Statistical Analysis of Networks

Bipartite Graphs/ Two-Mode Networks

Motivating Example

Journal of Contemporary Criminal Justice
2015, Vol. 31(3) 243–261

Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1043986214553380
ccj.sagepub.com
SAGE

Diffusion of Ideas and Technology: The Role of Networks in Influencing the Endorsement and Use of On-Officer Video Cameras

Jacob T. N. Young[1] and Justin T. Ready[1]

- <u>Questions</u>:
  - How do police officers “frame” body-worn cameras?
  - Is the meaning officers attribute to cameras created and transmitted in groups?

Empirical Example

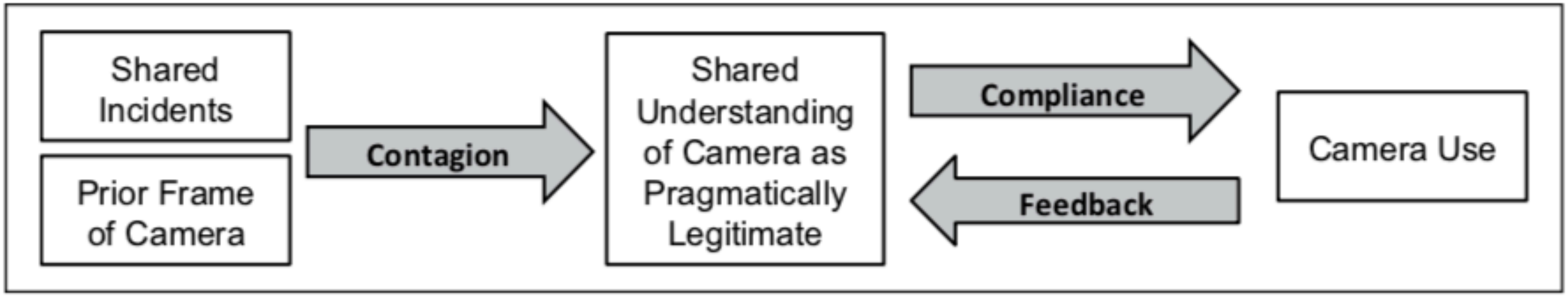
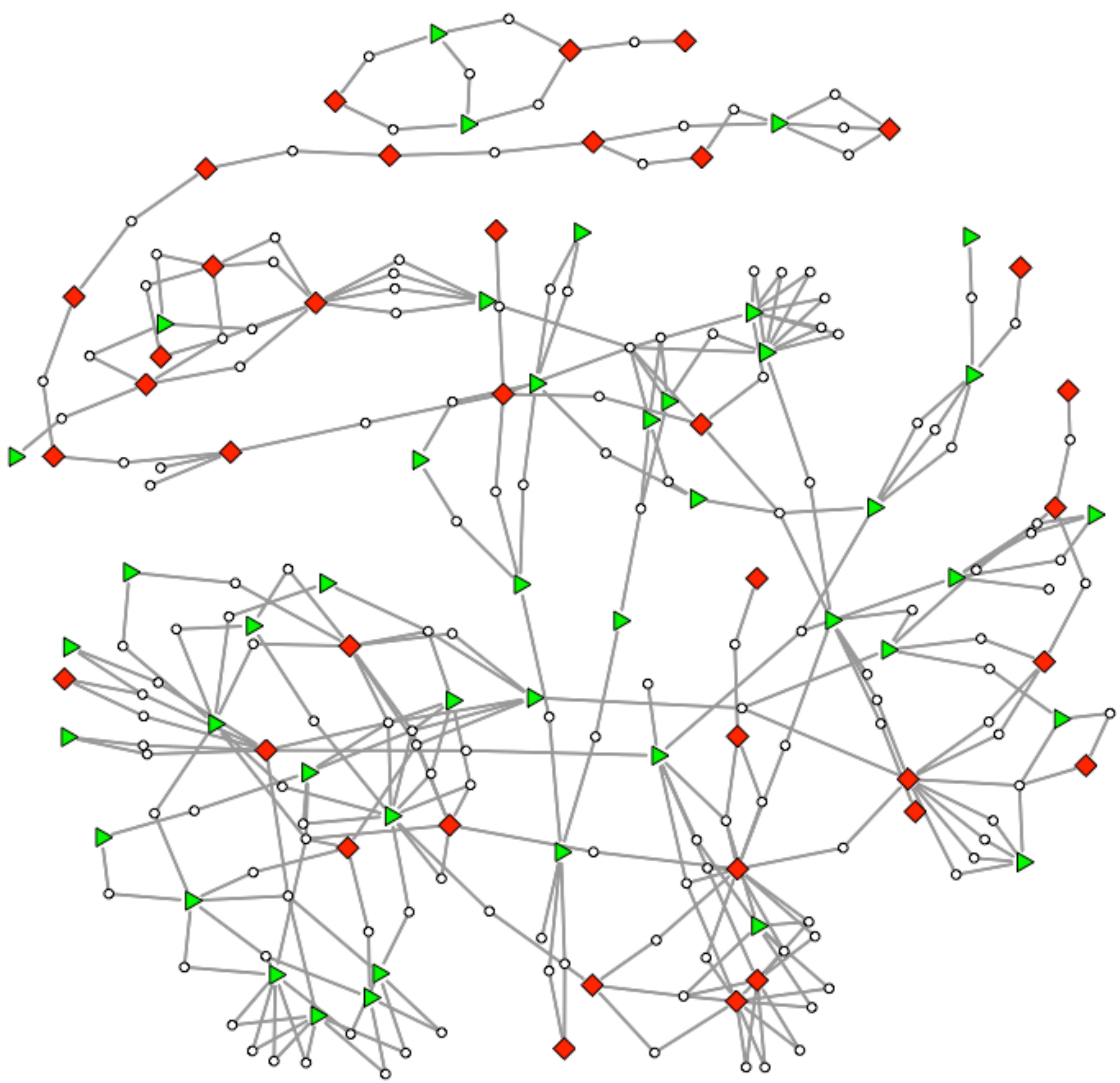


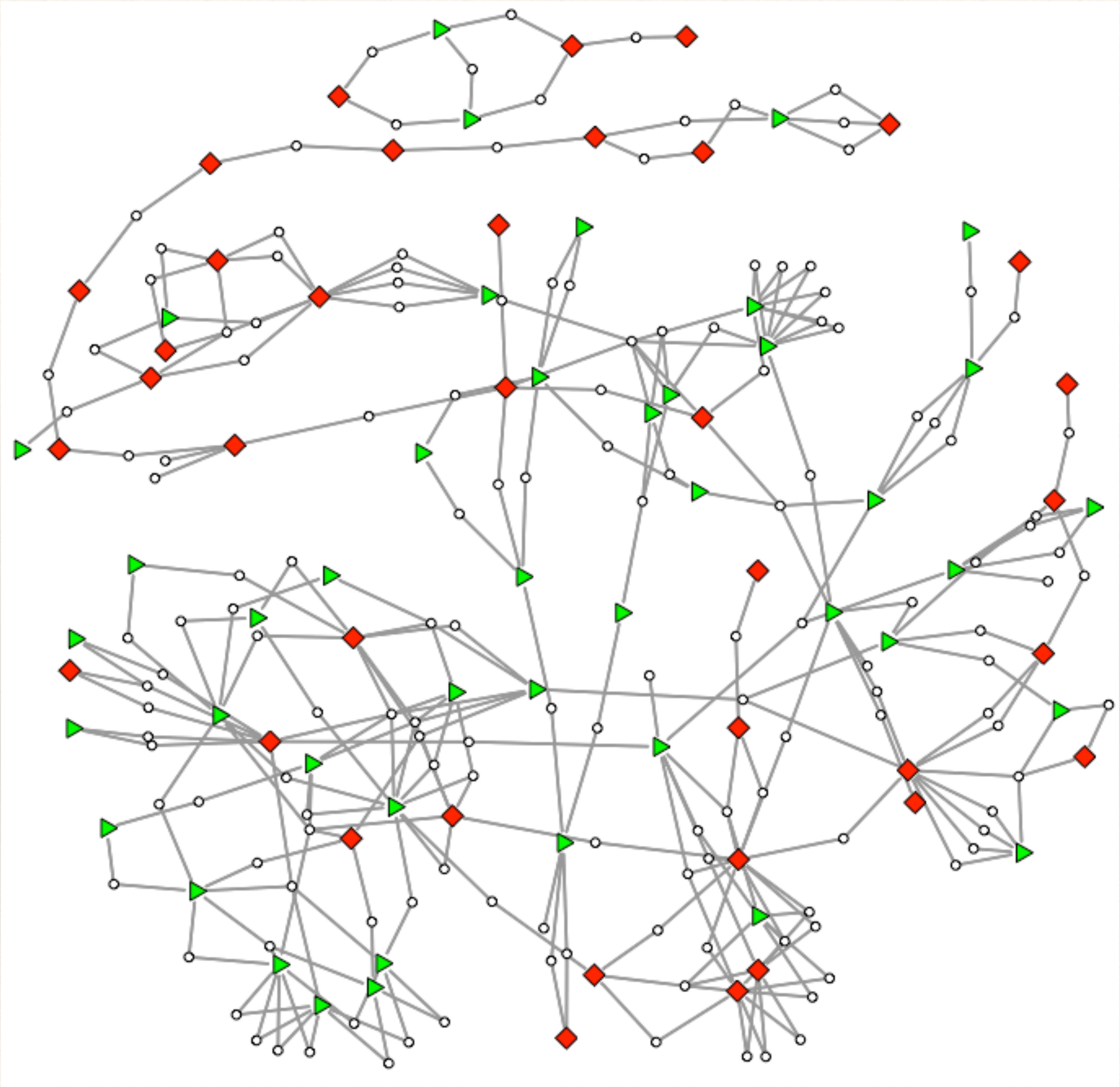
Figure 1. Diffusion of pragmatic legitimacy frame and compliance.

Bipartite Graph of Incidents and Officers by Treatment or Control Condition



Red/Square=Treatment Condition
Green/Triangle=Control Condition
White/Circle=Incidents

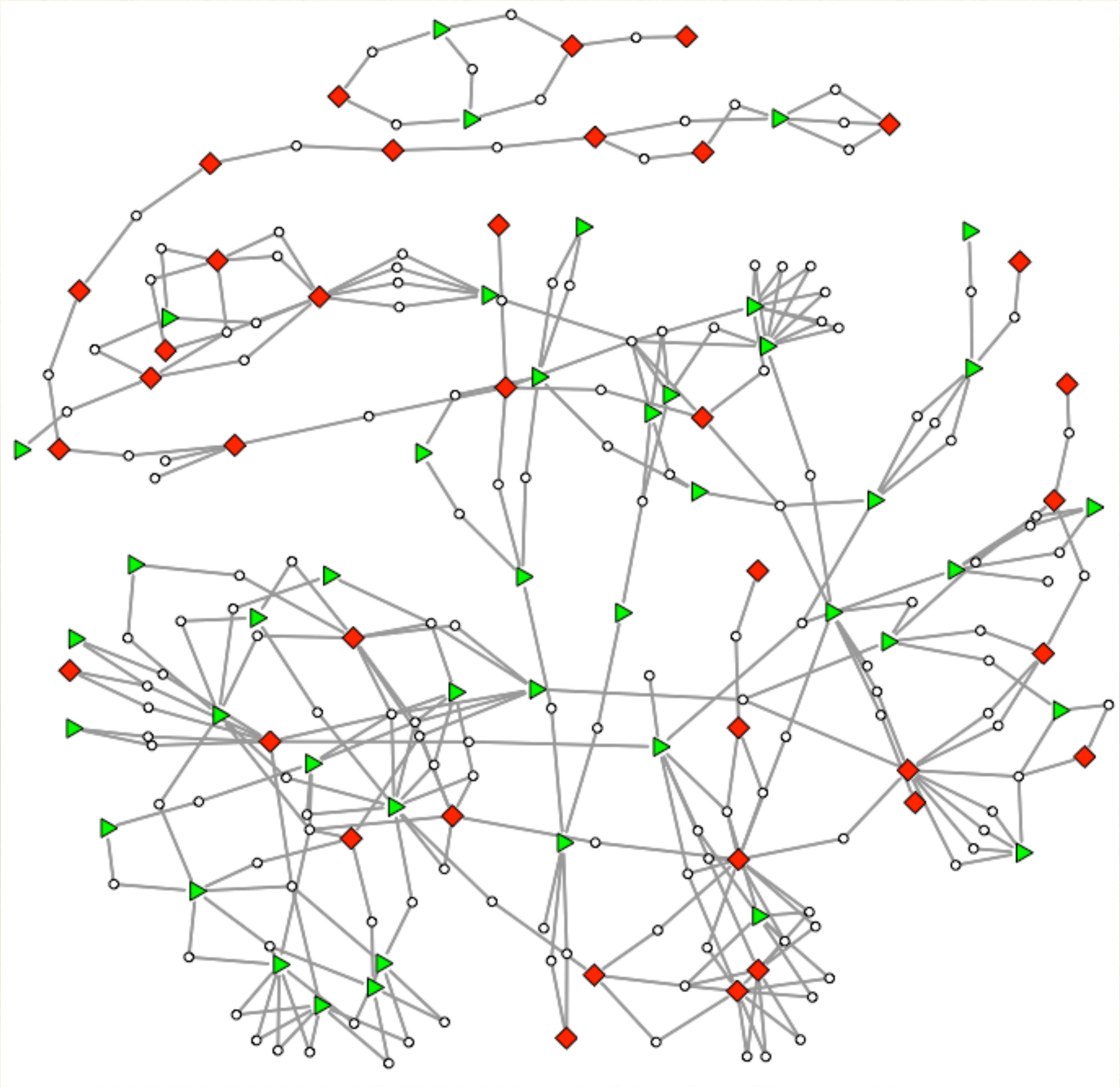
<u>Findings</u>:
Officers views of cameras changed based on who they interacted with through the network



What is the **concept** of interest?

How is it **conceptualized**?

How is it **operationalized**?



Learning Goals

- At the end of the lecture, you should be able to answer these questions:
  - How are **bipartite** graphs different from **unipartite** graphs?
  - What are some structural properties of bipartite graphs that we can examine?

Introduction

- So far, we have examined graphs that are:
  - Unipartite (i.e. one partition of the node set).
- We want to look at graph structures that:
  - Have multiple partitions of node sets (i.e. *n*-mode).

