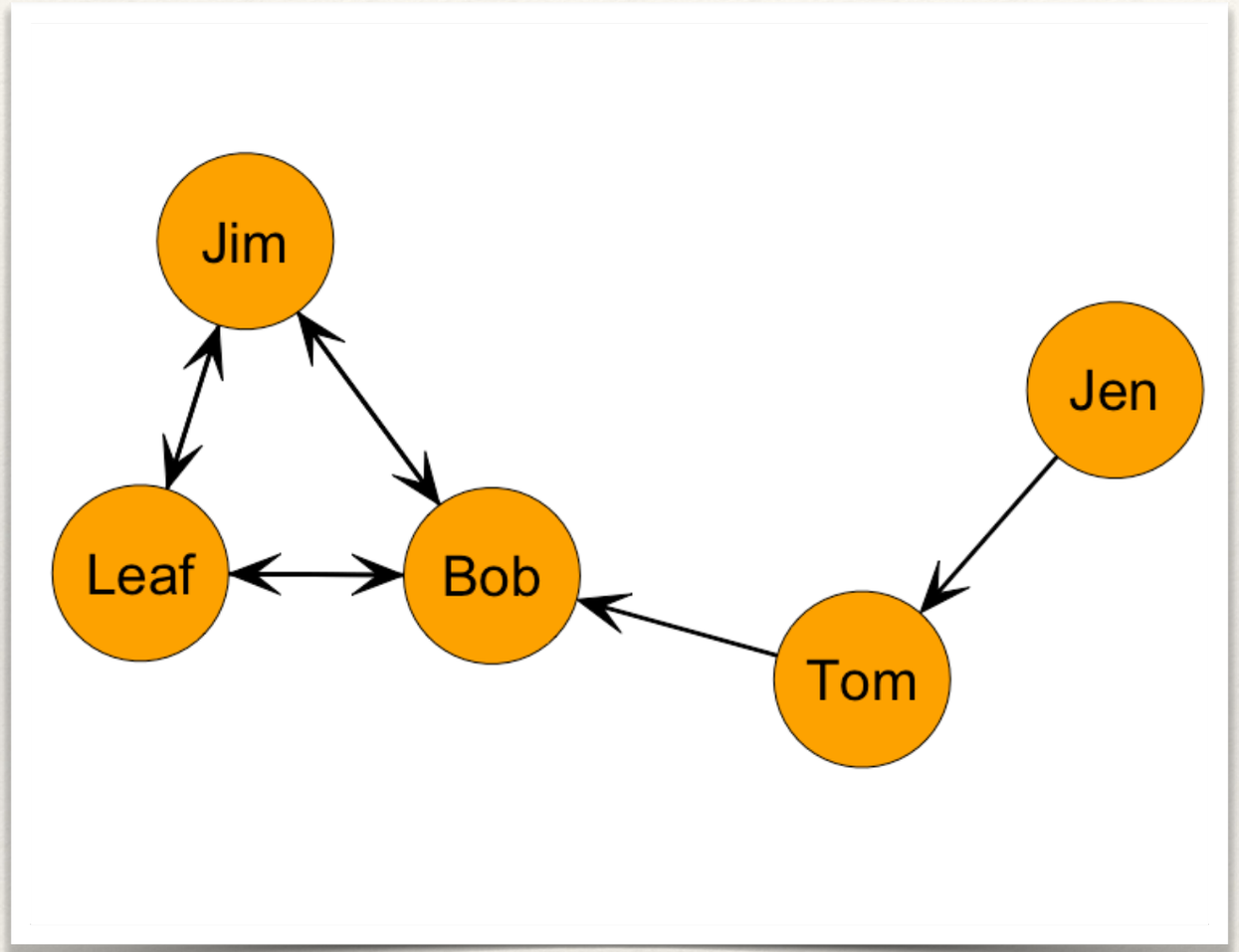
Example: Closeness Centrality for Directed Binary Network



Closeness centrality in directed graphs usually focuses on the send network (i.e. outgoing ties).

Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen					
Tom					
Bob					
Leaf					
Jim					

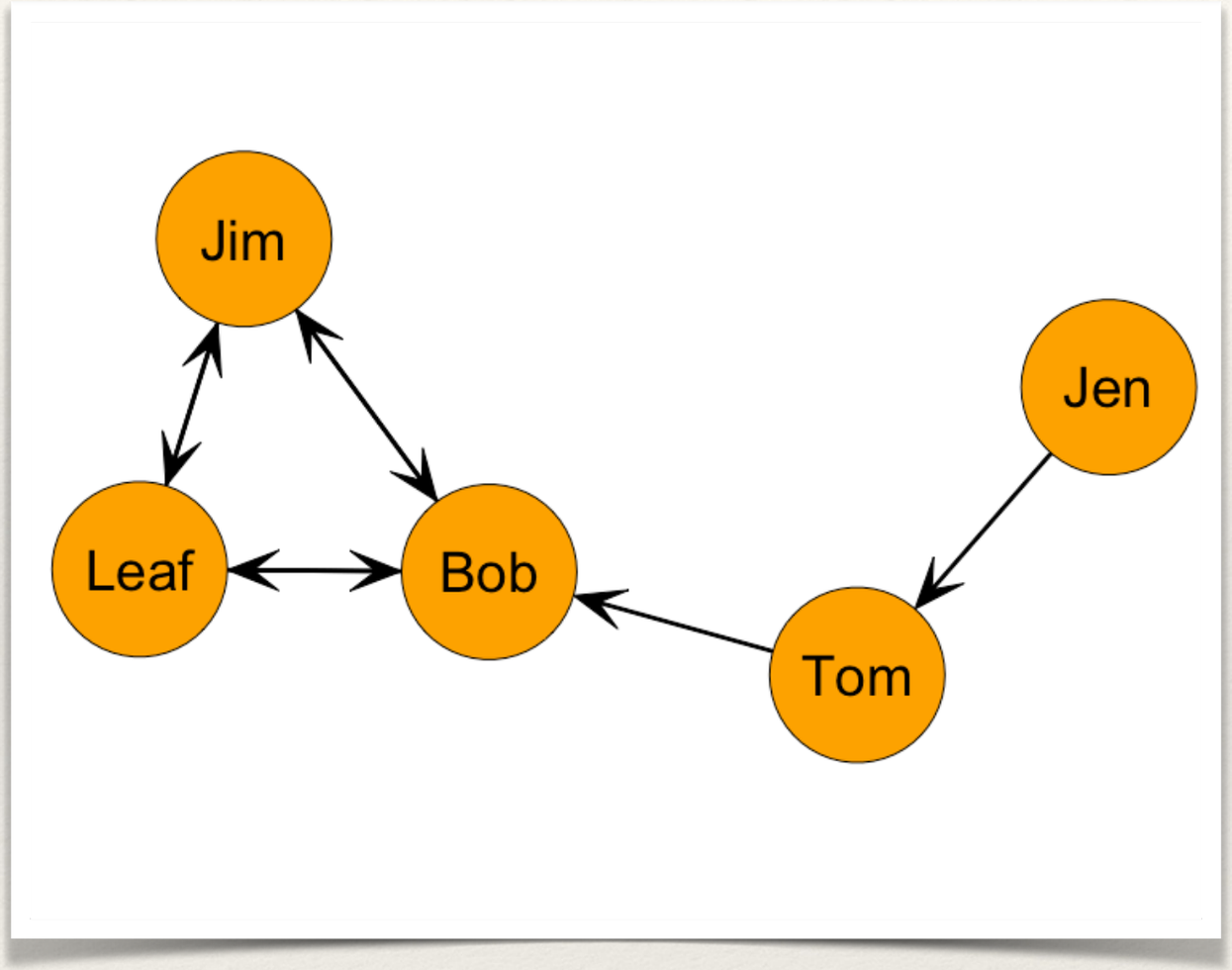
Example: Closeness Centrality for Directed Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen					
Tom					
Bob					
Leaf					
Jim					

Remember, ***row*** *sums are the* ***outdegree,*** *so row distance tells you how close <u>you</u> are to others.*

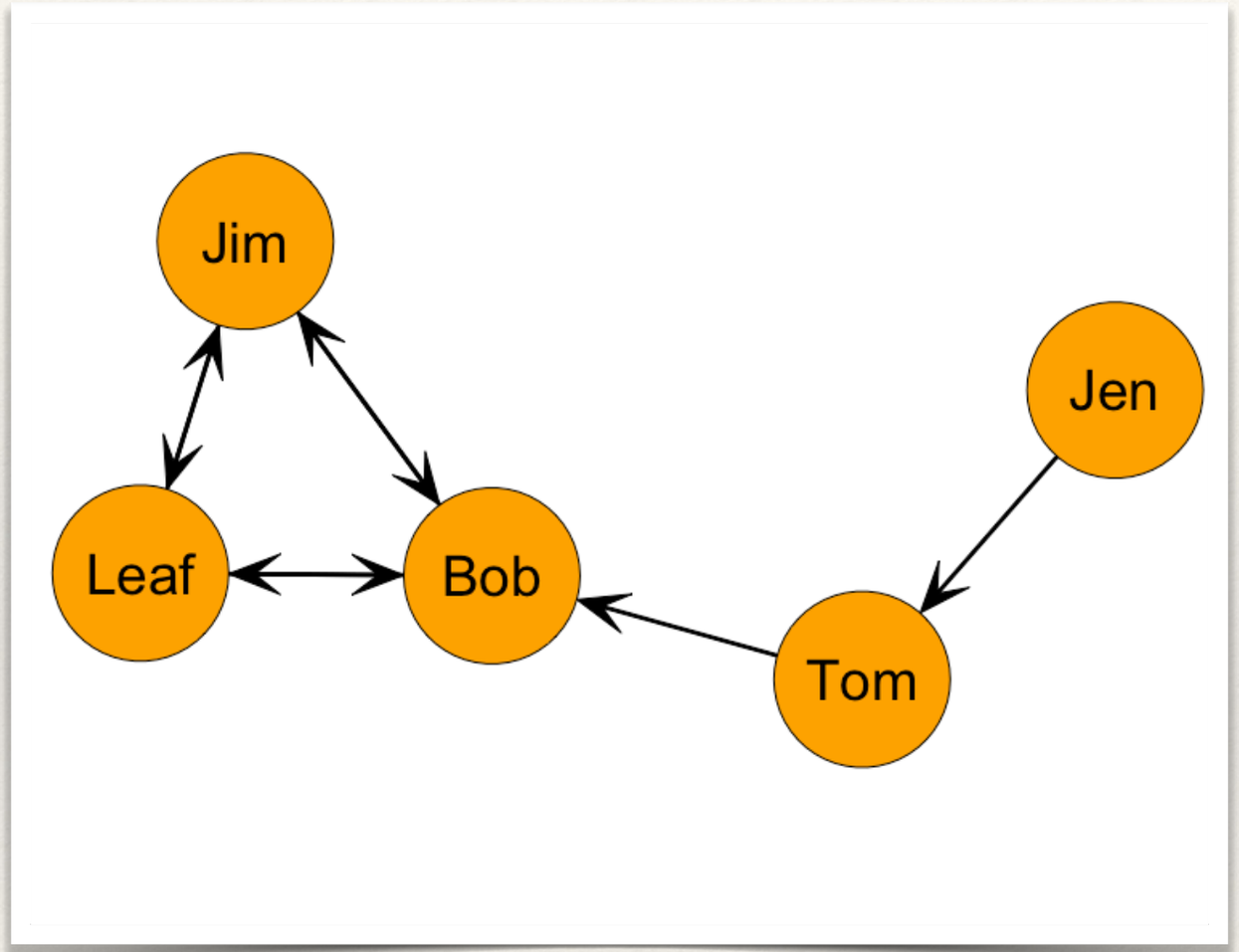
Example: Closeness Centrality for Directed Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen		?	?		
Tom					
Bob					
Leaf					
Jim					

*How far is **Jen** from Tom? From Bob?*

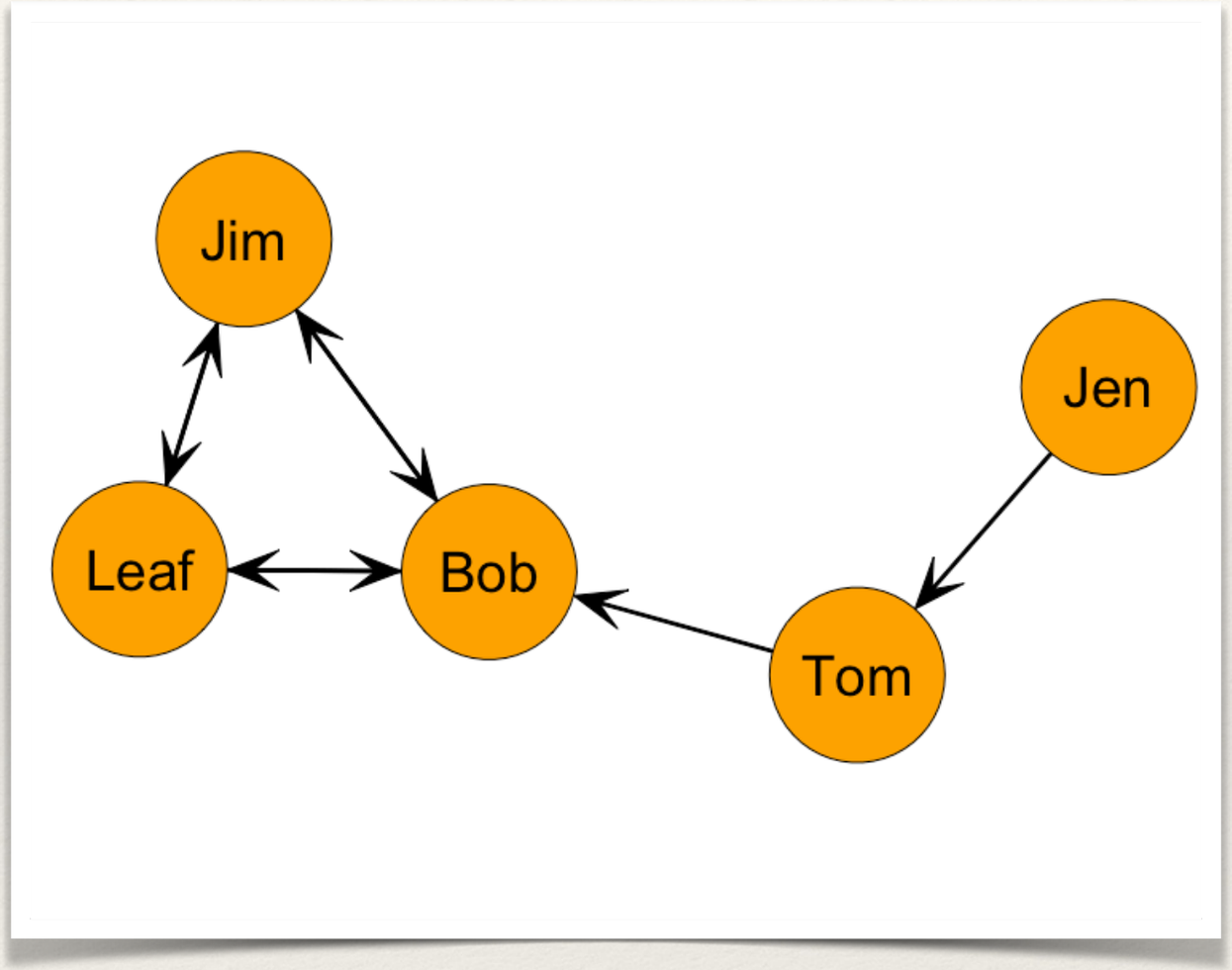
Example: Closeness Centrality for Directed Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen		1	2		
Tom					
Bob					
Leaf					
Jim					

*How far is **Jen** from Tom? From Bob?*

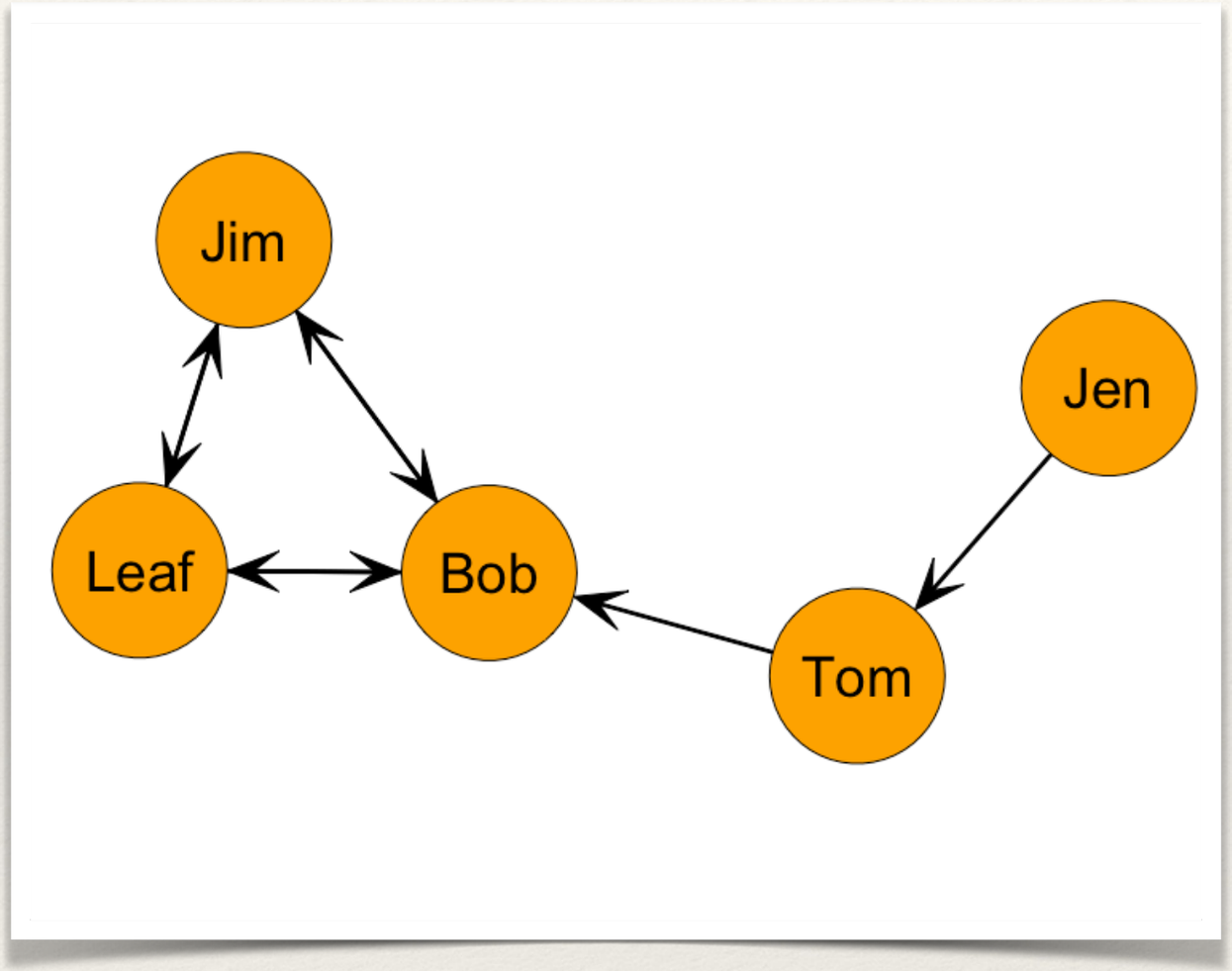
Example: Closeness Centrality for Directed Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen		1	2	?	?
Tom					
Bob					
Leaf					
Jim					

What about the rest?

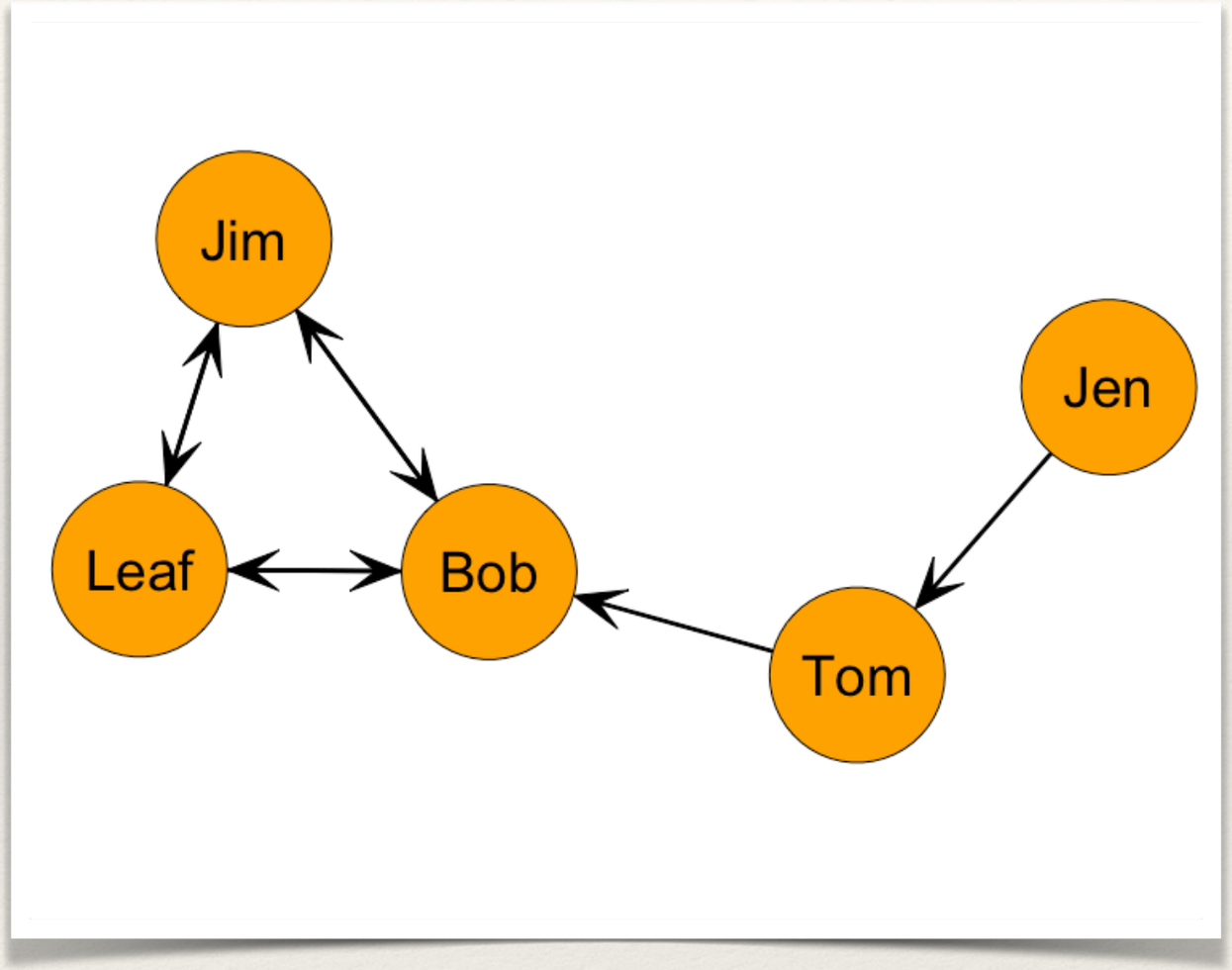
Example: Closeness Centrality for Directed Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen		1	2	3	3
Tom					
Bob					
Leaf					
Jim					

*What is **Jen's** closeness in the graph?*

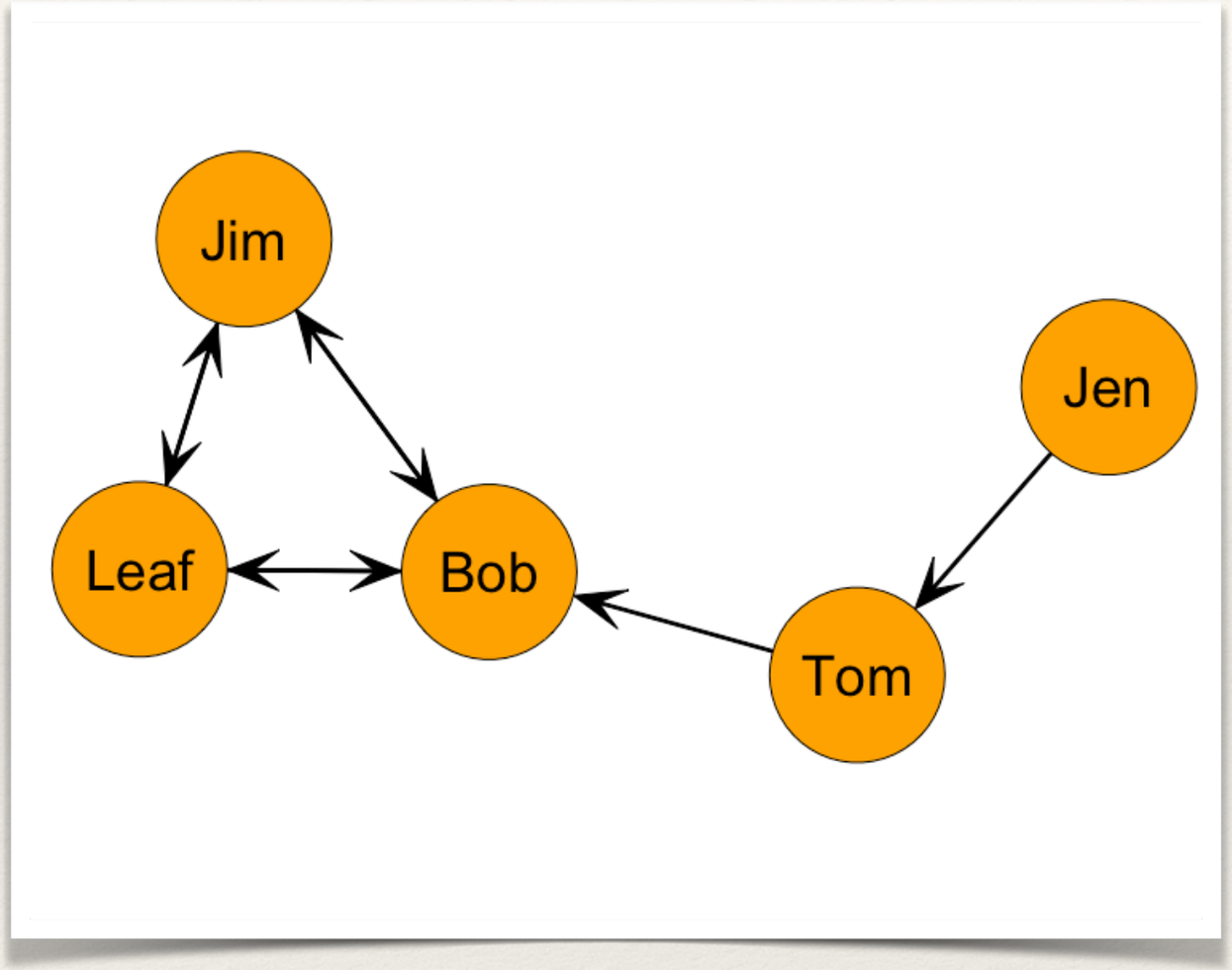
Example: Closeness Centrality for Directed Binary Network



Distance Matrix						
	Jen	Tom	Bob	Leaf	Jim	Sum
Jen		1	2	3	3	9
Tom						
Bob						
Leaf						
Jim						

*What is **Jen's** closeness in the graph?*

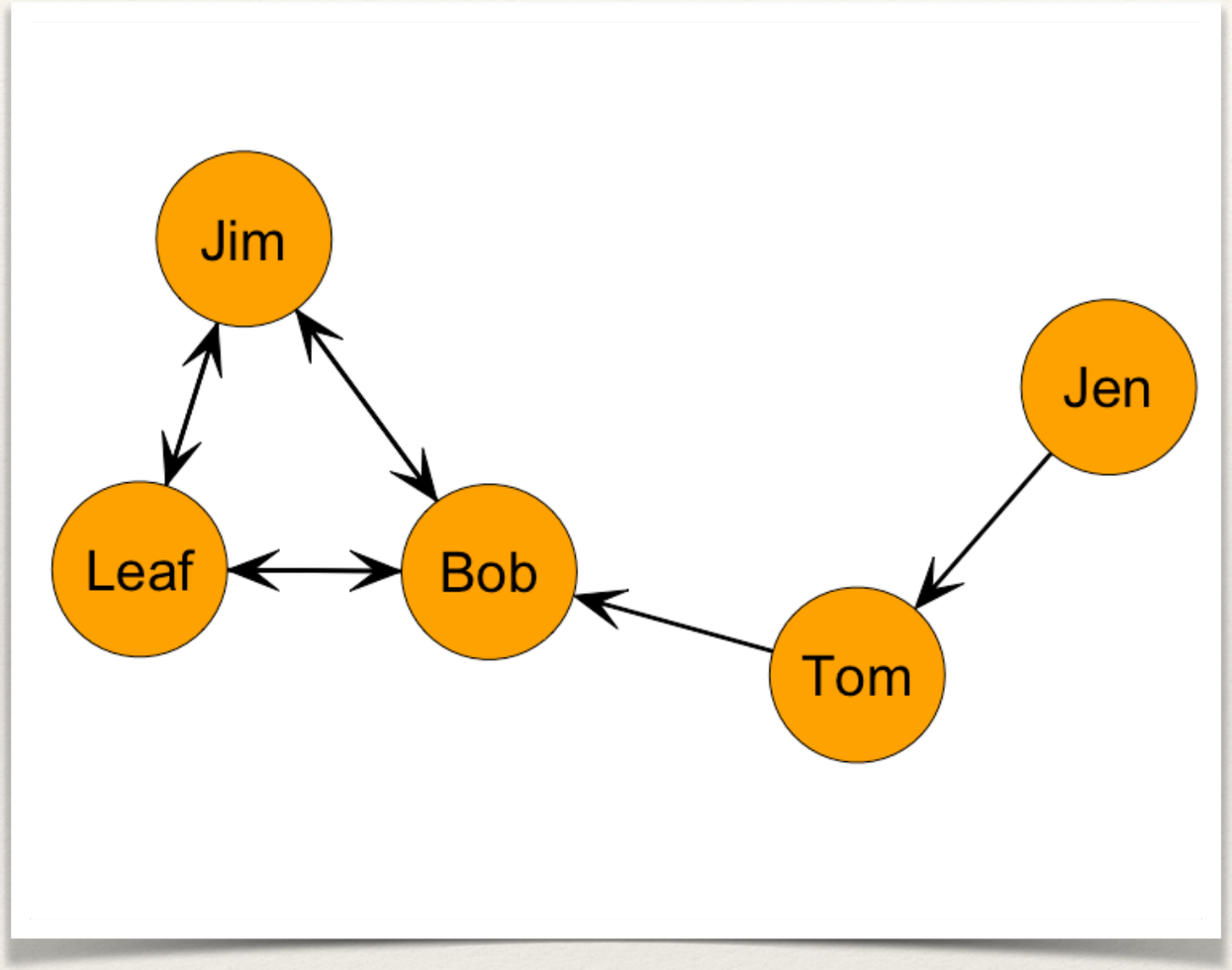
Example: Closeness Centrality for Directed Binary Network



Distance Matrix						
	Jen	Tom	Bob	Leaf	Jim	Sum
Jen		1	2	3	3	9
Tom						
Bob						
Leaf						
Jim						

Just take the inverse, which is 1/9 = 0.111.

