Social Network Analysis

Two-mode Network

Dr. Chun-Hsiang Chan

Department of Geography,
National Taiwan Normal University

Photo credit: midjourney

Outline

– Network-level Measures
– Network Density
– Dyad
– Triad
– Transitivity
– Clustering
– Clustering Coefficient
– Centralization
– Core-peripheral Structure
– Two-mode Network
– Paper Reading
– References

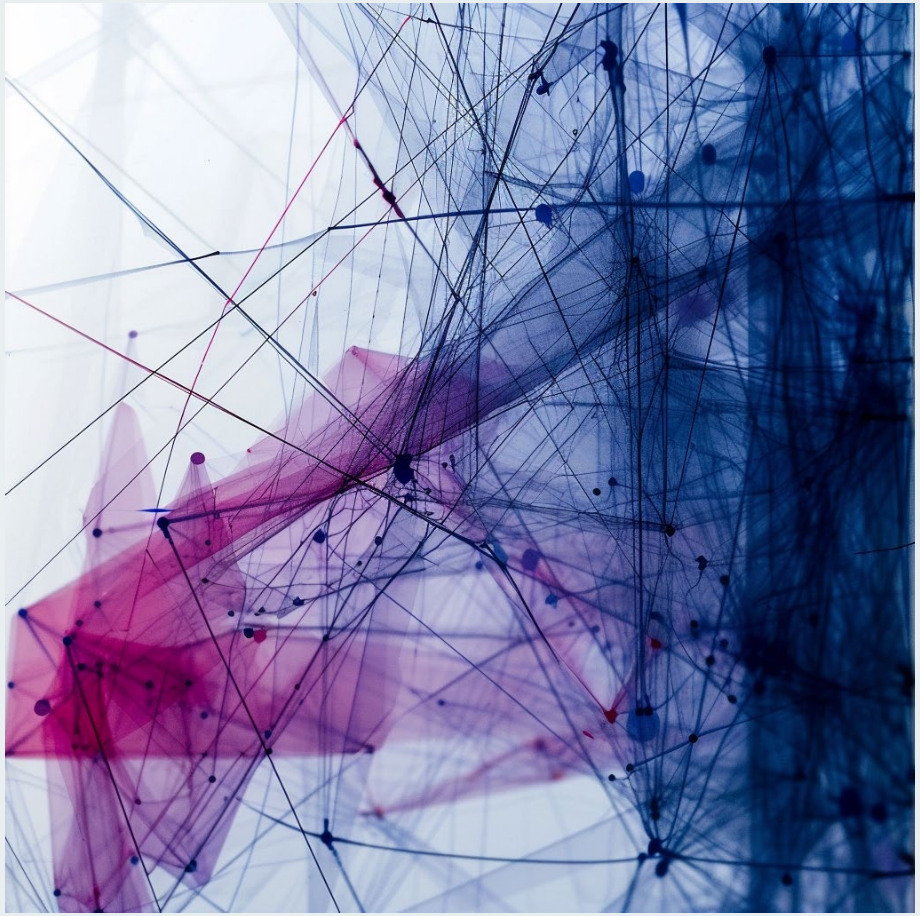


Photo credit: midjourney

Network-level Measures

– Descriptive features: size and density
– Dyadic relation: reciprocity/ mutuality
– Triadic relation: triadic census/ transitivity
– Network scale: average path distance
– Network pattern: clustering and centralization
– Network structure: core-periphery structure

Network Density

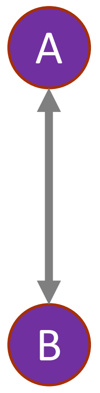
– **Network density** describes the portion of the potential connections in a network that are actual connections.

$$\binom{n}{2} = \frac{n(n-1)}{2}$$

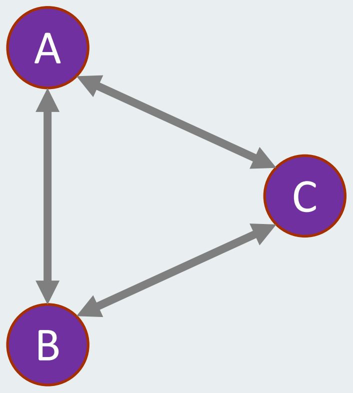
$$density = \frac{actual\ connections}{n(n-1)/2}$$

Dyad

- To characterize the relation of a pair…
  - Common analysis unit in sociology
  - Exist relation between two actors? Is it reciprocal?
  - Asymmetrical?
  - …
  - …
  - …



Direct Reciprocity



Direct and Indirect Reciprocity

Dyad – Mutuality/ Reciprocity

– Network-level measure

$$R = \frac{(A_{ij} = 1)\ and\ (A_{ji} = 1)}{(A_{ij} = 1)\ or\ (A_{ji} = 1)}$$

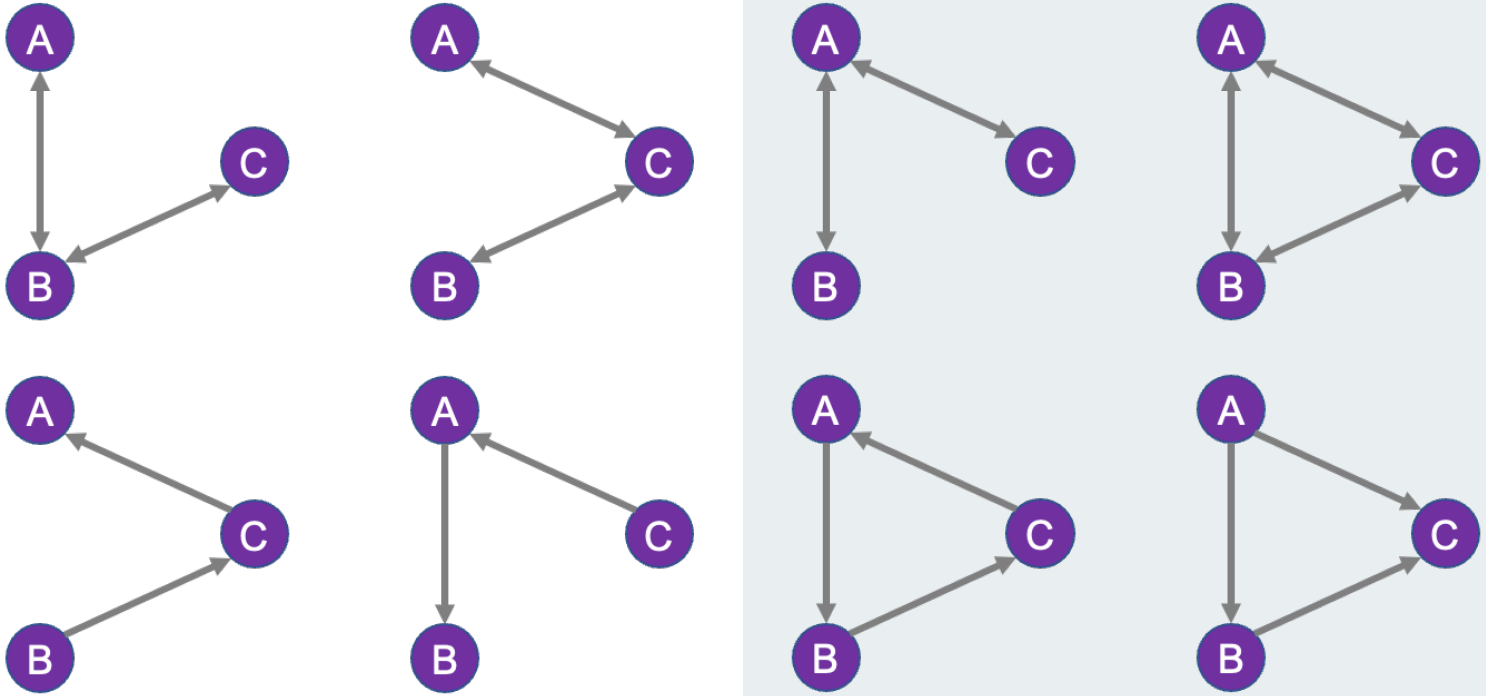
where A_{ij} indicates a link from i to j.