Two-Mode Networks

- Data are structured such that nodes come from two separate classes.
  - Examples:
    - Members of various groups, authors of papers, students in courses, participants in an event.

- A very different way of **conceptualizing** and **operationalizing** social structure.

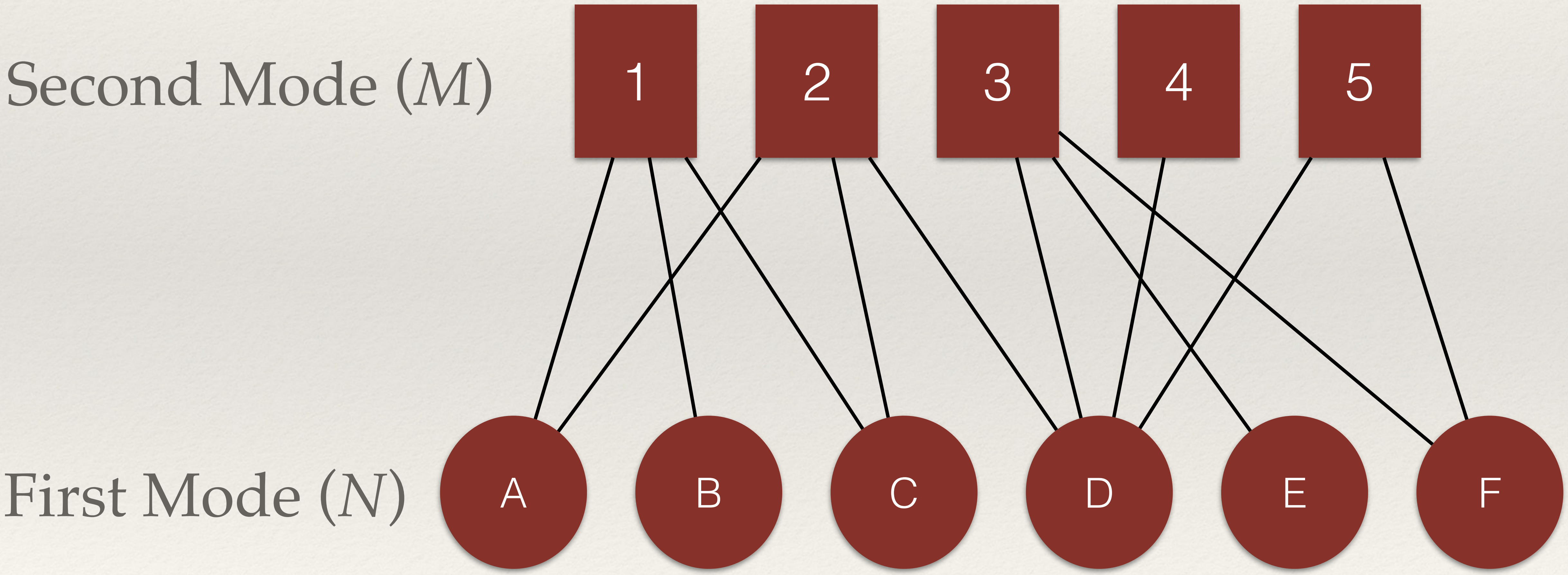
Bipartite Graphs

- Two-mode data can be represented by *bipartite* graphs:
    - A graph, such that there are two partitions of nodes (called modes), and edges only occur between these partitions (i.e. not within).

Graph Notation

- Definition of a **bipartite graph**: $G = (N, M, L)$
  - Node/Vertex set: $N = \{n_1, n_2 \ldots, n_g\}$
  - Node/Vertex set: $M = \{m_1, m_2 \ldots, m_g\}$
  - Line/Edge set: $L = \{l_1, l_2 \ldots, l_L\}$
    - There are N nodes/vertices in the first set and M nodes/vertices in the second set.
    - There are L lines/edges in the graph.

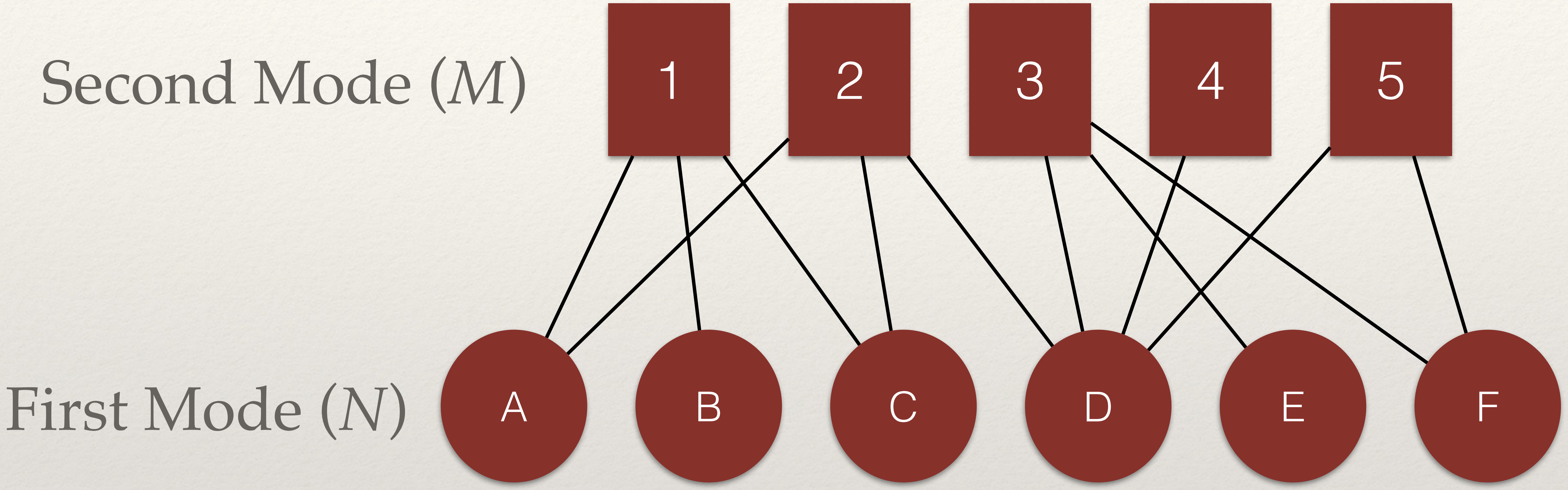
Bipartite Graphs



Sociometric Notation

- We can continue to use an *adjacency matrix*, to represent relations where each node/vertex is listed on the row and the column.

Bipartite Graphs



	1	2	3	4	5
A	1	1	0	0	0
B	1	0	0	0	0
C	1	1	0	0	0
D	0	1	1	1	1
E	0	0	1	0	0
F	0	0	1	0	1

Adjacency Matrix

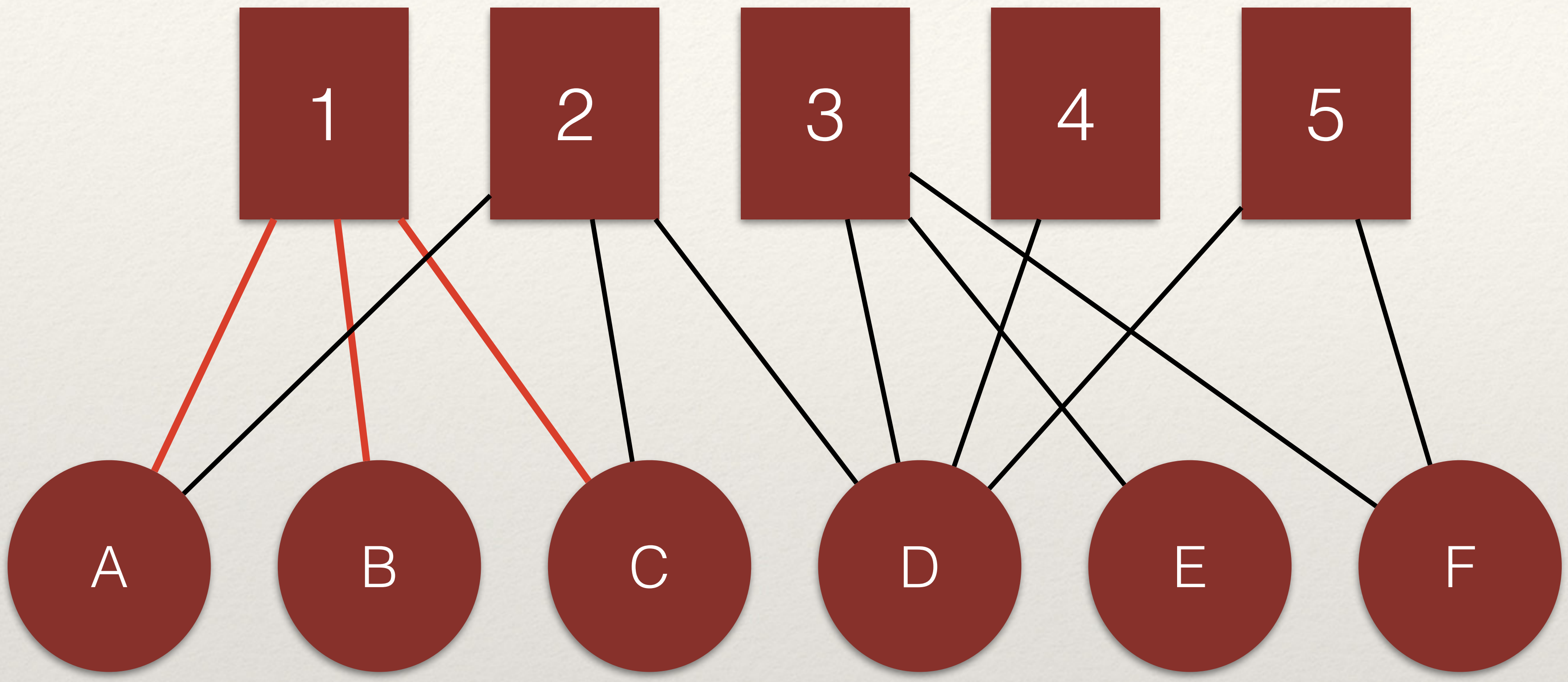
Second Mode (*M*)

First Mode (*N*)

	1	2	3	4	5
A	1	1	0	0	0
B	1	0	0	0	0
C	1	1	0	0	0
D	0	1	1	1	1
E	0	0	1	0	0
F	0	0	1	0	1

The **order** of the matrix is *NxM*. It is rectangular.

Bipartite Graphs

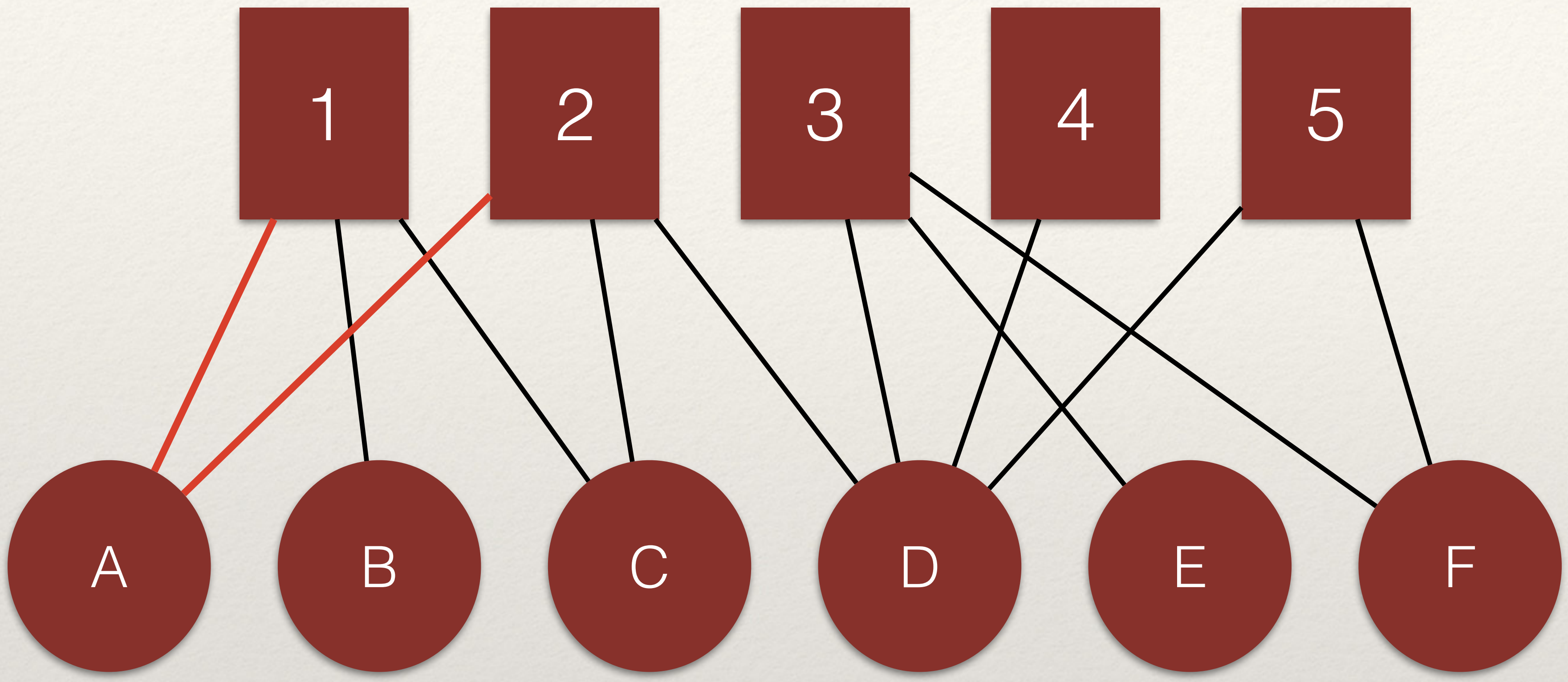


	1	2	3	4	5
A	1	1	0	0	0
B	1	0	0	0	0
C	1	1	0	0	0
D	0	1	1	1	1
E	0	0	1	0	0
F	0	0	1	0	1

Each column corresponds to the edges incident on a node, M_i, from the set M.

M usually corresponds to the event, group, etc.