Example: Closeness Centrality for Directed Binary Network

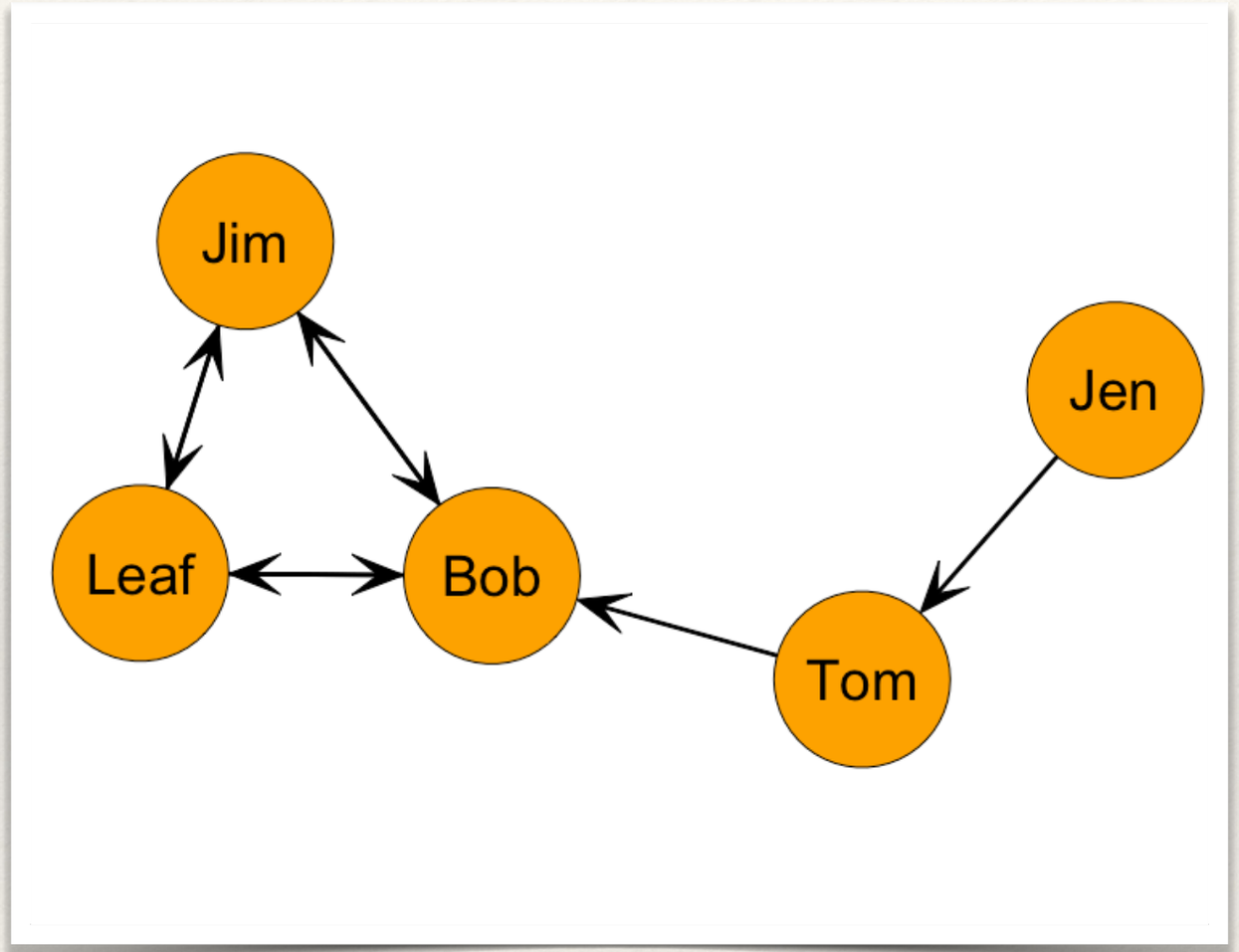


Distance Matrix						
	Jen	Tom	Bob	Leaf	Jim	Sum
Jen		1	2	3	3	9
Tom						
Bob						
Leaf						
Jim						

*Then, **Jen's** standardized score is*

(1/9)(g-1) = 0.111*4 = 0.444*

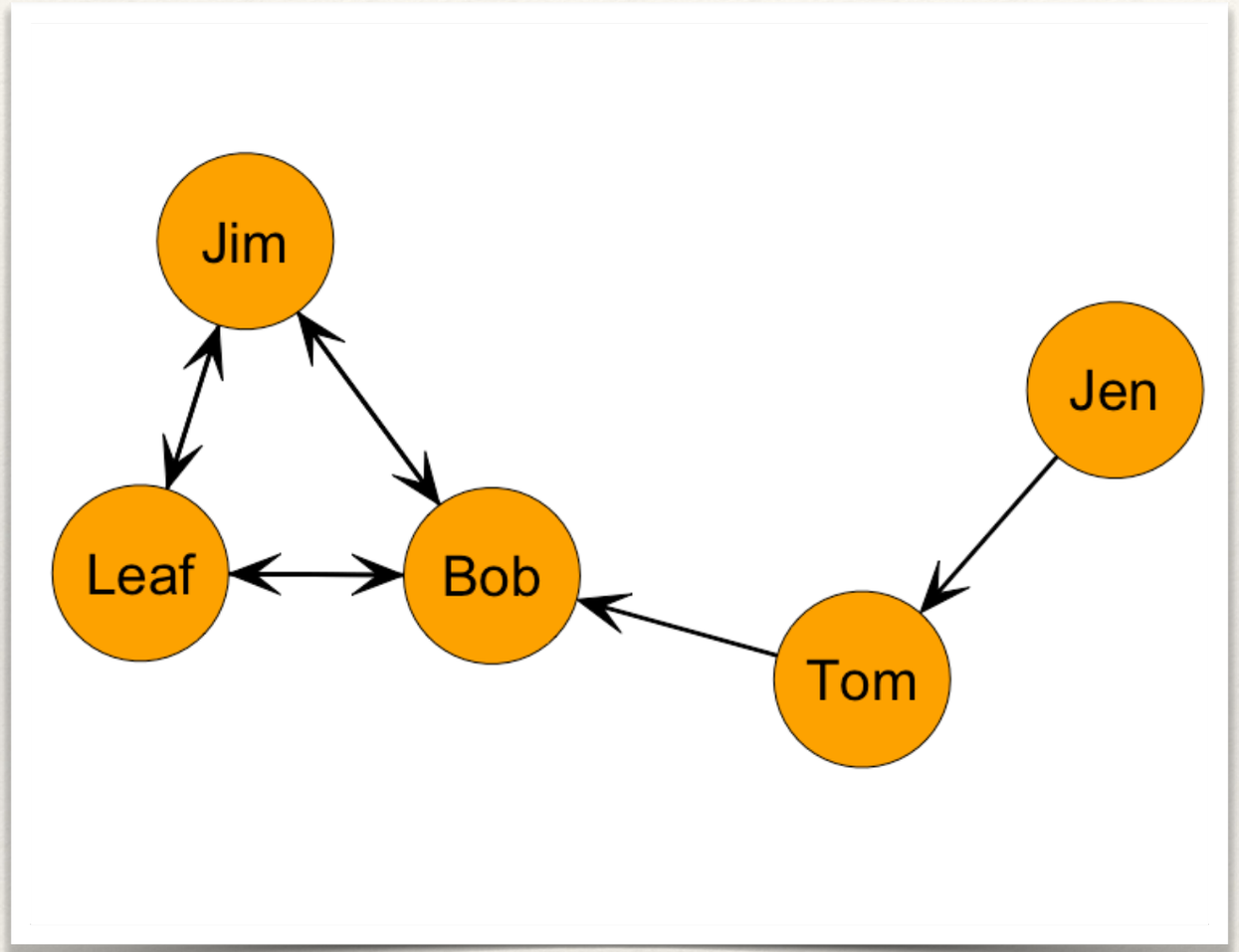
Example: Closeness Centrality for Directed Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen					
Tom					
Bob	?	?		?	?
Leaf					
Jim					

*What about **Bob**? How far is **Bob** from everyone?*

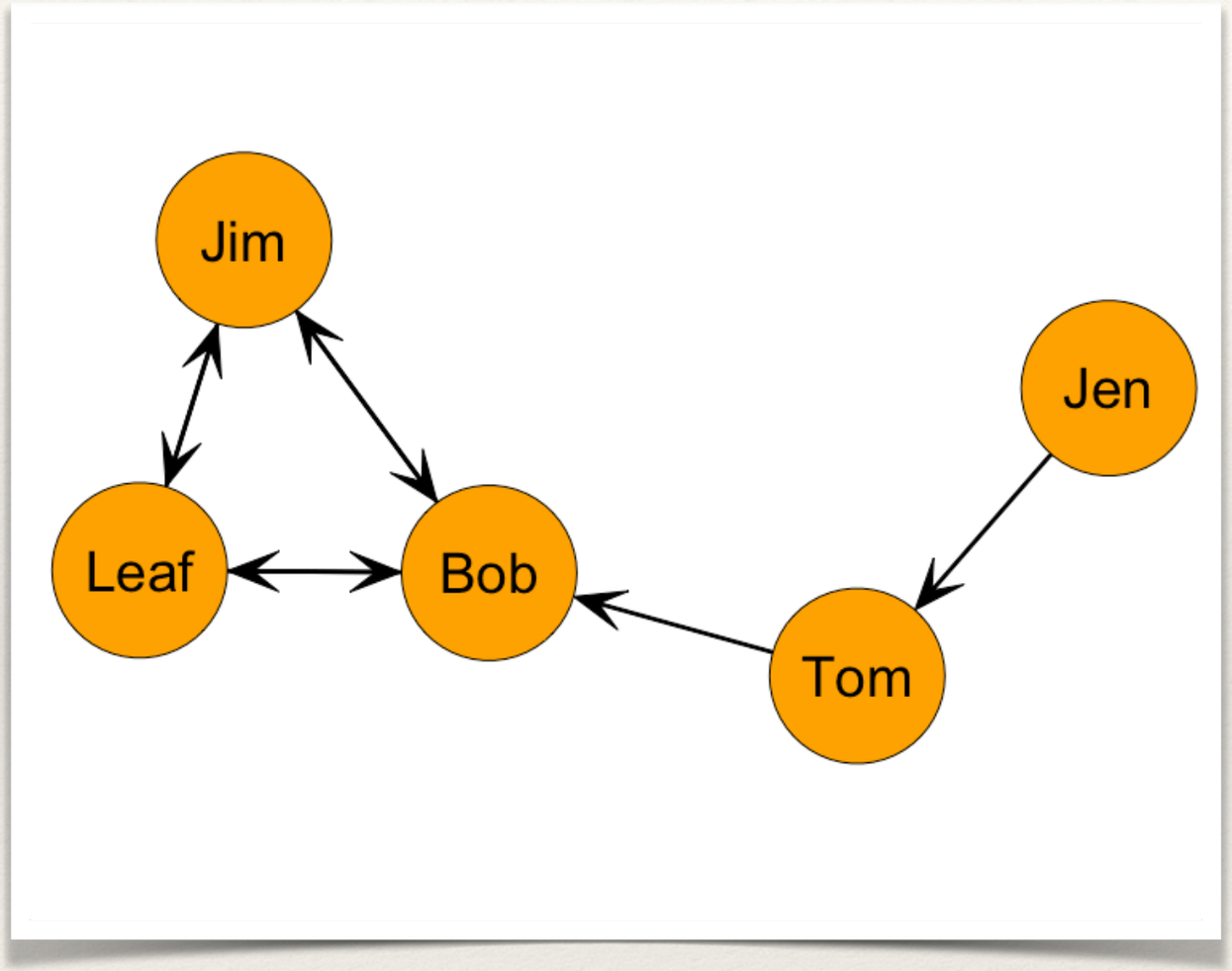
Example: Closeness Centrality for Directed Binary Network



Distance Matrix					
	Jen	Tom	Bob	Leaf	Jim
Jen					
Tom					
Bob	Inf	Inf		?	?
Leaf					
Jim					

*Why is **Bob's** distance to **Tom** infinite?*

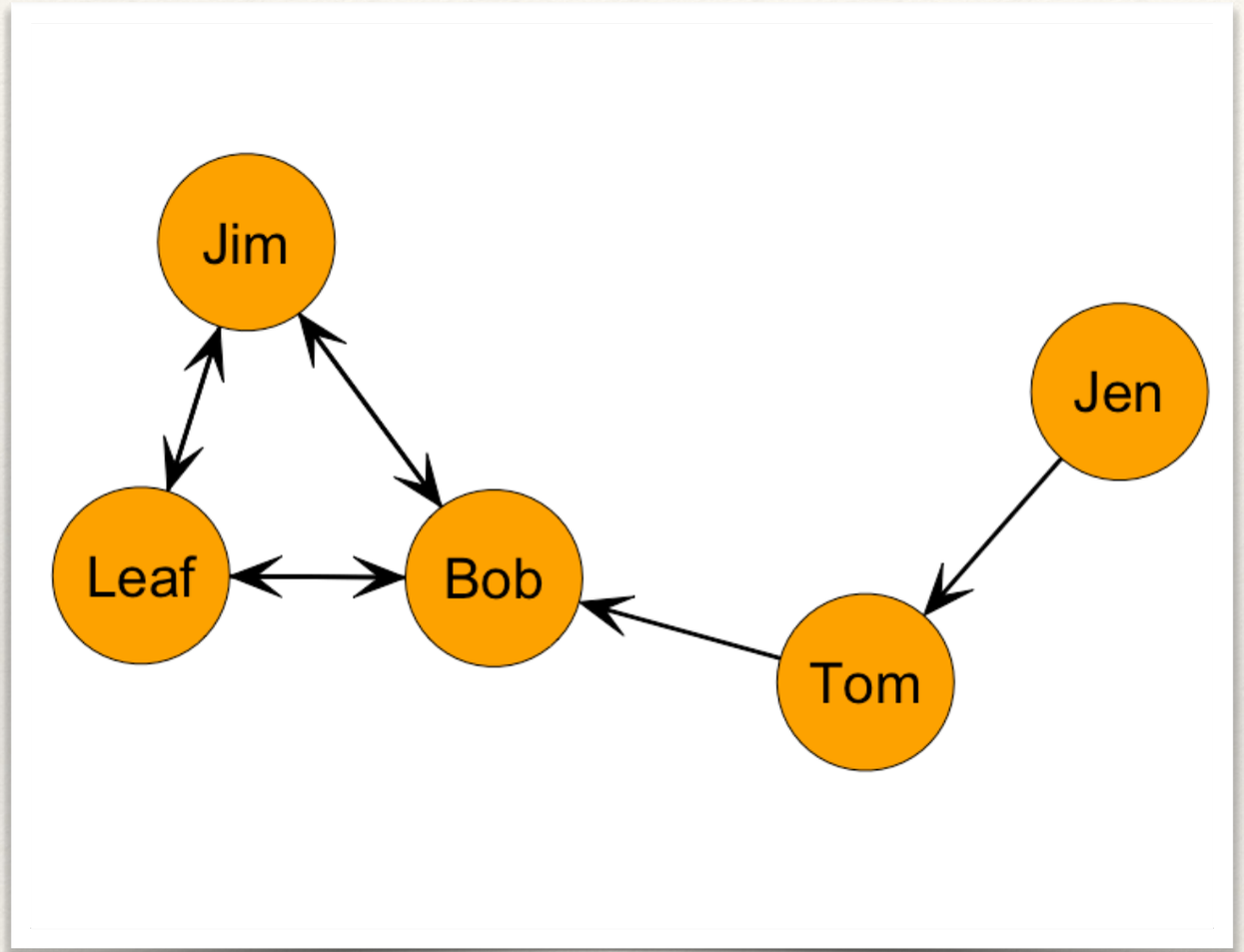
Example: Closeness Centrality for Directed Binary Network



Distance Matrix						
	Jen	Tom	Bob	Leaf	Jim	Sum
Jen						
Tom						
Bob	Inf	Inf		1	1	2
Leaf						
Jim						

*What is **Bob's** closeness?*

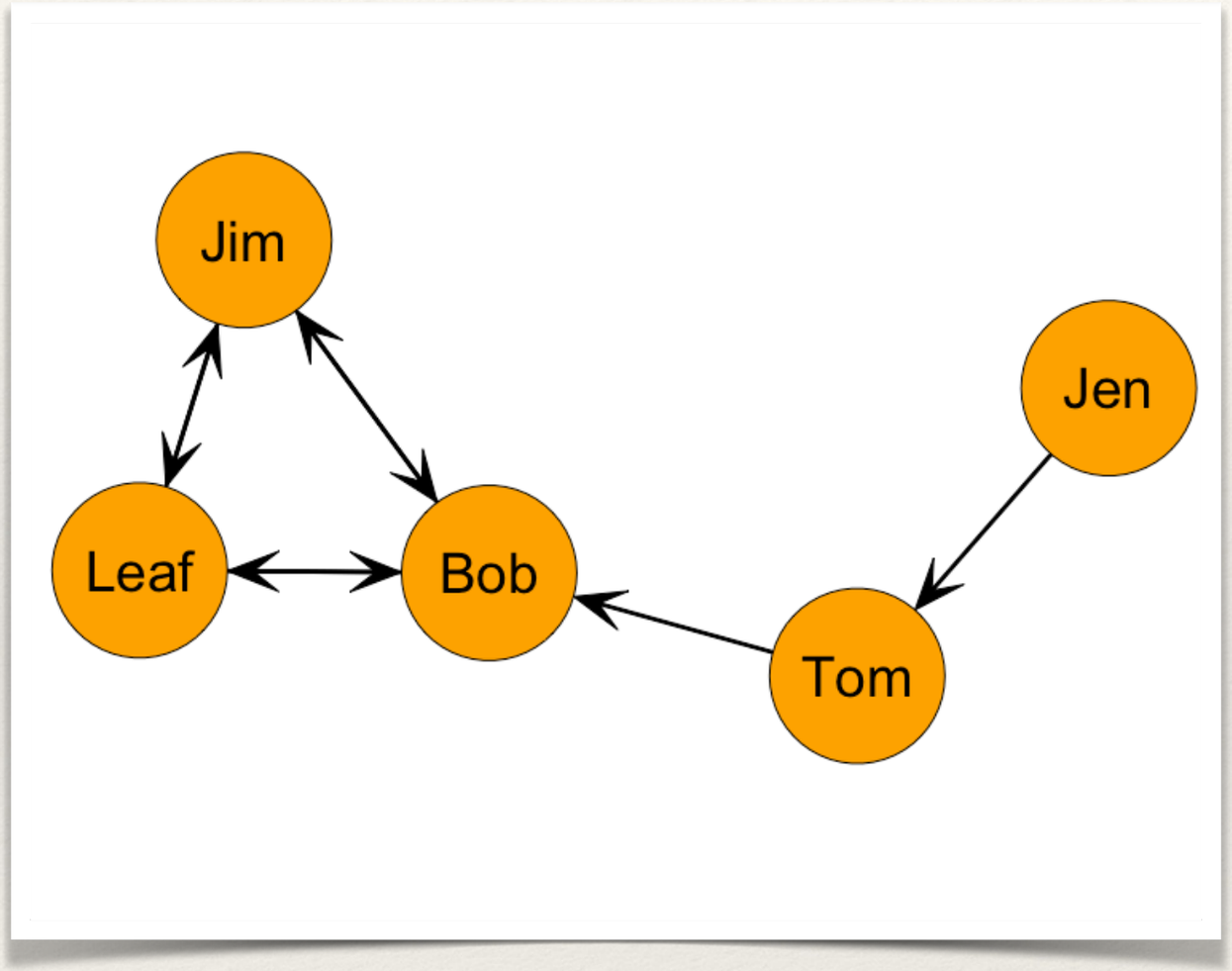
Example: Closeness Centrality for Directed Binary Network



Distance Matrix						
	Jen	Tom	Bob	Leaf	Jim	Sum
Jen						
Tom						
Bob	Inf	Inf		1	1	2
Leaf						
Jim						

Take the inverse and ***Bob's*** *score is:* $1/2 = 0.5$

Example: Closeness Centrality for Directed Binary Network



Distance Matrix						
	Jen	Tom	Bob	Leaf	Jim	Sum
Jen						
Tom						
Bob	Inf	Inf		1	1	2
Leaf						
Jim						

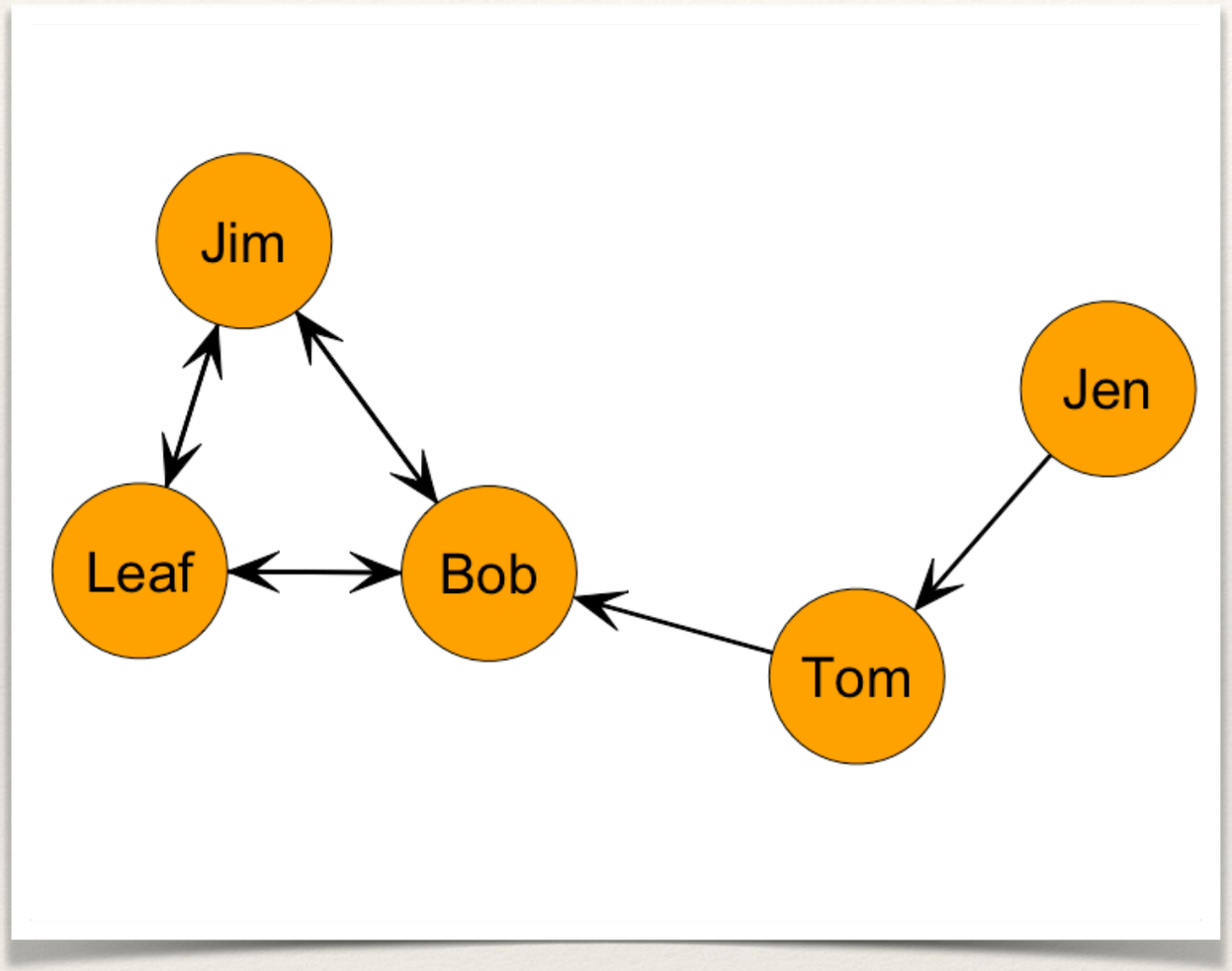
*Then, **Bob's** standardized score is*

$(1/2)*(g-1) = 0.5*4 = 0.2$

Group Closeness Centralization: Directed Binary Graphs

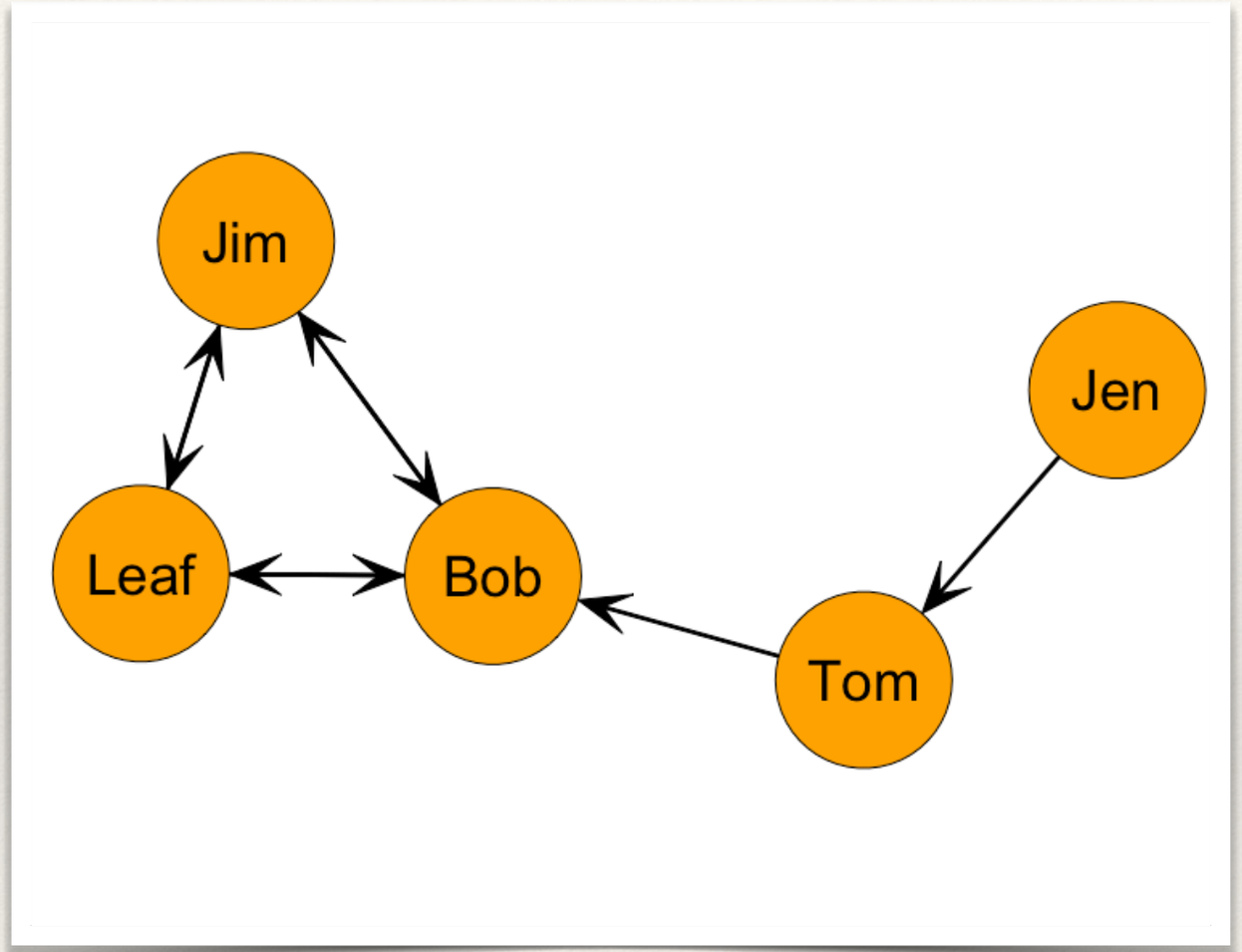
- The denominator for group closeness centralization is not defined for directed graphs (see Wasserman & Faust 1994: p. 200).
  - But, we could just treat it as undirected and calculate a score.

Example: Undirected, Binary Network



Approximately 0.556

Example: Undirected, Binary Network



Compare the centralization scores:

Indegree = 0.438

Outdegree = 0.125

Closeness = 0.556

What can we say about the differences in the centralization scores for each type of centrality?

Betweenness Centrality

- We have seen how centrality can be conceptualized as:
  - Having a high number of ties
  - Being close to others in the network
  - We can also conceptualize centrality as a node that lies on a particular path between other nodes.
    - *Betweenness centrality* is based on the number of **shortest** paths between *j* and *k* that actor *i* resides on.

Betweenness Centrality

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i)/g_{jk}$$

Betweenness Centrality

The number of geodesics linking j to k.

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i)/g_{jk}$$

Betweenness Centrality

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i)/g_{jk}$$

The number of geodesics linking j to k.

The number of geodesics linking j and k that contain i.