– High reciprocity network is easy to formulate a group and increases network distance.

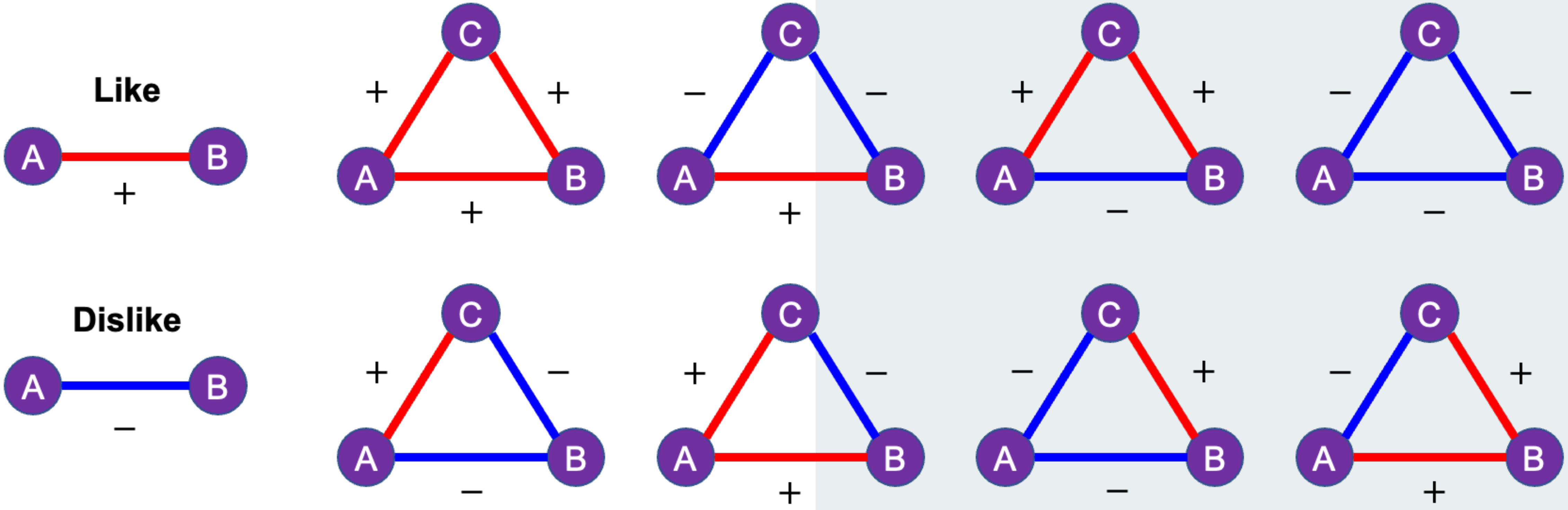
– Reciprocated ties may suggest "stronger ties", e.g., in friendship network, best friend ties (strong ties) are more likely to be reciprocated.