Bipartite Graphs



	1	2	3	4	5
A	1	1	0	0	0
B	1	0	0	0	0
C	1	1	0	0	0
D	0	1	1	1	1
E	0	0	1	0	0
F	0	0	1	0	1

Each row corresponds to the edges incident on a node, N_i, from the set N.

N usually corresponds to the person.

Examining Bipartite Graphs

- There are several approaches to examining bipartite graphs:
  - Keep the graph bipartite and examine the properties.
  - *Project* the graph to one mode (either N or M) and examine the properties (we will do this next week).

Bipartite Graph Properties

- As with unipartite graphs or one-mode networks, we can examine various properties of the data to tell us about the structure of the object.
  - <u>Examples</u>:
    - How dense is the graph? (Density)
    - How are the edges distributed over nodes? (Degree Centrality)
    - How "clustered" are the data? (Dyadic clustering)

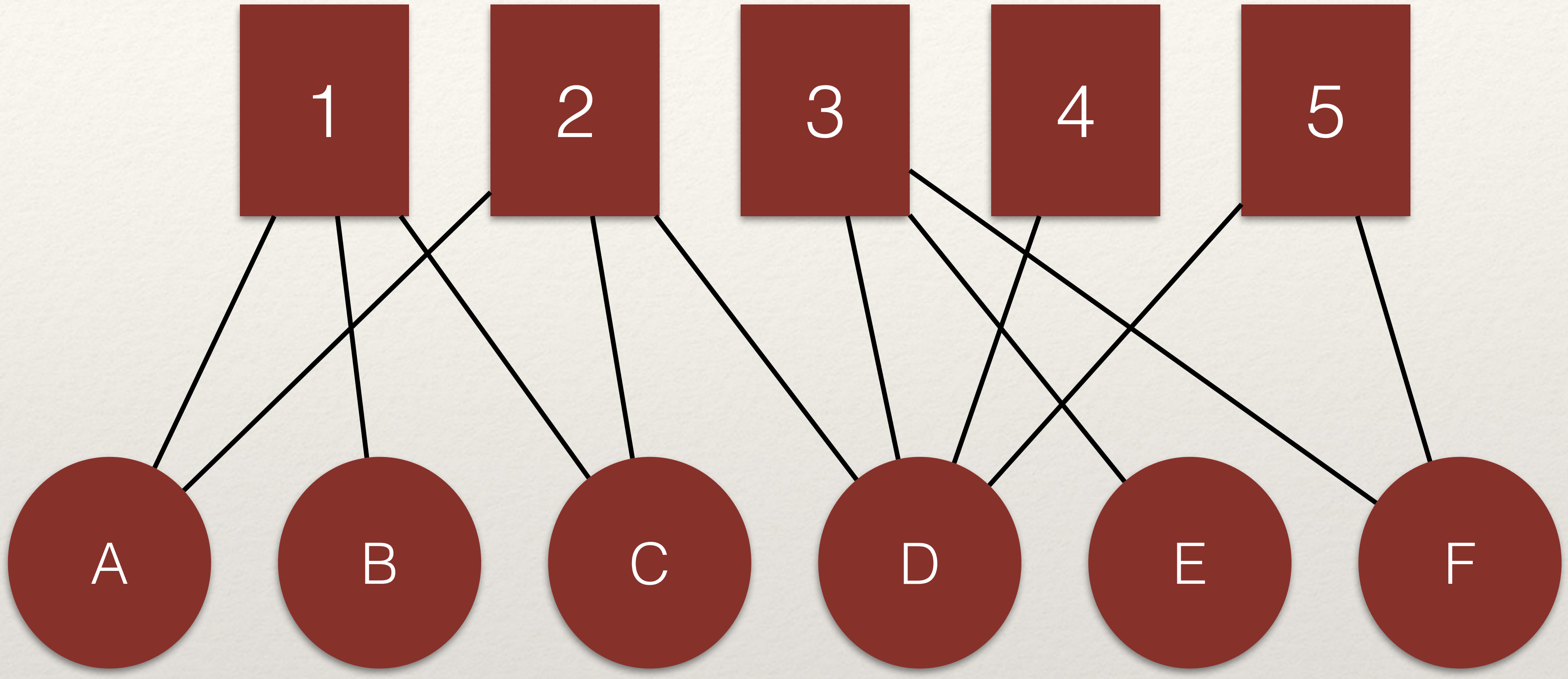
Density: Bipartite Graphs

- The *density* of a two-mode network is the number of edges observed *L*, divided by the number of possible pairwise relations between the vertex sets.
    - The number of possible connections between the vertices is N x M.
    - So, the density is:

$$\frac{L}{N \times M}$$

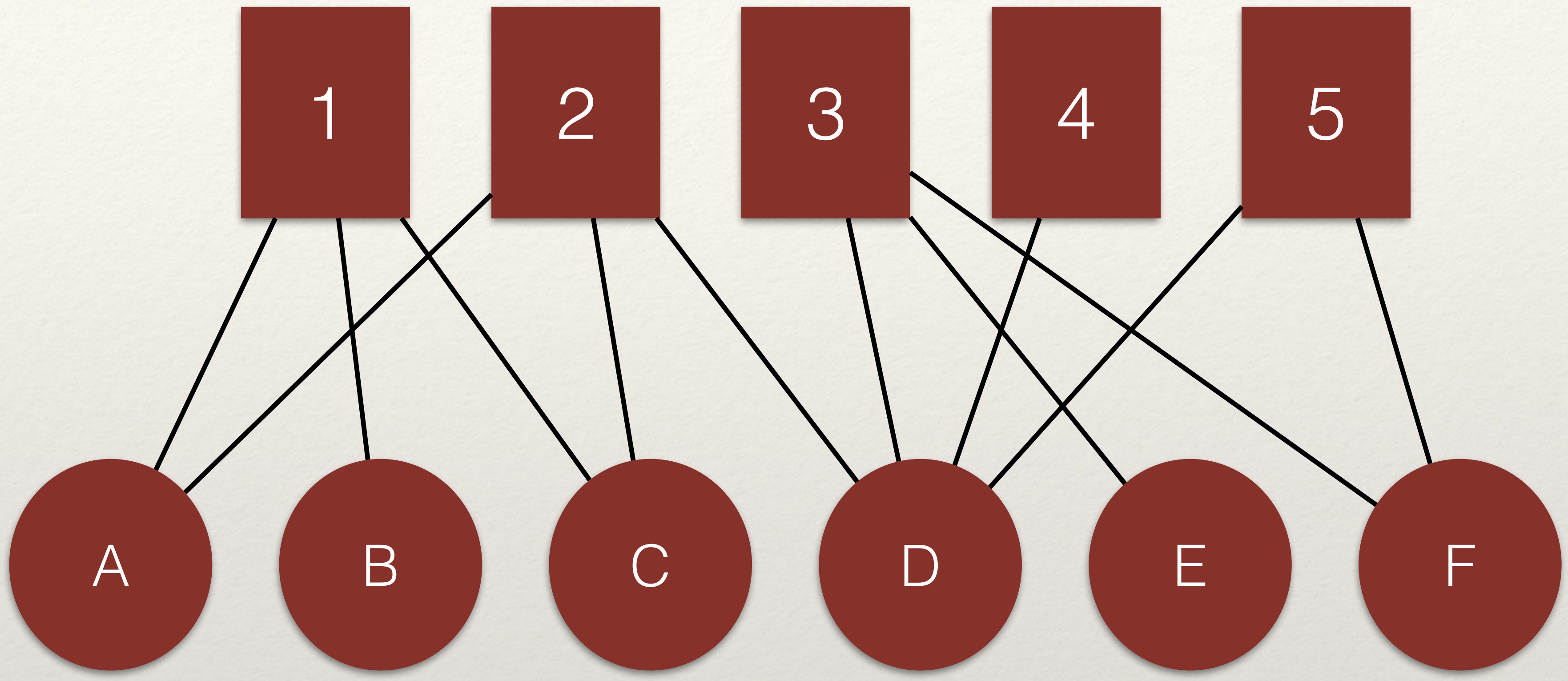
Example

What is the density of this network?



Example

What is the density of this network?



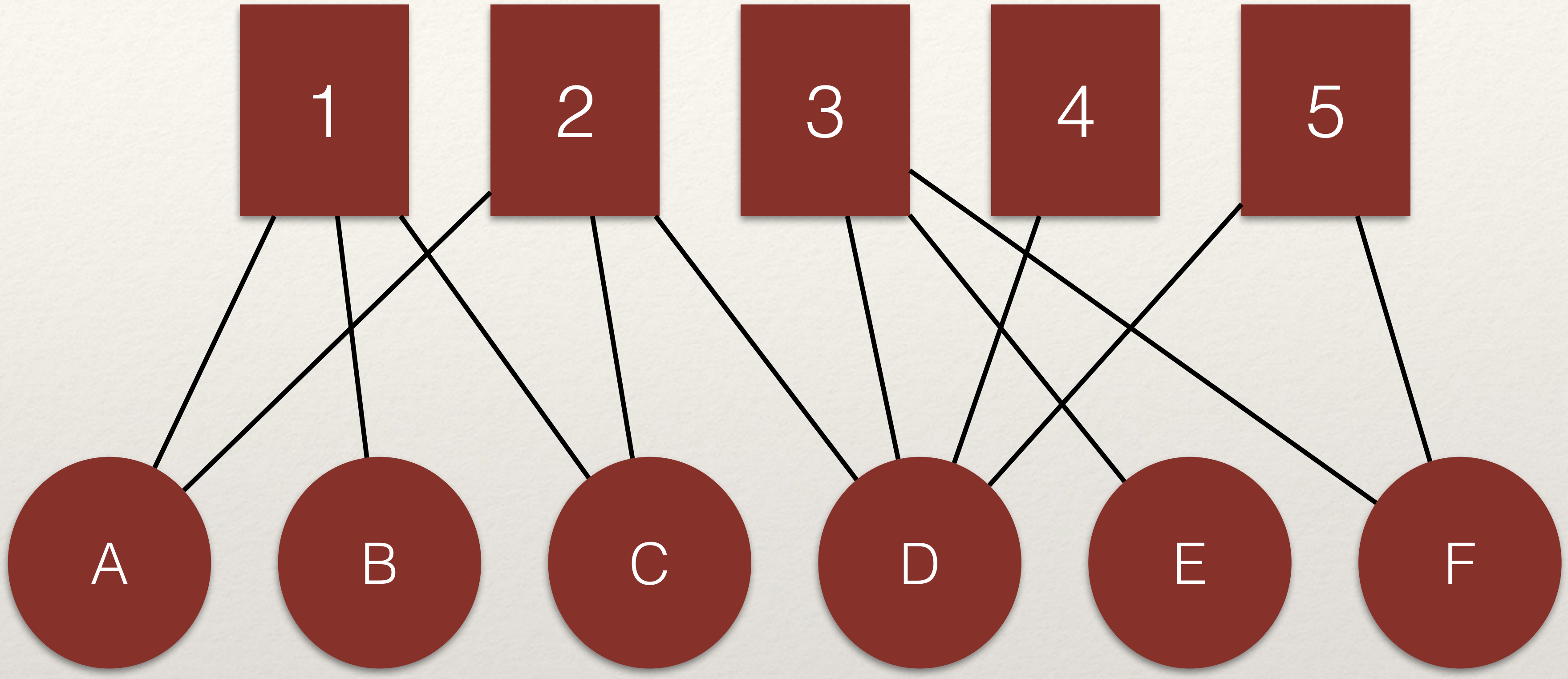
First, calculate the number of edges.

Then, calculate *N* x *M*

		M				
		1	2	3	4	5
N	A	1	1	0	0	0
	B	1	0	0	0	0
	C	1	1	0	0	0
	D	0	1	1	1	1
	E	0	0	1	0	0
	F	0	0	1	0	1

Example

What is the density of this network?