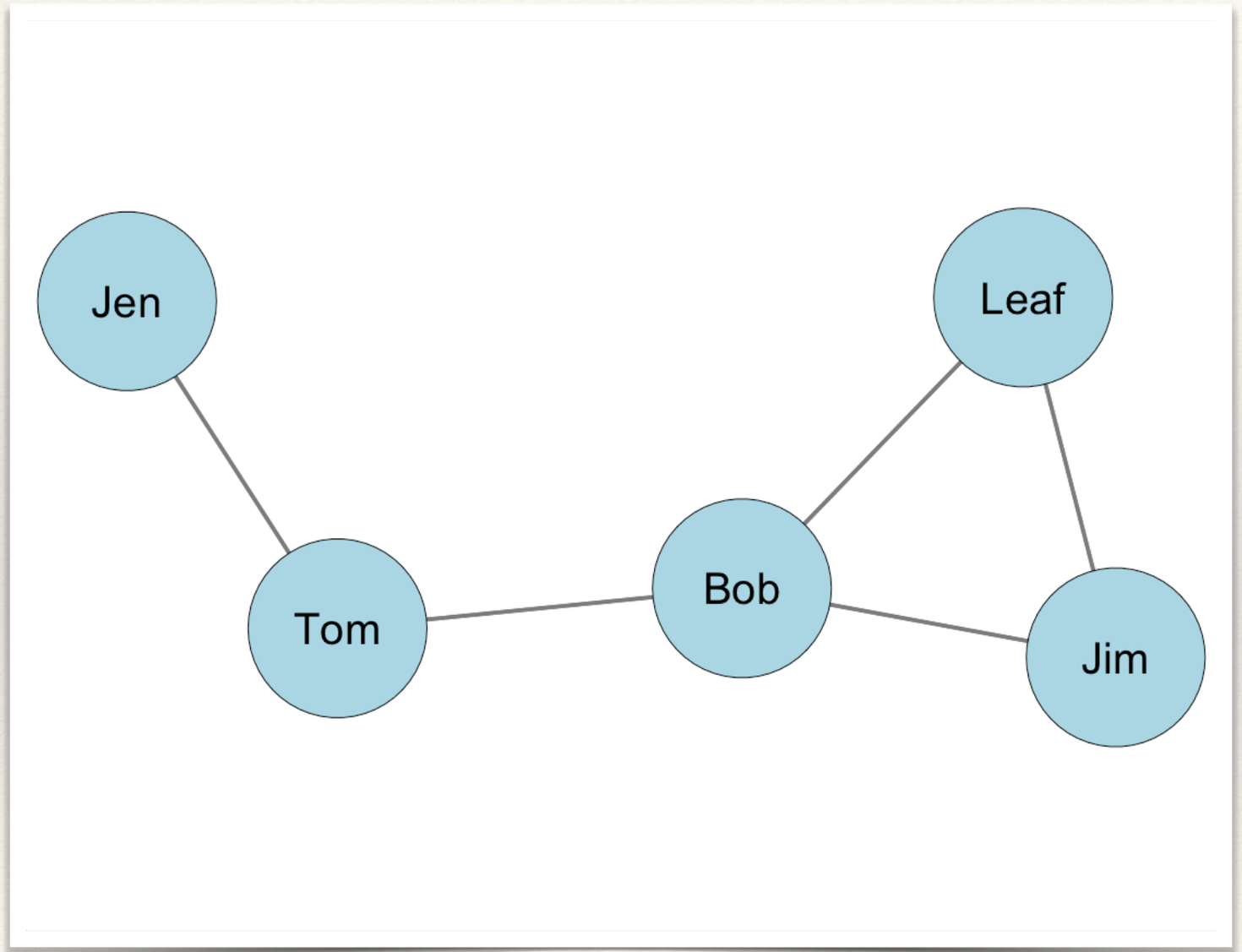
Betweenness Centrality

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i)/g_{jk}$$

So, betweenness centrality is the ratio of the geodesics between *j* and *k* that contain *i*.

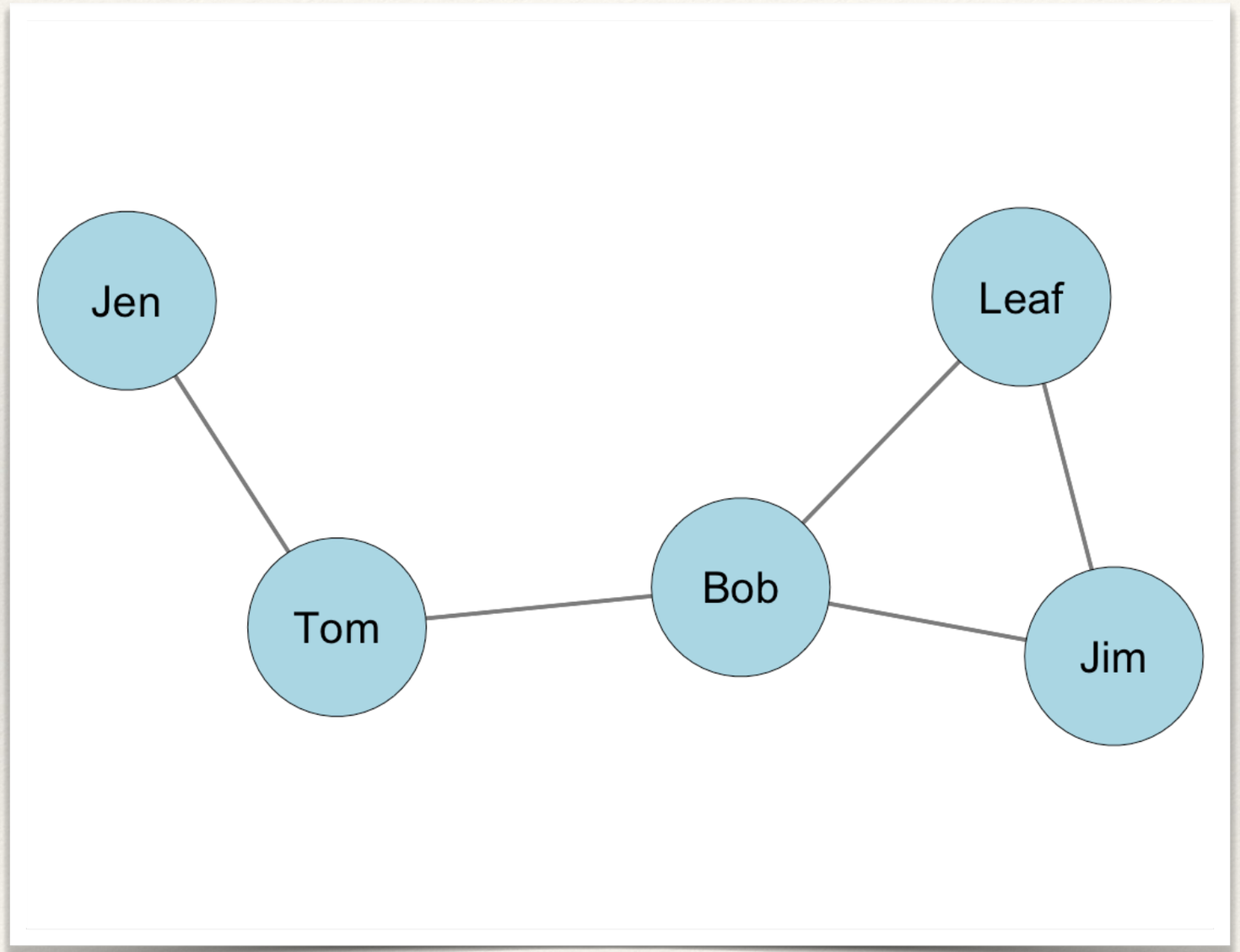
Undirected Networks

Example: Undirected, Binary Network



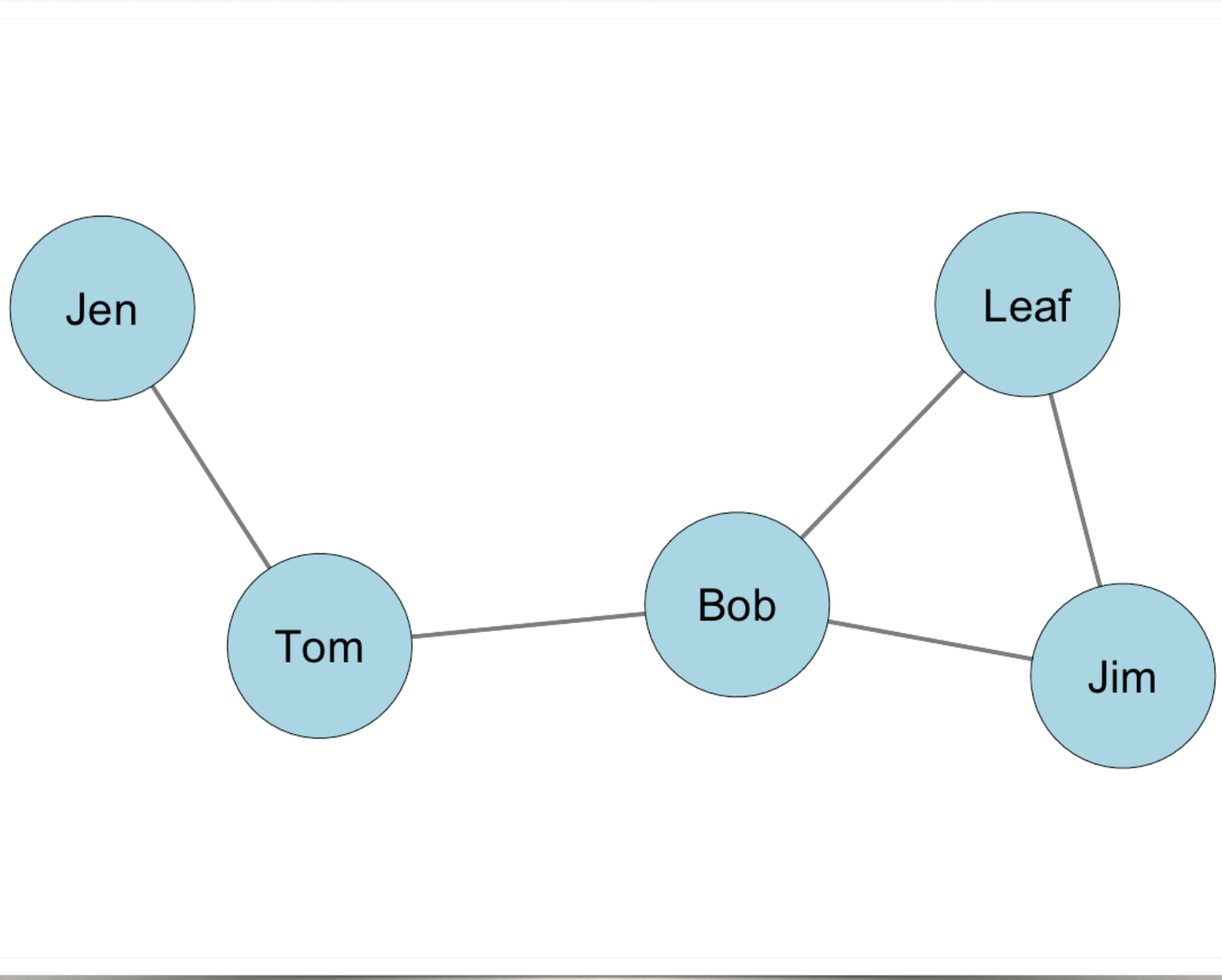
What are the paths between **Jen** and **Jim**?

Example: Undirected, Binary Network



Let's calculate the betweenness centrality for **Bob**.

Example: Undirected, Binary Network



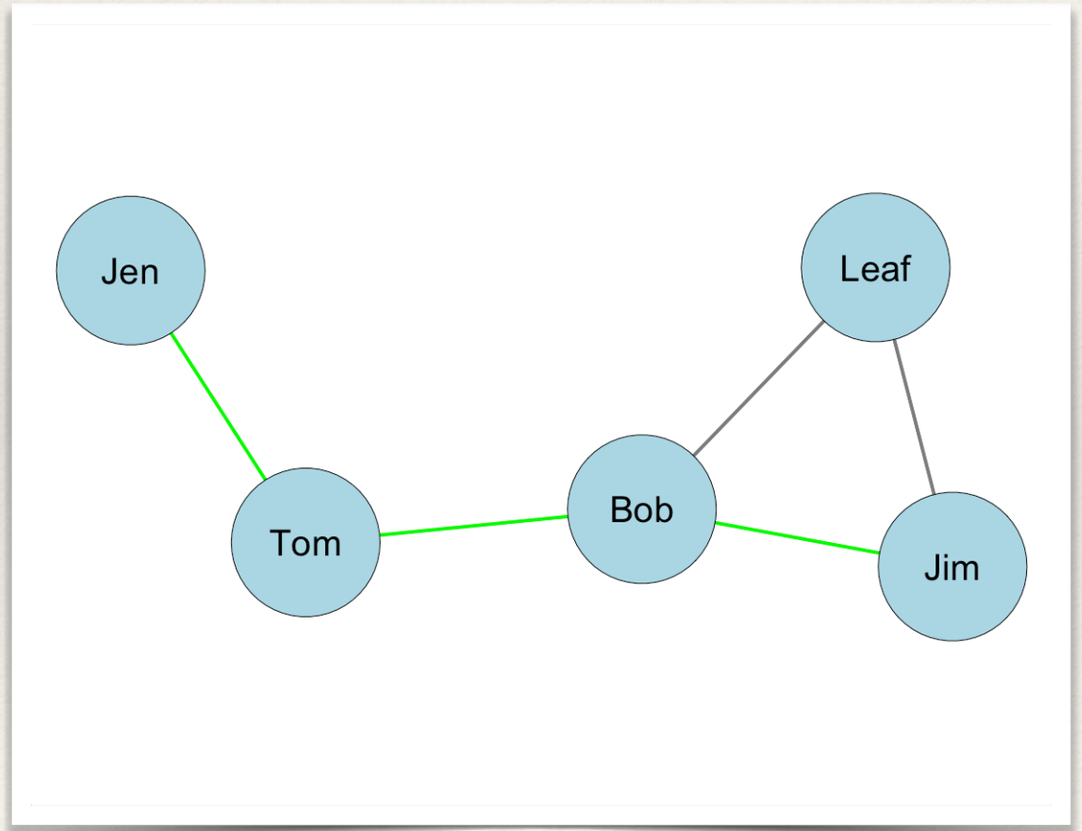
Let's calculate the betweenness centrality for **Bob**.

First, we need to find the geodesics for other nodes.

Then, figure out how many **Bob** occupies.

Example: Undirected, Binary Network

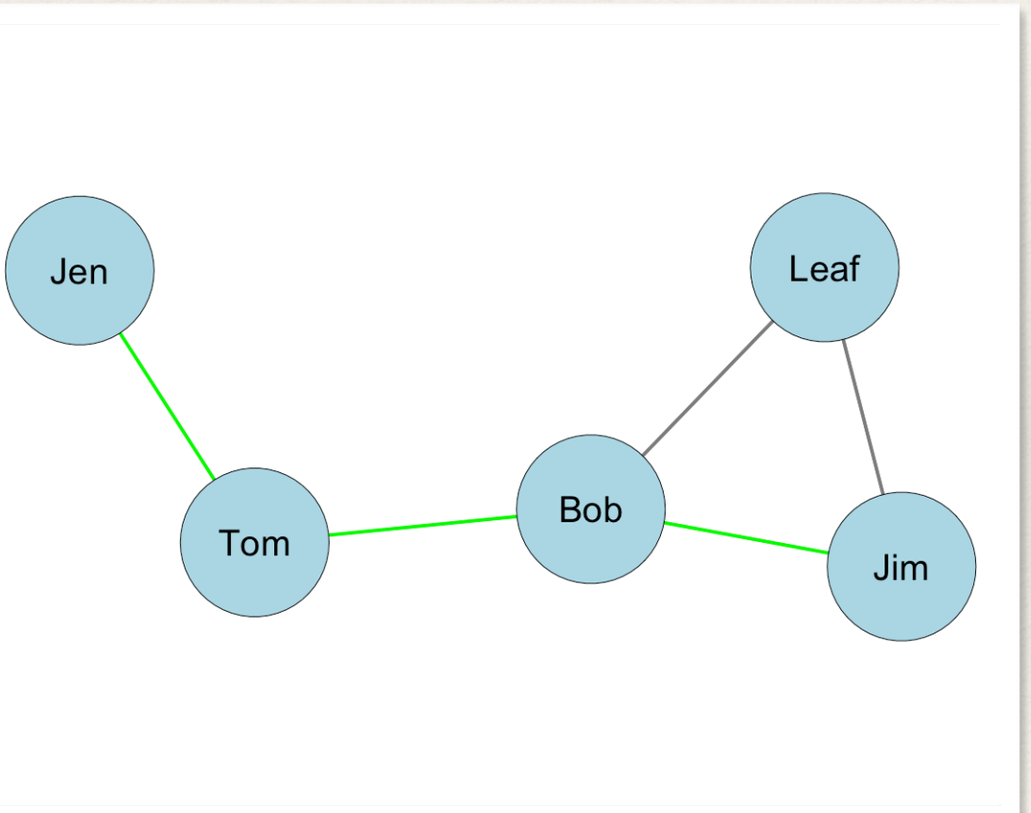
What are the paths between **Jen** and **Jim**?



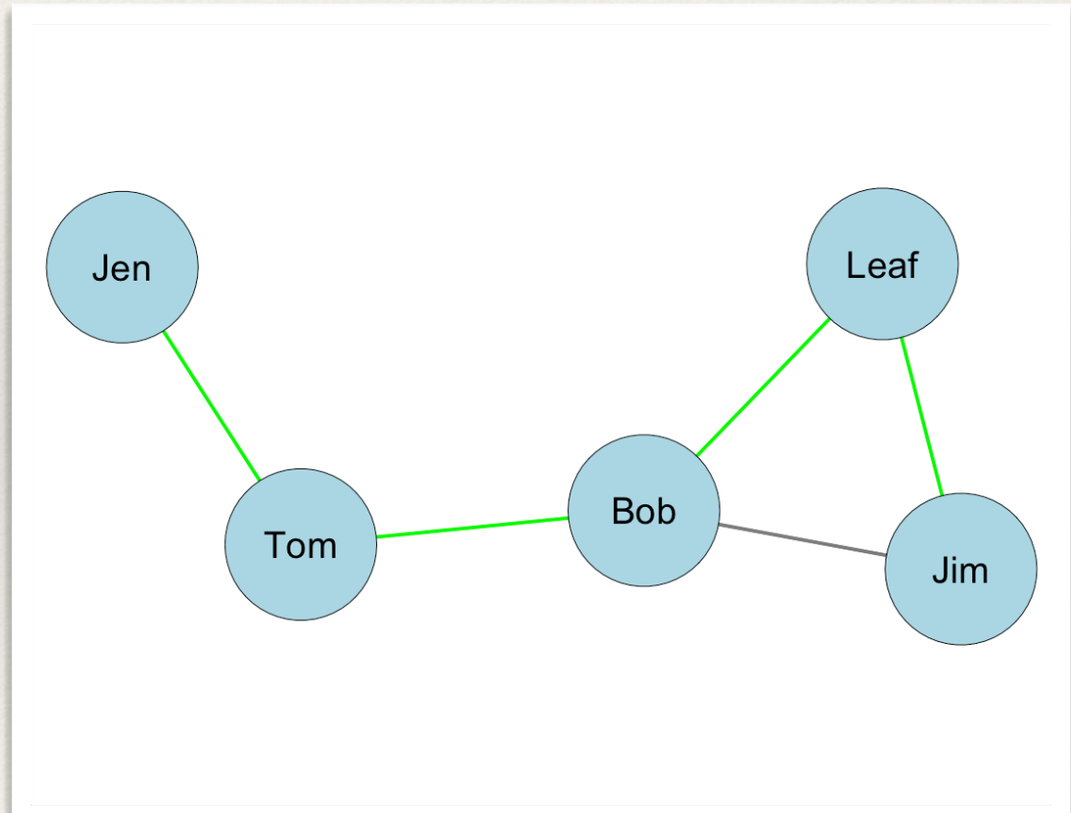
Jen-Tom-Bob-Jim

Example: Undirected, Binary Network

What are the paths between **Jen** and **Jim**?

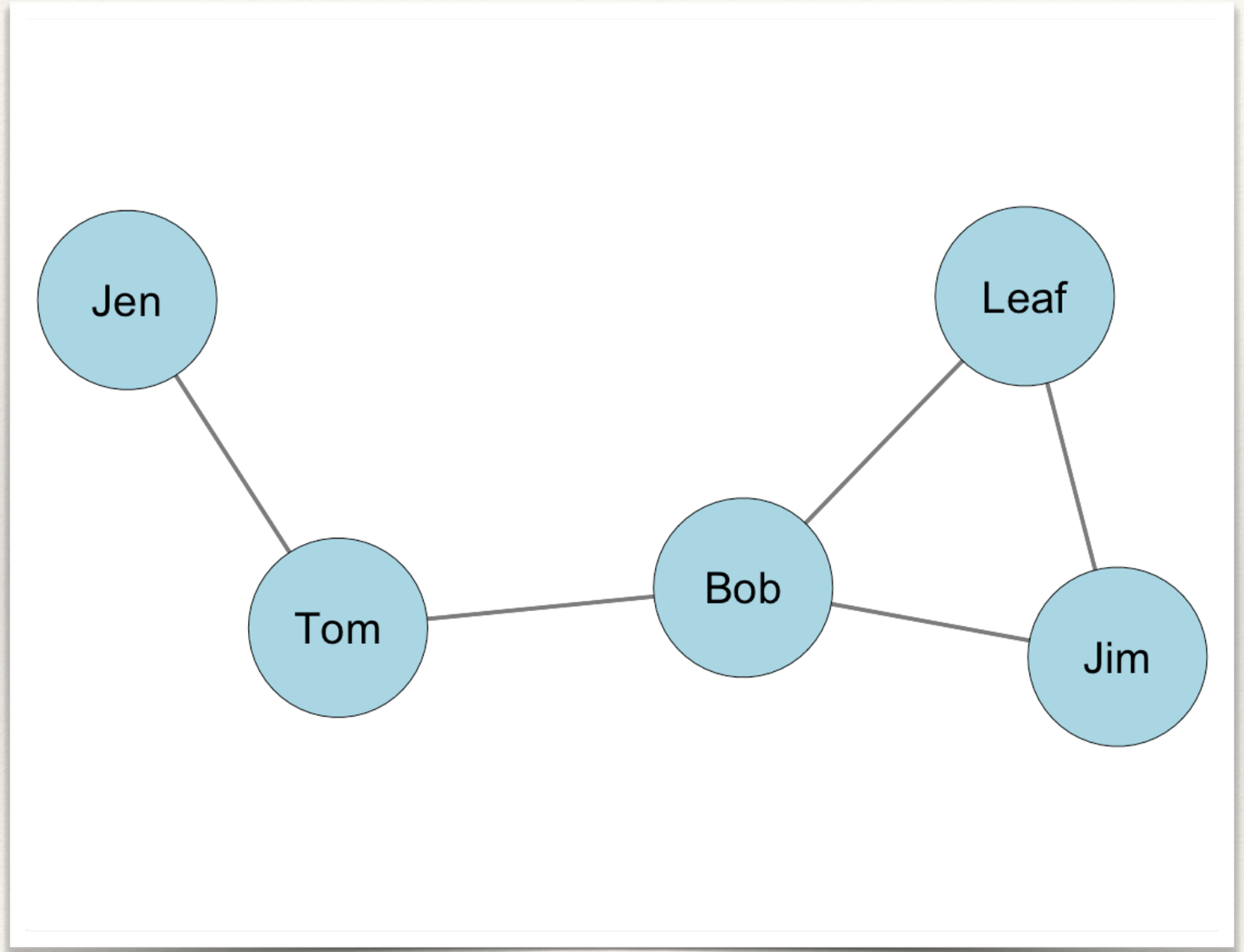


Jen-Tom-Bob-Jim



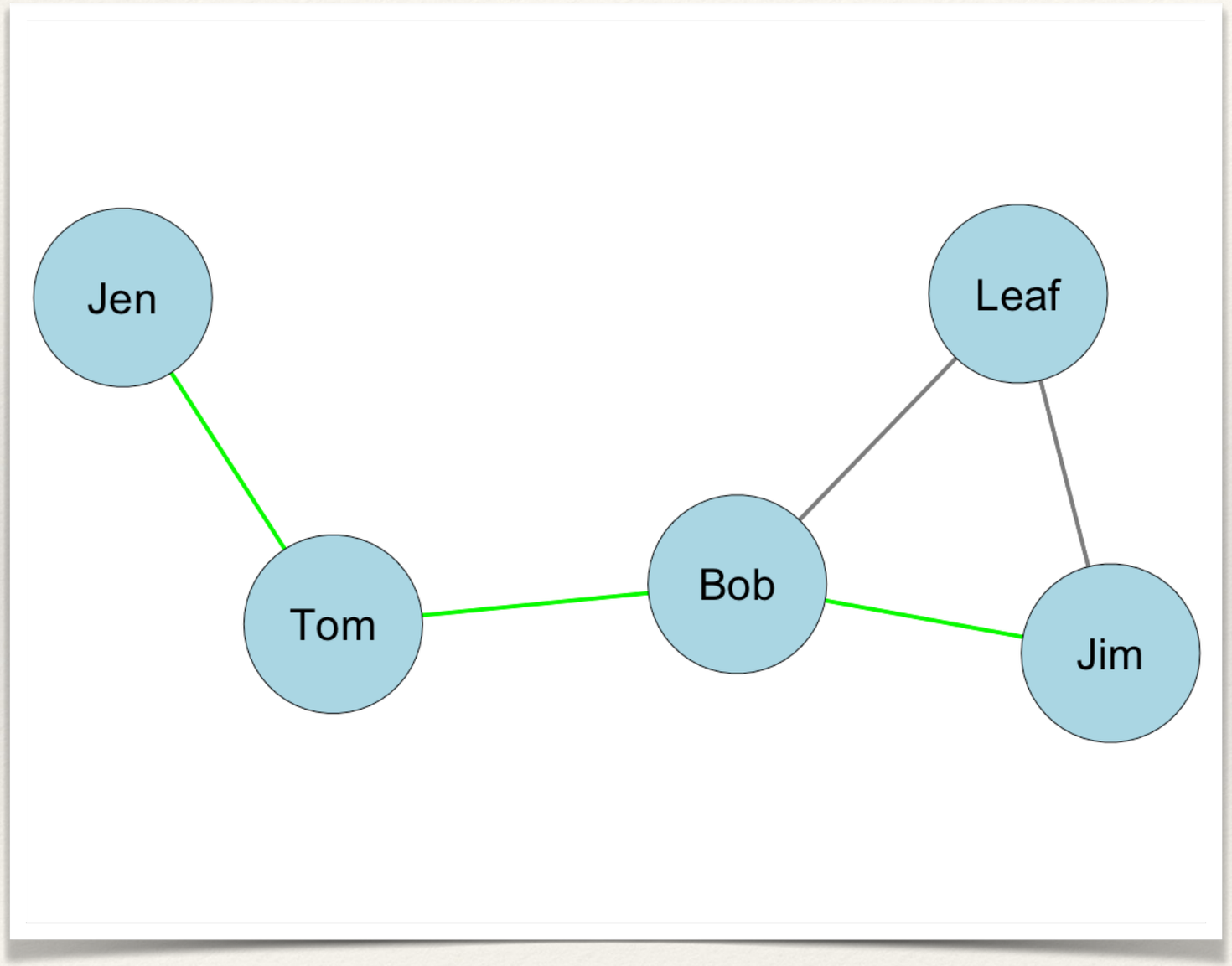
Jen-Tom-Bob-Leaf-Jim

Example: Undirected, Binary Network



What are the ***geodesic*** paths between **Jen** and **Jim**?

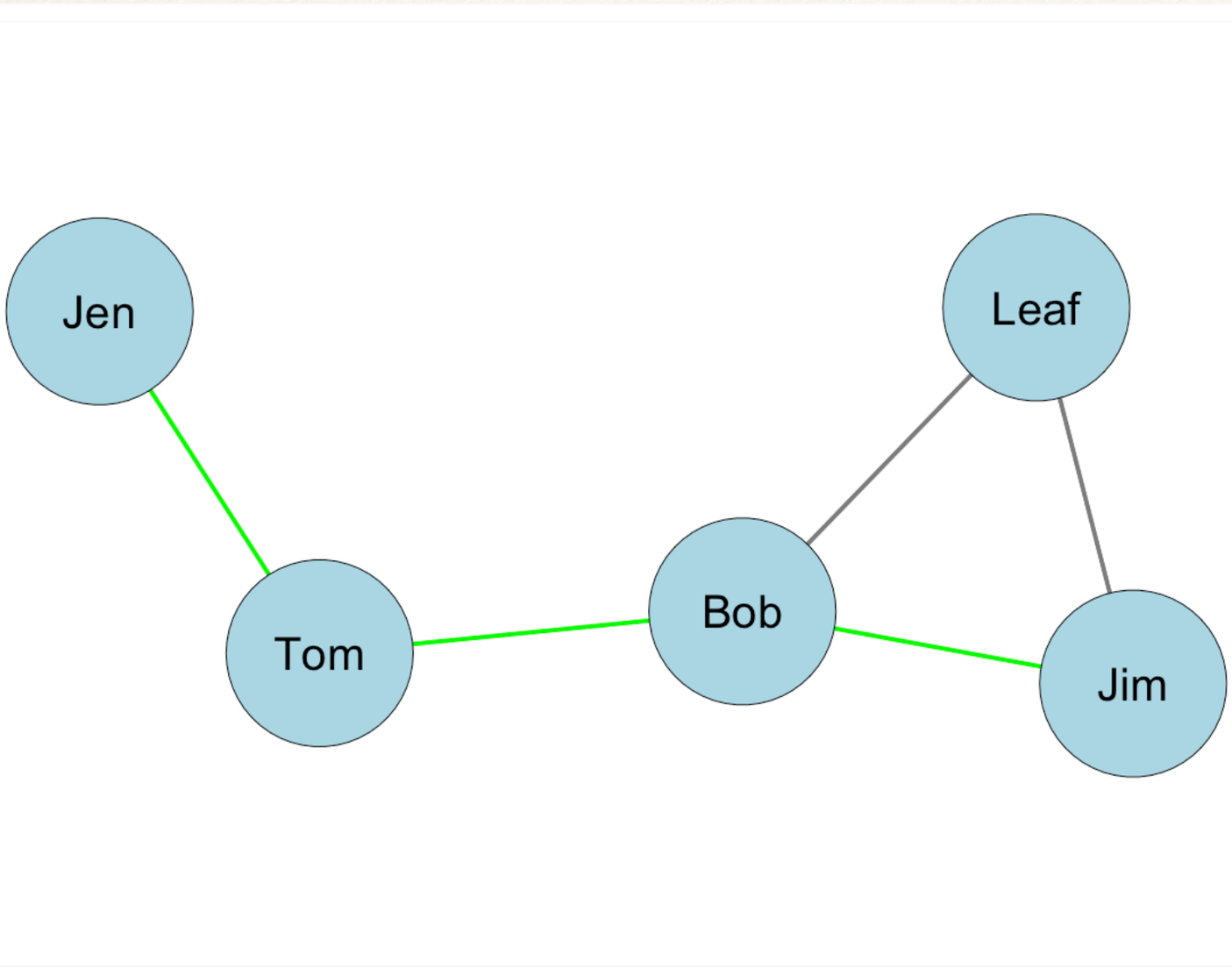
Example: Undirected, Binary Network



What are the ***geodesic*** paths between **Jen** and **Jim**?