Triad

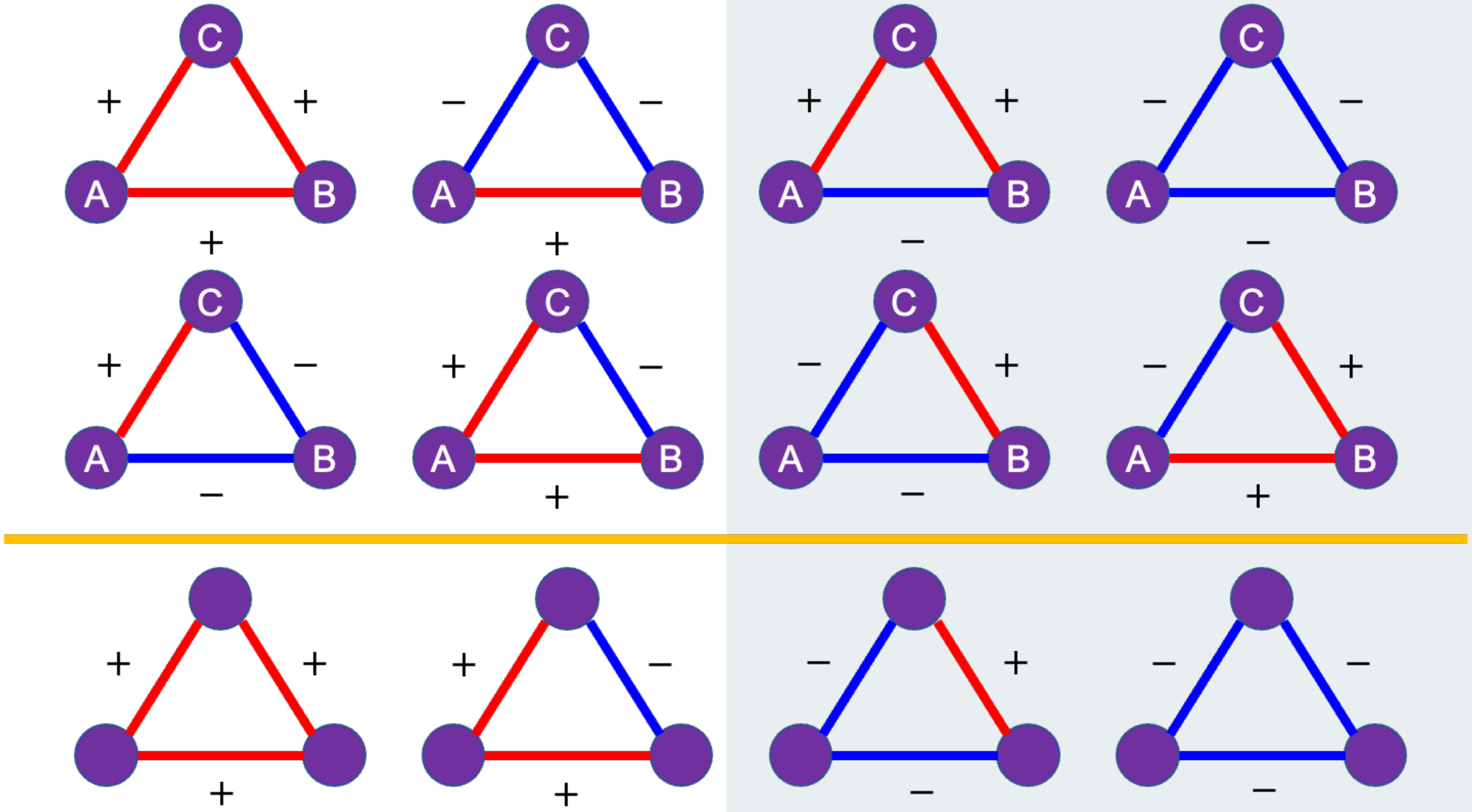
– From dyadic to triadic relations, …



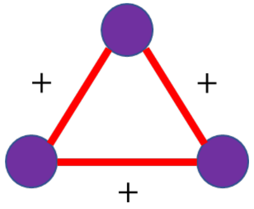
Triadic Relation



Triadic Relation



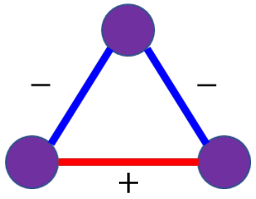
Triad and Balance



$(+)(+)(+) = (+)$

Balance

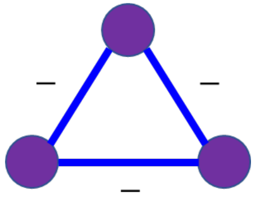
A friend of a friend is a friend.



$(-)(+)(-) = (-)$

Balance

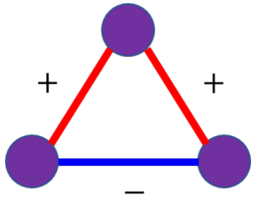
A enemy of my enemy is a friend.



$(-)(-)(-) = (-)$

Unbalance

A enemy of a my enemy is a enemy.

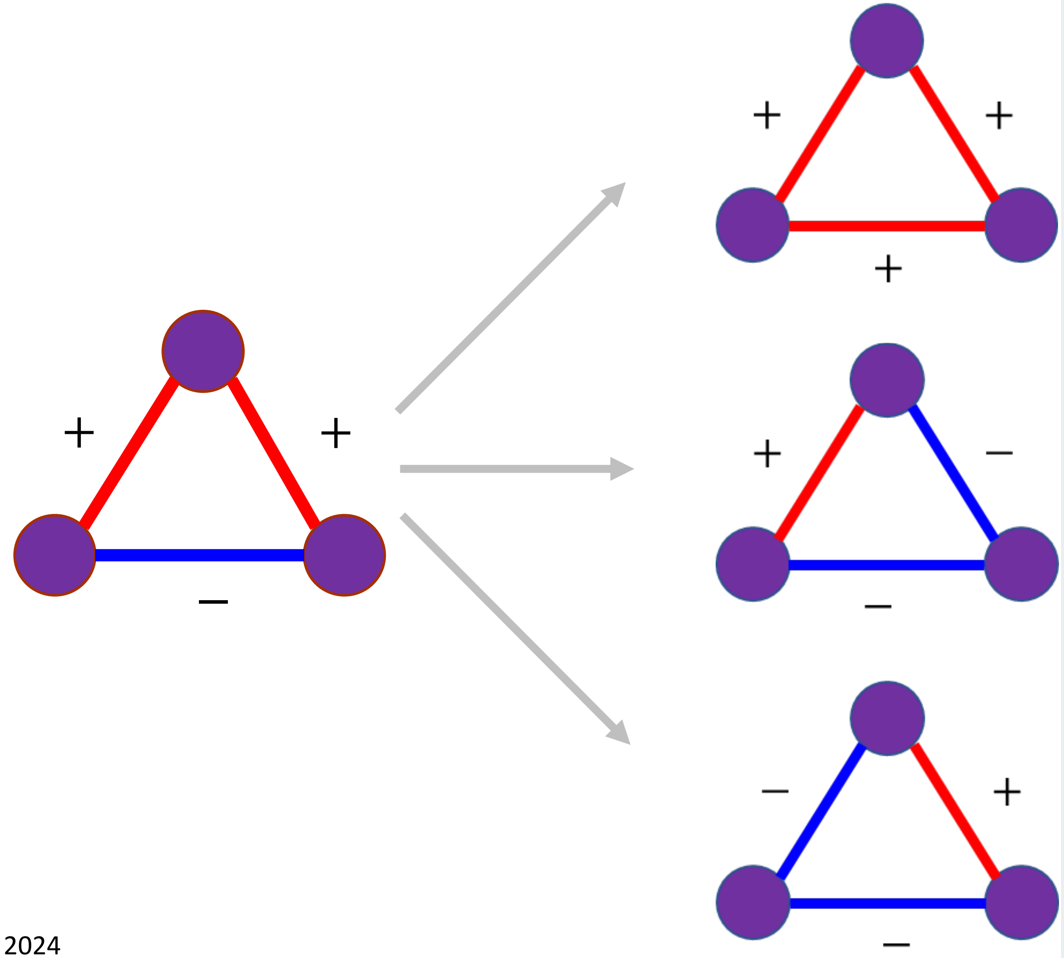


$(+)(-)(+) = (-)$

Unbalance

A friend of a friend is a enemy.

Triad and Balance



Become friends

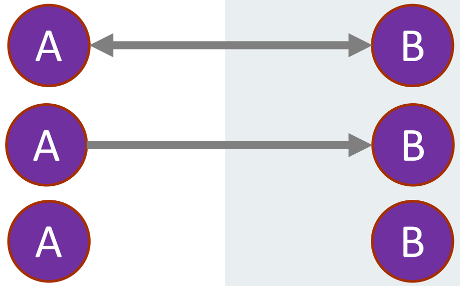
Become enemies

Become enemies

Directed Dyadic Relations

– Directed dyadic relations:

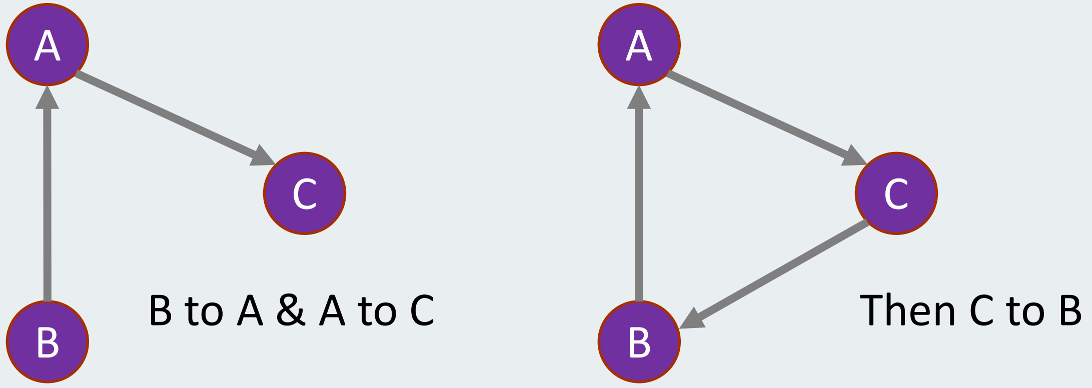
1) **Mutual**
2) **Asymmetric**
3) **Null**



– Every triad contains 3 dyads.