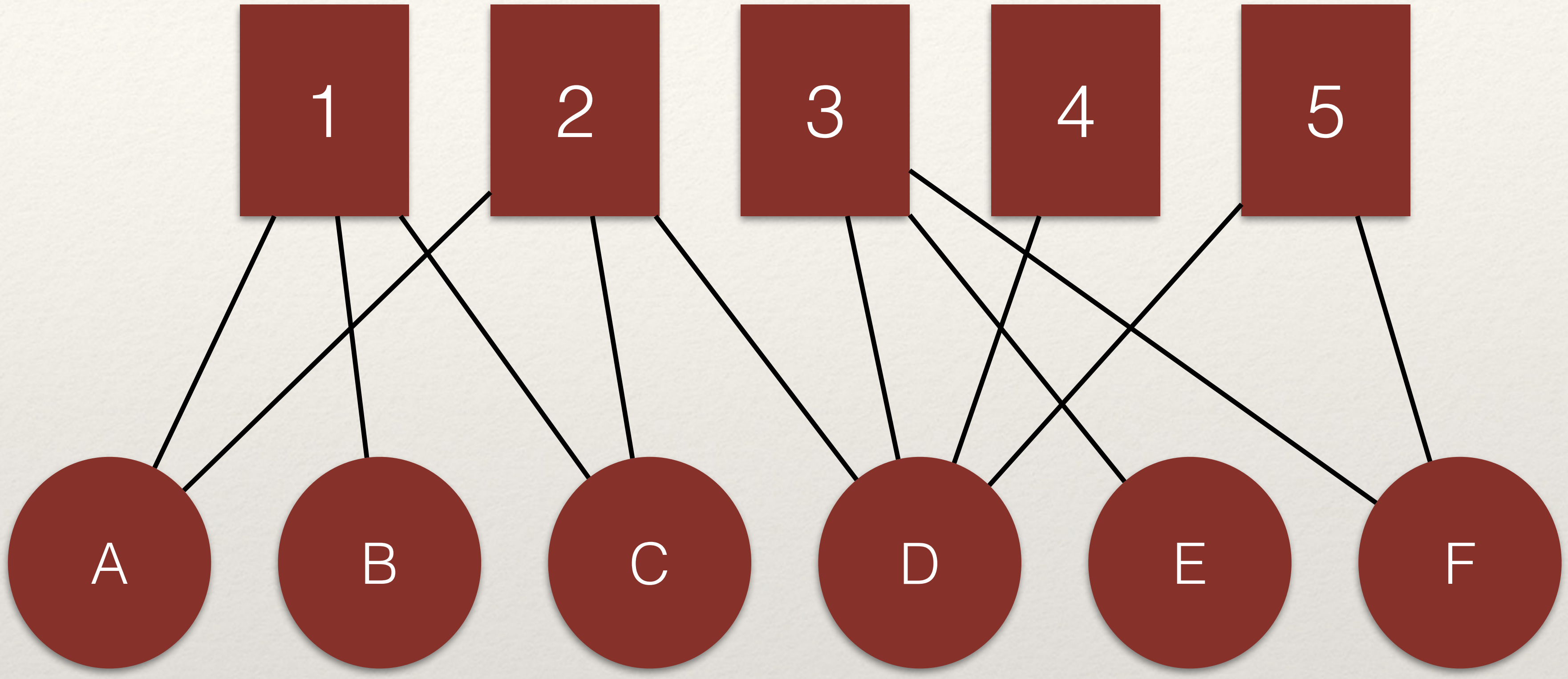


$$\frac{L}{N \times M} = \frac{12}{6 \times 5} = \frac{12}{30} = 0.4$$

		M				
		1	2	3	4	5
N	A	1	1	0	0	0
	B	1	0	0	0	0
	C	1	1	0	0	0
	D	0	1	1	1	1
	E	0	0	1	0	0
	F	0	0	1	0	1

Example

What does a ***density*** *of 0.4 mean?*



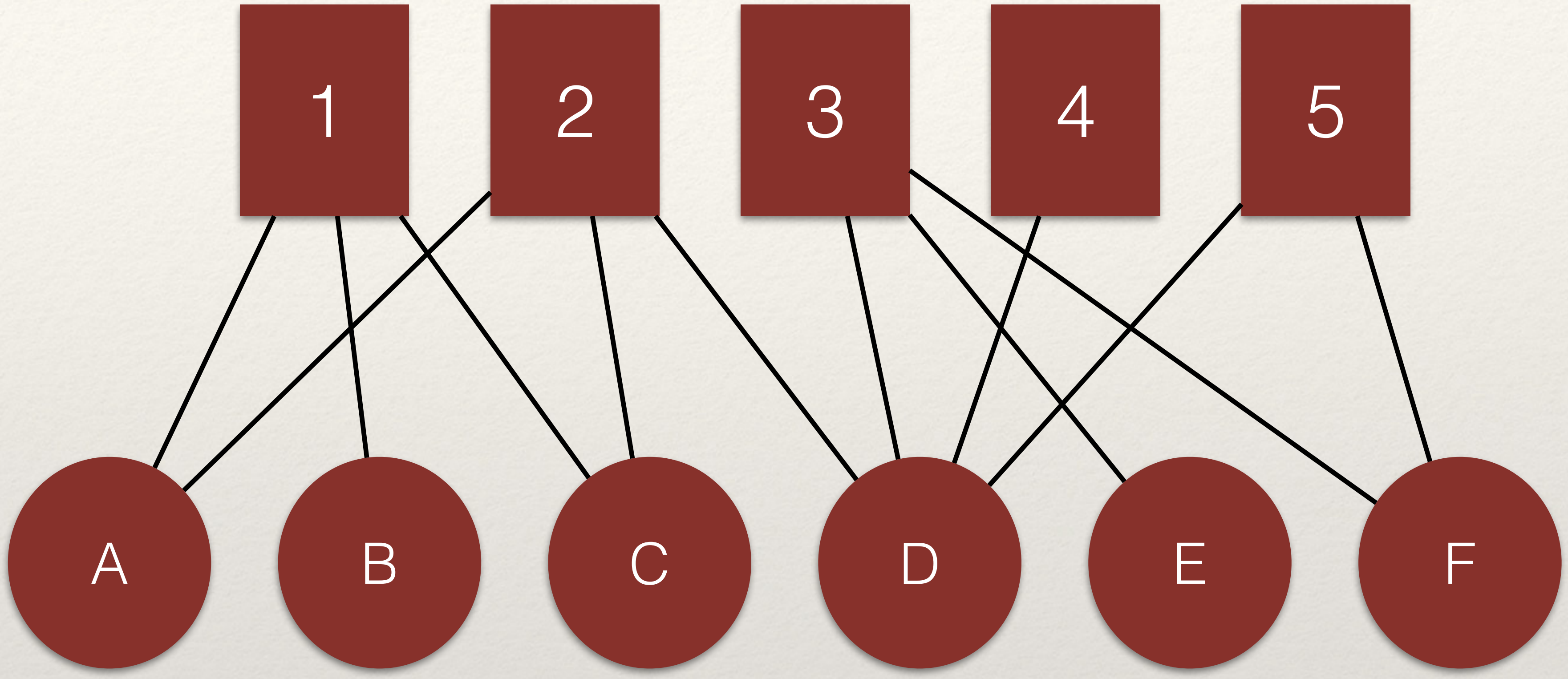
		M				
		1	2	3	4	5
N	A	1	1	0	0	0
	B	1	0	0	0	0
	C	1	1	0	0	0
	D	0	1	1	1	1
	E	0	0	1	0	0
	F	0	0	1	0	1

Degree Centrality: Bipartite Graphs

- For a bipartite graph there are *two* degree distributions:
  - The distribution of ties in the first mode (N).
  - The distribution of ties in the second mode (M).
    - The *row sum* for the adjacency matrix gives the degree centrality scores for the first mode, N.
    - The *column sum* for the adjacency matrix gives the degree centrality scores for the second mode, M.

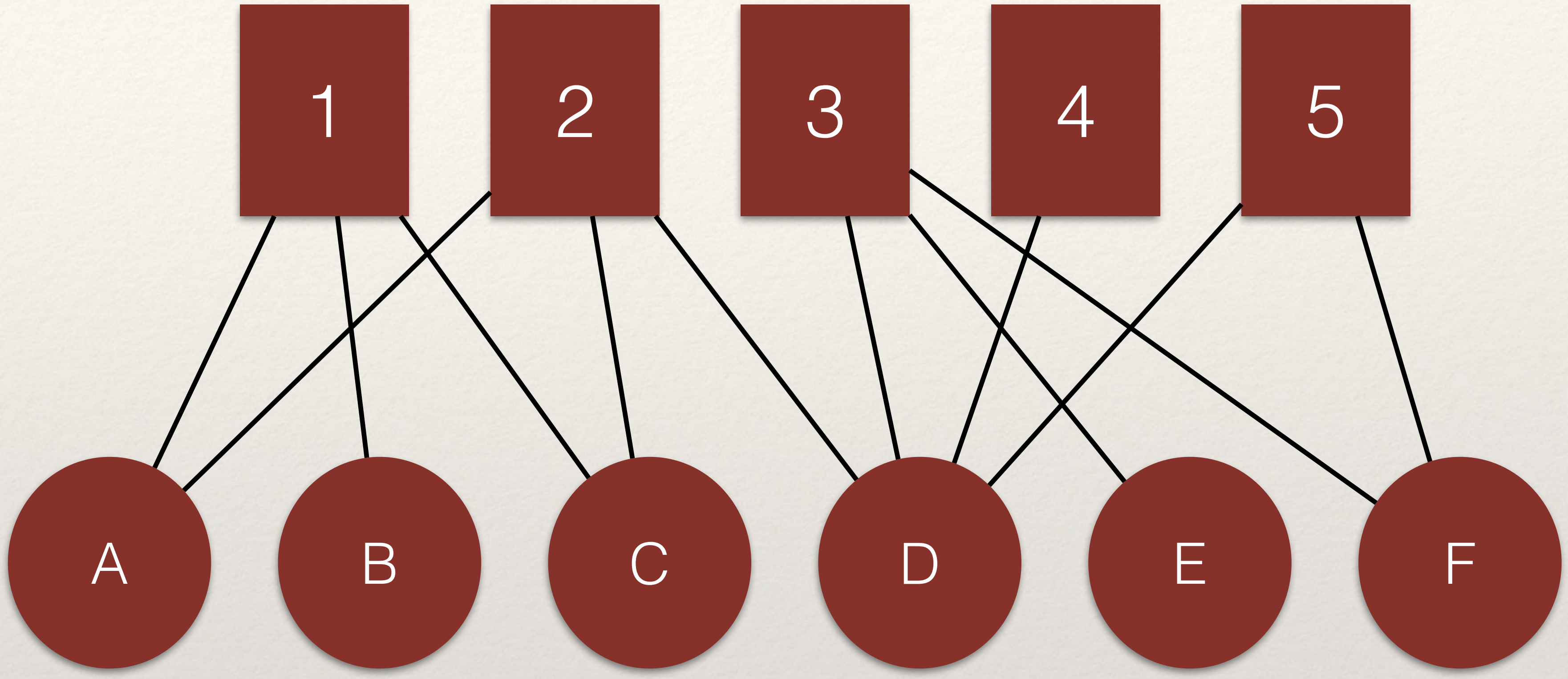
Example

What are the degree centrality scores for each vertex set in this example?



Example

What are the degree centrality scores for each vertex set in this example?

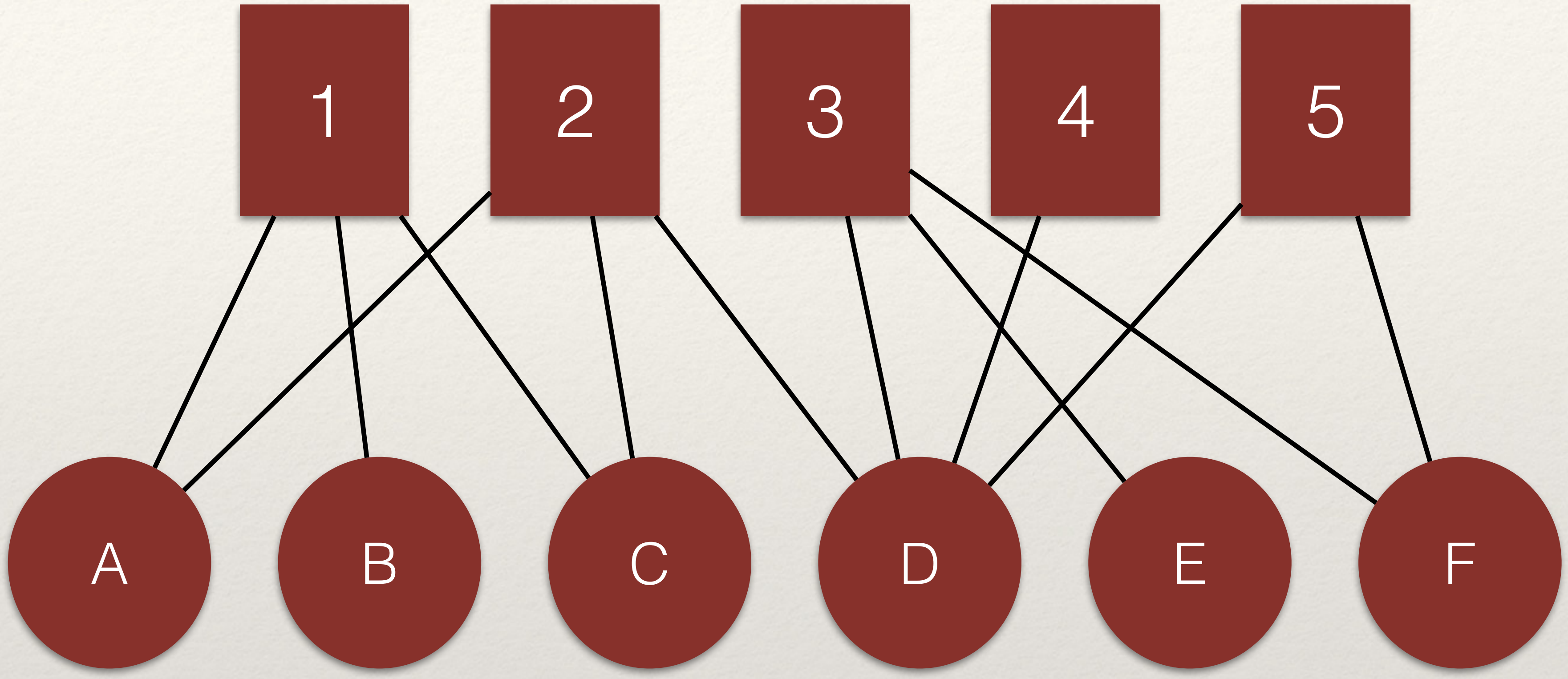


First, get the row sums.

		M					
		1	2	3	4	5	
N	A	1	1	0	0	0	2
	B	1	0	0	0	0	1
	C	1	1	0	0	0	2
	D	0	1	1	1	1	4
	E	0	0	1	0	0	1
	F	0	0	1	0	1	2

Example

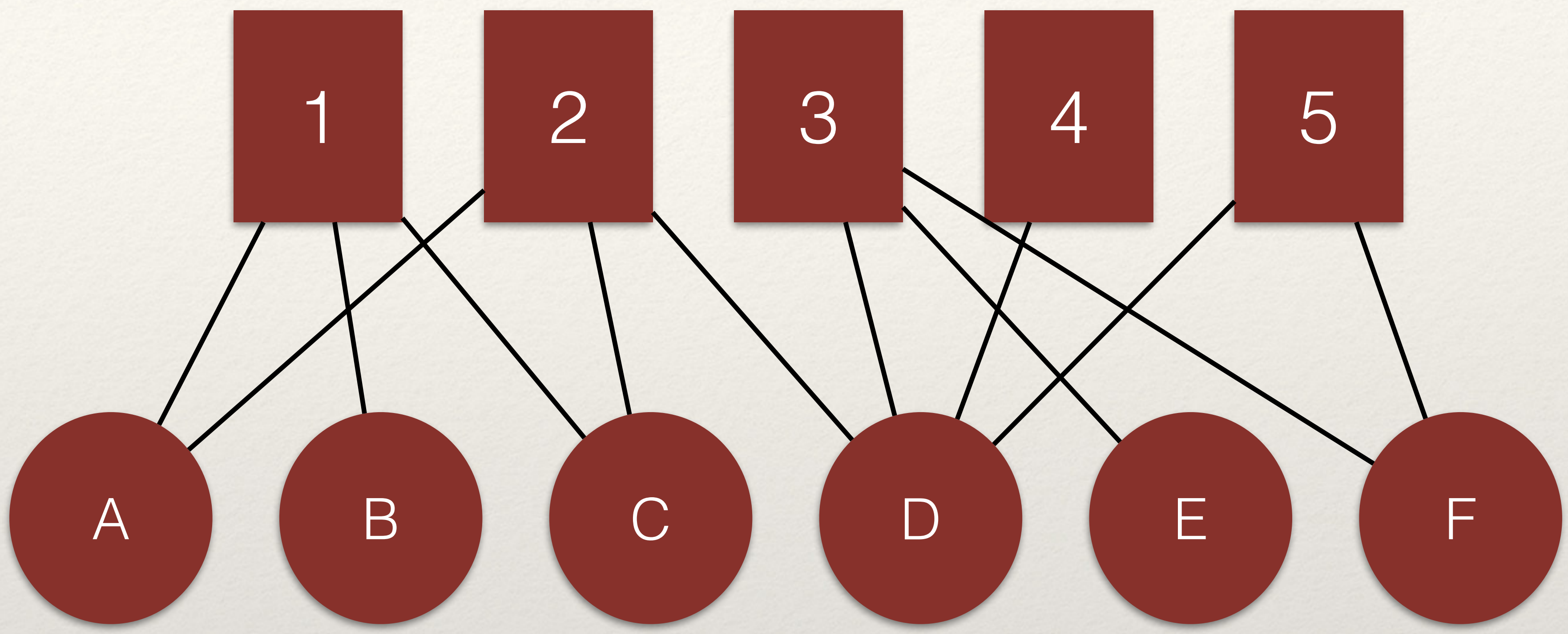
What are the degree centrality scores for each vertex set in this example?