Jen-Tom-Bob-Jim

Example: Undirected, Binary Network

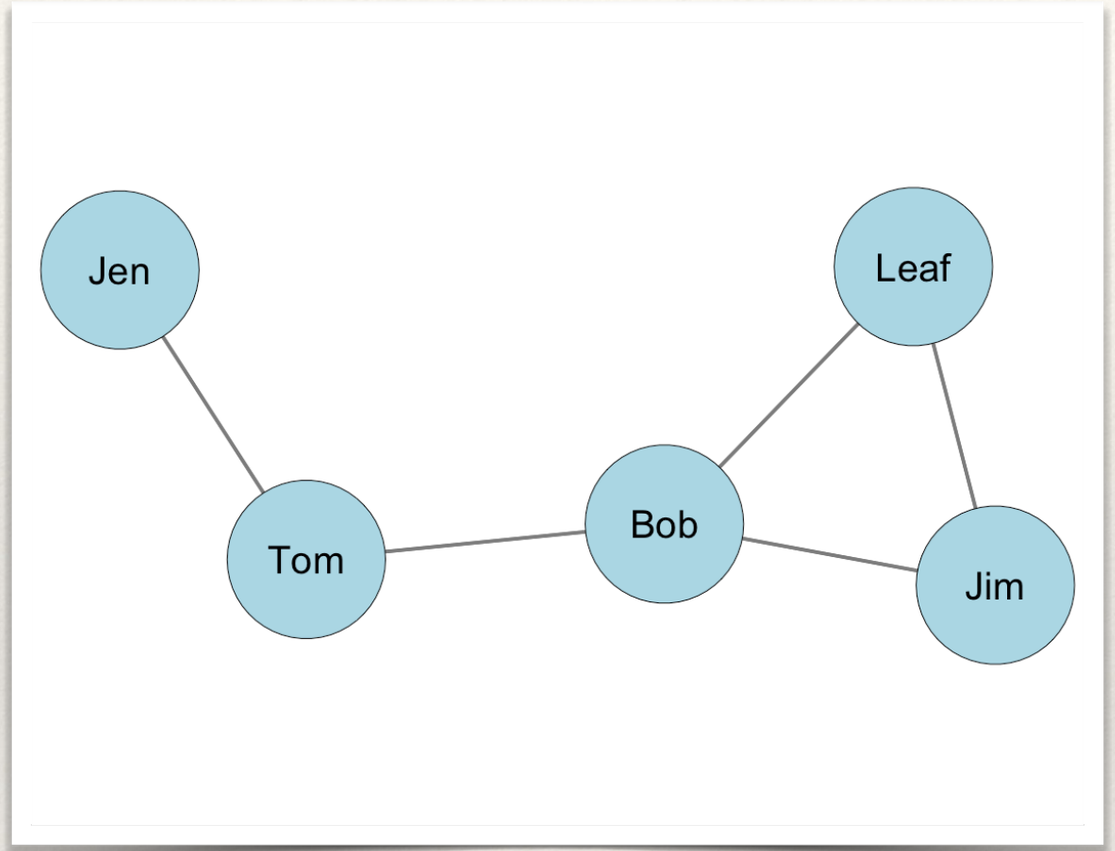


What are the ***geodesic*** paths between **Jen** and **Jim**?

Jen-Tom-Bob-Jim

Is there anyone on the geodesic between Jen and Jim?

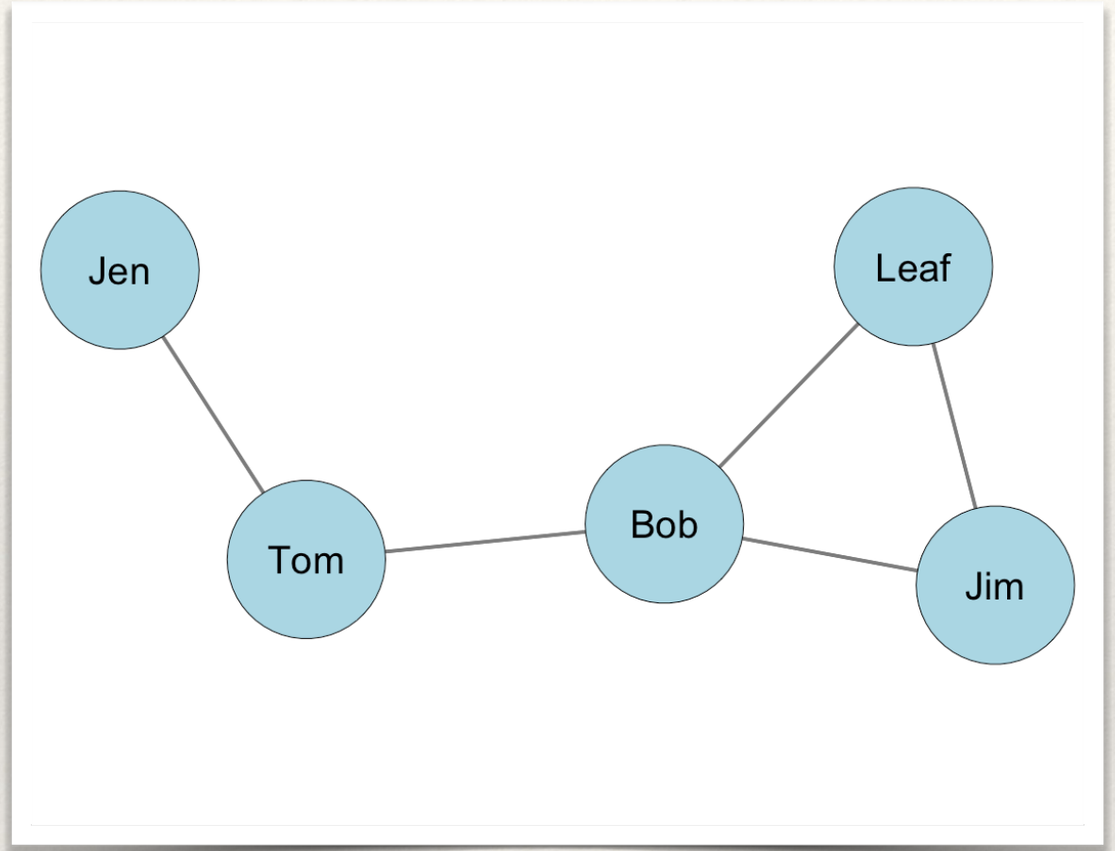
Example: Undirected, Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen				/?
Tom				
Leaf				
Jim				

How many geodesics from **Jen** to **Jim**?

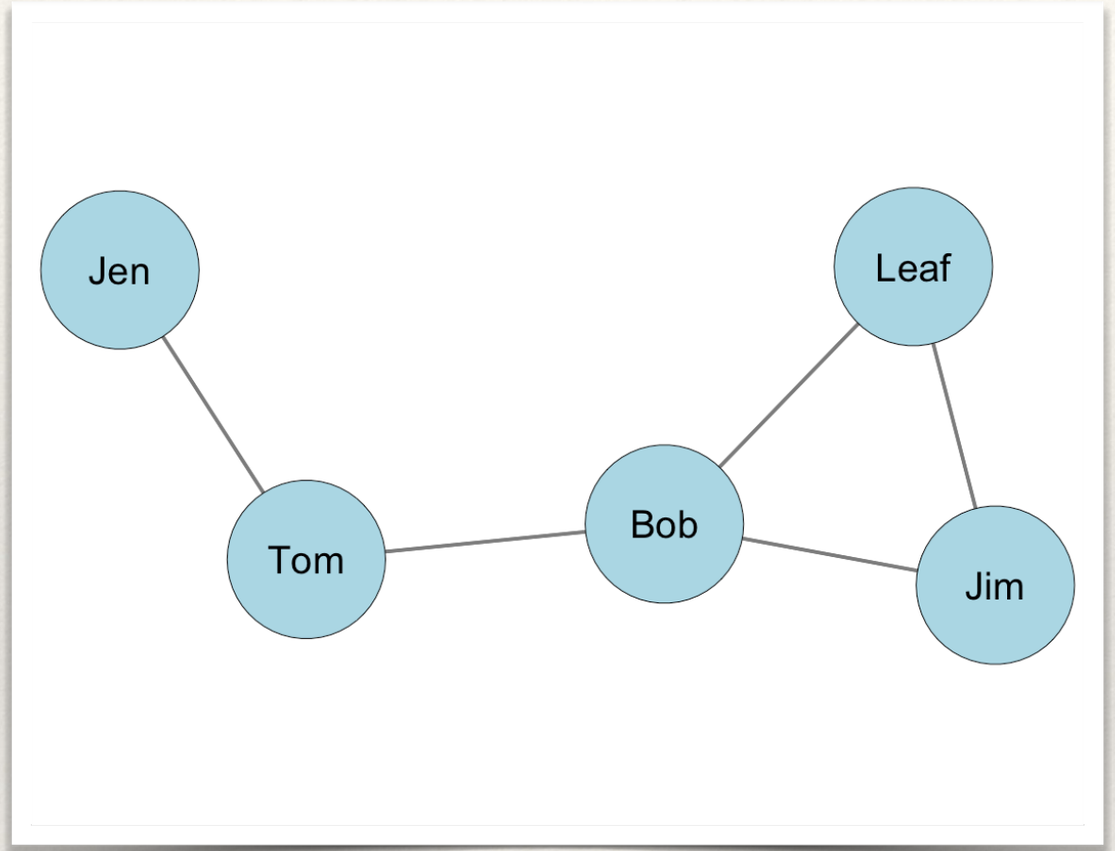
Example: Undirected, Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen				/1
Tom				
Leaf				
Jim				

There is 1 geodesic from **Jen** to **Jim**

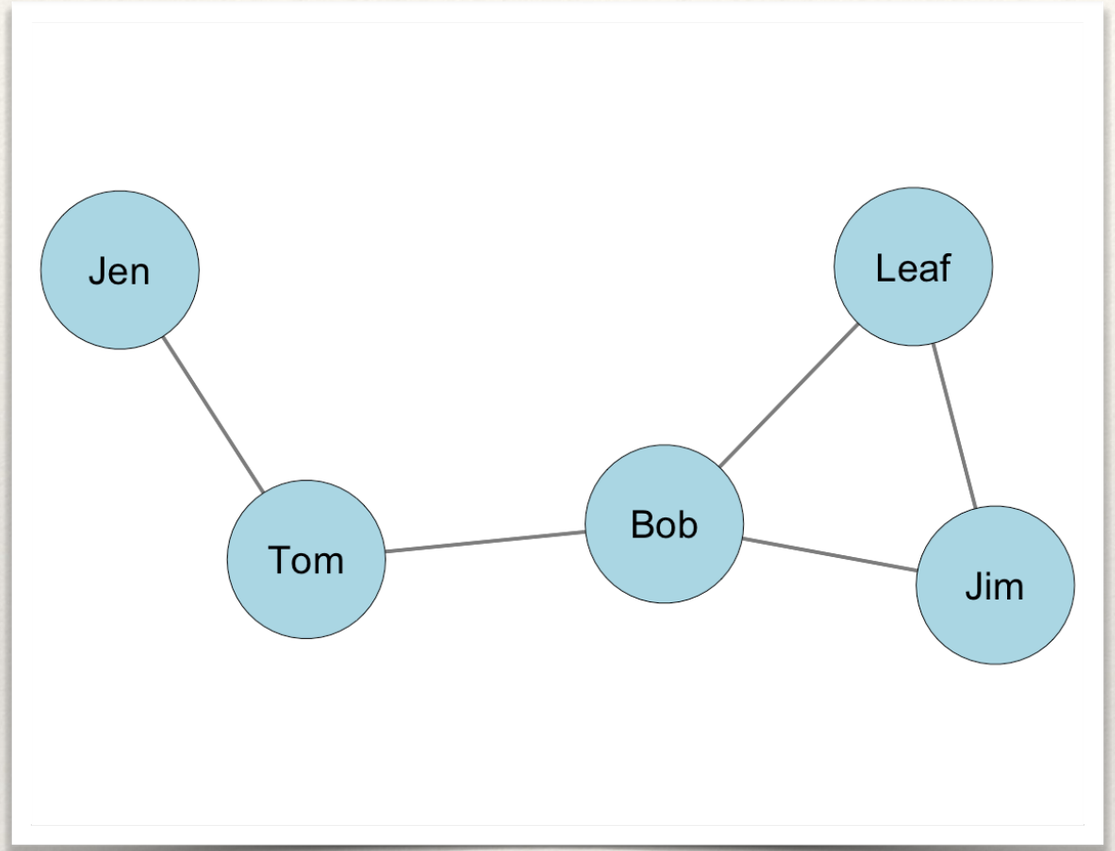
Example: Undirected, Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen				?/1
Tom				
Leaf				
Jim				

How many geodesics from **Jen** to **Jim** include **Bob**?

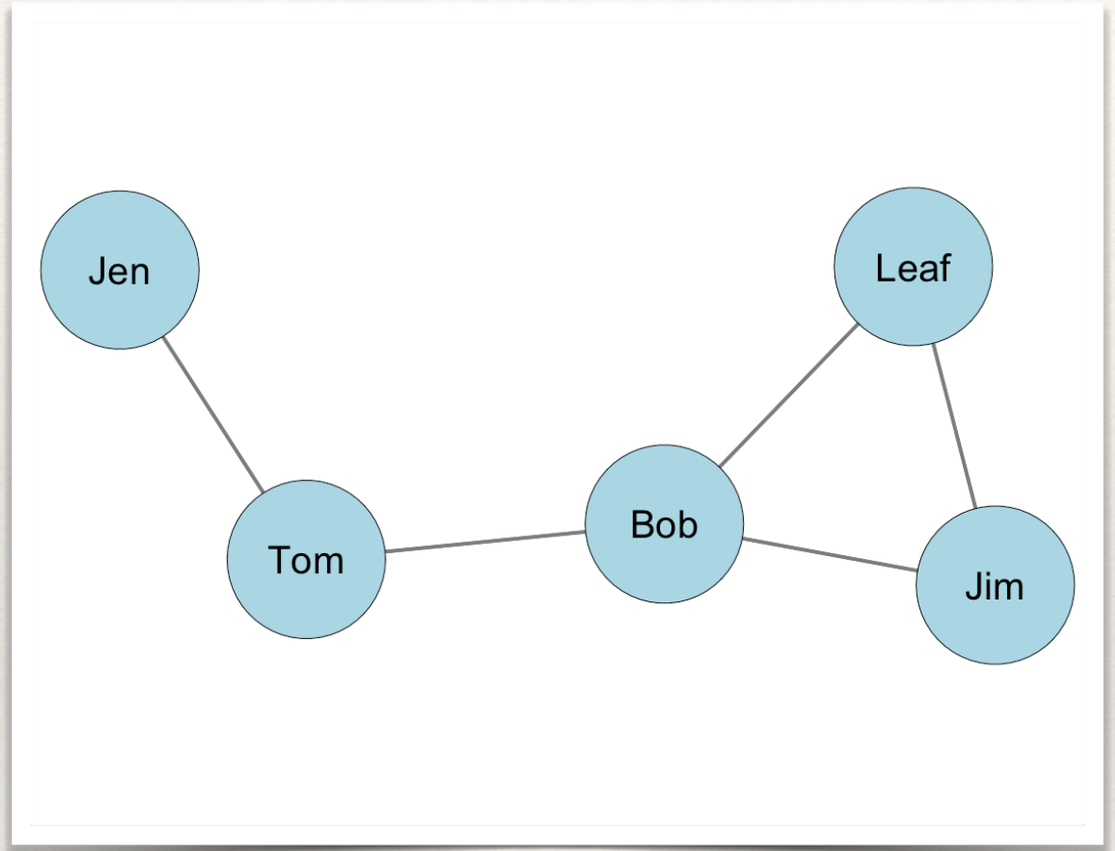
Example: Undirected, Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen				1/1
Tom				
Leaf				
Jim				

Bob is on the only geodesic from **Jen** to **Jim**

Example: Undirected, Binary Network

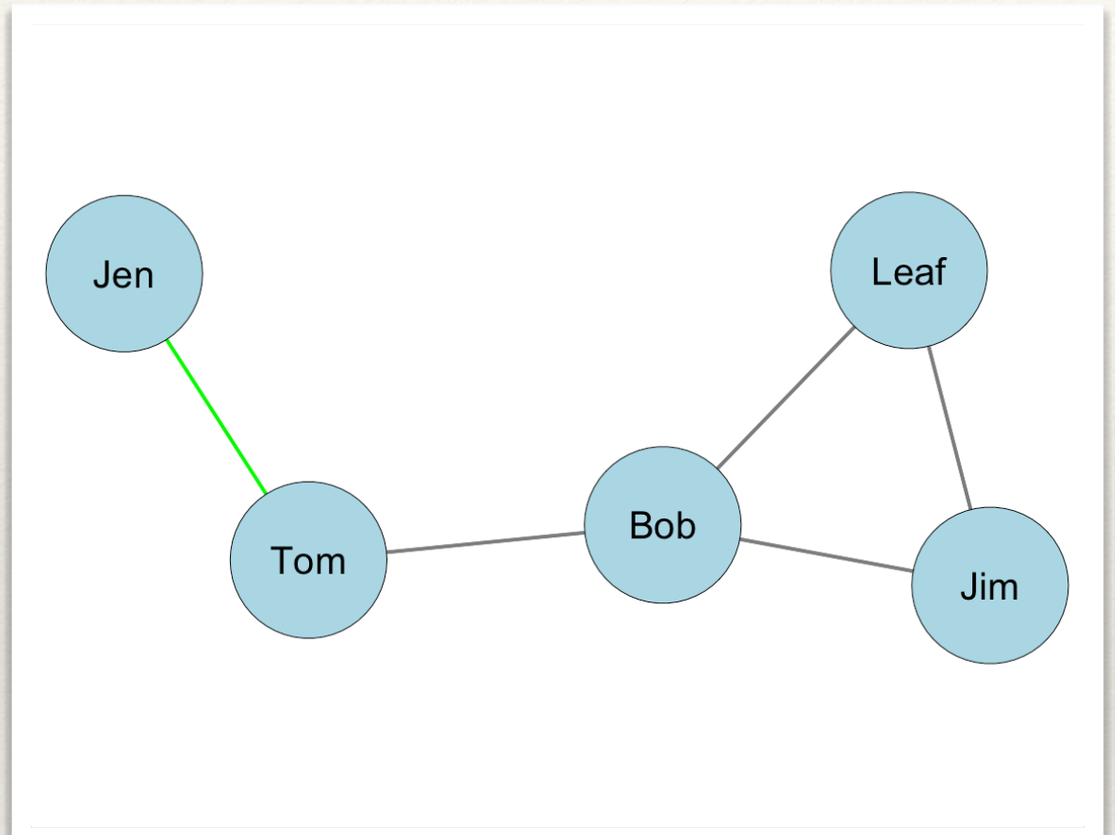


Geodesic Proportions for *Bob*

	Jen	*Tom*	*Leaf*	*Jim*
Jen		?/?		1/1
Tom				
Leaf				
Jim				

What about **Jen** to **Tom**?

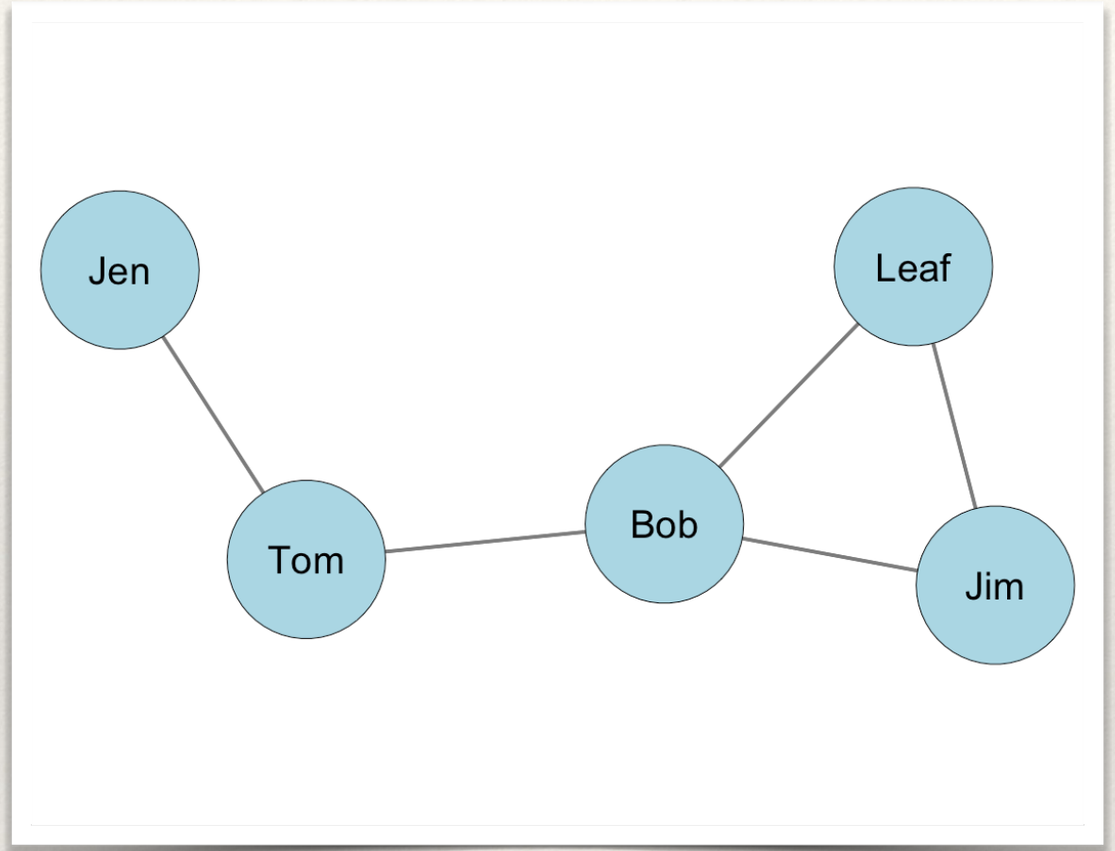
Example: Undirected, Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1		1/1
Tom				
Leaf				
Jim				

Bob is not on the only geodesic from **Jen** to **Tom**

Example: Undirected, Binary Network

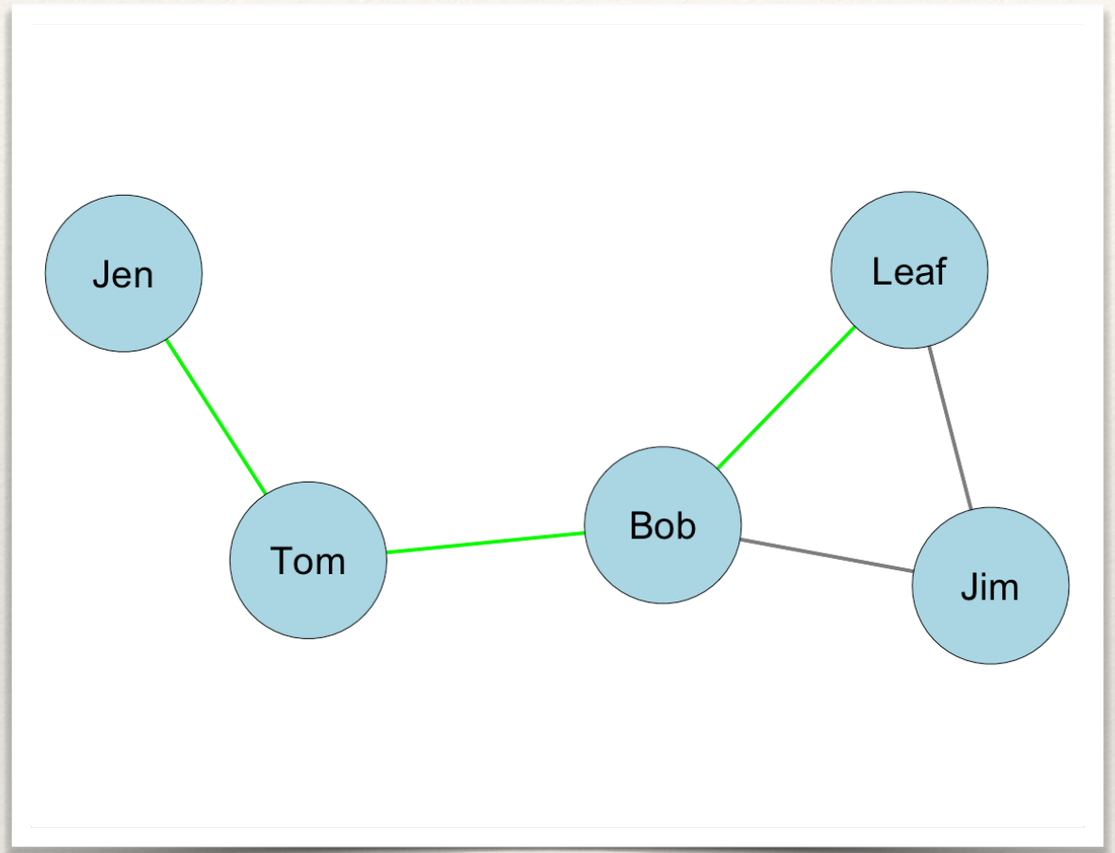


Geodesic Proportions for *Bob*

	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	?/?	1/1
Tom				
Leaf				
Jim				

What about **Jen** to **Leaf**?

Example: Undirected, Binary Network

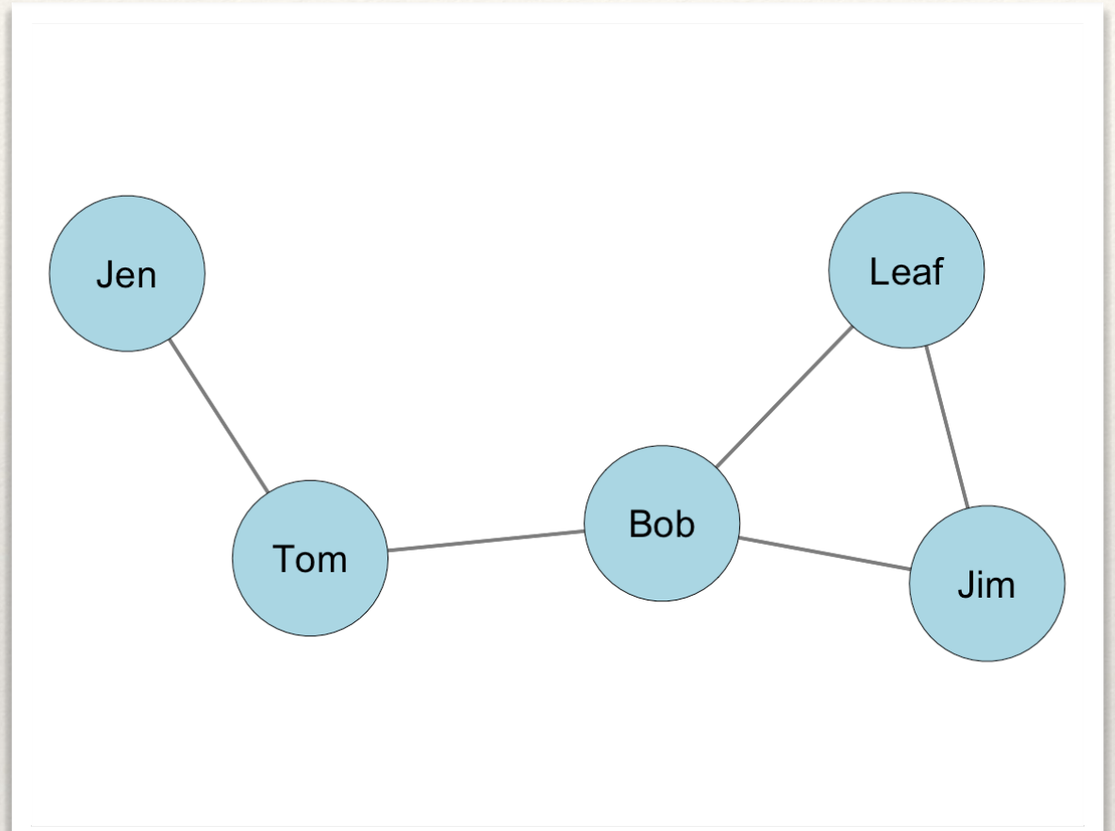


Geodesic Proportions for *Bob*

	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom				
Leaf				
Jim				

Bob is on the geodesic from **Jen** to **Leaf**

Example: Undirected, Binary Network

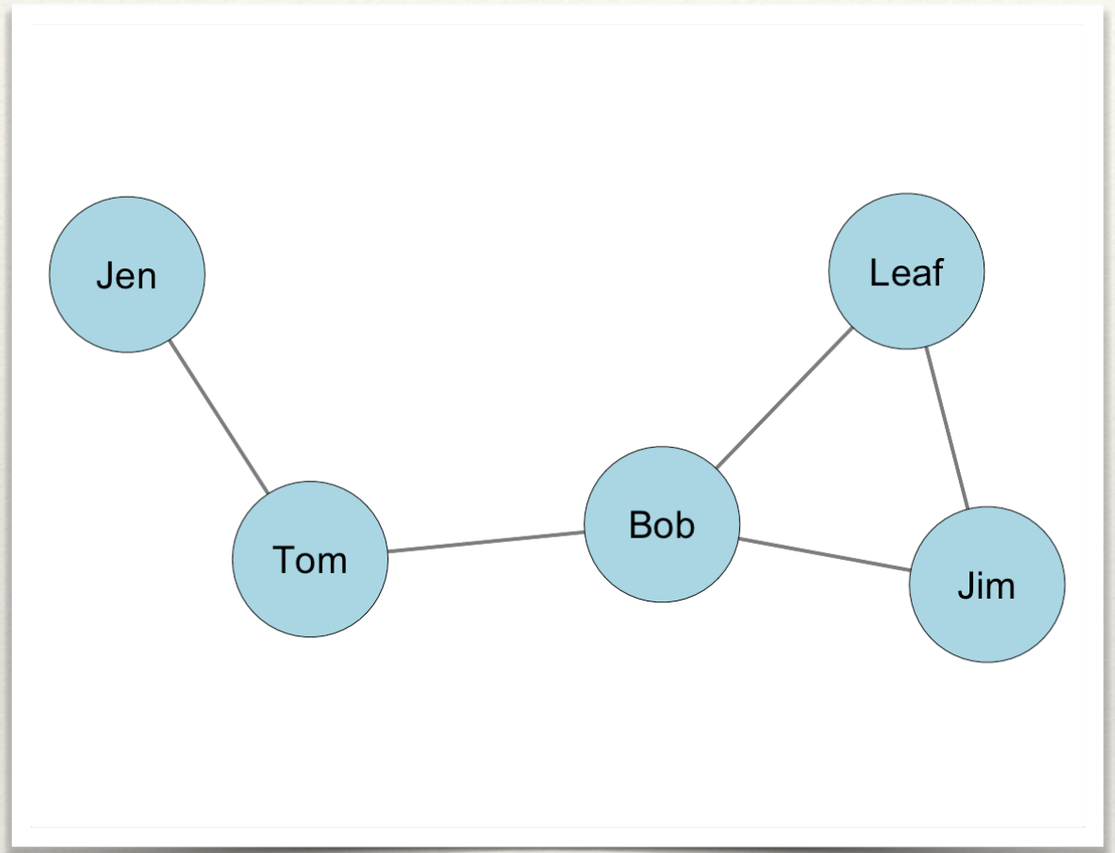


Of the geodesics between
Jen, Tom
Jen, Leaf
Jen, Jim
how many include **Bob**?