Transitivity

- Transitivity is a very important property in social networks.
- It refers to the degree of transitivity of the relationship between two nodes connected by an edge in the network.



- A property of triad
- Assume directed relations
- The saliency of a triad may differ for each other, depending on their position within the triad

Triad Types

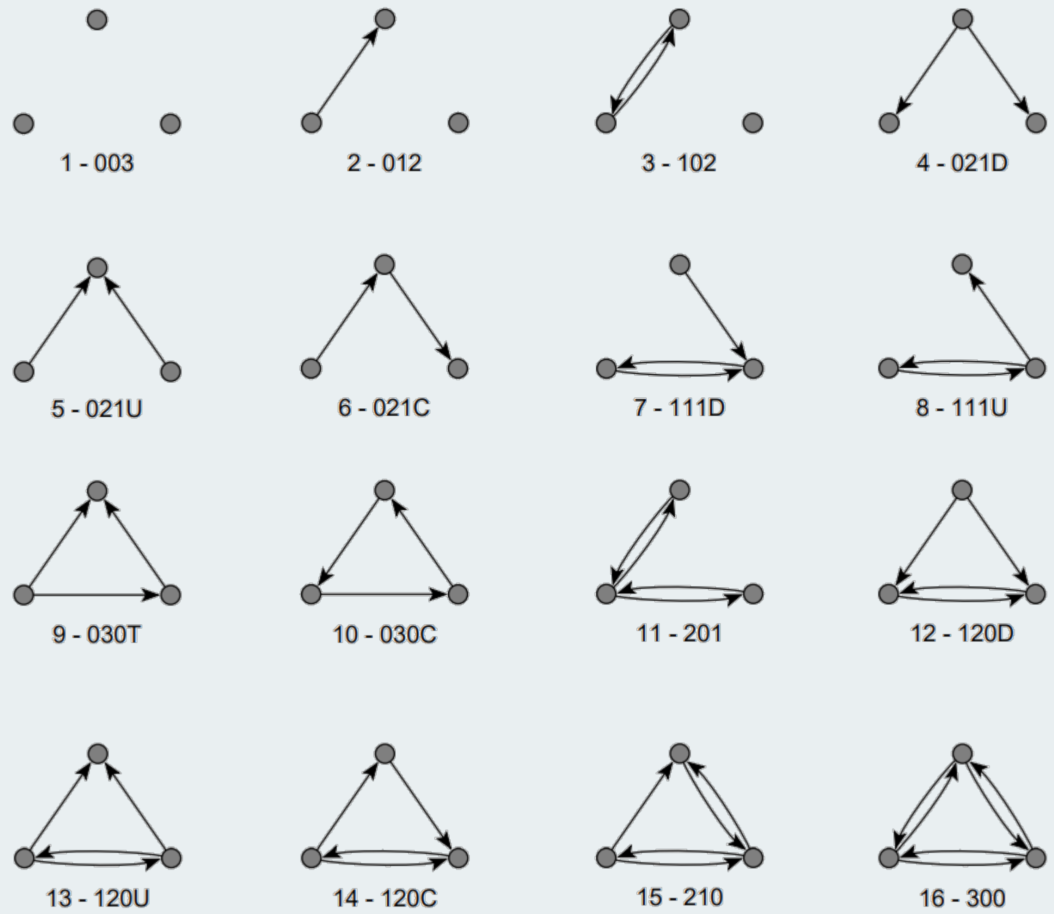
– **U:** up

– **D:** down

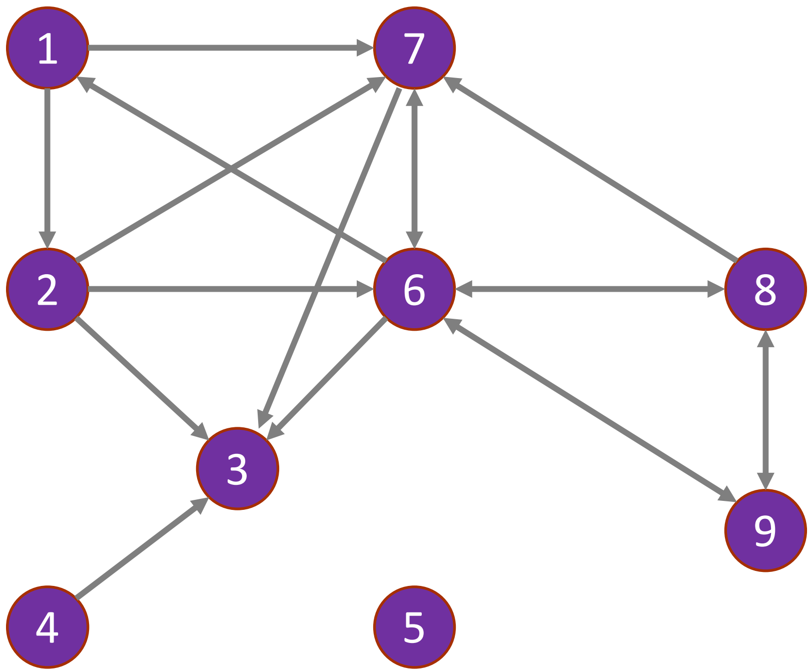
– **C:** cyclical

– **T:** transitive

http://vlado.fmf.uni-lj.si/pub/networks/doc/triads/triads.pdf



Triad Census



In [102]:
```
nx.triadic_census(g3)
```

```
{'003': 25,
 '012': 26,
 '102': 13,
 '021D': 1,
 '021U': 4,
 '021C': 2,
 '111D': 3,
 '111U': 4,
 '030T': 2,
 '030C': 0,
 '201': 1,
 '120D': 1,
 '120U': 0,
 '120C': 0,
 '210': 1,
 '300': 1}
```

Transitivity

[NetworkX]

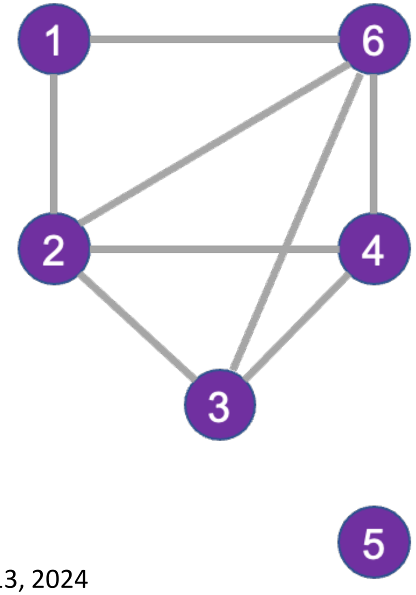
– Compute graph transitivity, the fraction of all possible triangles present in G.