Second, get the column sums.

		M				
		1	2	3	4	5
N	A	1	1	0	0	0
	B	1	0	0	0	0
	C	1	1	0	0	0
	D	0	1	1	1	1
	E	0	0	1	0	0
	F	0	0	1	0	1
		3	3	3	1	2

Example



		M					
		1	2	3	4	5	
N	A	1	1	0	0	0	2
	B	1	0	0	0	0	1
	C	1	1	0	0	0	2
	D	0	1	1	1	1	4
	E	0	0	1	0	0	1
	F	0	0	1	0	1	2
		3	3	3	1	2	

Degree Centrality: Bipartite Graphs

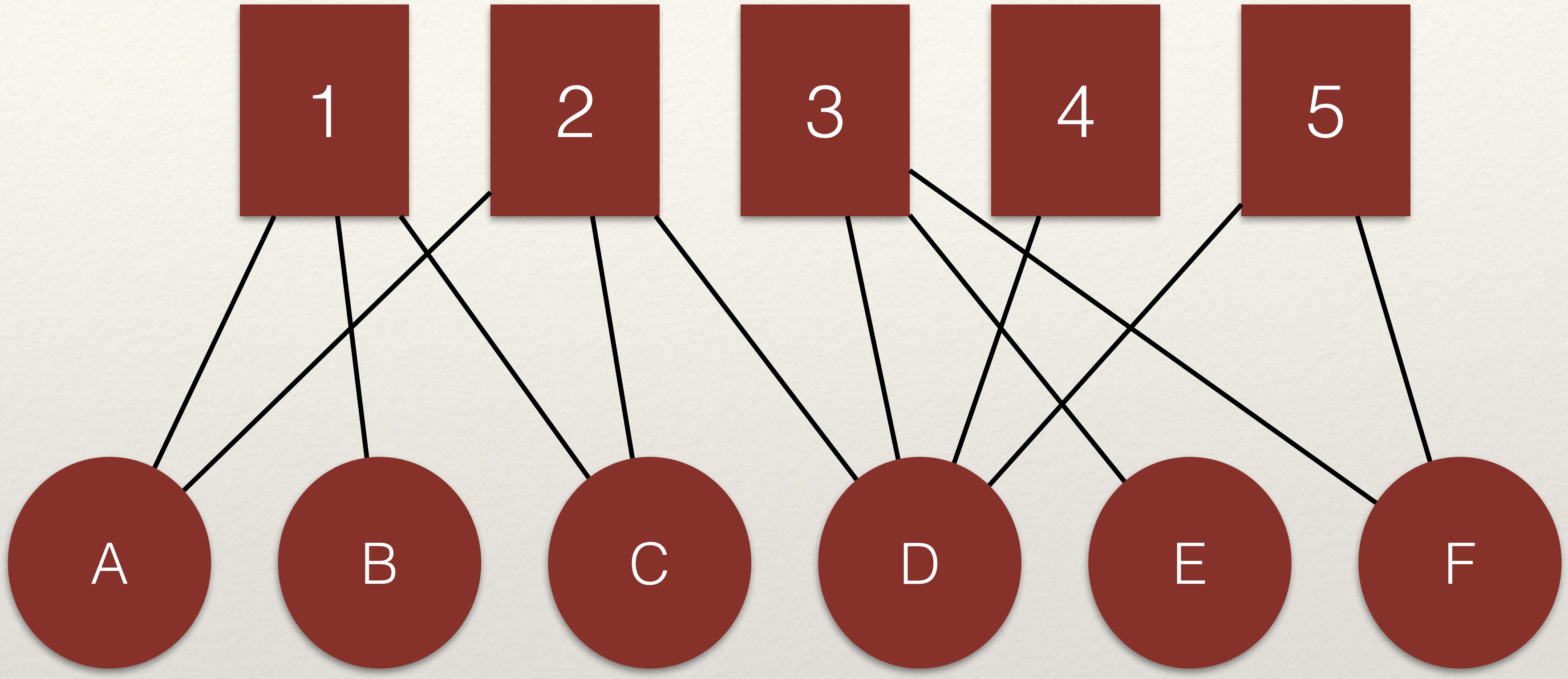
- Degree centrality scores for each node/vertex set not only reflects each node's connectivity to nodes in the other set, but also depend on the size of that set.
  - Larger networks will have a higher maximum possible degree centrality value.
    - *Solution?*

Standardized Degree Centrality: Bipartite Graphs

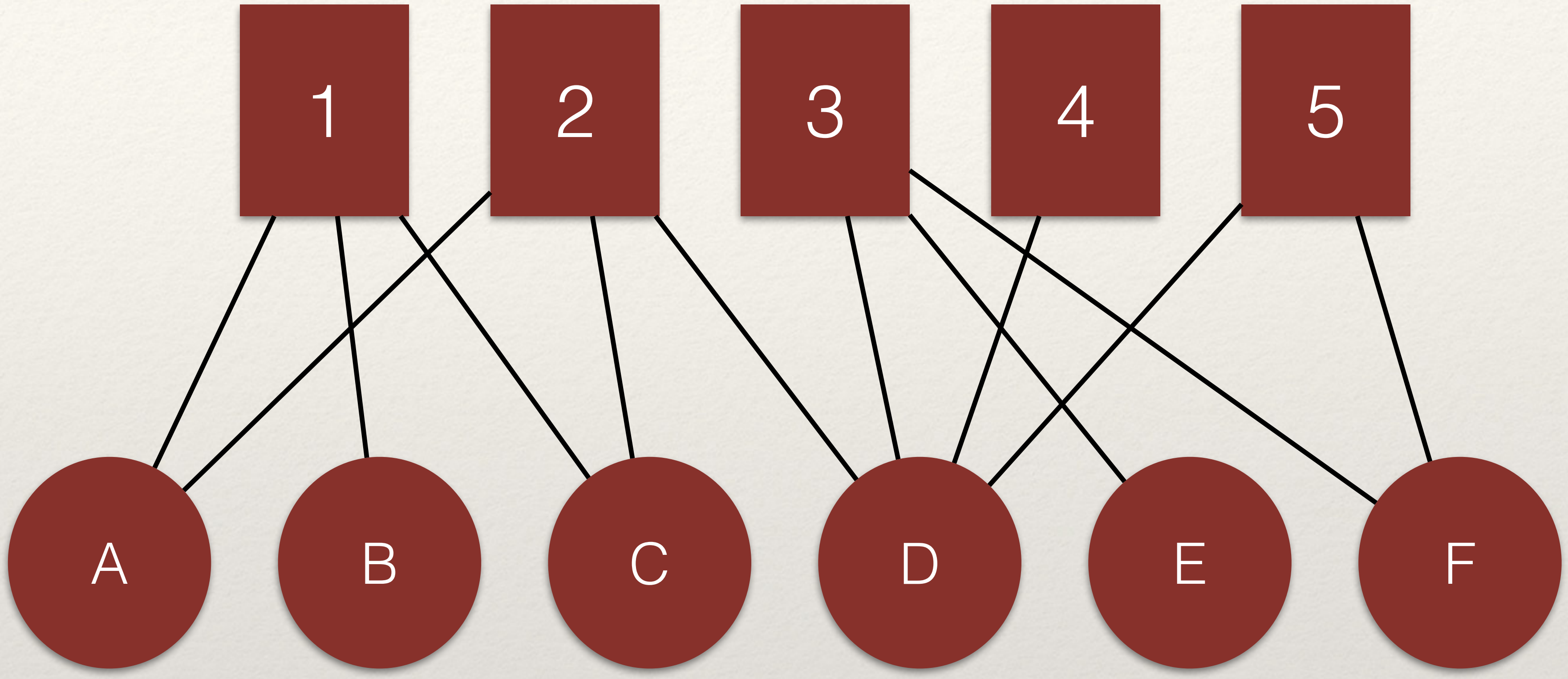
- Standardize!
  - We can account for differences across networks by dividing each degree centrality score by the number of nodes/vertices in the opposite set.
  - For N, we divide by M.
  - For M, we divide by N.

Example

What are the standardized degree centrality scores for each vertex set in this example?



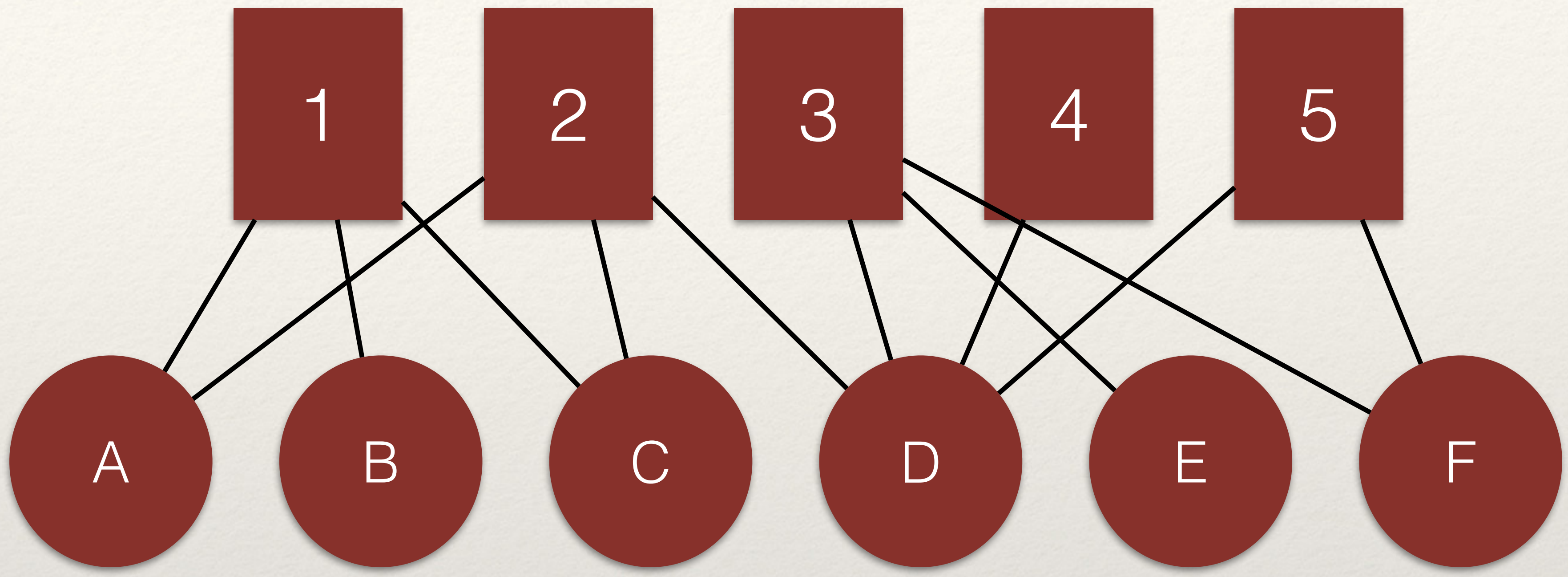
Example



Divide the row sums by M (i.e. 5).

		M						
		1	2	3	4	5	Raw	Stand.
N	A	1	1	0	0	0	2	0.4
	B	1	0	0	0	0	1	0.2
	C	1	1	0	0	0	2	0.4
	D	0	1	1	1	1	4	0.8
	E	0	0	1	0	0	1	0.2
	F	0	0	1	0	1	2	0.4

Example



Second, divide the column sums by *N* (i.e. 6).

