Social Groups

Saptarshi Pyne
Assistant Professor
Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur, Rajasthan, India 342030

CSL7390 Social Network Analysis Lectures 9-12
January 29th-February 7th, 2024

What we discussed in the last class

- How to measure the importance or ‘centrality’ of a vertex in a given network?

Groups of vertices

In social networks, vertices (i.e. social actors) form groups with other vertices to meet their needs.

Examples of groups of vertices

- Family
- Friends
- Neighbours
- Coworkers
- Collaborators
- Business partners
- Communities
- Political parties
- Social media channels belonging to the same domain
- Functional modules of smart sensors
- Fleet of autonomous vehicles
- etc.

Types of groups in an undirected network

- k-clique
- k-clan
- k-core
- k-component

k-clique vs k-clan

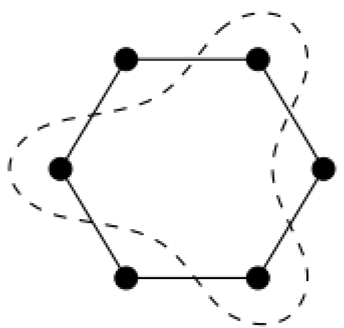
A k-clique is a maximal subset of vertices such that each vertex is at most k hops away from other vertices in the subset (we are only considering the shortest paths between two vertices).

A k-clan is a k-clique where the said hops go through only the vertices within the k-clique.

The figure shows a 2-clan of 6 vertices which is also a 3-clique of 6 vertices.

The vertices within the dashed area is no longer a 2-clan. However, it is a 2-clique.

Q. Is an 1-clique always an 1-clan?



1-clique or simply ‘clique’

A maximal subset of vertices such that each vertex is connected to all other vertices in the subset.

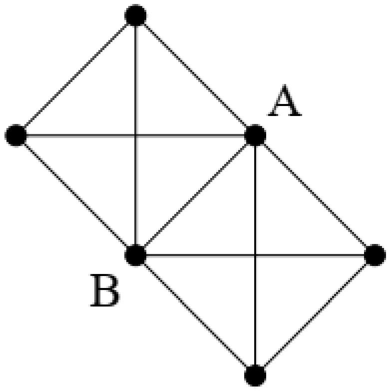
Maximal = If we add any vertex (from the remaining network) to the clique, the clique will no longer be a clique.

Example:

A friend circle where everyone is direct friends with everyone else. No mutual friendships.

Overlapping cliques

Cliques can overlap i.e. share common vertices.



Two overlapping cliques. Vertices A and B in this network both belong to two cliques of four vertices.

k-core

A k-core is a maximal subset of vertices such that each vertex is connected to at least k other vertices in the subset.

Maximal = If we add any vertex (from the remaining network) to the k-core, the k-core will no longer be a k-core.

Example: The students of an elective course where every student may not know all other students but he/she knows at least a few other students.

Two k-cores can not overlap

If a k-core of size n_1 overlaps with another k-core of size n_2, it will simply form a larger k-core of size $(n_1 + n_2 – c)$ where c = the number of common vertices.

A simple algorithm to find all k-cores in a network

Input arguments: Graph G = (V, E) and value of k

Do

    Remove all vertices with degree < k;

    // It may introduce new vertices with degree < k,   // hence we need to reiterate

While (no vertices were removed in the current iteration);

Do

    Randomly select a vertex;

    Find all the vertices in its k-core; // k = User input

    Remove these vertices from the network;

While (no vertices are remaining);