Geodesic Proportions for *Bob*

	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom				
Leaf				
Jim				

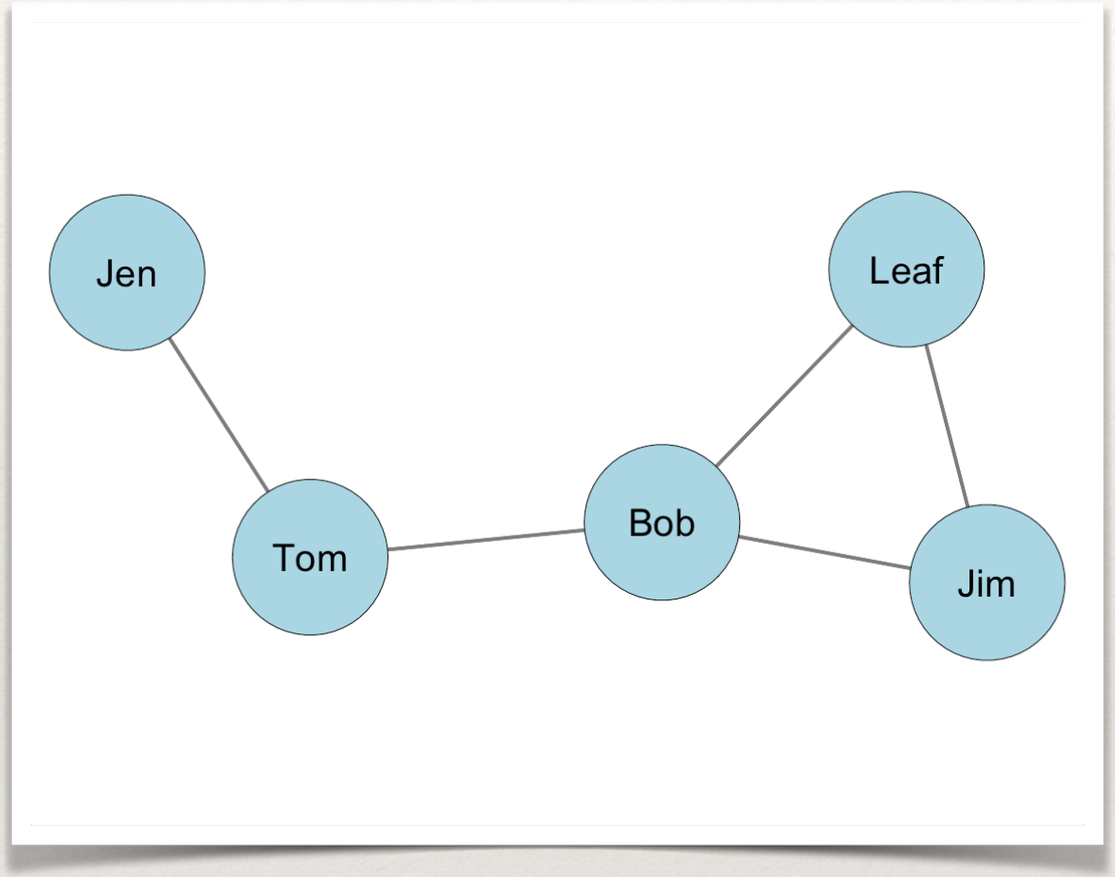
Example: Undirected, Binary Network



Geodesic Proportions for *Bob*

	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom				
Leaf				
Jim				

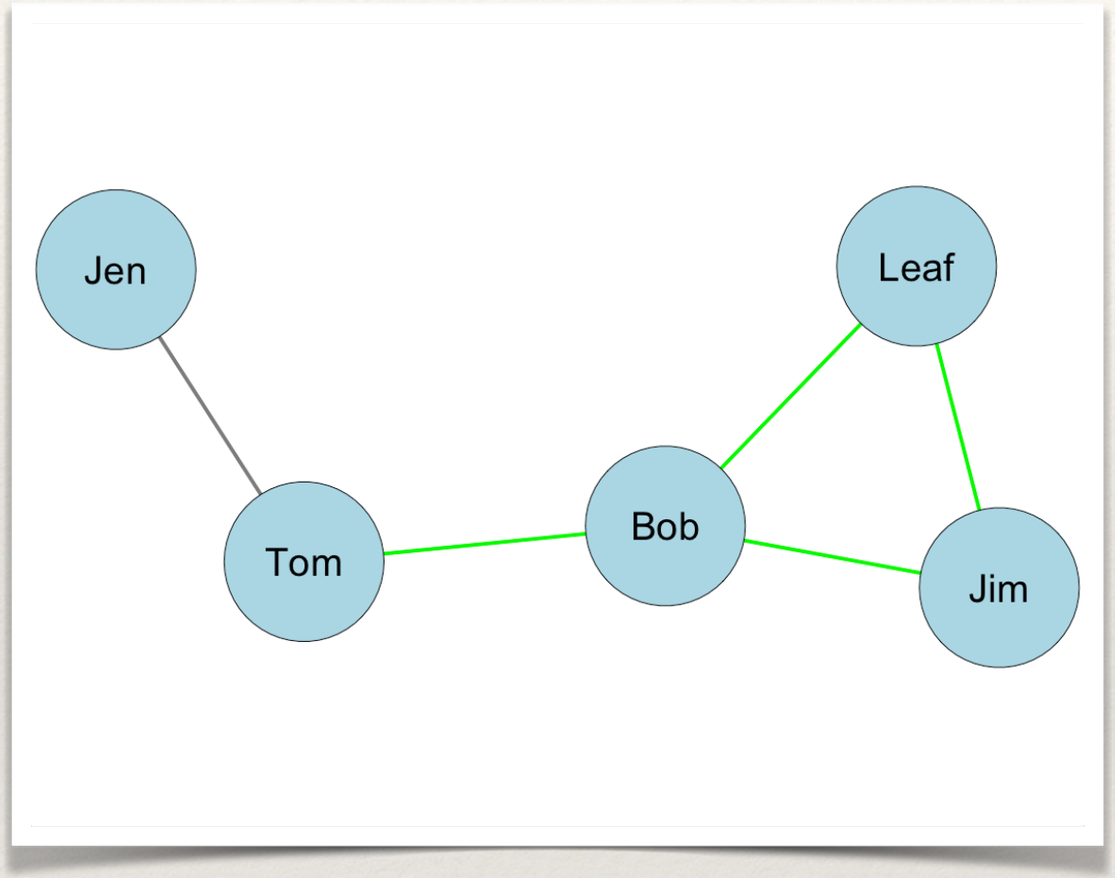
Example: Undirected, Binary Network



To finish, we need to calculate the geodesics for the rest of the matrix for **Bob**

Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom			?/?	?/?
Leaf				?/?
Jim				

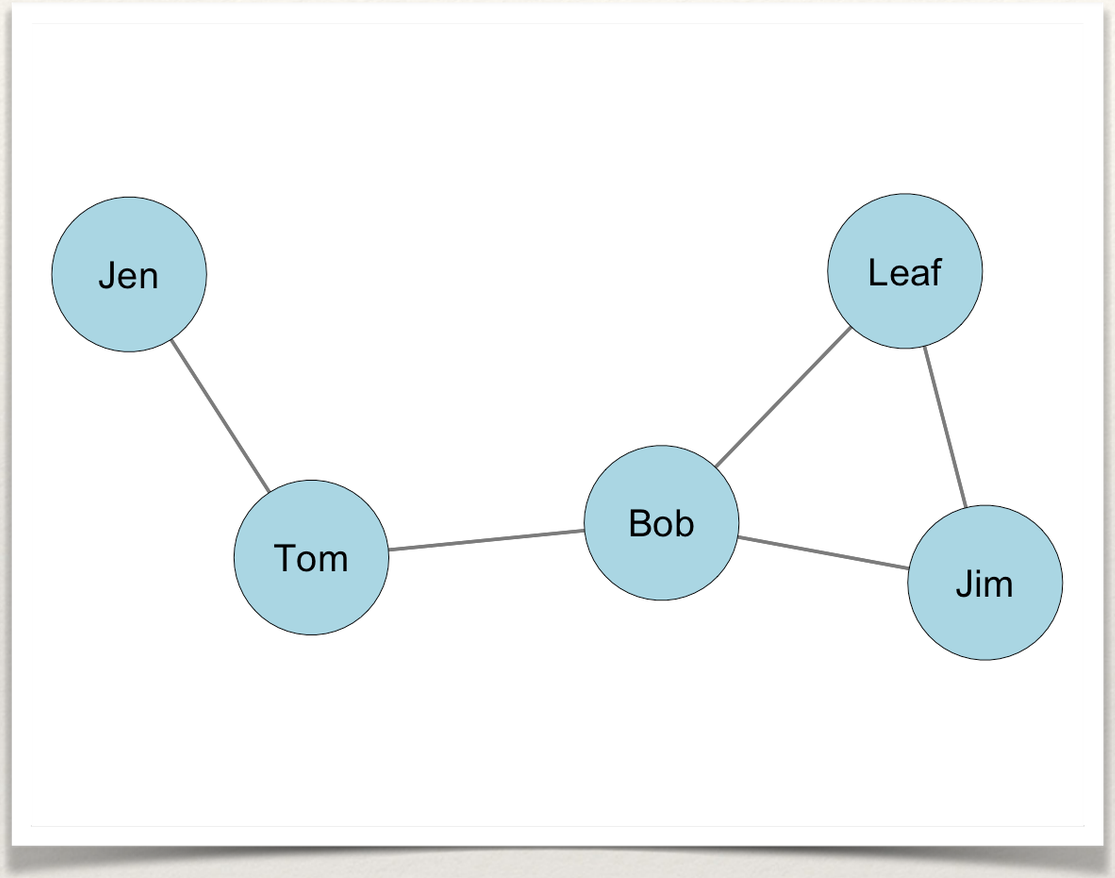
Example: Undirected, Binary Network



To finish, we need to calculate the geodesics for the rest of the matrix for **Bob**

Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom			1/1	1/1
Leaf				0/1
Jim				

Example: Undirected, Binary Network

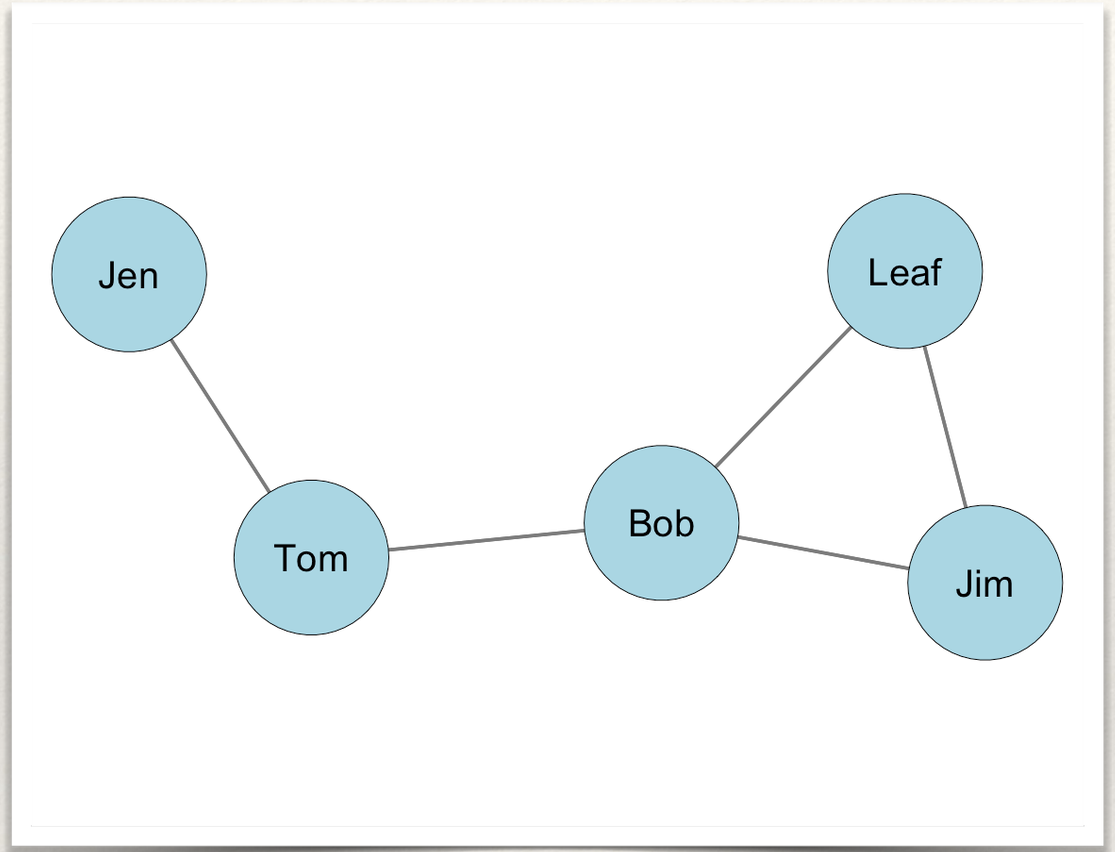


Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom			1/1	1/1
Leaf				0/1
Jim				

The sum of all these ratios is **Bob's** betweenness centrality score.

*What is **Bob's** score?*

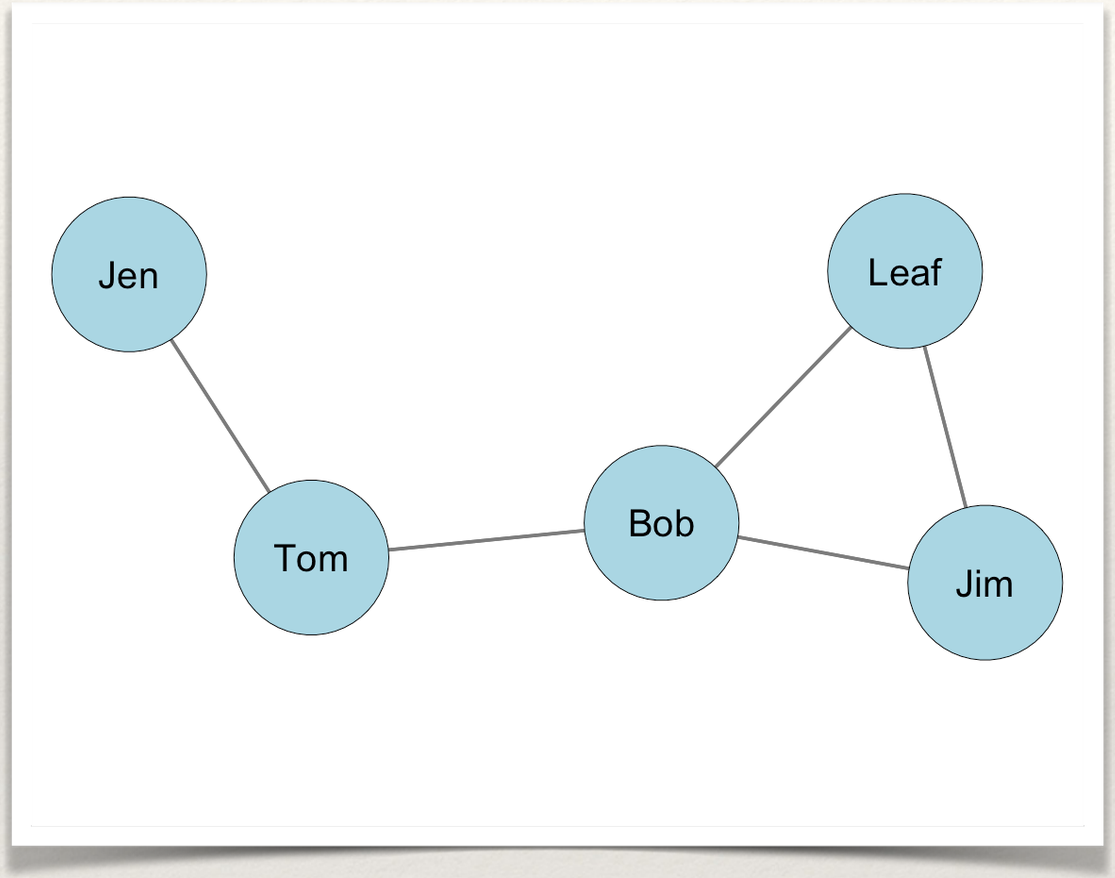
Example: Undirected, Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom			1/1	1/1
Leaf				0/1
Jim				

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i)/g_{jk}$$

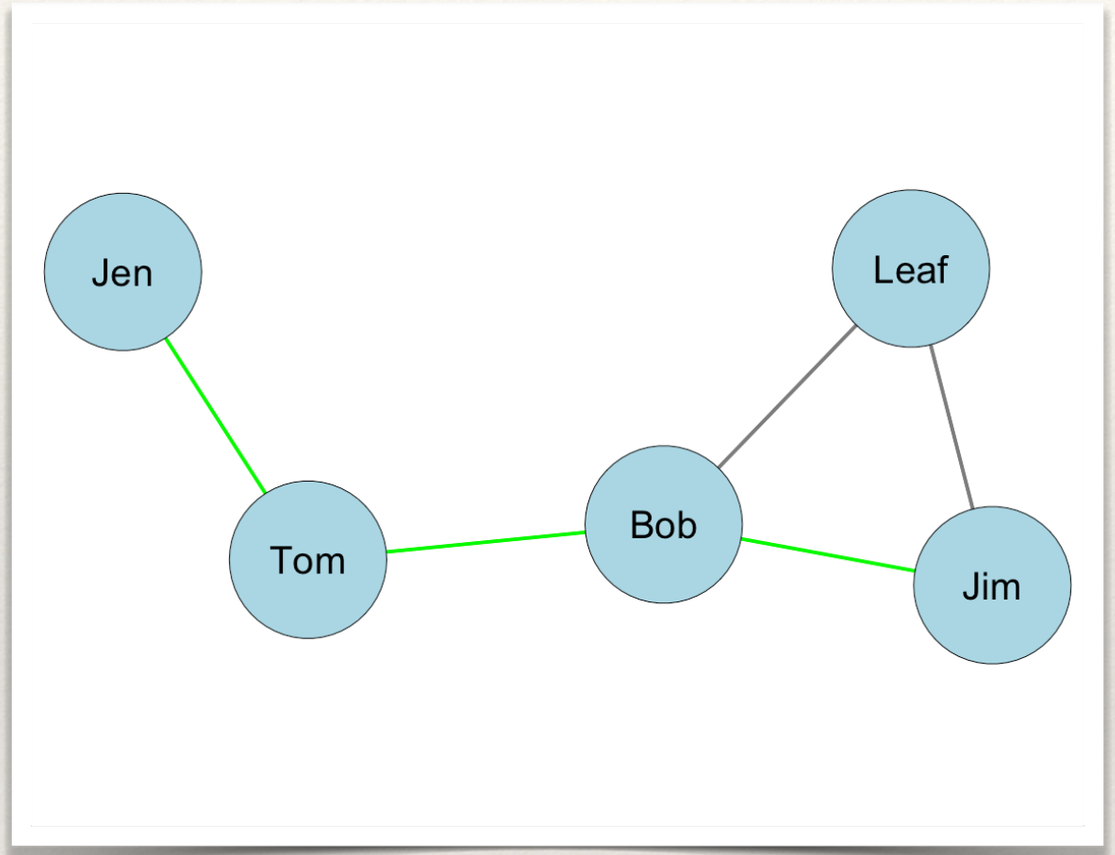
Example: Undirected, Binary Network



Geodesic Proportions for *Leaf*				
	Jen	*Tom*	*Bob*	*Jim*
Jen		?/?	?/?	?/?
Tom			?/?	?/?
Bob				?/?
Jim				

Let's do this now for **Leaf**. *What is* ***Leaf's*** *betweenness centrality?*

Example: Undirected, Binary Network



Geodesic Proportions for *Leaf*				
	Jen	*Tom*	*Bob*	*Jim*
Jen		0/1	0/1	0/1
Tom			0/1	0/1
Bob				0/1
Jim				

Leaf's betweenness centrality is 0

Betweenness Centrality: Undirected Binary Graphs

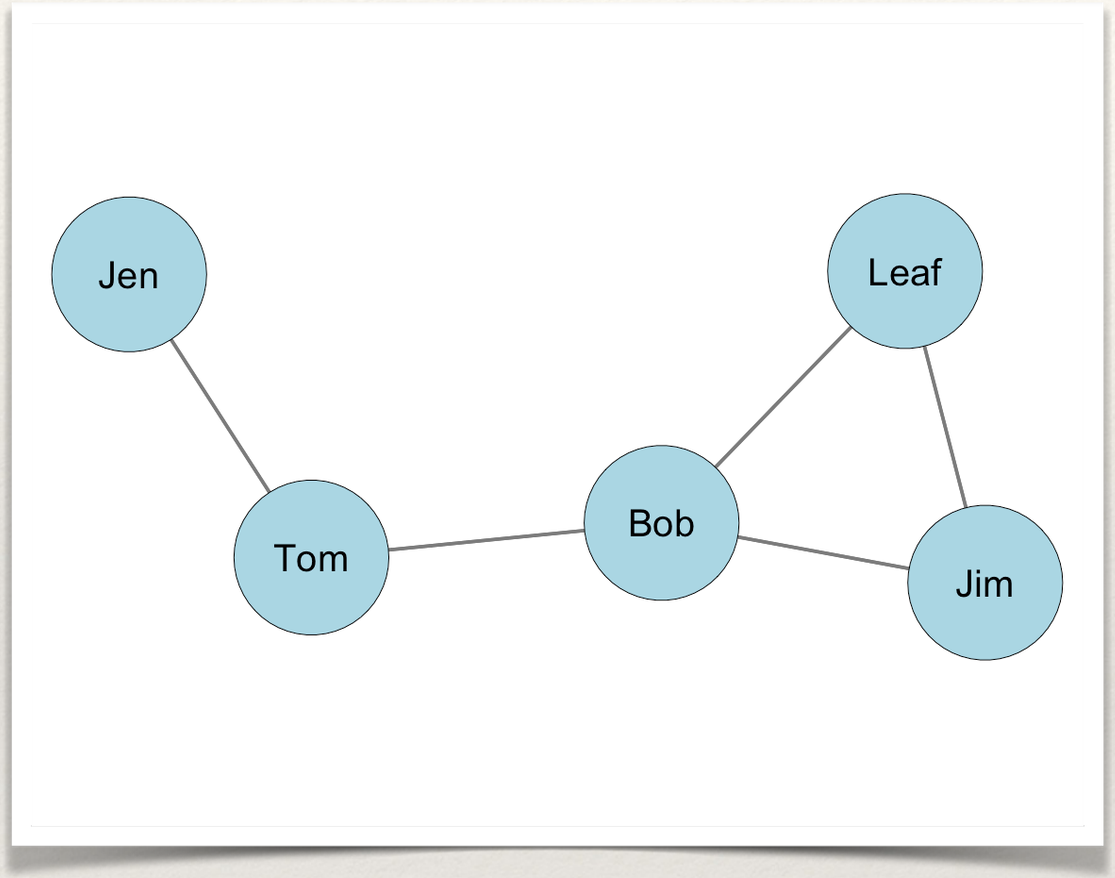
- Actor betweenness centrality not only reflects each node's connectivity to other nodes but also depends on the size of the network, *g*.
- Suppose we wanted to compare across networks of different sizes…
  - *Solution?*

Standardized Betweenness Centrality: Undirected Binary Graphs

$$C'_B(n_i) = \frac{\sum_{j<k} g_{jk}(n_i)/g_{jk}}{[(g-1)(g-2)/2]} = \frac{C_B(n_i)}{[(g-1)(g-2)/2]}$$

The maximum number of pairs of actors not including n_i

Example: Undirected, Binary Network



Unstandardized (raw) for **Bob**: 4

Standardized for **Bob**:
4 / [(5-1)(5-2)/2] = 0.667

Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom			1/1	1/1
Leaf				0/1
Jim				

Group Betweenness Centralization: Undirected Binary Graphs

- We can also summarize the entire network, in terms of how close nodes are to each other.
- *Group betweenness centralization* tells us how much variation there is in the closeness scores.

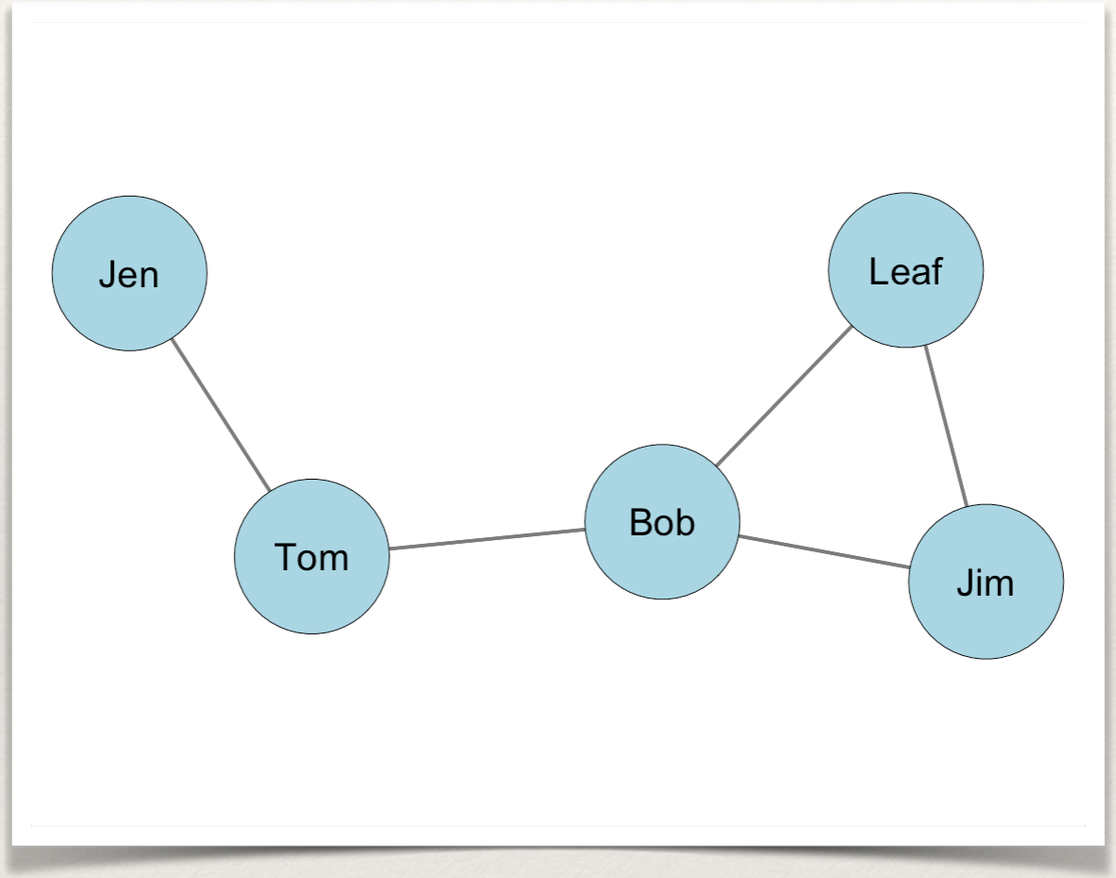
Group Betweenness Centrality: Undirected Binary Graphs

Largest value of the standardized betweenness for the network

Standardized betweenness score for actor i

$$C_B = \frac{\sum_{i=1}^{g} [C'_B(n^*) - C'_B(n_i)]}{(g-1)}$$

Example: Undirected, Binary Network



Standardized Betweenness Centrality Scores

Jen = 0.000

Tom = 0.500

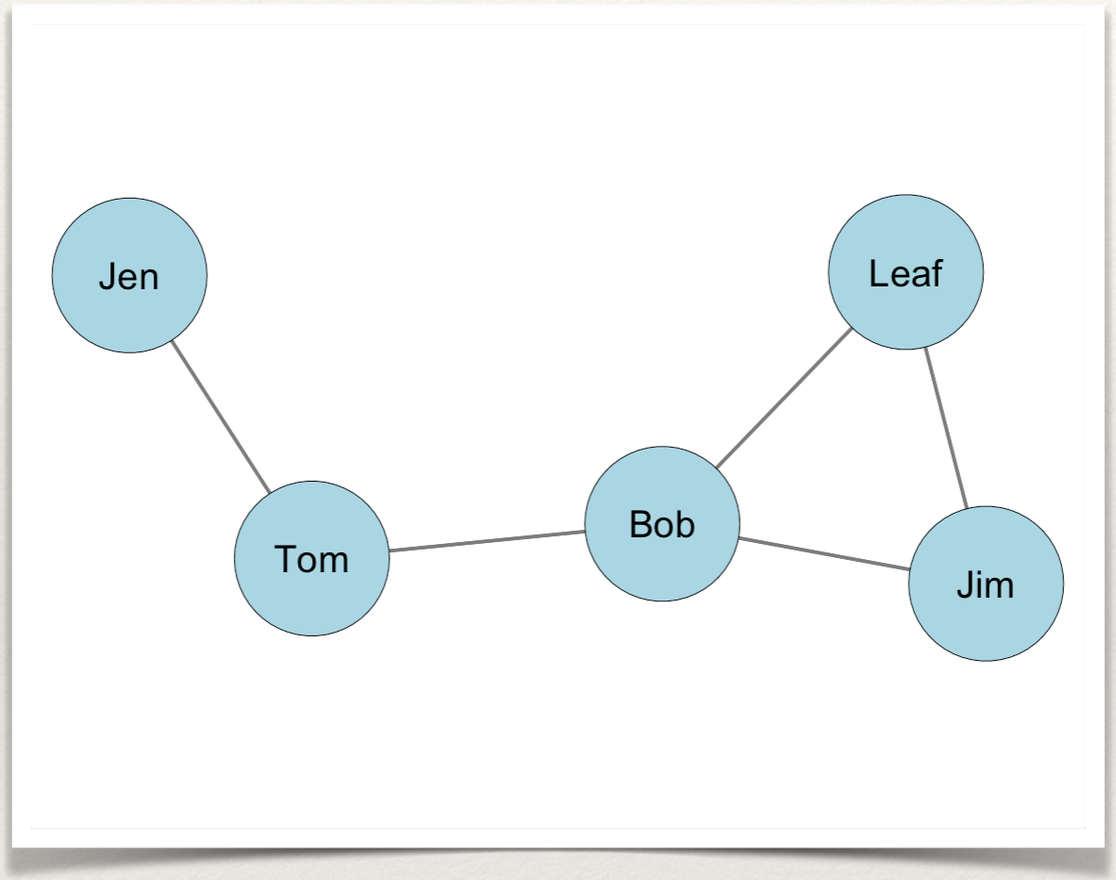
Bob = 0.667

Leaf = 0.000

Jim = 0.000

What is the group betweenness centralization score for this graph?