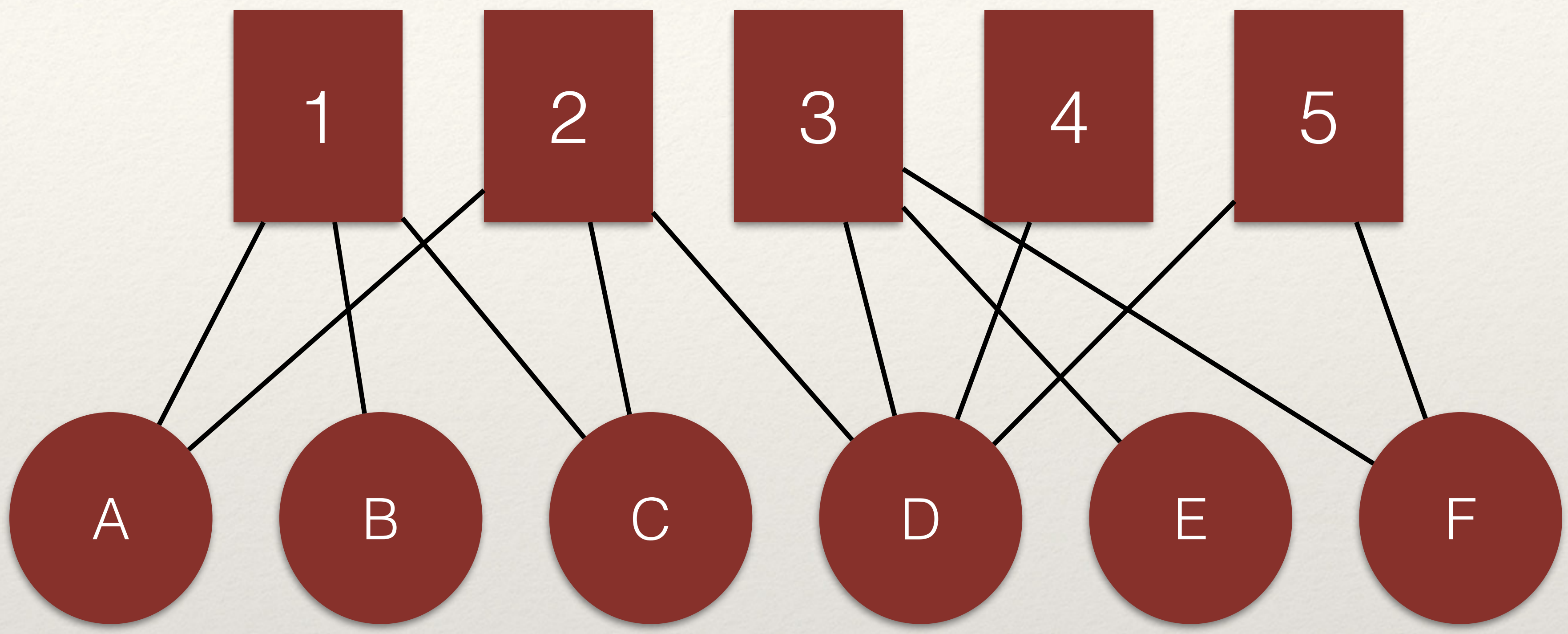
		M				
		1	2	3	4	5
N	A	1	1	0	0	0
	B	1	0	0	0	0
	C	1	1	0	0	0
	D	0	1	1	1	1
	E	0	0	1	0	0
	F	0	0	1	0	1
	Raw	3	3	3	1	2
	Stand	0.5	0.5	0.5	0.167	0.334

Mean Degree Centrality: Bipartite Graphs

- As before, we could examine the central tendency by examining the mean degree for each node/vertex set.
  - For N, we divide by L/N.
  - For M, we divide by L/M.
    - **Note:** for the mean we use the number of nodes in the corresponding vertex set, for standardizing we use the opposite vertex set.

Example

What is the mean degree centrality score for each vertex set in this example?

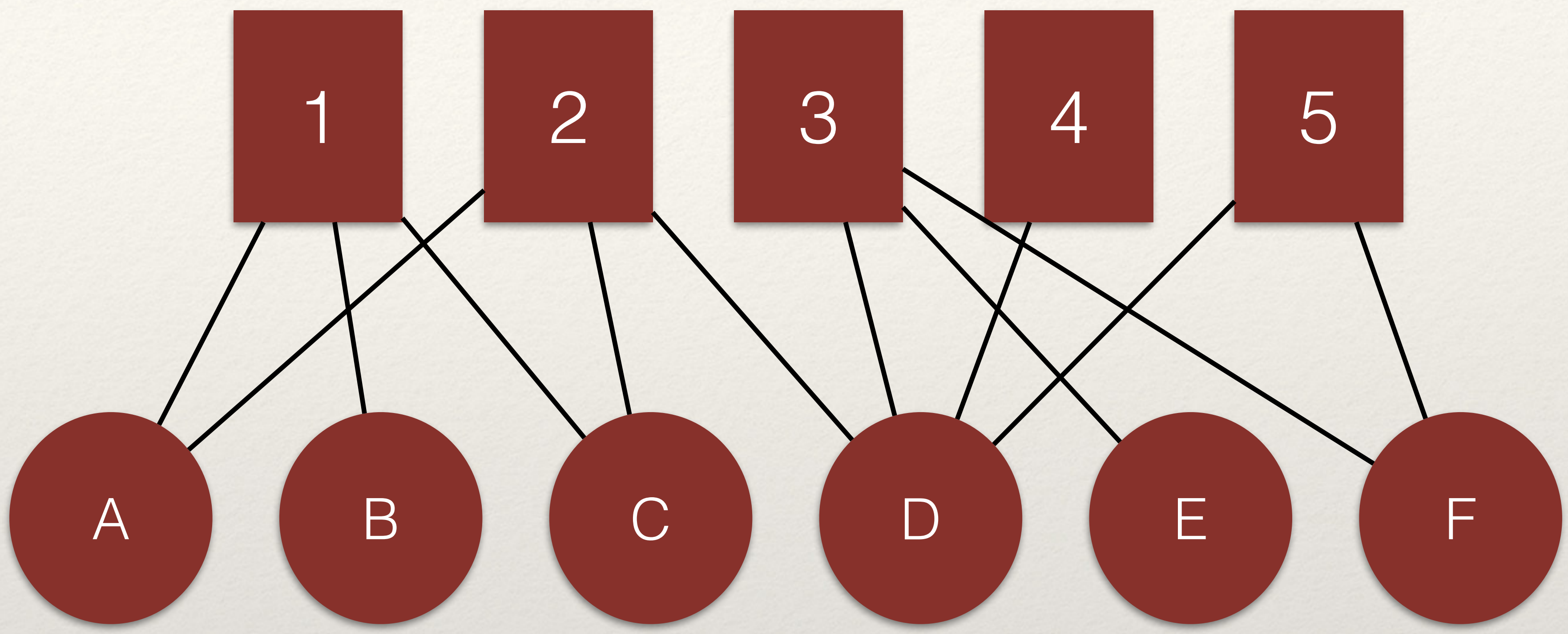


		M					
		1	2	3	4	5	
N	A	1	1	0	0	0	2
	B	1	0	0	0	0	1
	C	1	1	0	0	0	2
	D	0	1	1	1	1	4
	E	0	0	1	0	0	1
	F	0	0	1	0	1	2
		3	3	3	1	2	

Example

What is the mean degree centrality score for each vertex set in this example?

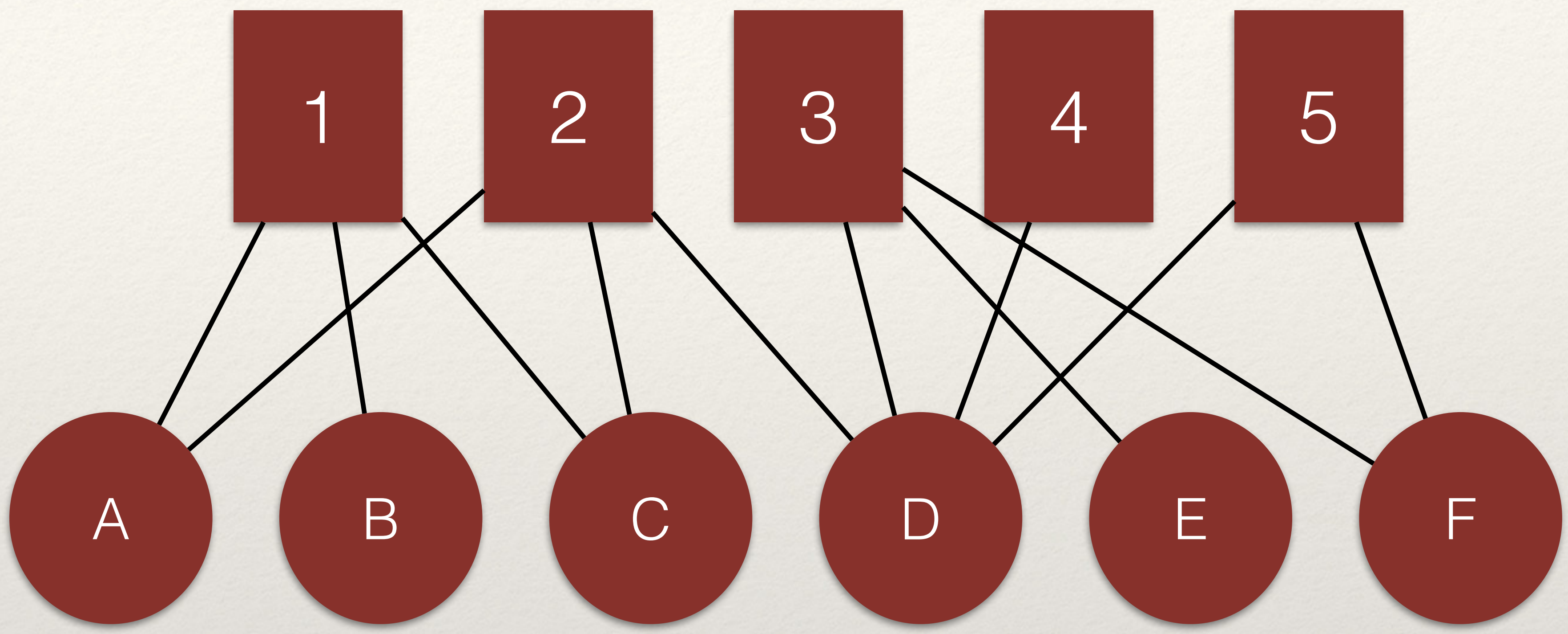
For N, it is $12/6 = 2$



		M					
		1	2	3	4	5	
N	A	1	1	0	0	0	2
	B	1	0	0	0	0	1
	C	1	1	0	0	0	2
	D	0	1	1	1	1	4
	E	0	0	1	0	0	1
	F	0	0	1	0	1	2
		3	3	3	1	2	

Example

What is the mean degree centrality score for each vertex set in this example?



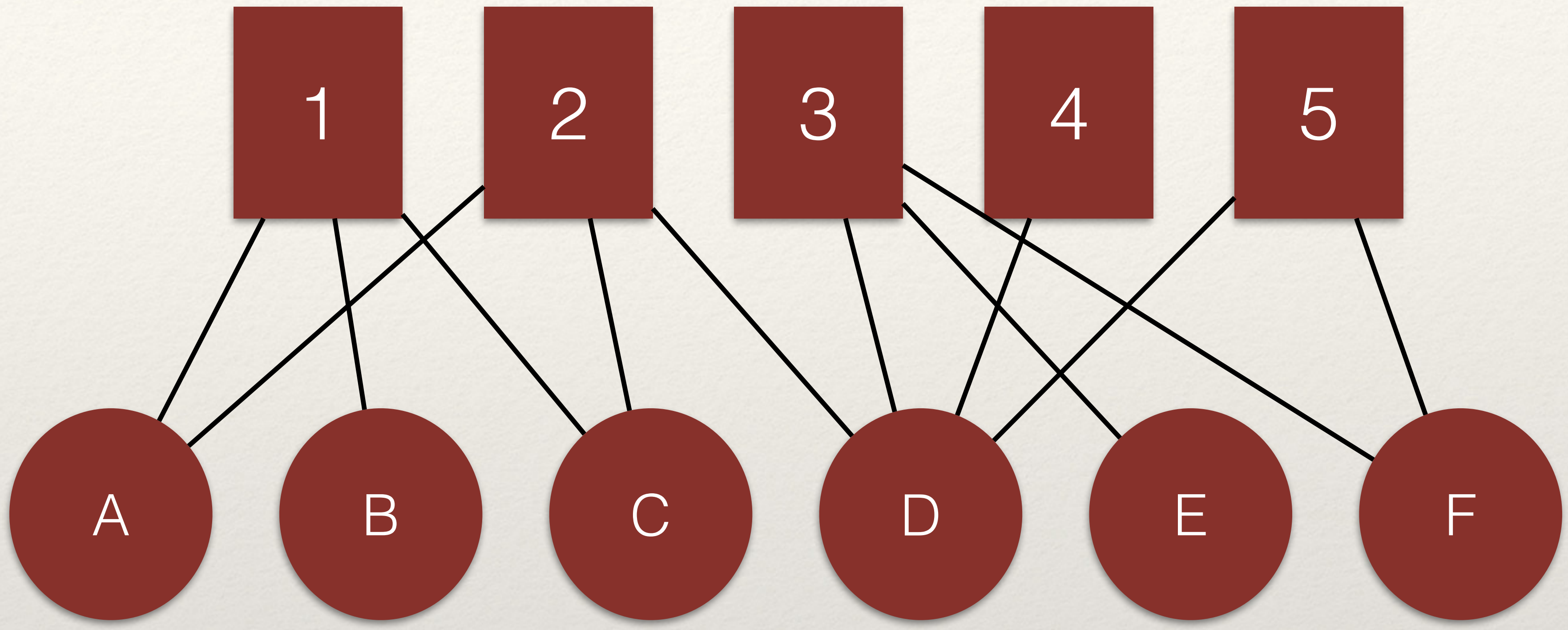
		M					
		1	2	3	4	5	
N	A	1	1	0	0	0	2
	B	1	0	0	0	0	1
	C	1	1	0	0	0	2
	D	0	1	1	1	1	4
	E	0	0	1	0	0	1
	F	0	0	1	0	1	2
		3	3	3	1	2	

For N, it is $12/6 = 2$

For M, it is $12/5 = 2.4$

Example

What does the difference between the means tell us?



For N, it is $12/6 = 2$

For M, it is $12/5 = 2.4$

		M					
		1	2	3	4	5	
N	A	1	1	0	0	0	2
	B	1	0	0	0	0	1
	C	1	1	0	0	0	2
	D	0	1	1	1	1	4
	E	0	0	1	0	0	1
	F	0	0	1	0	1	2
		3	3	3	1	2	

Dyadic Clustering: Bipartite Graphs

- The density tells us about the overall level of ties between the node/vertex sets in the graph.
- Degree centrality tells us about how many edges are incident on a node in each node/vertex set.
- *What about the overlap in ties?*
  - In other words, do nodes in N tend to "share" nodes in M?
    - This is the notion of **clustering** in a graph.