– Possible triangles are identified by the number of “triads” (two edges with a shared vertex).

$$transitivity = T = 3 \times \frac{number\ of\ triangles}{number\ of\ triads}$$

Average Path Distance

– Average path distance is calculated by the average distance of arbitrary two nodes in the graph.



	1	2	3	4	5	6	total
1	0	1	2	2		1	6
2	1	0	1	1		1	4
3	2	1	0	1		1	7
4	2	1	1	0		1	6
5							
6	1	1	1	1		0	4

$= 27$

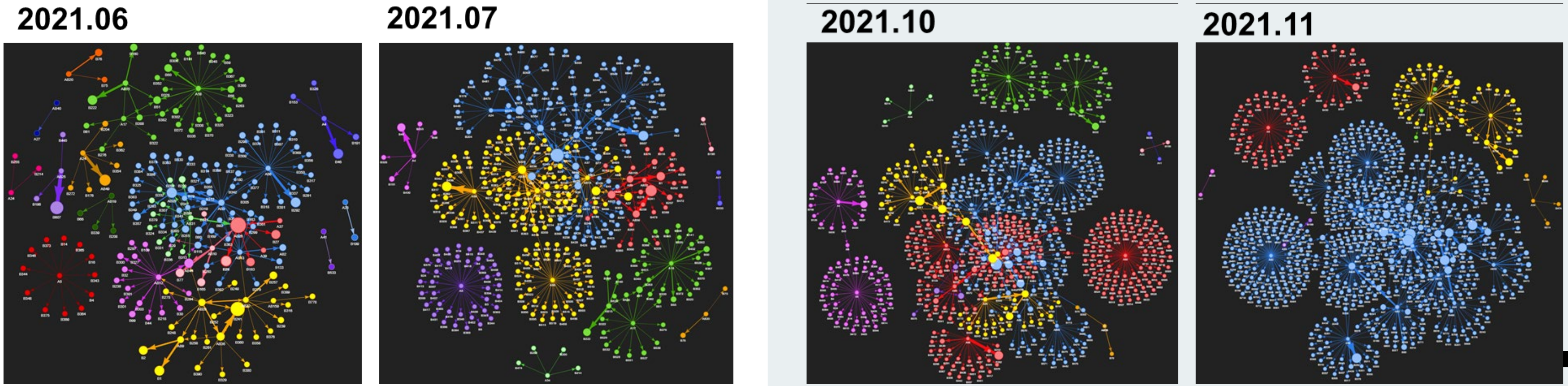
$$\frac{5 \times (5-1)}{2} = 10\ pair\ distance \Rightarrow average\ dist = \frac{27}{10} = 2.7$$

Problem: Density vs Average Distance

- Please generate two graphs (low and high density) and calculate the network density and average distance.
- Then, discuss the relationship between two network-level parameters.

Clustering

– **Clustering:** a set of nodes that are interconnected.



Clustering Coefficient

– For **unweighted** graphs, the clustering of a node u is the fraction of possible triangles through that node that exist,

$$c_u = \frac{2T(u)}{\deg(u)\,(\deg(u) - 1)}$$

– where $T(u)$ is the number of triangles through node u and $\deg(u)$ is the degree of u.

Clustering Coefficient

– For **weighted** graphs, there are several ways to define clustering. the one used here is defined as the geometric average of the subgraph edge weights,

$$c_u = \frac{2T(u)}{\deg(u)\,(\deg(u) - 1)} \sum_{vm} (\widehat{w_{uv}}\widehat{w_{uw}}\,\widehat{w_{vw}})^{\frac{1}{3}}$$

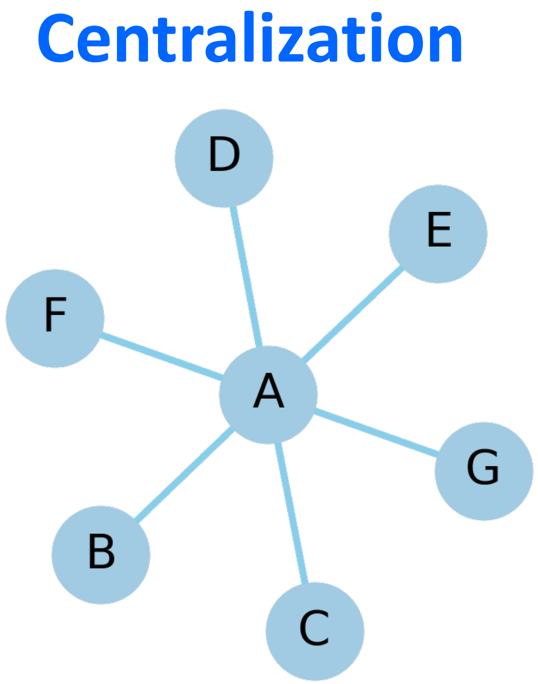
– The edge weights $\widehat{w_{uv}}$ are normalized by the maximum weight in the network $\widehat{w_{uv}} = w_{uv}/\max(w)$.

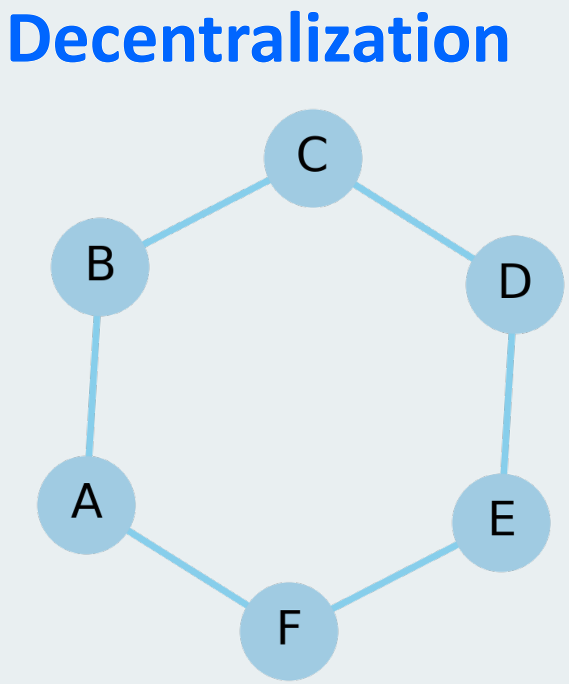
– The value of c_u is assigned to 0 if $\deg(u) < 2$.