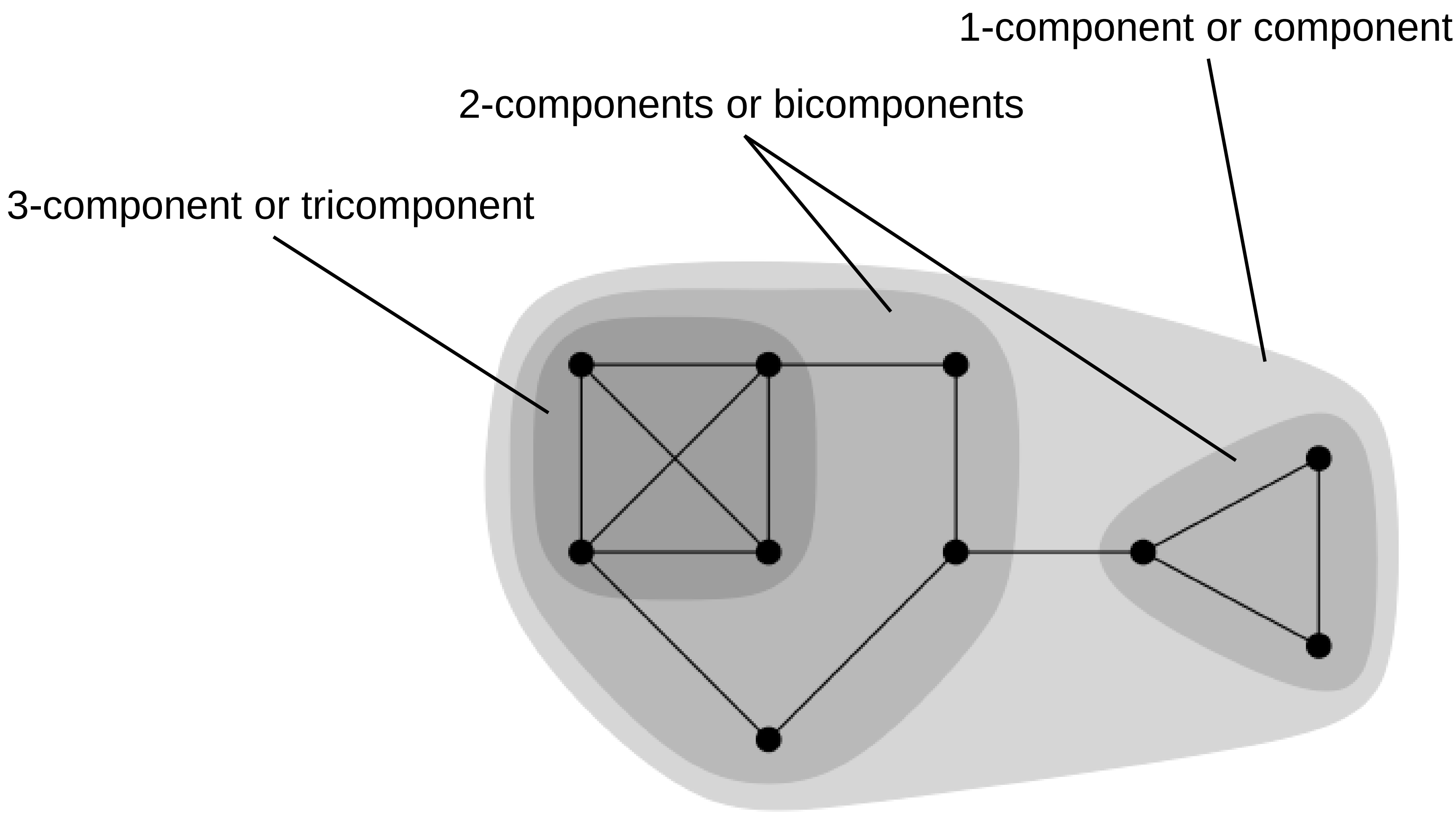
Component and k-component

A component is a maximal subset of vertices such that each vertex is reachable from all other vertices.

A k-component is a maximal subset of vertices such that each vertex is reachable from each of the other vertices by at least k ‘vertex-independent’ paths.

Vertex-independent paths = Paths that share no common vertices except the starting and ending vertices.

Examples of k-components



Application of k-component

Q. How many vertices do we need to remove to disconnect two given vertices?

=

Q. What is the size of the 'vertex cut set' between the two given vertices?

≥

Q. What is the value of k in a k-component that includes these two vertices?

Application of k-component (contd.)

Q. How robust is a given network?

=

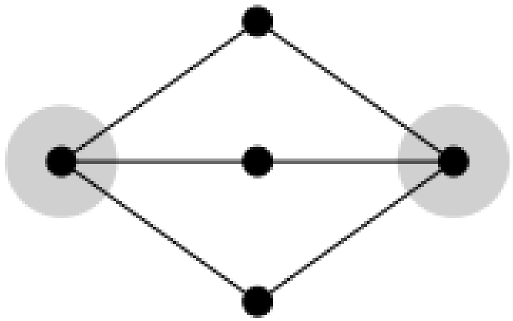
Q. How high is the value of k if the network is a k-component?

Examples:

Placing network routers. (defensive)

Breaching adversarial networks. (offensive)

k-component does not mandate internal paths



Unlike k-clan, we do not require the paths to go through the k-component.

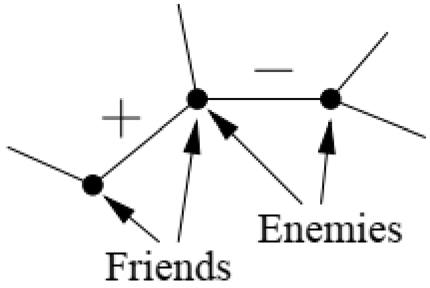
Hence, the highlighted vertices form a tricomponent.