Example: Undirected, Binary Network



What is the group betweenness centralization score for this graph?

Deviated Standardized Betweenness Centrality Scores

Jen = 0.667 - 0.000 = 0.667

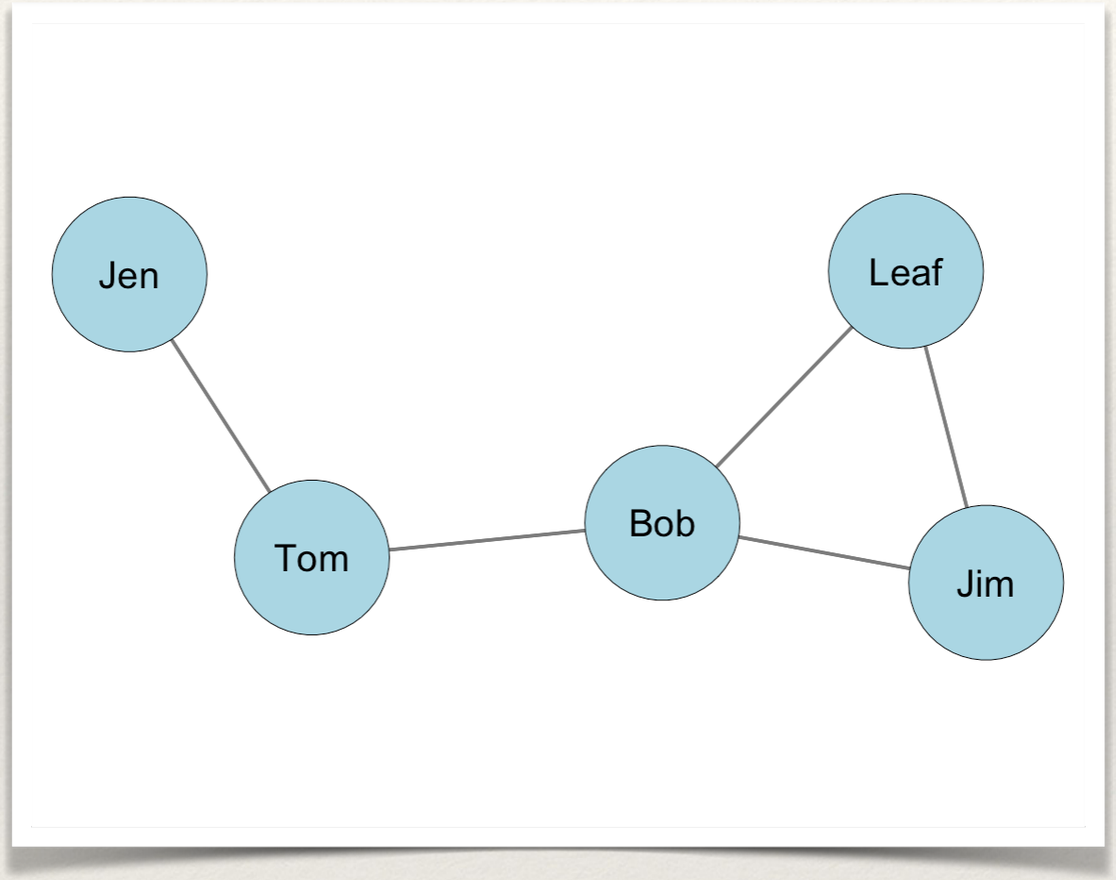
Tom = 0.667 - 0.500 = 0.167

Bob = 0.667 - 0.667 = 0.000

Leaf = 0.667 - 0.000 = 0.667

Jim = 0.667 - 0.000 = 0.667

Example: Undirected, Binary Network



What is the group betweenness centralization score for this graph?

Deviated Standardized Betweenness Centrality Scores

Jen = 0.667 - 0.000 = 0.667

Tom = 0.667 - 0.500 = 0.167

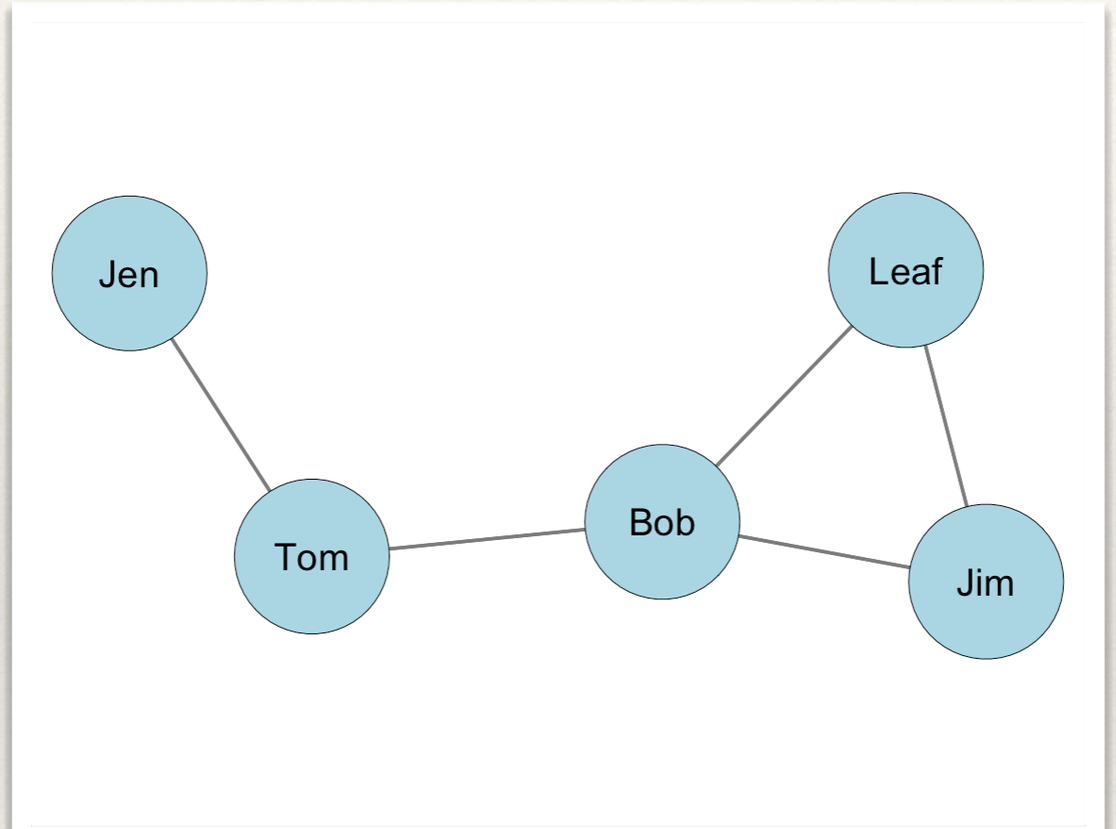
Bob = 0.667 - 0.667 = 0.000

Leaf = 0.667 - 0.000 = 0.667

Jim = 0.667 - 0.000 = 0.667

Sum = 2.168

Example: Undirected, Binary Network



$$C_B = \frac{\sum_{i=1}^{g} [C'_B(n^*) - C'_B(n_i)]}{(g-1)}$$

Deviated Standardized Betweenness Centrality Scores

Jen = 0.667 - 0.000 = 0.667

Tom = 0.667 - 0.500 = 0.167

Bob = 0.667 - 0.667 = 0.000

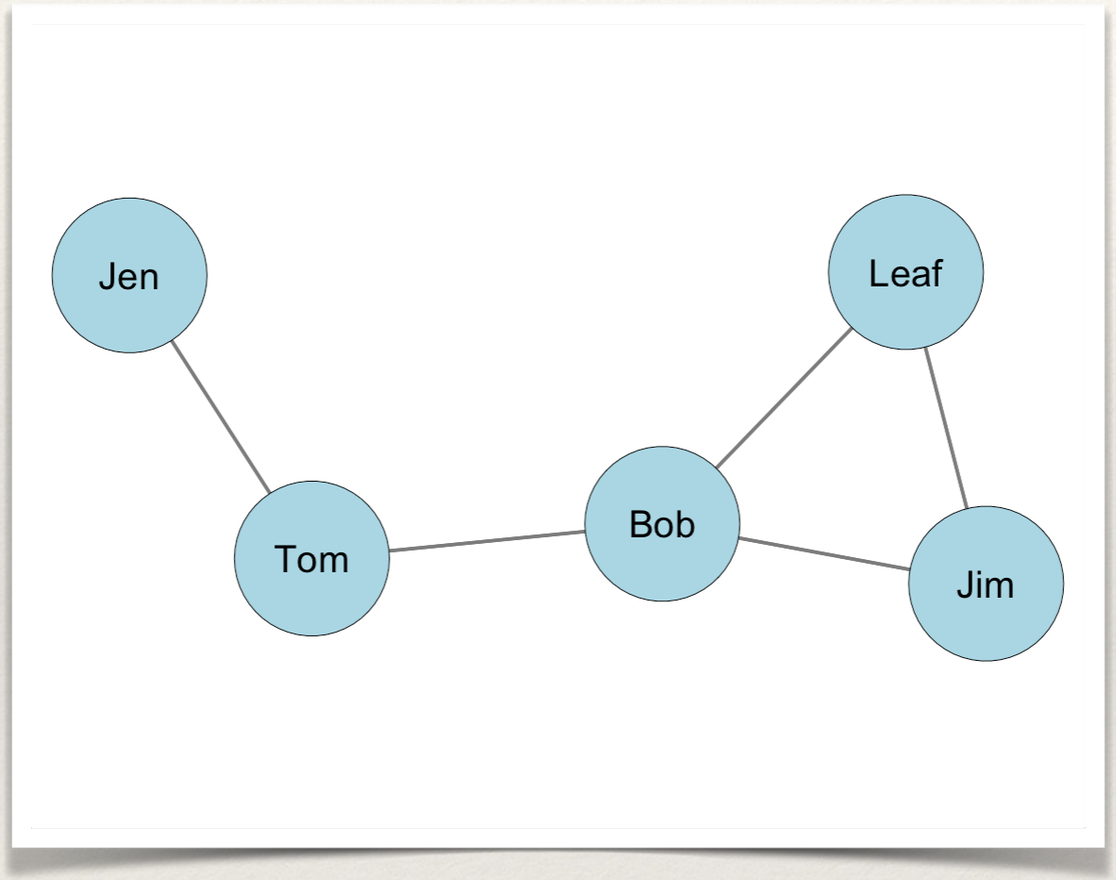
Leaf = 0.667 - 0.000 = 0.667

Jim = 0.667 - 0.000 = 0.667

Sum = 2.168

Denominator = 4

Example: Undirected, Binary Network



2.168 / 4 = **0.542**

Deviated Standardized Betweenness Centrality Scores

Jen = 0.667 - 0.000 = 0.667

Tom = 0.667 - 0.500 = 0.167

Bob = 0.667 - 0.667 = 0.000

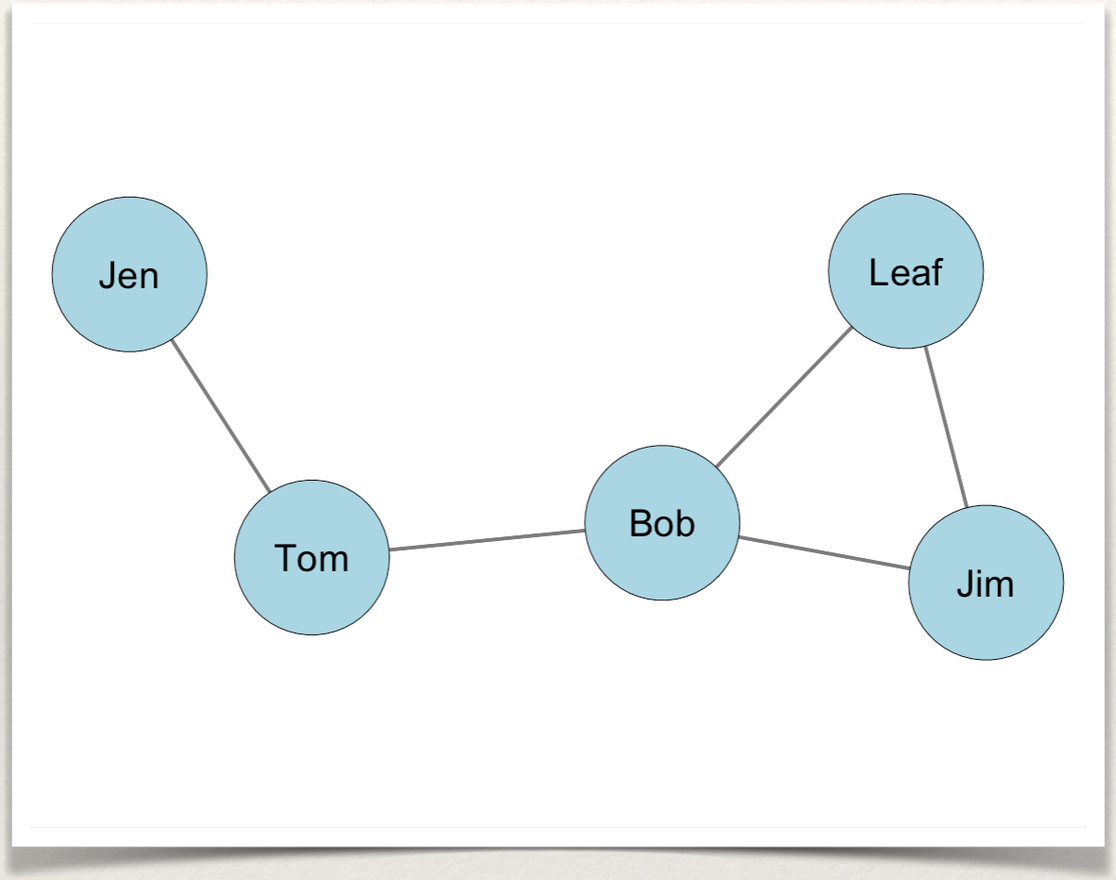
Leaf = 0.667 - 0.000 = 0.667

Jim = 0.667 - 0.000 = 0.667

Sum = 2.168

Denominator = 4

Example: Undirected, Binary Network



Compare the centralization scores:

Degree = 0.416
Closeness = 0.551
Betweenness = 0.542

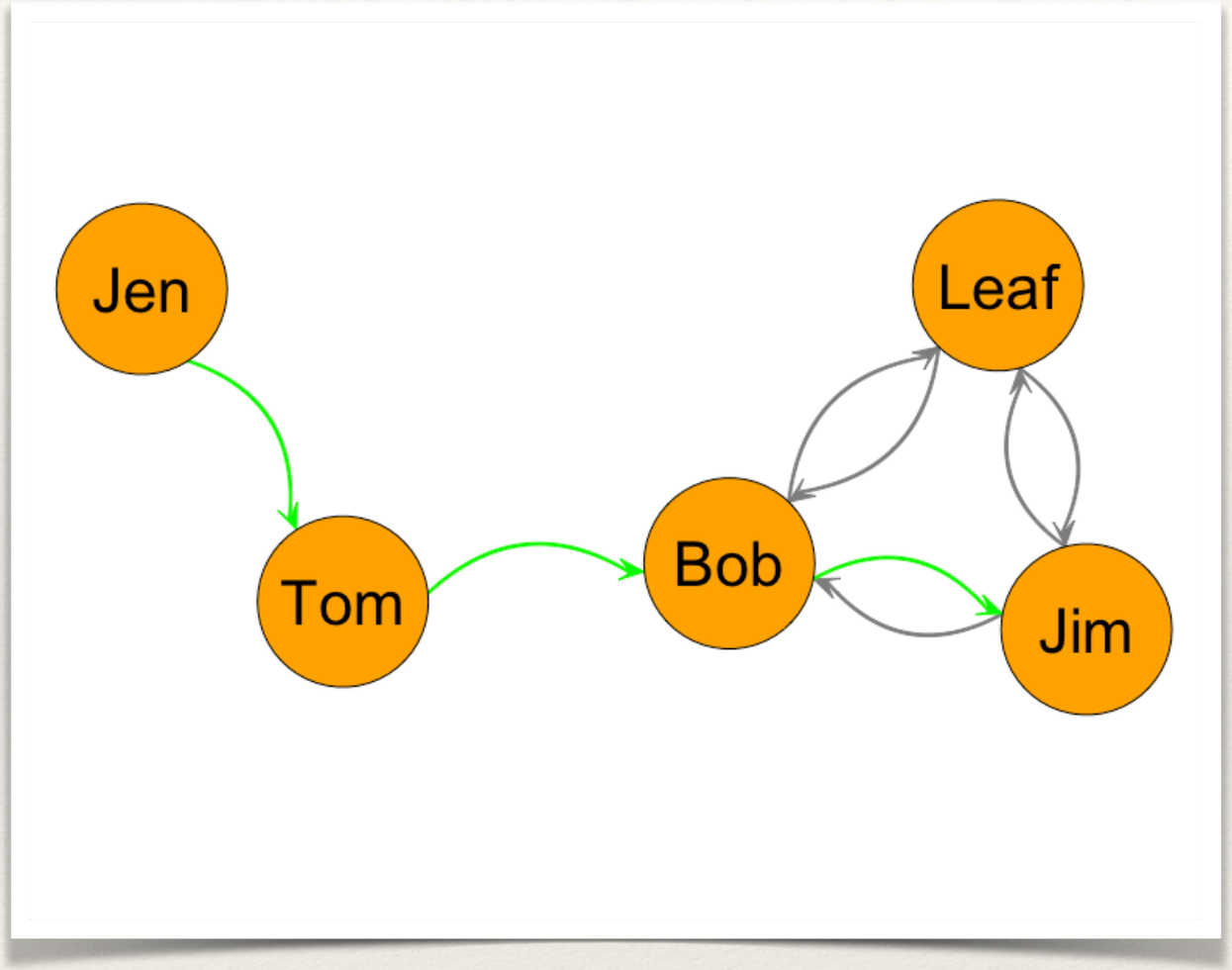
What can we say about the differences in the centralization scores for each type of centrality?

Directed Networks

Betweenness Centrality: Directed Binary Graphs

- Recall that in a directed network, the directionality matters.
- As a result, we have to consider how this might influence our measures.

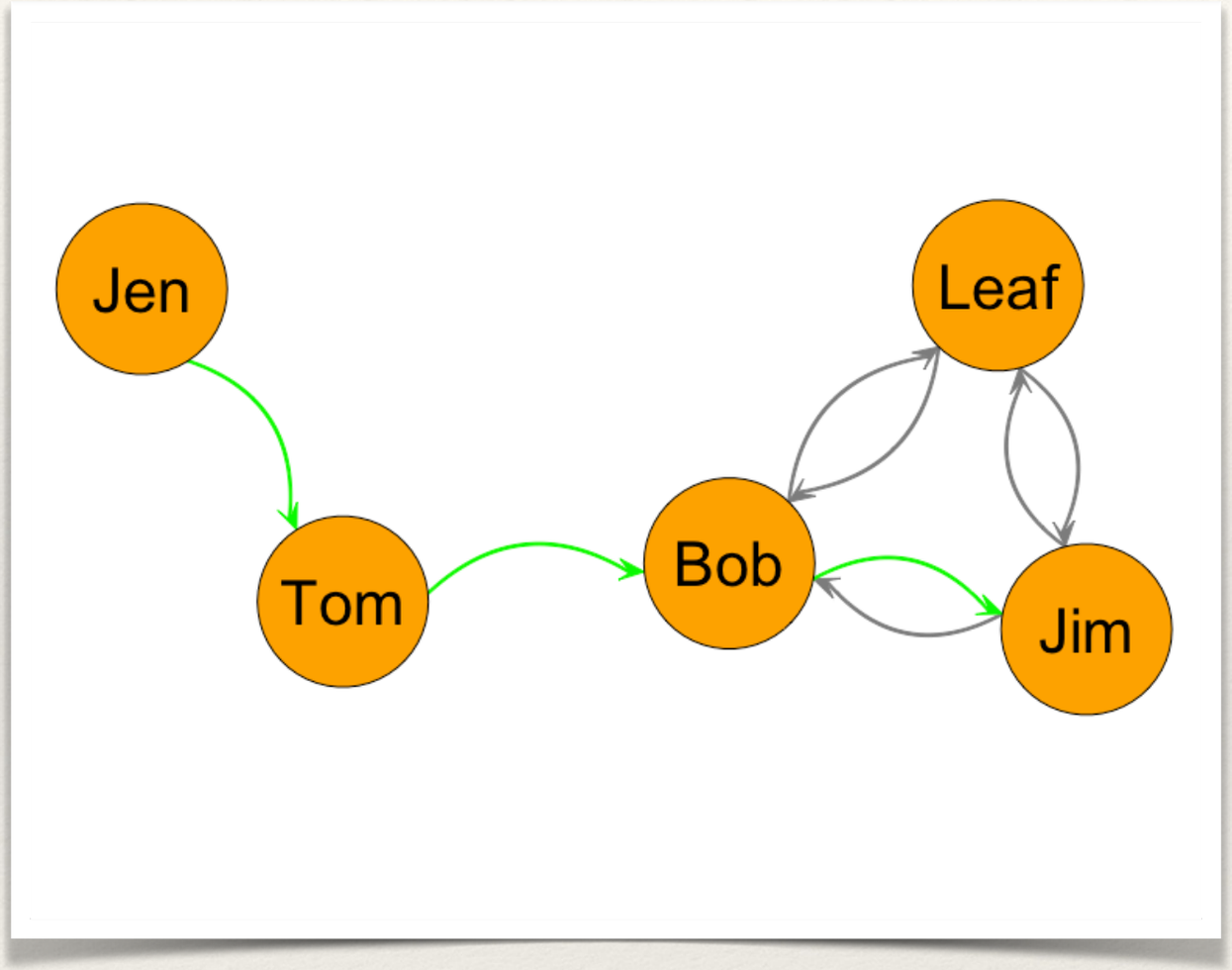
Example: Betweenness Centrality for Directed Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen				/?
Tom				
Leaf				
Jim				

As with closeness, we look at the rows

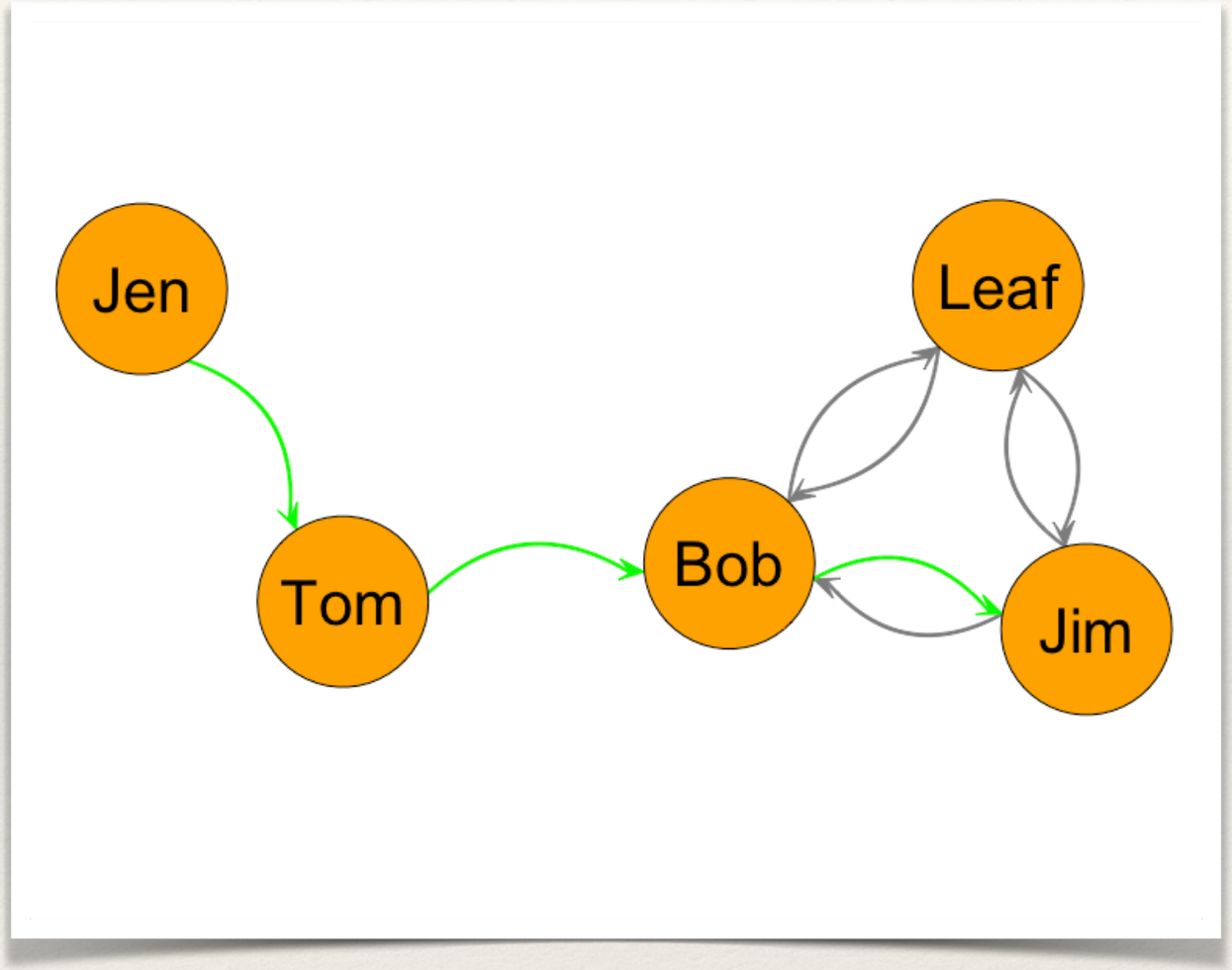
Example: Betweenness Centrality for Directed Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen				/?
Tom				
Leaf				
Jim				