– Additionally, this weighted definition has been generalized to support negative edge weights.

 https://networkx.org/documentation/stable/reference/algorithms/generated/networkx.algorithms.cluster.clustering.html#networkx.algorithms.cluster.clustering

Centralization





Centralization

- **Centralization** measurement:

$$CD = \frac{\sum(\max(C_{Di}) - C_{Di})}{n^2 - 3n + 2}$$

- Where $\max(C_{Di})$ is the maximum centrality score in the network, C_{Di} refers to the individual centrality score, and n represents the network size.
- Or by simply calculating the standard deviation of the centrality scores for the network.
- The higher the standard deviation ➔ more heterogenous centrality scores
- Central persons usually are "gatekeepers".

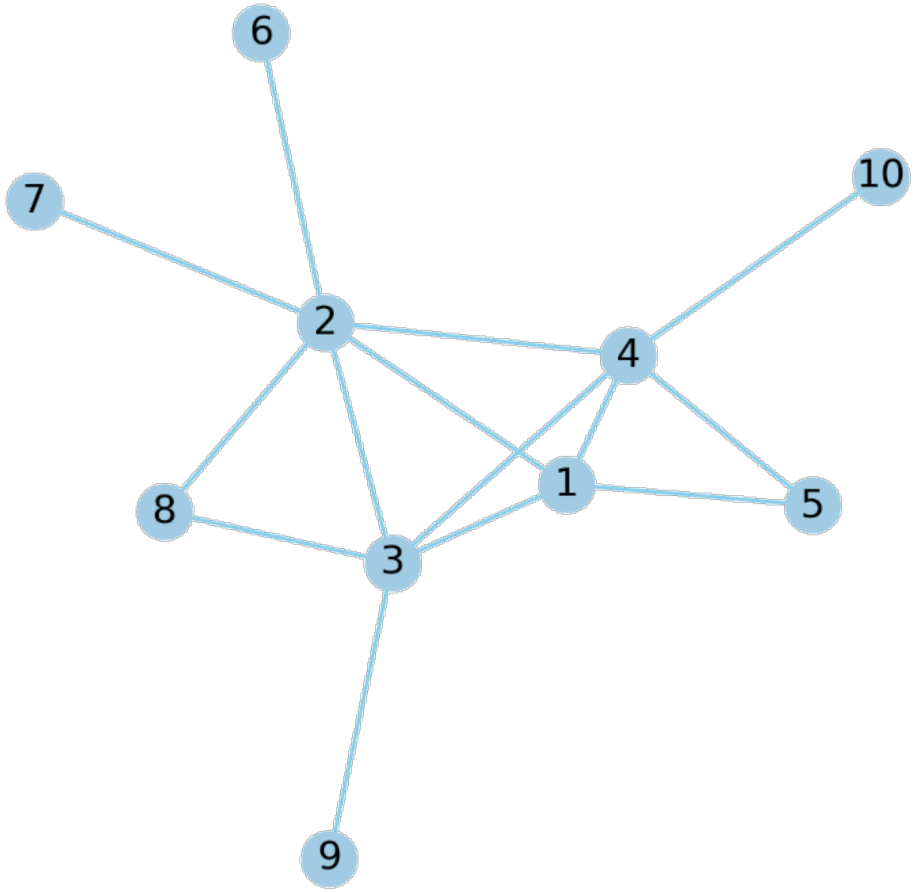
Core-periphery Structure

– In a core-periphery network, the nodes are partitioned into two categories:

– **Core** – every two actors are connected.

– **Periphery** – every two actors are **not** connected.

Core-periphery Structure



	1	2	3	4	5	6	7	8	9	10
1		1	1	1	1	0	0	0	0	0
2	1		1	1	0	1	1	1	0	0
3	1	1		1	0	0	0	1	1	0
4	1	1	1		1	0	0	0	0	1
5	1	0	0	1		0	0	0	0	0
6	0	1	0	0	0		0	0	0	0
7	0	1	0	0	0	0		0	0	0
8	0	1	1	0	0	0	0		0	0
9	0	0	1	0	0	0	0	0		0
10	0	0	0	1	0	0	0	0	0	

Two-mode Network

– Two-mode data are formulated from the co-events, co-organization, and co-situation.

– **Two-mode network (affiliation network):**
People are the rows, and the events or organizations are the columns.

	Event A	Event B	Event C
Person 1	1	1	0
Person 2	1	0	1
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
Person N	1	1	0

Two-mode Network

– Engaging in events together fosters interaction among members and boosts the likelihood of cultivating mutually beneficial relationships.

– Simultaneous participation in multiple events creates a connection between them. Shared group identities facilitate information exchange and offer opportunities for collaborative efforts, such as having a joint board of directors to promote coordination between firms.

– The two-mode network has duality characteristics.

Two-mode Network

– $a_{ik} = \begin{cases} 1 & actor\ i\ participate\ event\ k \\ 0 & \end{cases}$

– Given g actors and h events

	Event A	Event B	Event C
1	0	1	0
2	1	0	1
3	1	1	0
4	0	1	1

The event list of No. 2 actor

All participants of Event B

Please list all participants of each event

Event A

Event B

Event C

Two-mode Network

	Event A	Event B	Event C	Row Total
1	0	1	0	1
2	1	0	1	2
3	1	1	0	2
4	0	1	1	2
Column Total	2	3	2	7

$\{a_{i\cdot}\}$

Row Total represents the number of participated events for each person

7/4 = participate 2.33 events/ person

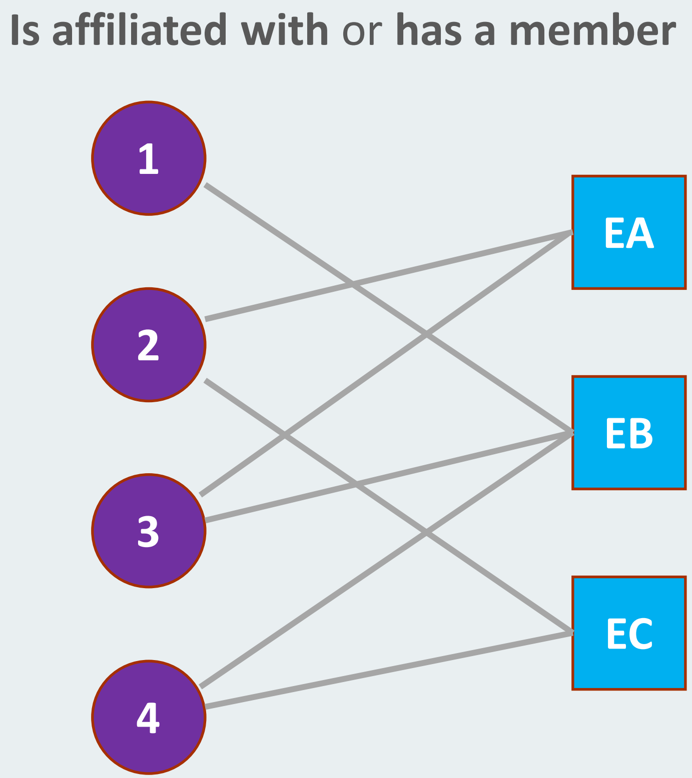
$\{a_{\cdot k}\}$

Column Total indicates the number of participants for each event

7/3 = overlapping 2.33 persons/ event

Two-mode Network

	Event A	Event B	Event C	Row Total
1	0	1	0	1
2	1	0	1	2
3	1	1	0	2
4	0	1	1	2
Column Total	2	3	2	7



Two-mode Network: find the co-participate events

– x_{ij}^{N} is the number of events, which actor i and actor j participated. N is the set of all actors.
– The minimum value is 0, where the maximum value is h.