Q. Is a k-component always a (k-1)-component where k ≥ 2?

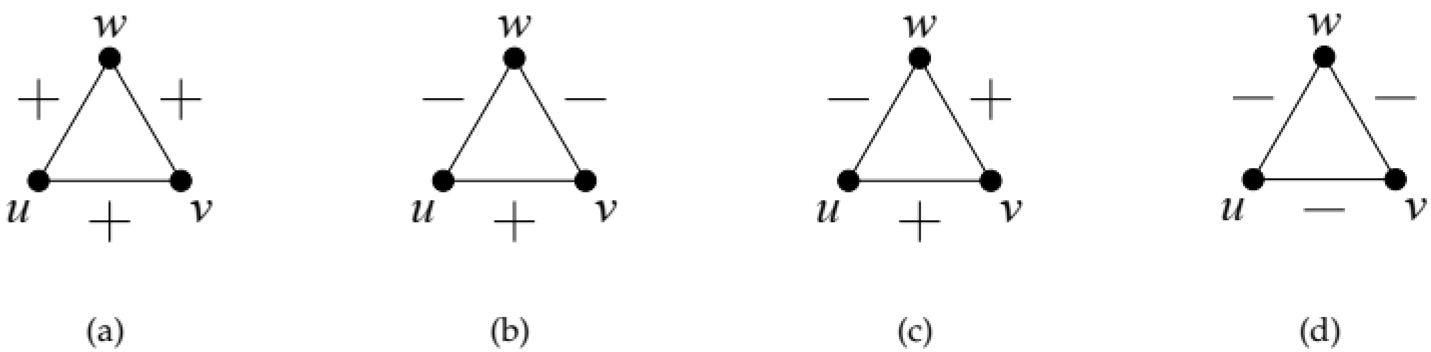
The effect of group formation (or groupism)

- Structural balance

‘Structural balance’ in signed undirected networks



A signed network



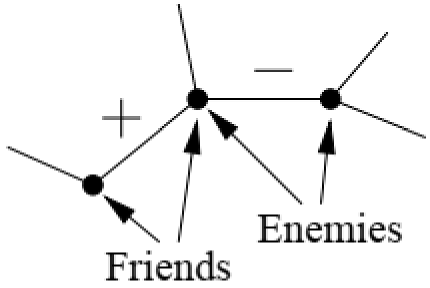
In 1953, Frank Harary proved that a structurally balanced or simply ‘balanced’ network can be divided into connected groups of vertices such that

(i) all connections between members of the **same group** are positive and

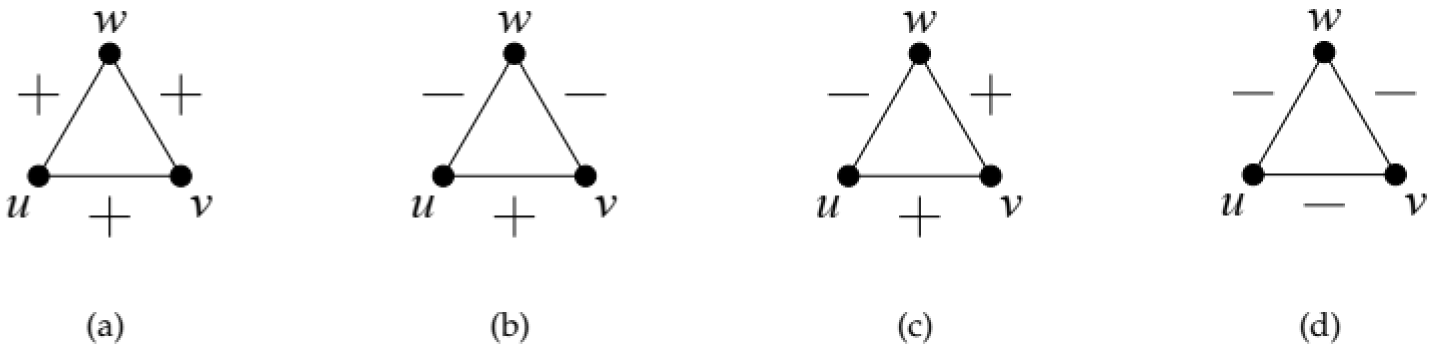
(ii) all connections between members of **different groups** are negative [1].