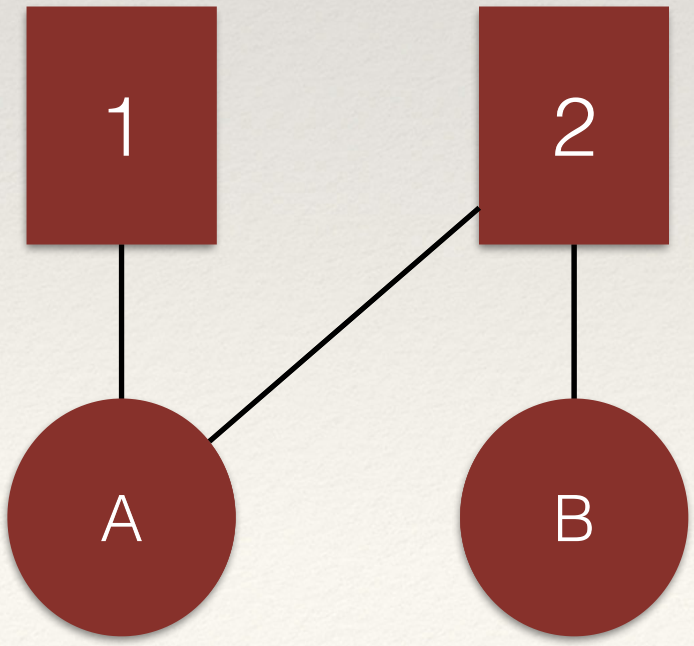
Dyadic Clustering: Bipartite Graphs

- In a bipartite graph, there are two interesting structures:
  - 3-paths (sometimes called L_3) and cycles (sometimes called C_4).

Dyadic Clustering: Bipartite Graphs

- In a bipartite graph, there are two interesting structures:
  - 3-paths (sometimes called L_3) and cycles (sometimes called C_4).

3-path



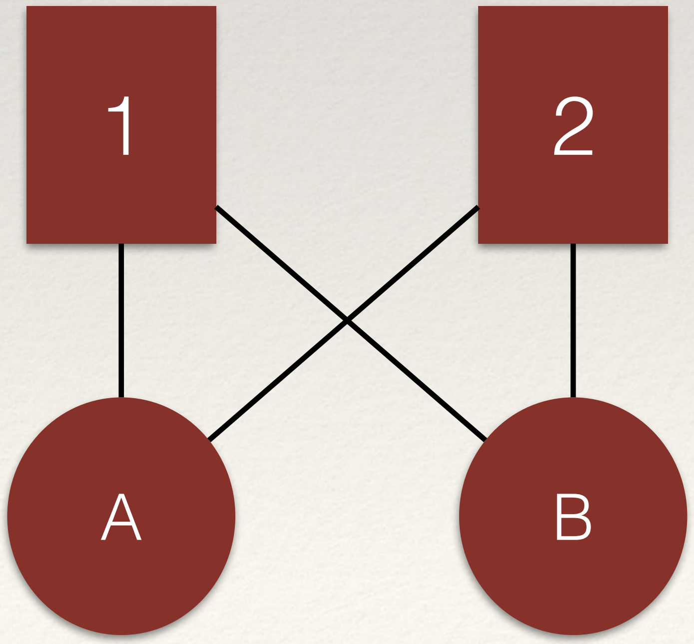
1-A-2-B

Dyadic Clustering: Bipartite Graphs

- In a bipartite graph, there are two interesting structures:
  - 3-paths (sometimes called L_3) and cycles (sometimes called C_4).

cycle

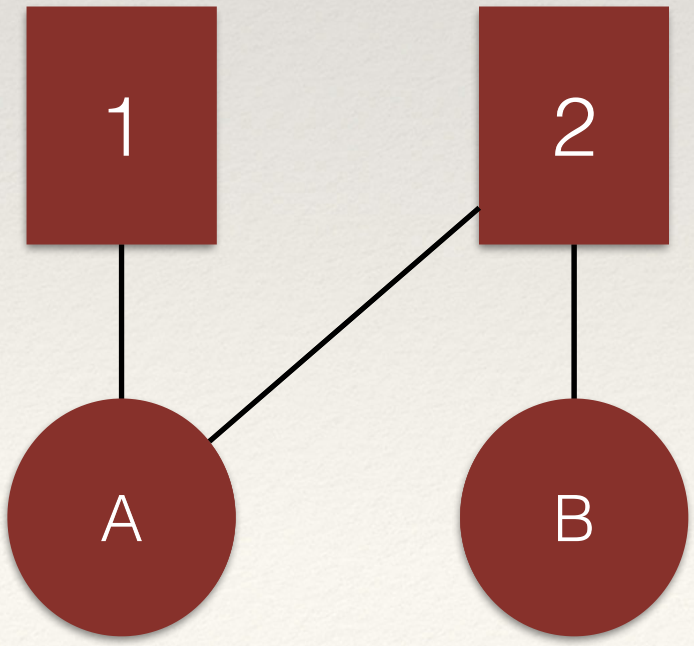
1-A-2-B-1



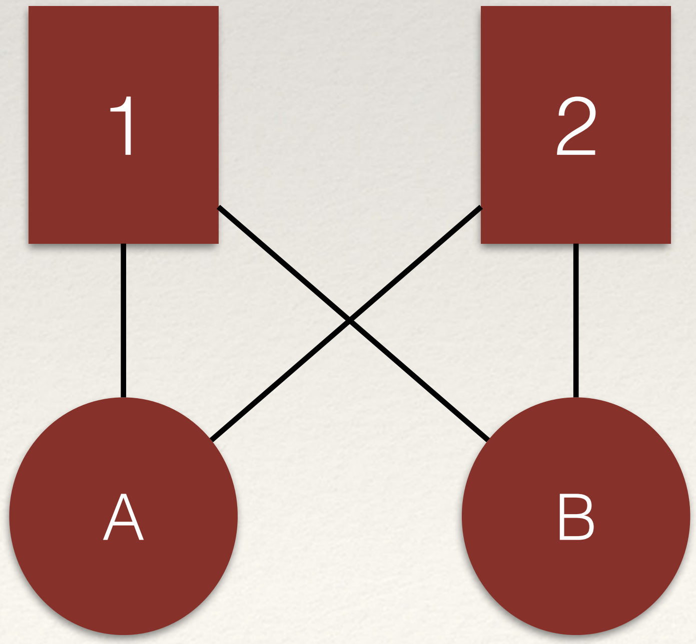
Dyadic Clustering: Bipartite Graphs

- In a bipartite graph, there are two interesting structures:
  - 3-paths (sometimes called L_3) and cycles (sometimes called C_4).

3-path



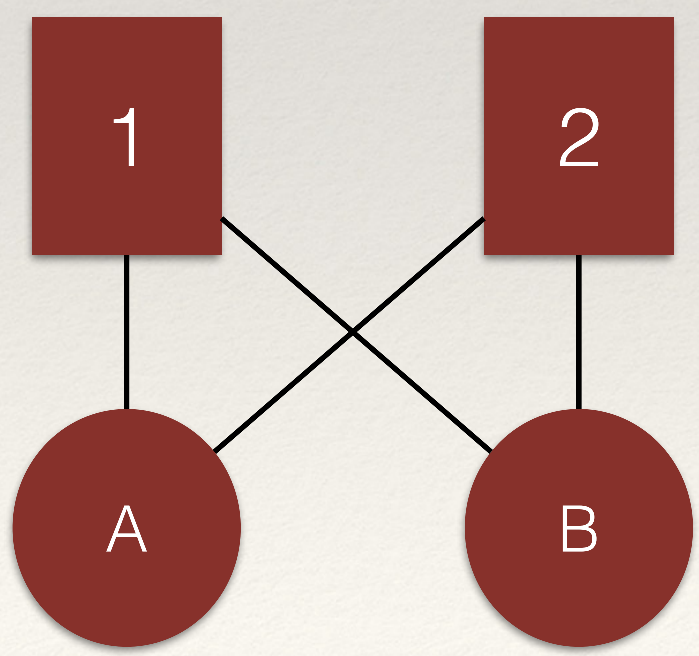
cycle



Dyadic Clustering: Bipartite Graphs

- Cycles in a graph create multiple ties between vertices in *both* modes.

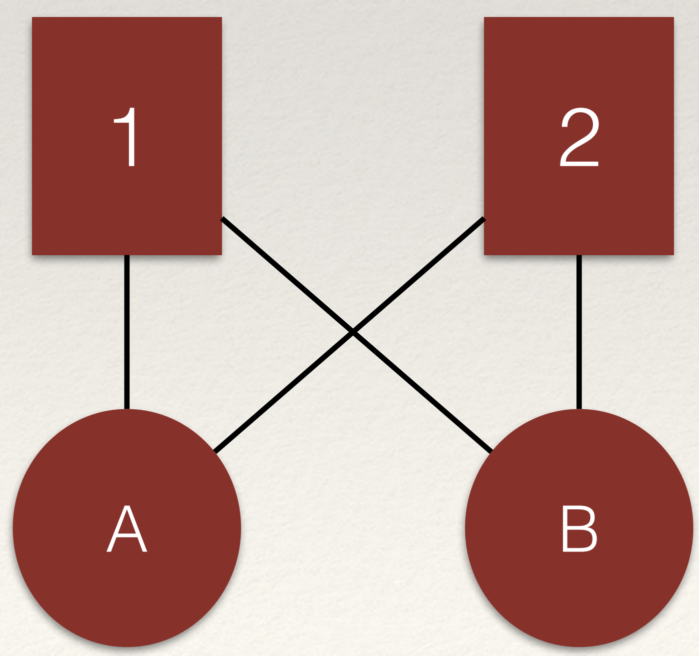
A and B
are both
linked
through
1 and 2



Dyadic Clustering: Bipartite Graphs

- Cycles in a graph create multiple ties between vertices in *both* modes.

A and B are both linked through 1 and 2



1 and 2 are both linked through A and B

Dyadic Clustering: Bipartite Graphs

- The ratio of cycles to 3-paths in a graph is proportional to the level of *dyadic clustering* (sometimes called *reinforcement*).
  - A value of 1 indicates that every 3-path is *closed* (i.e., is embedded in a cycle).
    - Networks with values at or close to 1 will have considerable redundancy in ties.

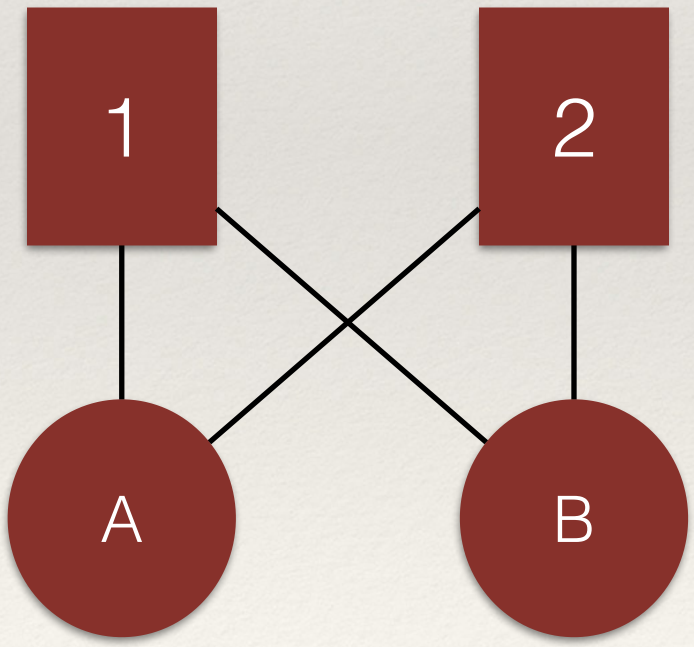
Dyadic Clustering: Bipartite Graphs

- Specifically, the dyadic clustering coefficient is the ratio of cycles X 4, divided by the number of 3-paths.

$$\frac{4 \times C_4}{L_3}$$

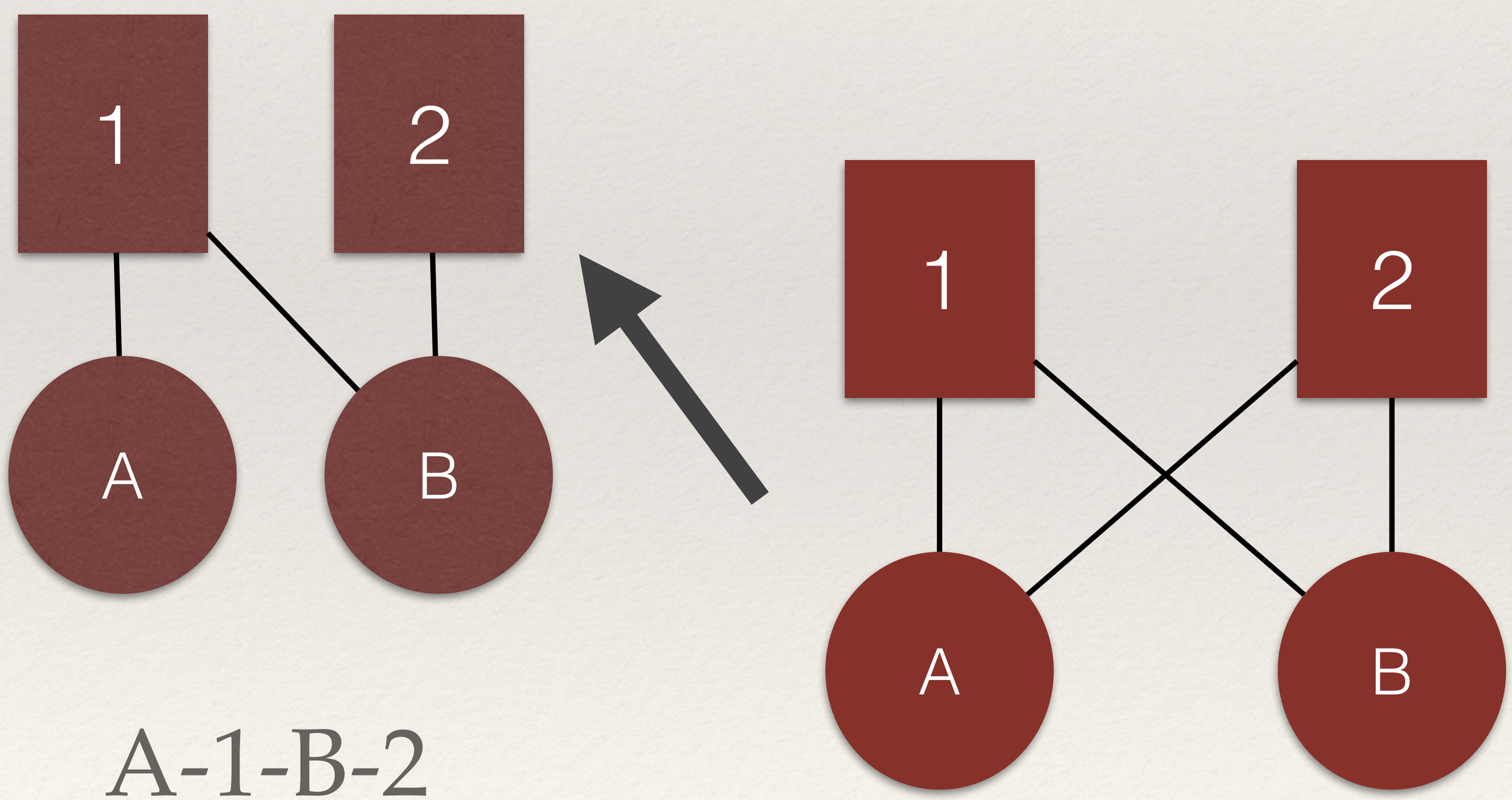
Dyadic Clustering: Bipartite Graphs

- In a cycle, there are four 3-paths.