$x_{ij}^{N} = x_{ji}^{N}$ symmetric relationship

$$x_{ij}^{N} = \sum_{k=1}^{h} a_{ik} a_{jk} \text{ , } where\ h\ is\ the\ number\ of\ events$$

Two-mode Network: find the co-participate events

	Event A	Event B	Event C
1	0	1	0
2	1	0	1
3	1	1	0
4	0	1	1

×

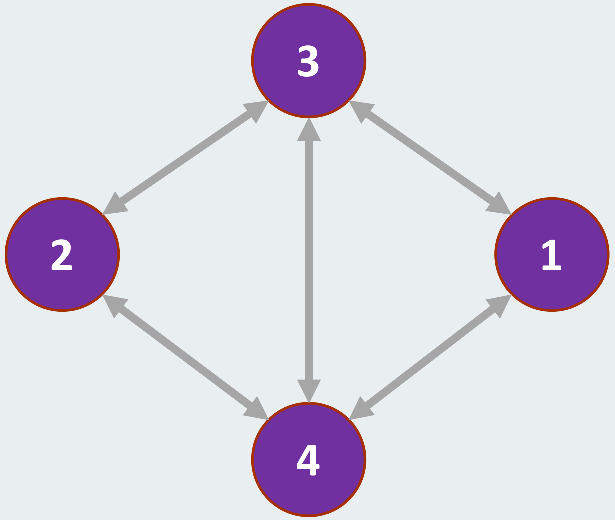
	1	2	3	4
Event A	0	1	1	0
Event B	1	0	1	1
Event C	0	1	0	1

$$X^N = AA^T$$

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

Co-event times

Event times



Two-mode Network: find the co-participate events

– x_{kl}^{M} is the participant who participate both event k and event l, where M is the set of events.
– Minimum value is 0.
– Maximum value is g.

$x_{kl}^{M} = x_{lk}^{M}$ symmetric relationship

$$x_{kl}^{M} = \sum_{i=1}^{g} a_{ik} a_{il} \text{, } where\ g\ is\ the\ number\ of\ actors$$

Two-mode Network: find the co-participate events

	1	2	3	4
Event A	0	1	1	0
Event B	1	0	1	1
Event C	0	1	0	1

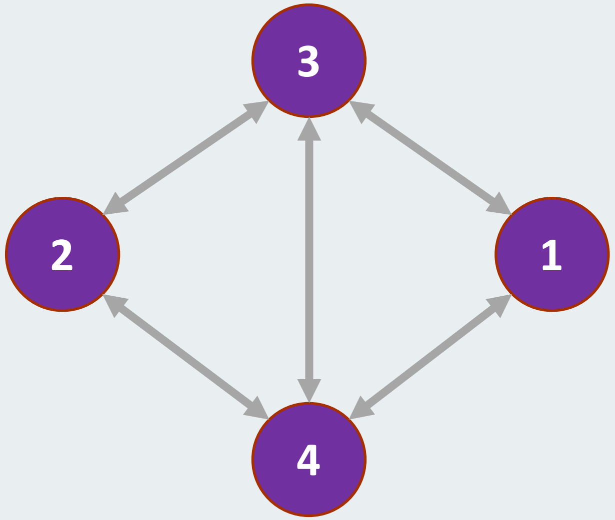
×

	Event A	Event B	Event C
1	0	1	0
2	1	0	1
3	1	1	0
4	0	1	1

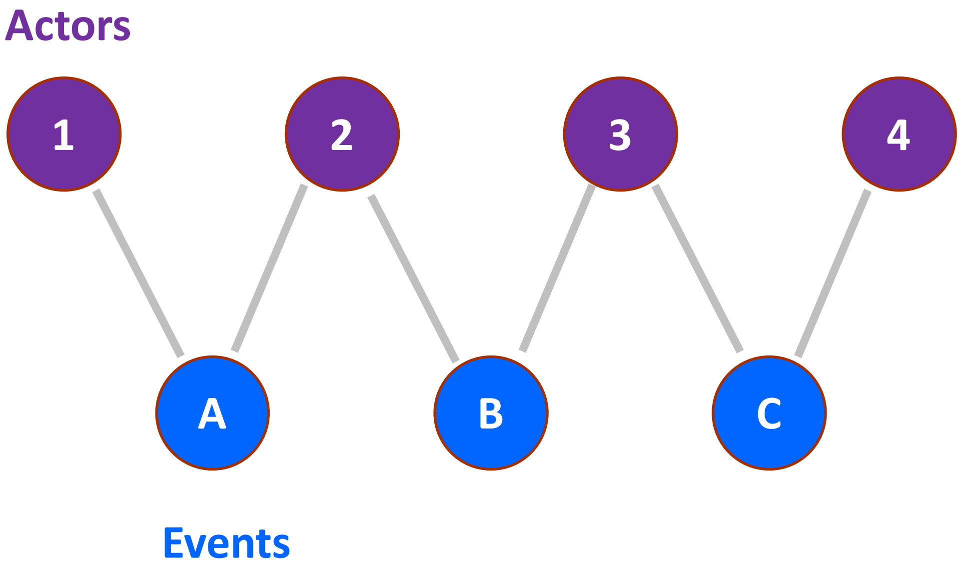
$$X^N = A^T A$$

	Event A	Event B	Event C
Event A	2	1	1
Event B	1	3	1
Event C	1	1	2

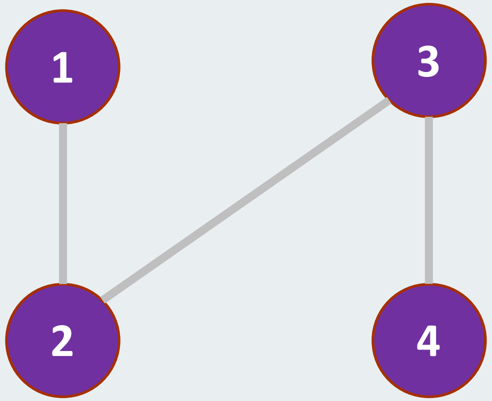
Participants in Event C



Two-mode Network: find the co-participate events



Two-mode network



One-mode network

Two-mode Network Measures

Rates of Participation

– Average number of events with which actors are affiliated.