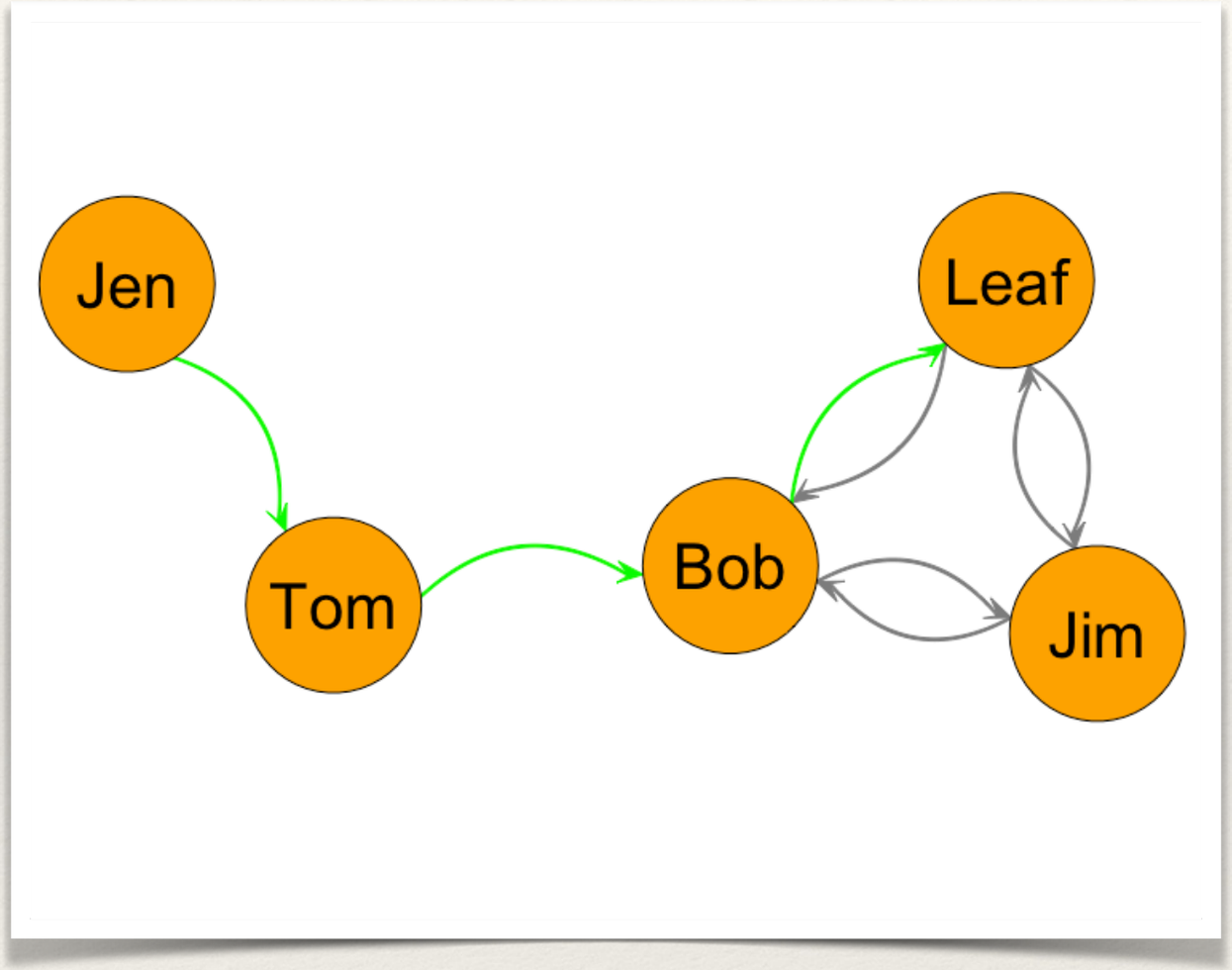
The rows show sending behavior

Example: Betweenness Centrality for Directed Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen				1/1
Tom				
Leaf				
Jim				

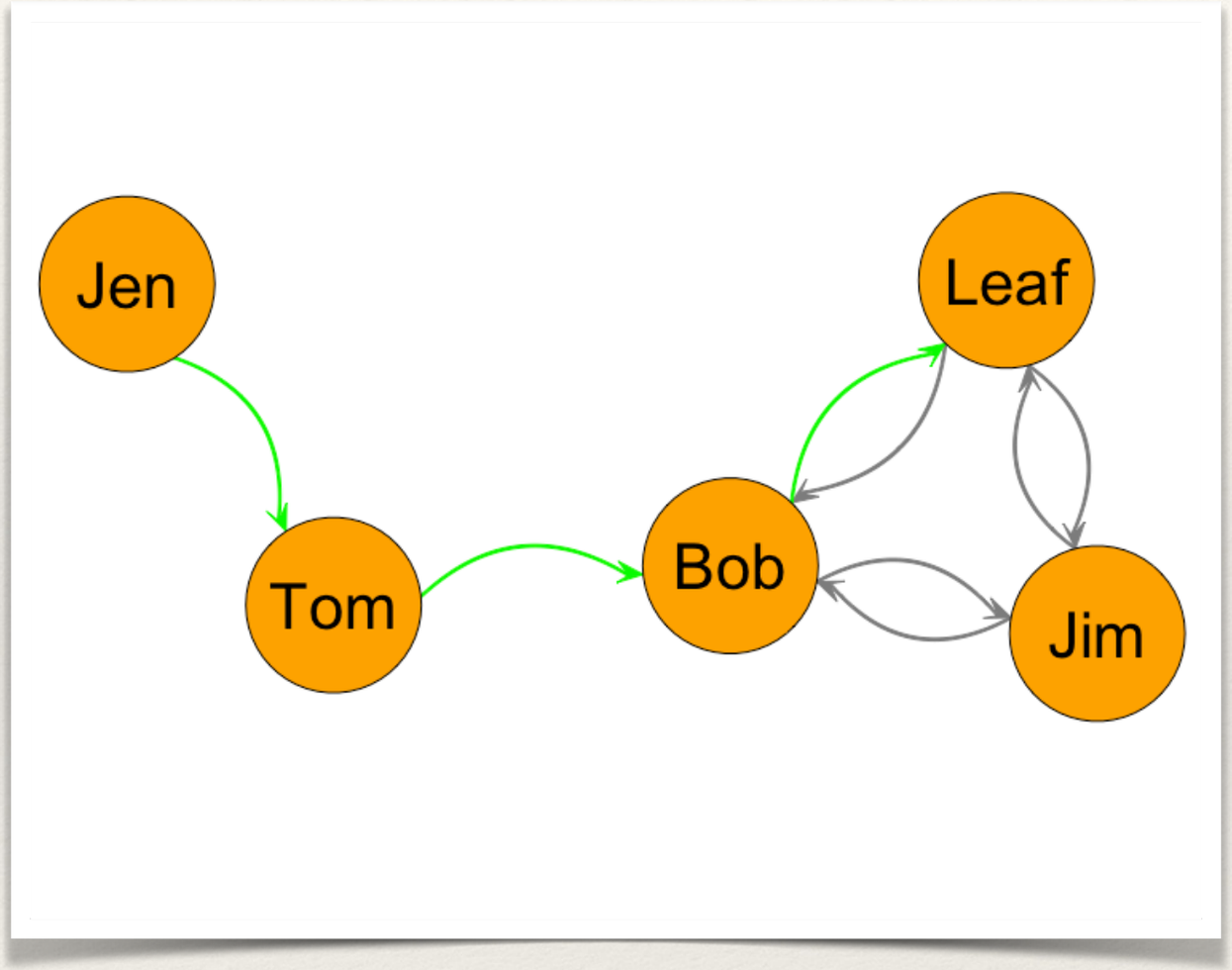
Example: Betweenness Centrality for Directed Binary Network



Geodesic Proportions for *Bob*

	Jen	*Tom*	*Leaf*	*Jim*
Jen			?	1/1
Tom				
Leaf				
Jim				

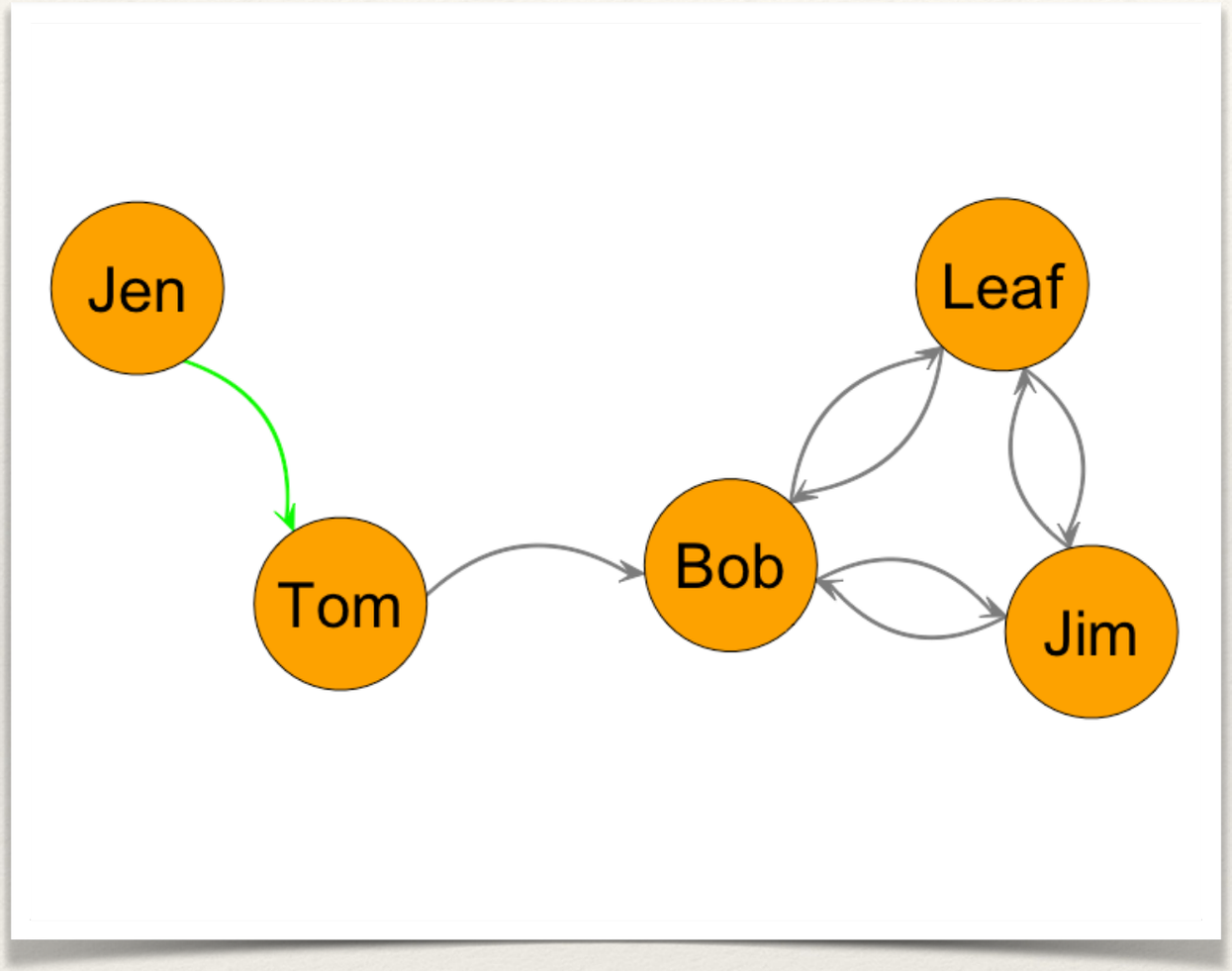
Example: Betweenness Centrality for Directed Binary Network



Geodesic Proportions for *Bob*

	Jen	*Tom*	*Leaf*	*Jim*
Jen			1/1	1/1
Tom				
Leaf				
Jim				

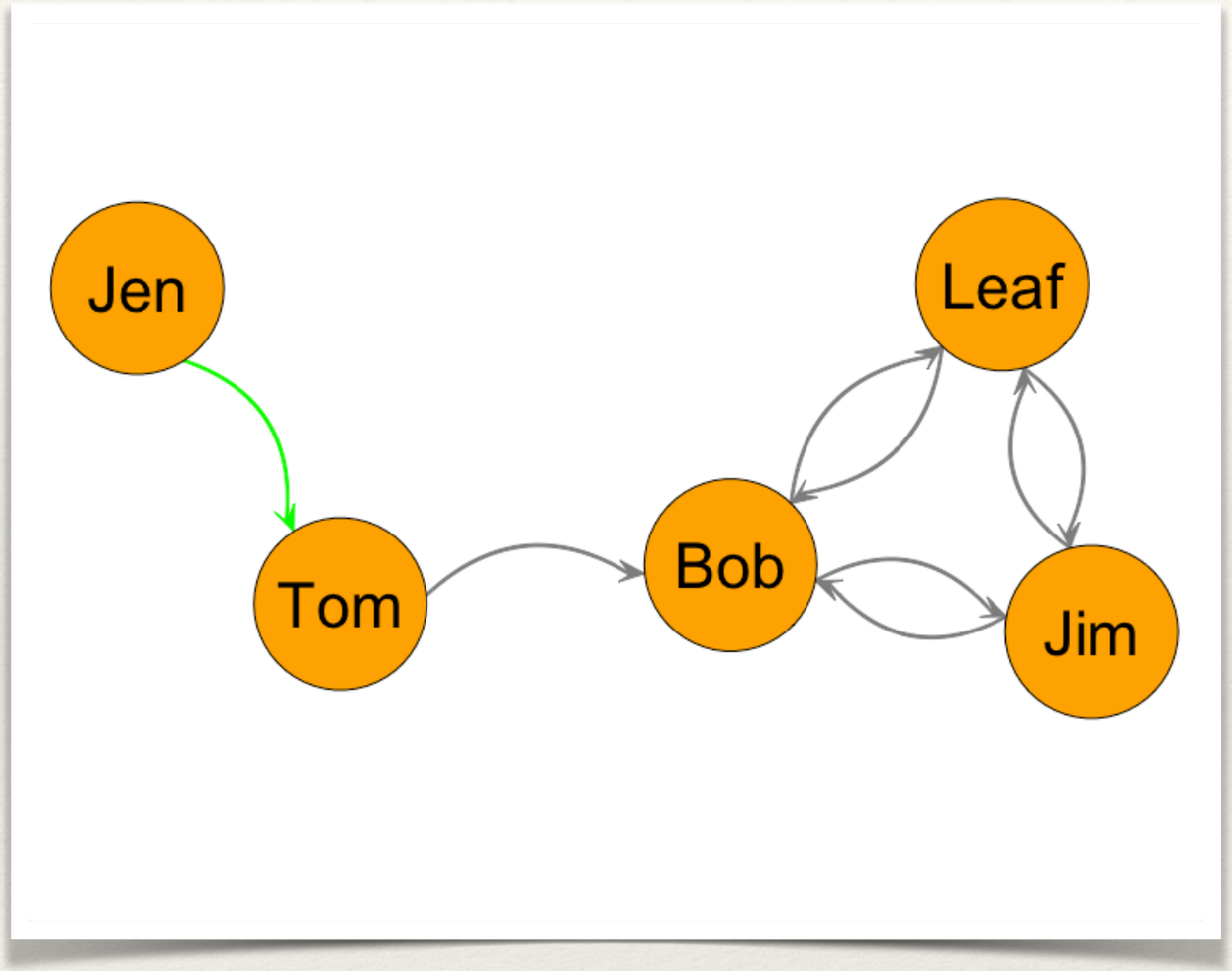
Example: Betweenness Centrality for Directed Binary Network



Geodesic Proportions for *Bob*

	Jen	*Tom*	*Leaf*	*Jim*
Jen		?	1/1	1/1
Tom				
Leaf				
Jim				

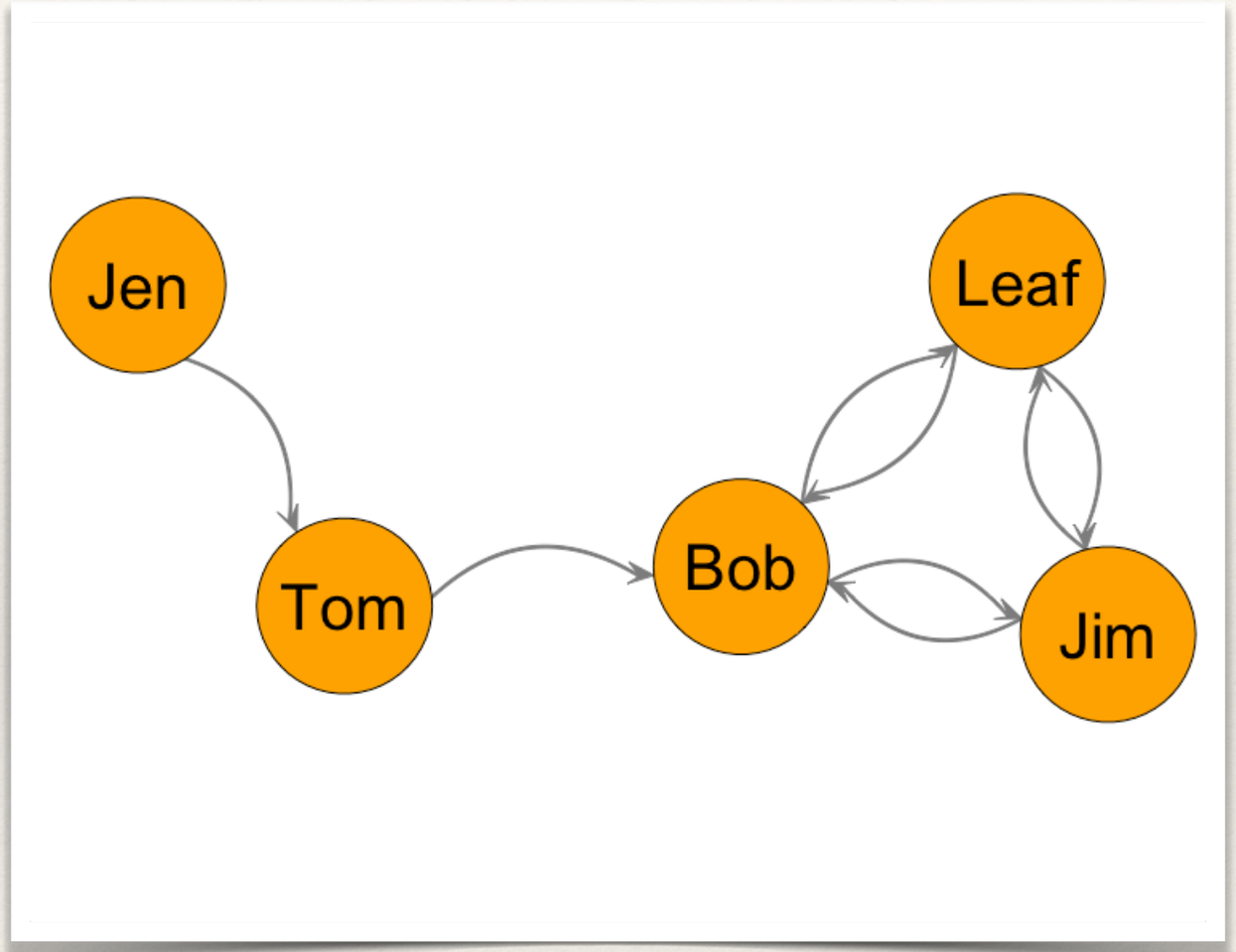
Example: Betweenness Centrality for Directed Binary Network



Geodesic Proportions for *Bob*

	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom				
Leaf				
Jim				

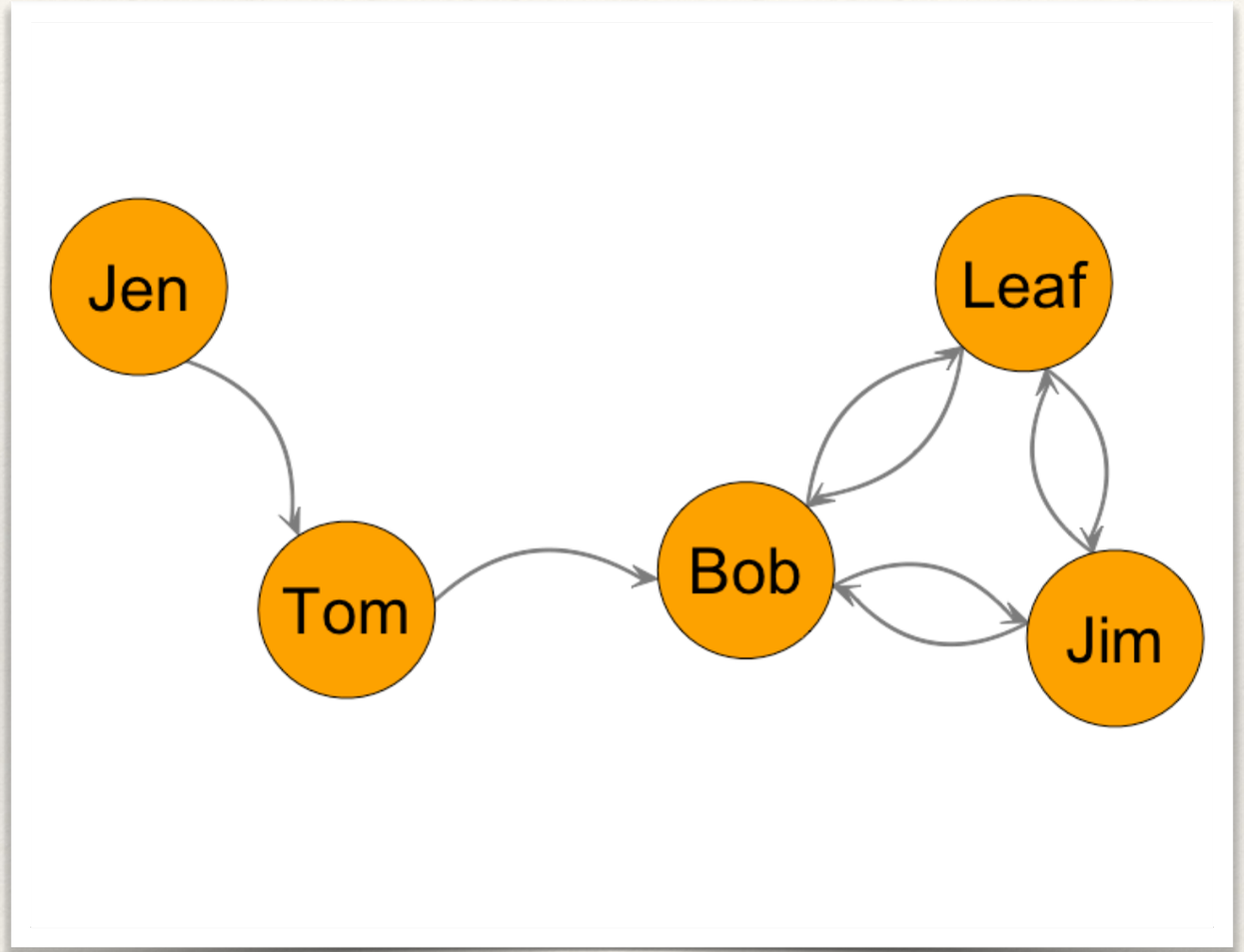
Example: Betweenness Centrality for Directed Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom	0/0		1/1	1/1
Leaf	0/0	0/0		0/1
Jim	0/0	0/0	0/1	

Then, we just
complete the matrix.

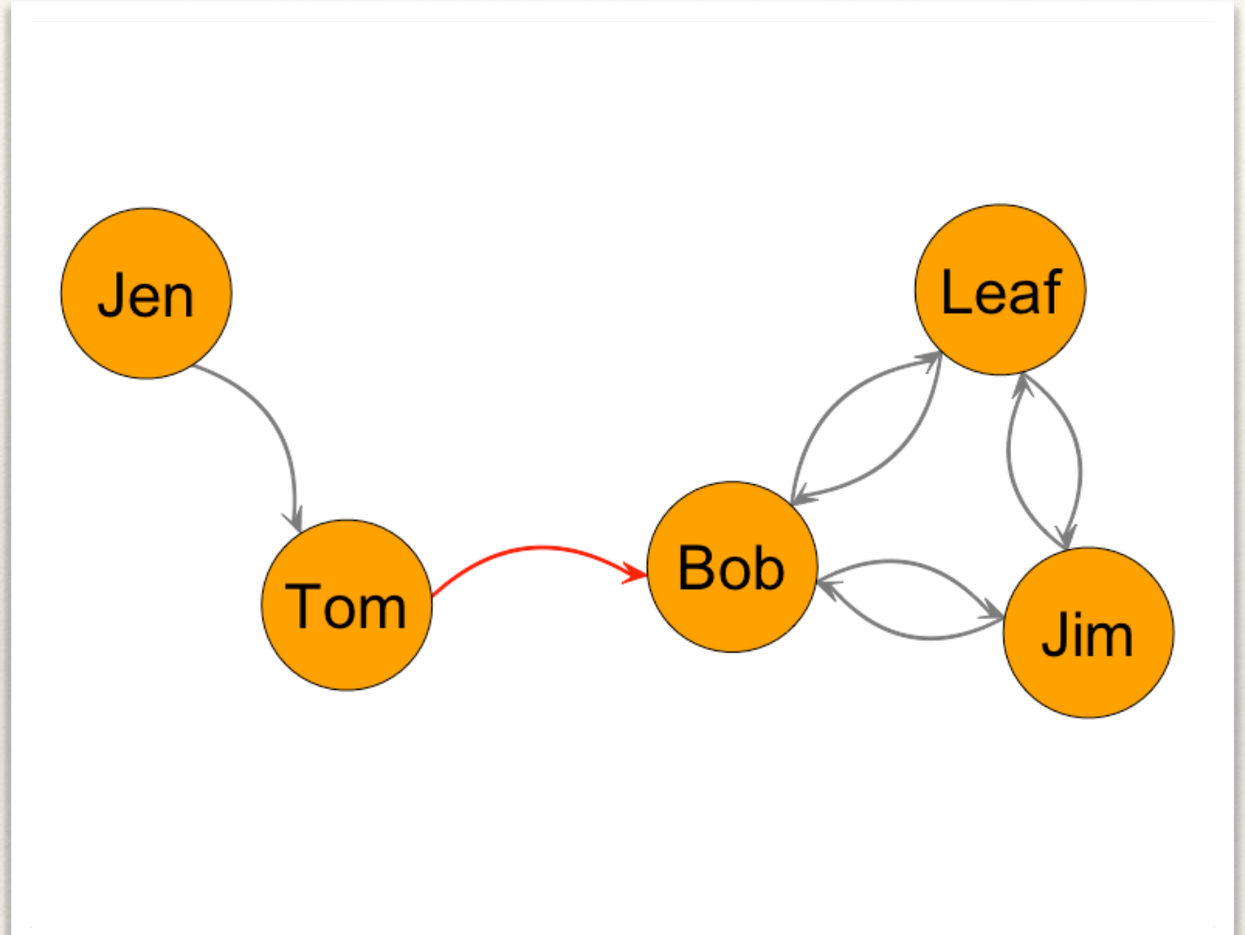
Example: Betweenness Centrality for Directed Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom	0/0		1/1	1/1
Leaf	0/0	0/0		0/1
Jim	0/0	0/0	0/1	

Why no geodesics?

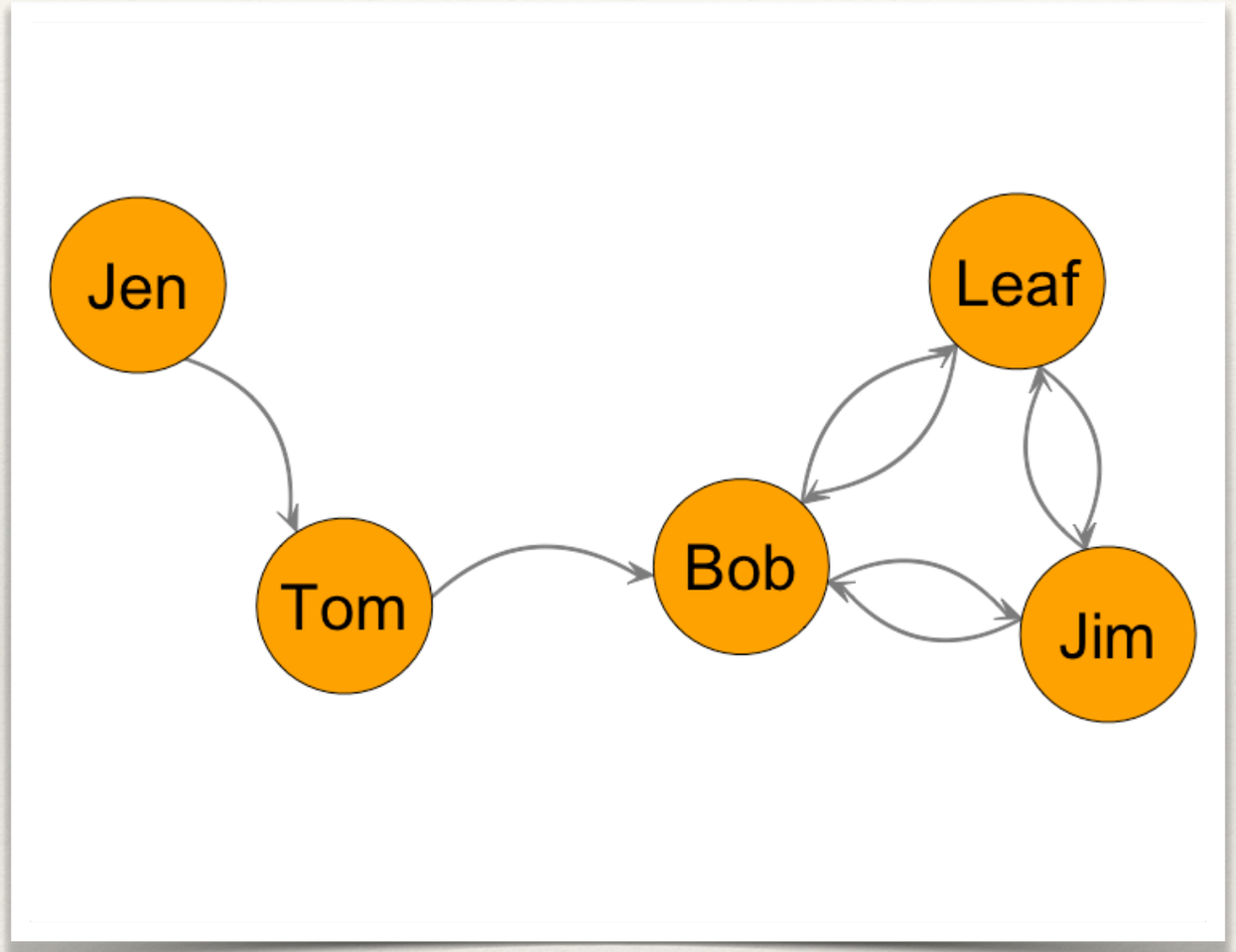
Example: Betweenness Centrality for Directed Binary Network



Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom	0/0		1/1	1/1
Leaf	0/0	0/0		0/1
Jim	0/0	0/0	0/1	

Tom and **Jen** cannot be reached past **Bob**, because there are no outgoing ties from **Bob**.