Dyadic Clustering: Bipartite Graphs

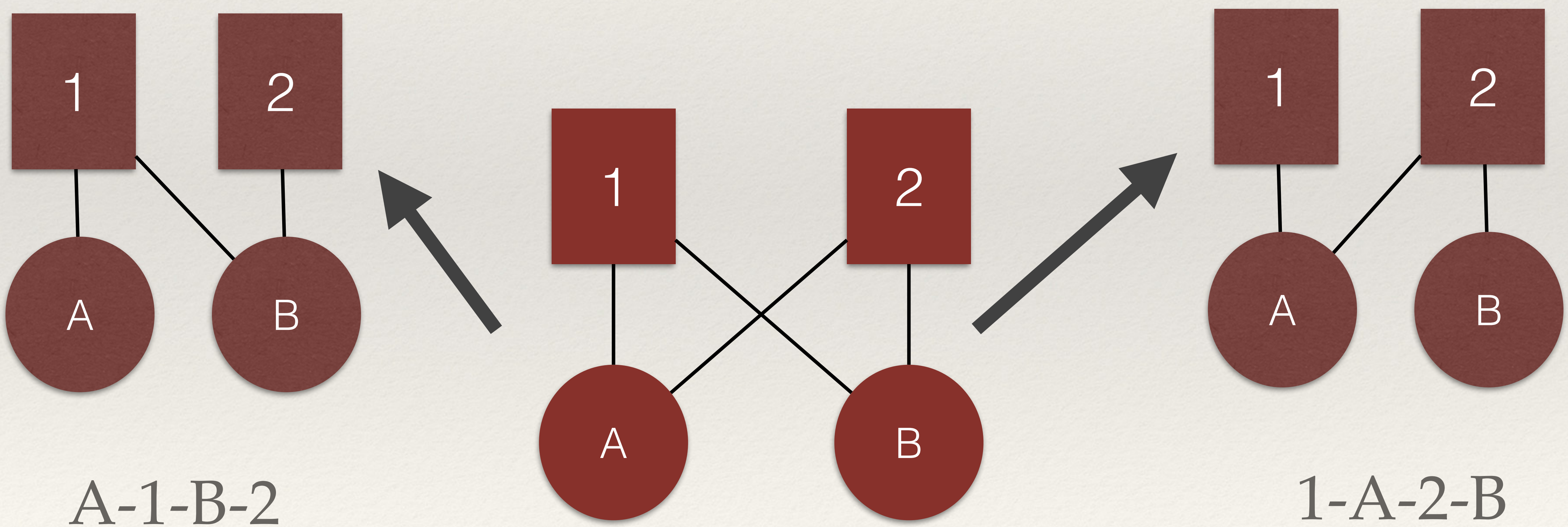
- In a cycle, there are four 3-paths.



A-1-B-2
2-B-1-A

Dyadic Clustering: Bipartite Graphs

- In a cycle, there are four 3-paths.

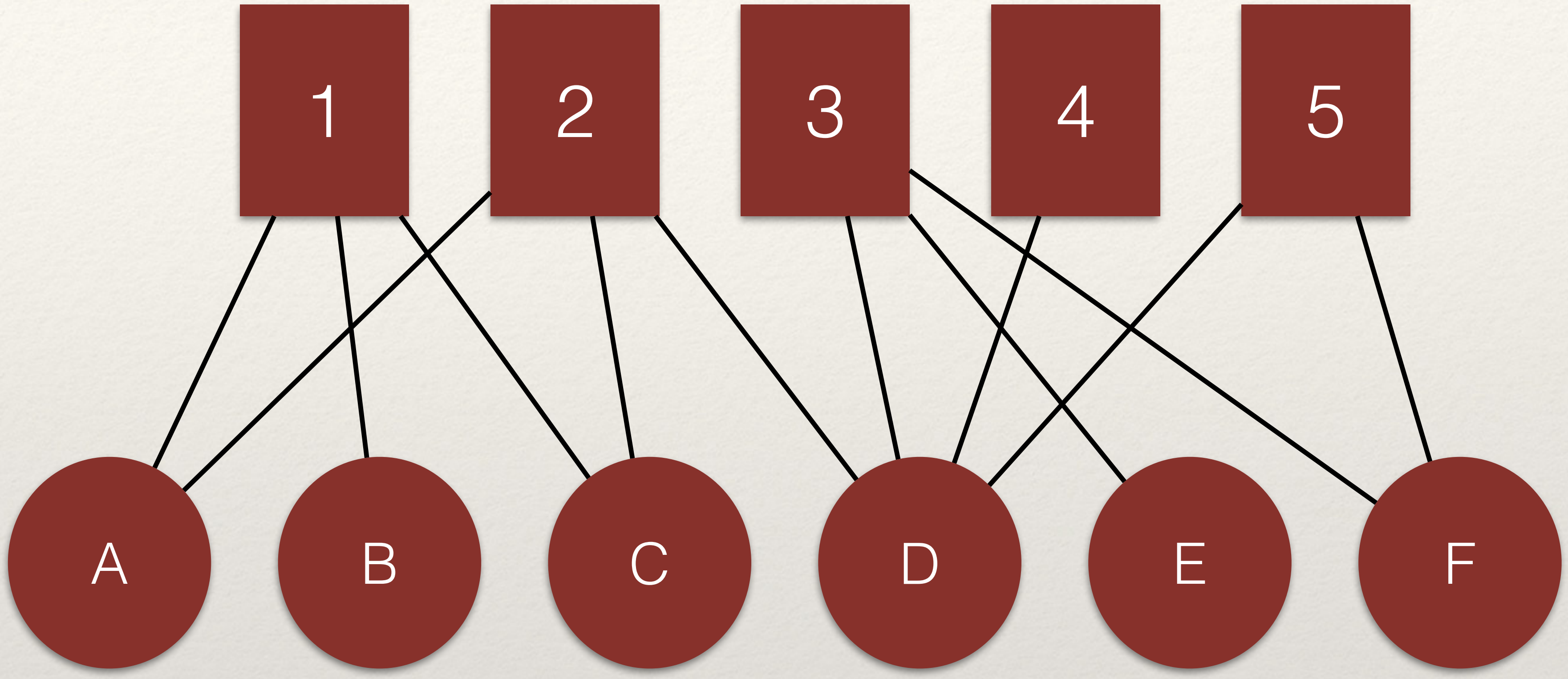


A-1-B-2
2-B-1-A

1-A-2-B
B-2-A-1

Example

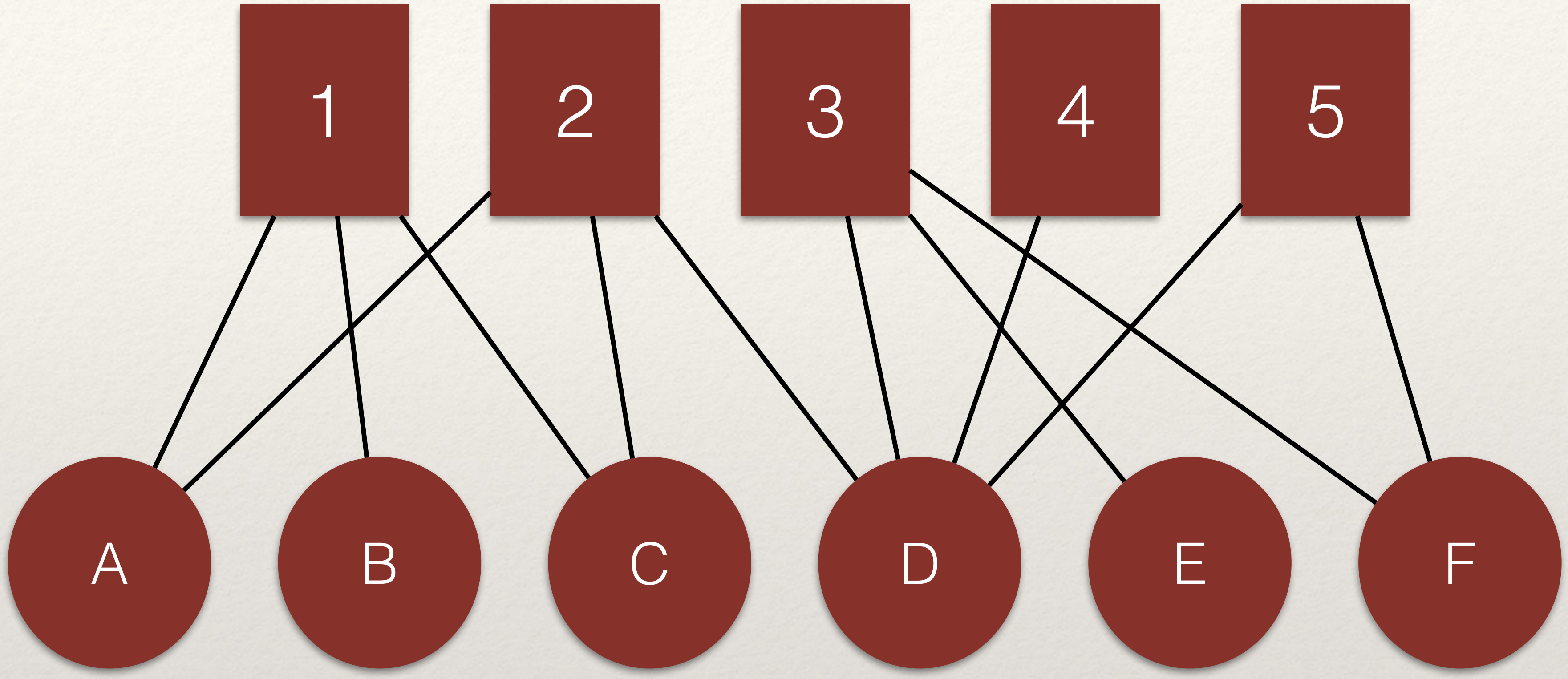
What is the ***dyadic clustering*** *for this graph?*



Example

*What is the **dyadic clustering** for this graph?*

0.307

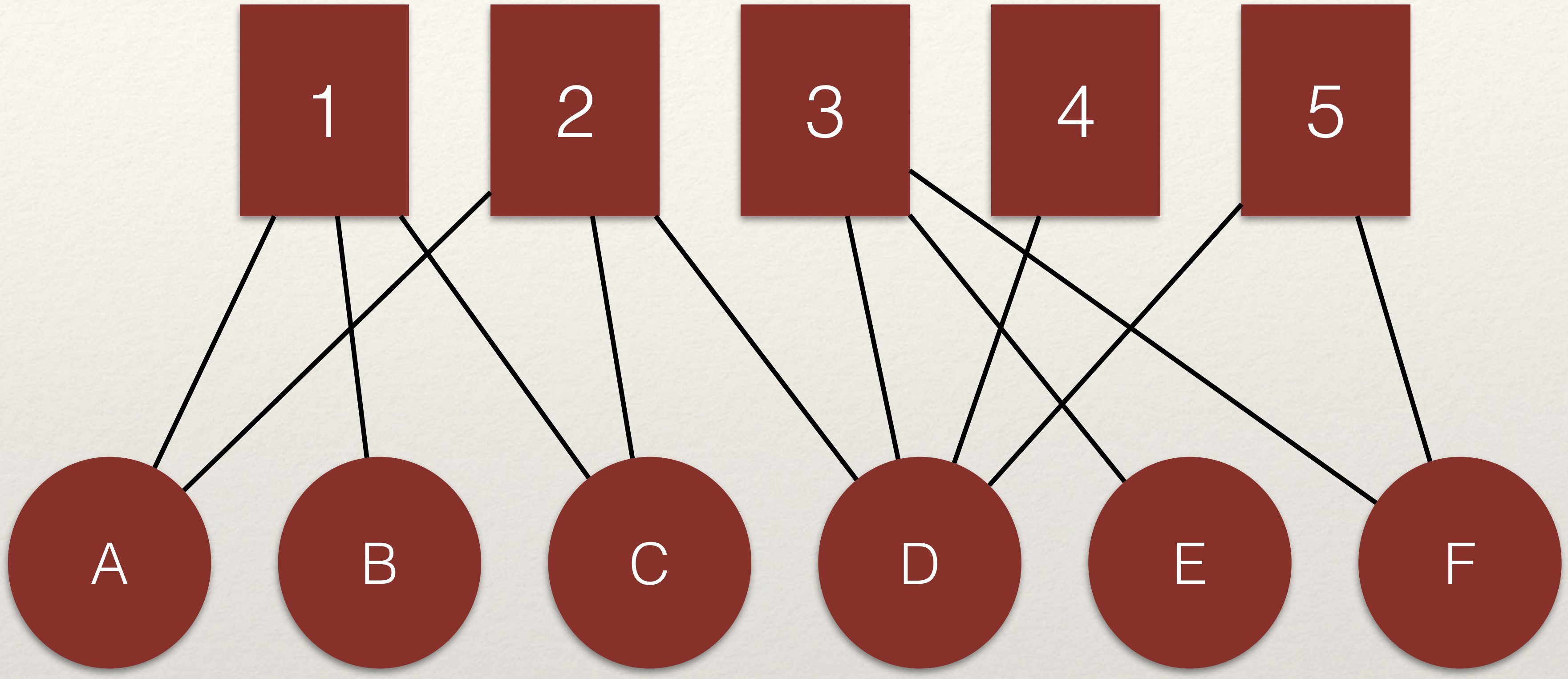


Example

What is the ***dyadic clustering*** *for this graph?*

0.307

What does a value of 0.307 mean?



Learning Goals

- At the end of the lecture, you should be able to answer these questions:
  - How are **bipartite** graphs different from **unipartite** graphs?
  - What are some structural properties of bipartite graphs that we can examine?

Questions?