– The diagonal summation of X^N or the row totals of A.

$$\overline{a_{i+}} = \frac{\sum_i^g \sum_j^h a_{ij}}{g} = \frac{a_{++}}{g} = \sum_i^g \frac{x_{ii}^N}{g}$$

Two-mode Network Measures

Size of Events

– Average number of events with which actors are affiliated.

– The diagonal summation of X^M or the column totals of A.

$$\overline{\boldsymbol{a_{+j}}} = \frac{\sum_i^g \sum_j^h \boldsymbol{a_{ij}}}{\boldsymbol{h}} = \frac{\boldsymbol{a_{++}}}{\boldsymbol{h}} = \sum_{\boldsymbol{ih}} \frac{\boldsymbol{x_{jj}^M}}{\boldsymbol{h}}$$

Two-mode Network Measures

– **Density of Person-by-Person Matrix**

$$\Delta_N = \frac{\sum_{i=1}^{g}\sum_{j=1}^{g} x_{ij}^N}{g(g-1)}, where\ i \neq j$$

Density: average co-participated events of each member pair

$$Min = 0;\ Max = h$$

If we dichotomize co-participated events, and then the physical meaning of density will be “the proportion of actors who share membership (co-membership) in any event”

Two-mode Network Measures

– **Density of Event-by-Event Matrix**

$$\Delta_M = \frac{\sum_{k=1}^{h}\sum_{l=1}^{h} x_{kl}^{M}}{h(h-1)}, where\ k \neq l$$

Density: average co-member of each event

$$Min = 0;\ Max = g$$

If we dichotomize co-member, and then the physical meaning of density will be "the proportion of events with co-members"

Paper Reading

ELSEVIER

Applied Mathematical Modelling

Volume 43, March 2017, Pages 207-220

Two-mode network modeling and analysis of dengue epidemic behavior in Gombak, Malaysia

Hafiz Abid Mahmood Malik [a], Abdul Waheed Mahesar [a], Faiza Abid [b], Ahmad Waqas [c], Mohamed Ridza Wahiddin [a]

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.apm.2016.10.060

Get rights and content

Malik, H. A. M., Mahesar, A. W., Abid, F., Waqas, A., & Wahiddin, M. R. (2017). Two-mode network modeling and analysis of dengue epidemic behavior in Gombak, Malaysia. Applied Mathematical Modelling, 43, 207-220.

Questions:

1. What is the objective of this paper?
2. How did the authors formulate the network and quantify the dengue epidemic behavior characteristics with nodal indicators?
3. What are the findings of this study?
4. If you want to achieve the same objective, how do you formulate the network?

References

– Wasserman, S., & Faust, K. (1994). Social network analysis: Methods and applications.
– Tsvetovat, M., & Kouznetsov, A. (2011). Social Network Analysis for Startups: Finding connections on the social web. " O'Reilly Media, Inc.".
– Hanneman, R. A., & Riddle, M. (2005). Introduction to social network methods.
– https://en.wikipedia.org/wiki/Social_network

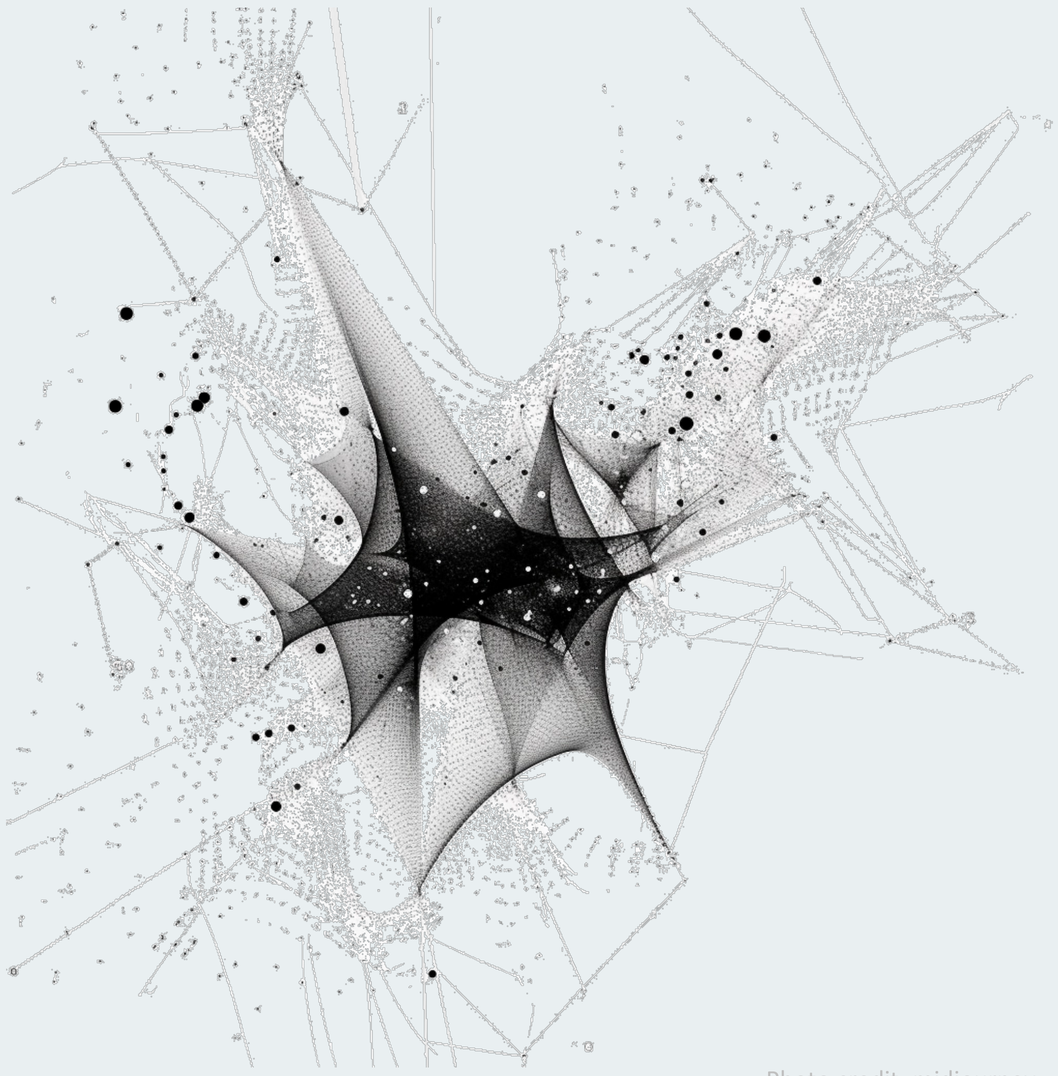


Photo credit: midjourney

The End

Thank you for your attention!

Email: chchan@ntnu.edu.tw
Website: toodou.github.io

Photo credit: midjourney