Example: Betweenness Centrality for Directed Binary Network

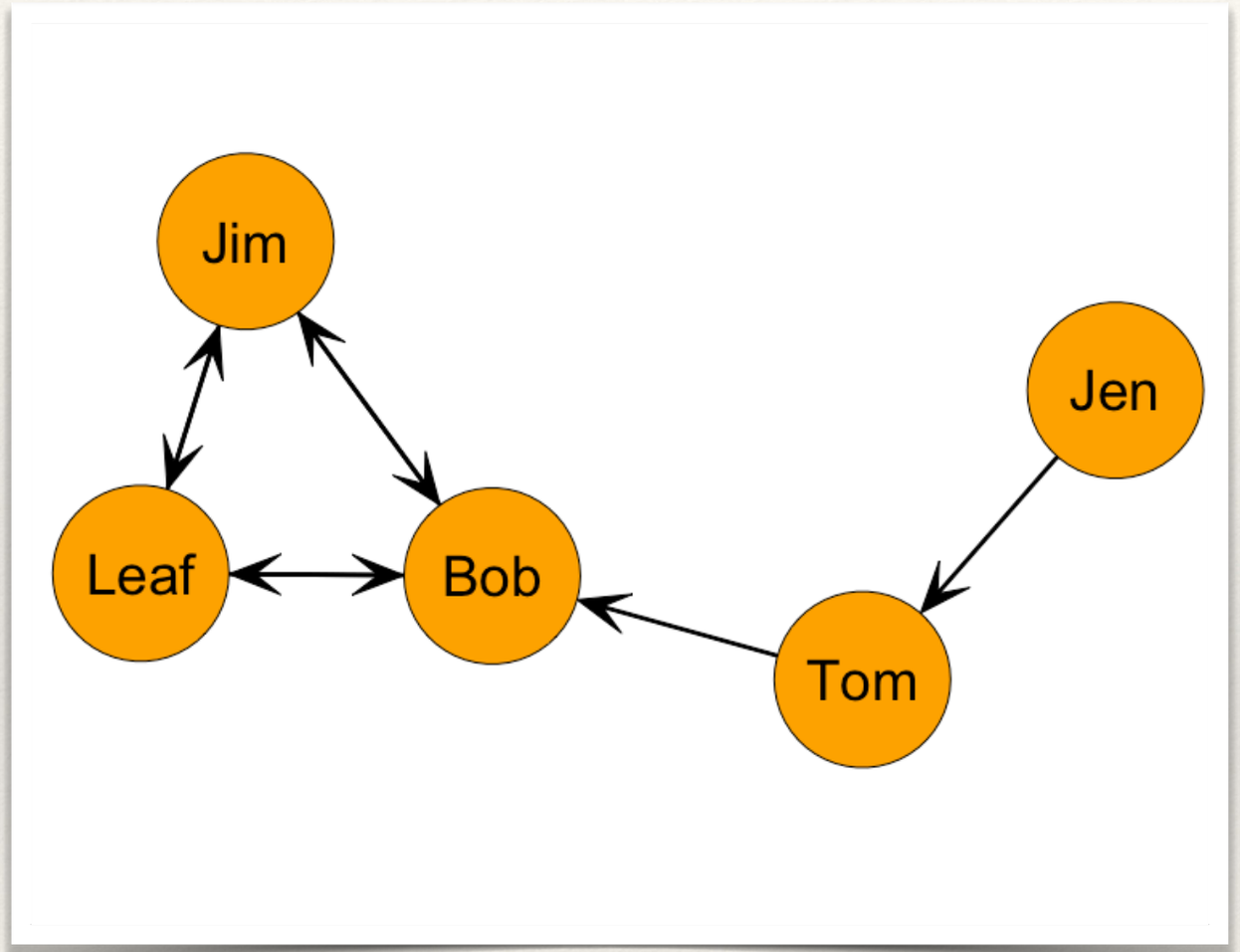


Unstandardized (raw) for **Bob**: 4

Standardized for **Bob**:
4 / [(5-1)(5-2)/2] = 0.667

Geodesic Proportions for *Bob*				
	Jen	*Tom*	*Leaf*	*Jim*
Jen		0/1	1/1	1/1
Tom	0/0		1/1	1/1
Leaf	0/0	0/0		0/1
Jim	0/0	0/0	0/1	

Example: Undirected, Binary Network



Compare the centralization scores:

Indegree = 0.438
Outdegree = 0.125
Closeness = 0.555
Betweenness = 0.270

What can we say about the differences in the centralization scores for each type of centrality?

Comparing Measures of Centrality

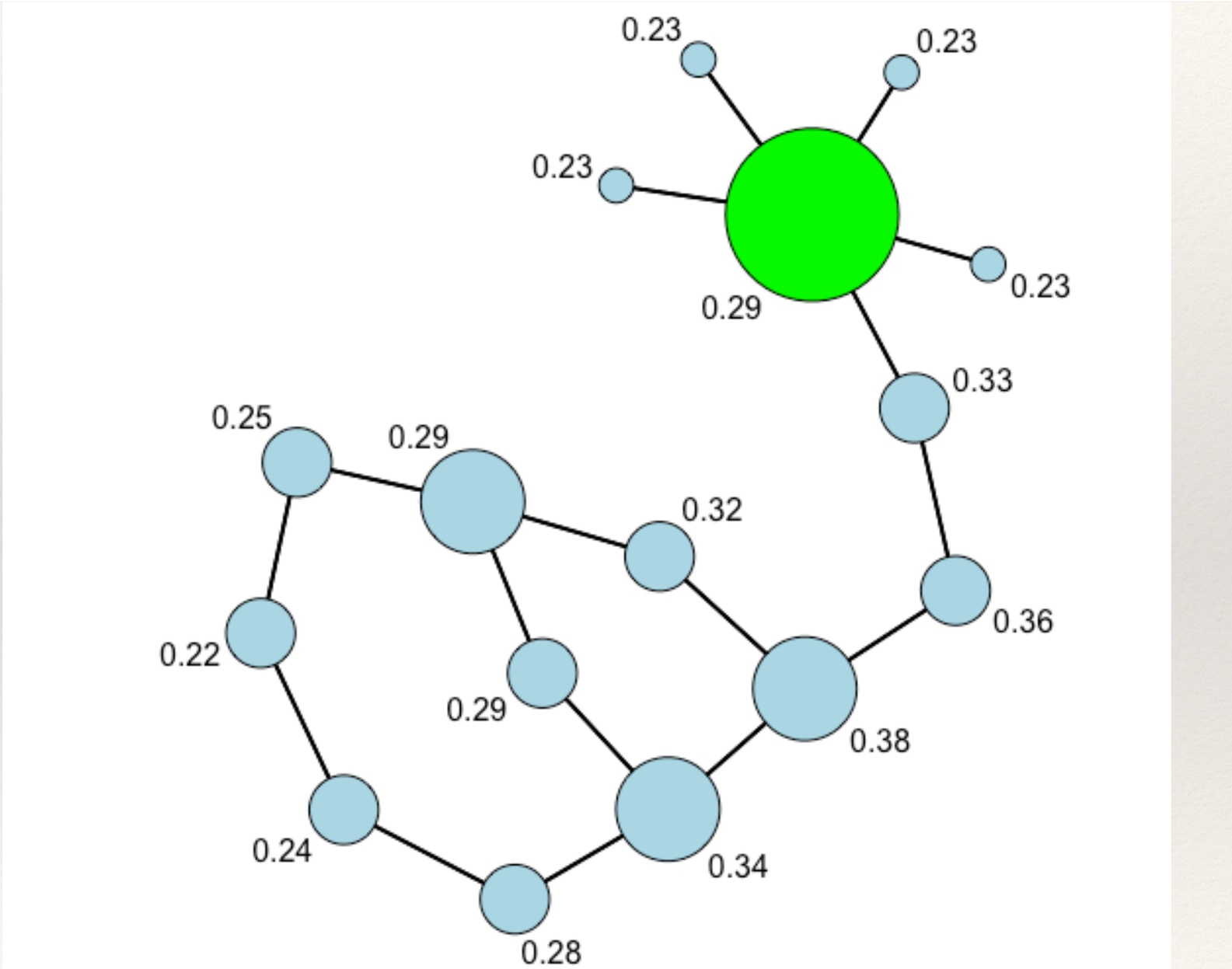
Comparing Measures of Centrality

	Low Degree	Low Closeness	Low Betweenness
High Degree		Embedded in cluster that is far from the rest of the network	Ego's connections are redundant - communication bypasses him/her
High Closeness	Key player tied to important/active alters		Probably multiple paths in the network, ego is near many people, but so are others
High Betweenness	Ego's few ties are crucial for network flow	Very rare. Would mean that ego monopolizes the ties from a small number of people to many others	

Comparing Measures of Centrality

	Low Degree	Low Closeness	Low Betweenness
High Degree		Embedded in cluster that is far from the rest of the network	Ego's connections are redundant - communication bypasses him/her
High Closeness	Key player tied to important/active alters		Probably multiple paths in the network, ego is near many people, but so are others
High Betweenness	Ego's few ties are crucial for network flow	Very rare. Would mean that ego monopolizes the ties from a small number of people to many others	

High Degree, Low Closeness



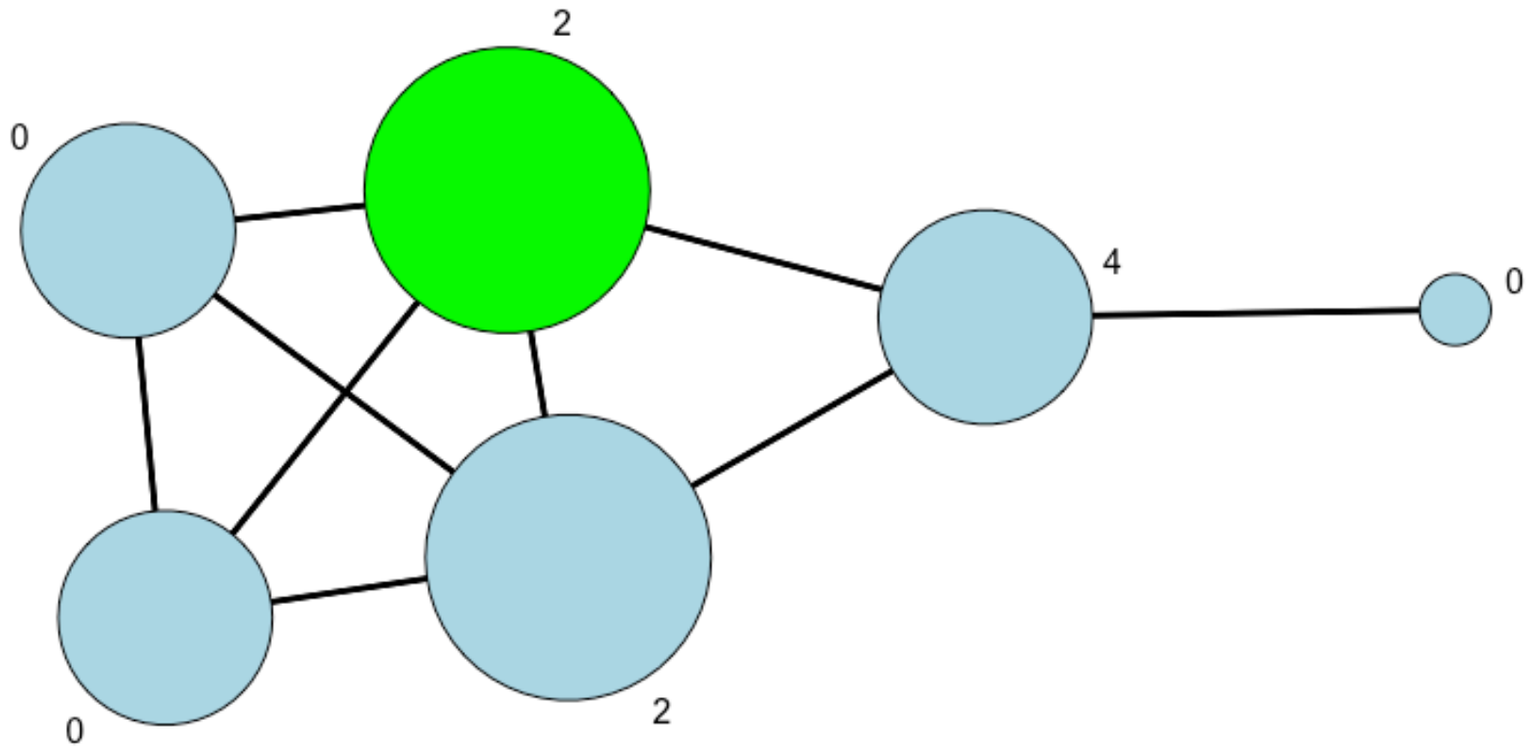
Nodes sized by degree

Nodes labeled by closeness centrality

Comparing Measures of Centrality

	Low Degree	Low Closeness	Low Betweenness
High Degree		Embedded in cluster that is far from the rest of the network	Ego's connections are redundant - communication bypasses him/her
High Closeness	Key player tied to important/active alters		Probably multiple paths in the network, ego is near many people, but so are others
High Betweenness	Ego's few ties are crucial for network flow	Very rare. Would mean that ego monopolizes the ties from a small number of people to many others	

High Degree, Low Betweenness



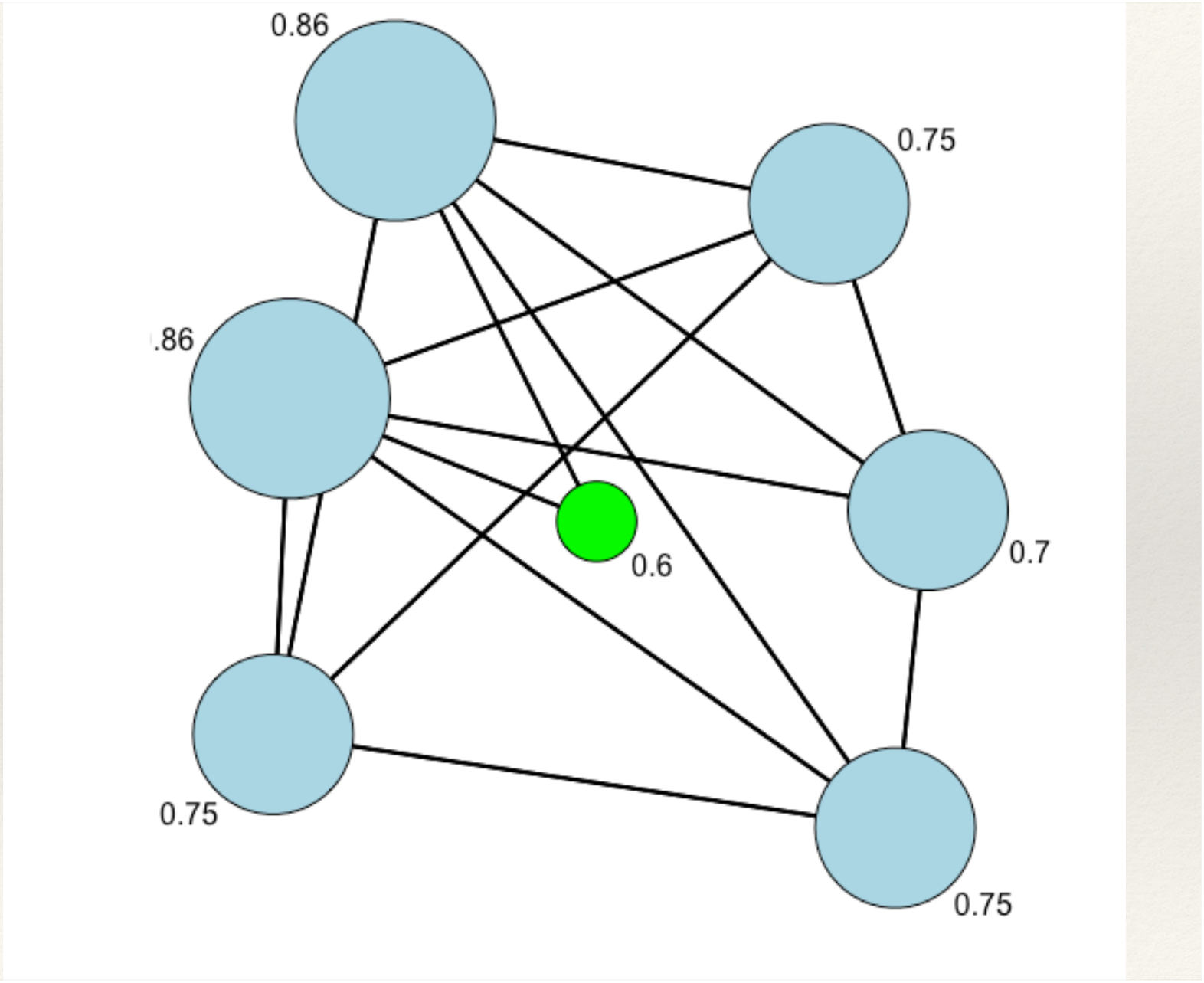
Nodes sized by degree

Nodes labeled by betweenness centrality

Comparing Measures of Centrality

	Low Degree	Low Closeness	Low Betweenness
High Degree		Embedded in cluster that is far from the rest of the network	Ego's connections are redundant - communication bypasses him/her
High Closeness	Key player tied to important/active alters		Probably multiple paths in the network, ego is near many people, but so are others
High Betweenness	Ego's few ties are crucial for network flow	Very rare. Would mean that ego monopolizes the ties from a small number of people to many others	

High Closeness, Low Degree



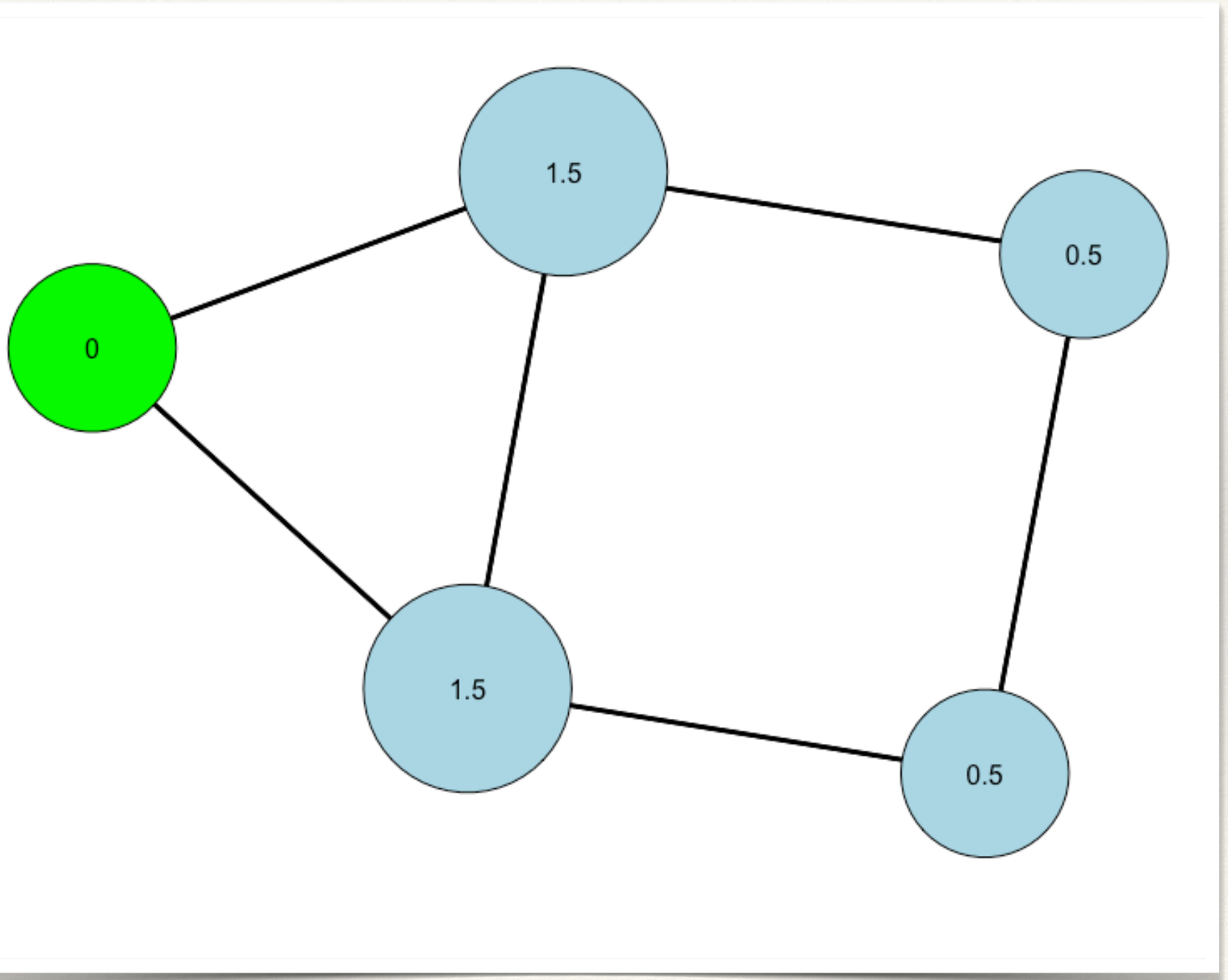
Nodes sized by degree centrality

Nodes labeled by closeness centrality

Comparing Measures of Centrality

	Low Degree	Low Closeness	Low Betweenness
High Degree		Embedded in cluster that is far from the rest of the network	Ego's connections are redundant - communication bypasses him/her
High Closeness	Key player tied to important/active alters		Probably multiple paths in the network, ego is near many people, but so are others
High Betweenness	Ego's few ties are crucial for network flow	Very rare. Would mean that ego monopolizes the ties from a small number of people to many others	

High Closeness, Low Betweenness



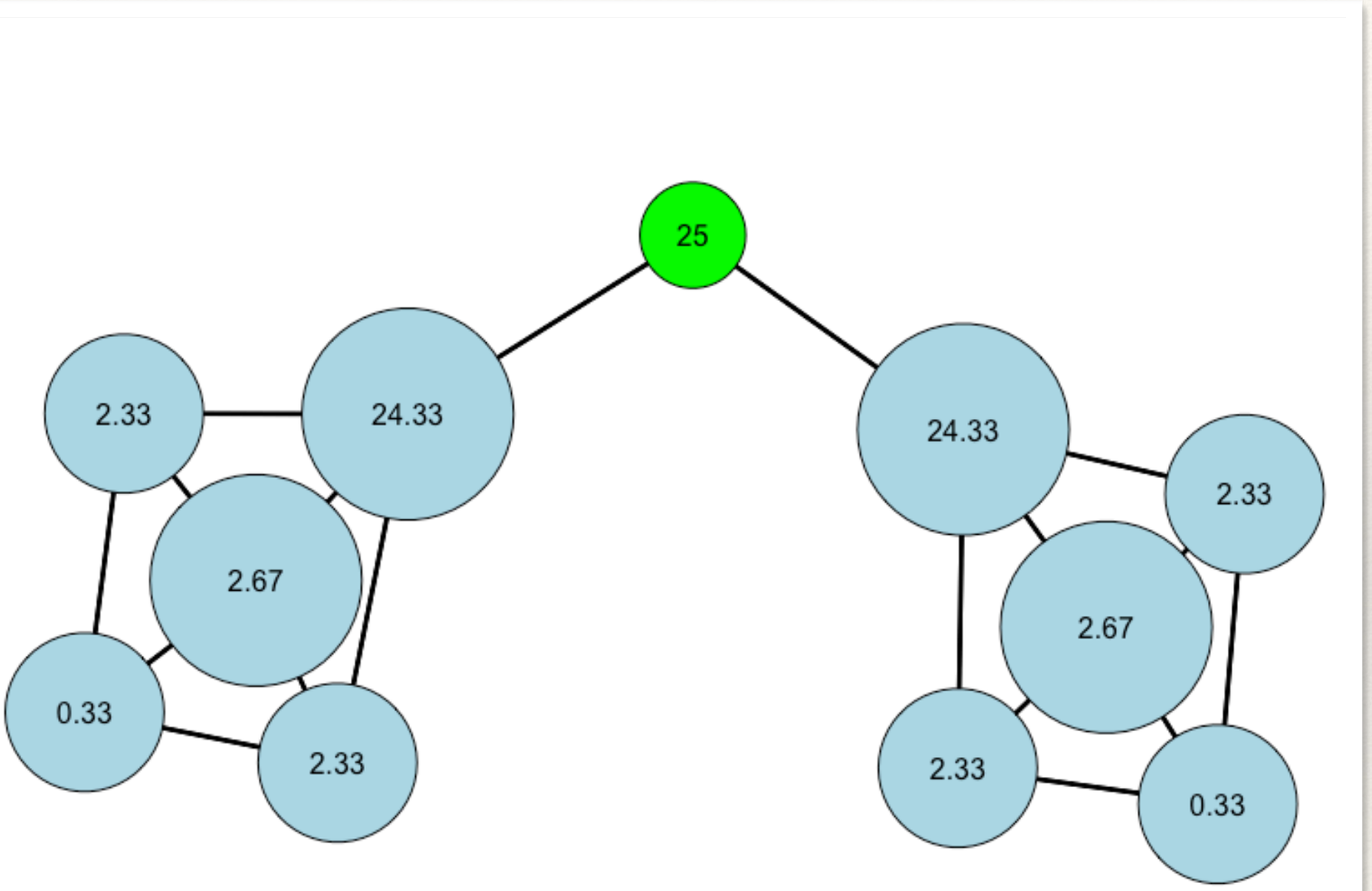
Nodes sized by closeness centrality

Nodes labeled by betweenness centrality

Comparing Measures of Centrality

	Low Degree	Low Closeness	Low Betweenness
High Degree		Embedded in cluster that is far from the rest of the network	Ego's connections are redundant - communication bypasses him/her
High Closeness	Key player tied to important/active alters		Probably multiple paths in the network, ego is near many people, but so are others
High Betweenness	Ego's few ties are crucial for network flow	Very rare. Would mean that ego monopolizes the ties from a small number of people to many others	

High Betweenness, Low Degree



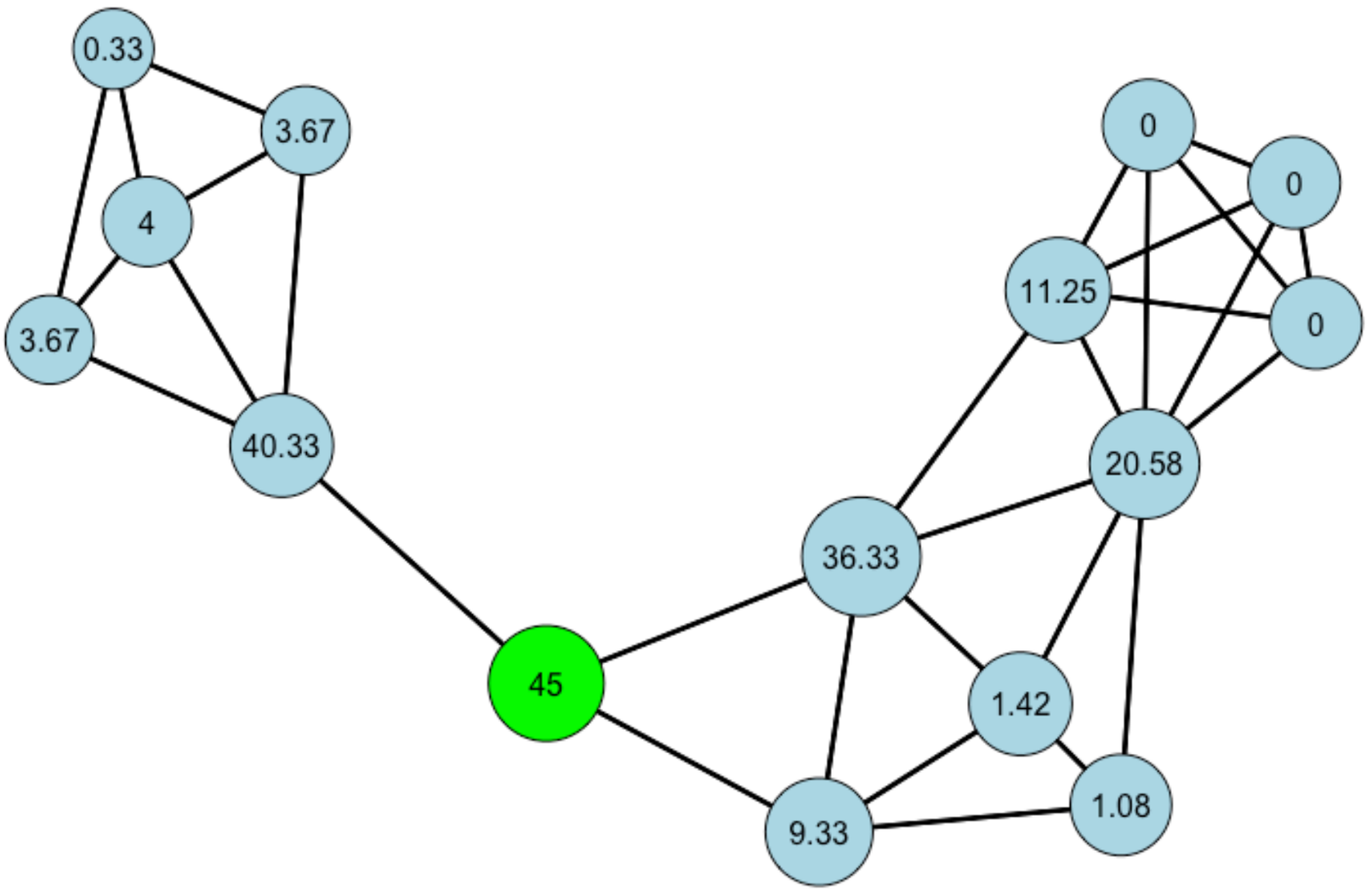
Nodes sized by degree

Nodes labeled by betweenness centrality

Comparing Measures of Centrality

	Low Degree	Low Closeness	Low Betweenness
High Degree		Embedded in cluster that is far from the rest of the network	Ego's connections are redundant - communication bypasses him/her
High Closeness	Key player tied to important/active alters		Probably multiple paths in the network, ego is near many people, but so are others
High Betweenness	Ego's few ties are crucial for network flow	Very rare. Would mean that ego monopolizes the ties from a small number of people to many others	

High Betweenness, Low Closeness



Nodes sized by closeness centrality

Nodes labeled by betweenness centrality

Learning Goals

- At the end of the lecture, you should be able to answer these questions:
  - What are some different ways we can conceptualize "centrality"?
  - What is *closeness* and *betweenness* centrality?
  - How do we calculate these measures for undirected and directed graphs?
  - What do comparing these measures tell us about the structure of a network?

Questions?