[1] Frank Harary, Michigan Math. Journal, 1953

‘Structural balance’ in signed undirected networks

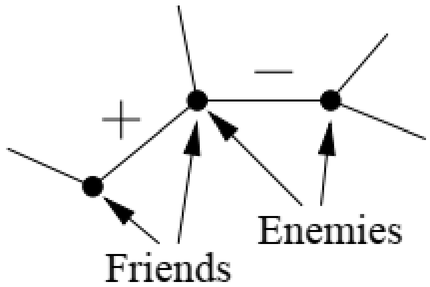


A signed network

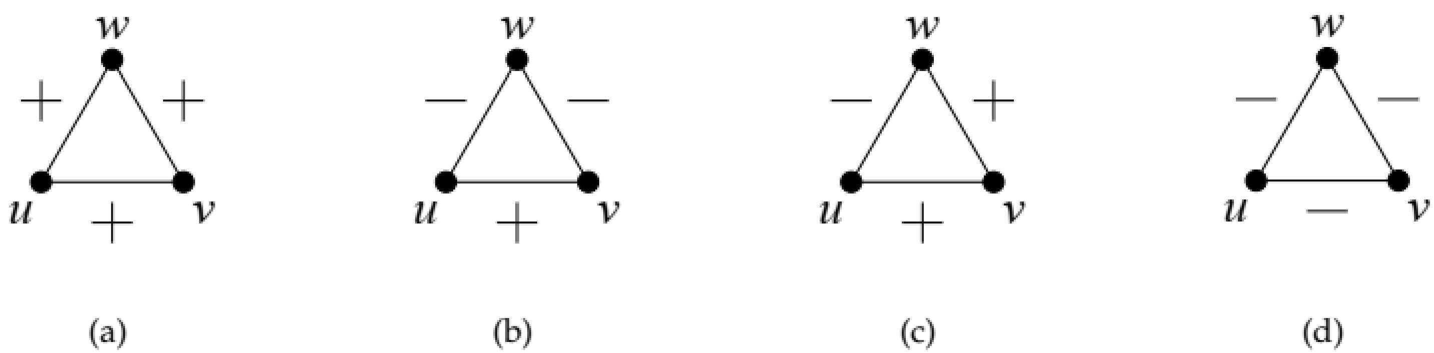


Loops with even number of negative signs are always balanced.

‘Structural balance’ in signed undirected networks



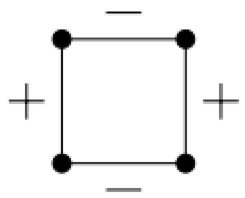
A signed network

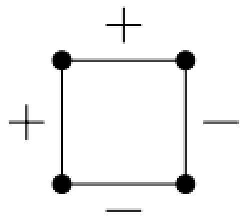


In network (d), any two vertices could become temporary allies to defeat the third vertex.

Unbalanced structures are extremely unreliable. The sign of an edge could change any time. The network may lose some or all of the edges anytime.

Signed loops with four vertices





Some stable configurations

Exemplary application:
Geopolitics is a rich domain where structural balance analysis is critical.
- Not only analysing the balance of the current international relations but
- also to analyse how adding, removing, or changing the sign of an edge may affect the overall balance.

It informs our decisions related to foreign policies, defence, etc.

Groups and structural balance

As Harary has pointed out, forming groups heavily influence the overall balance of a structure. Hence, structural balance analysis indicates

- how stable are the existing groups

(Am I safe with my current group?)

and

- if unstable, which group configurations will be stable

(Who should be in my group?)

- Moreover, structural balance analysis can be used to strategically create imbalance in an association structure.

How does someone select his/her group?

- Assortative mixing (homophily)
- Disassortative mixing (heterophily)

Homophily or assortative mixing

How does a person choose his/her group?

“Birds of a feather flock together”.

In the sociology literature, it has been long observed that most people associate with other people who are similar to them w.r.t. some parameters.

Exemplary parameters are academic programmes, gender, age, nationality, language, income.

Heterophily or disassortative mixing

How does a person choose his/her group?

“Opposites attract”.

In rare cases, people associate with other people who are dissimilar to them w.r.t. some parameters.

Exemplary parameters are academic programmes, gender, age, nationality, language, income.

Heterophily or disassortative mixing (contd.)

However, assortative or disassortative mixing can depend upon the type of association we are studying.

For some types of association, disassortative mixing is not rare, rather the norm.

Examples:

Research collaborations (skillset),
business partnerships (skillset, personalities),
marriages (gender, physical attributes, personalities).

Complex combinations of assortative and disassortative mixing

Marriages

- assortative w.r.t. nationality, religion, profession, etc.
- disassortative w.r.t. gender, physical attributes, personalities, etc.

International relations

- assortative w.r.t. political beliefs, religious beliefs, etc.
- disassortative w.r.t. natural resources, cultures, etc. (Is India better off grouping up with countries that produce petroleum?)

Complex combinations of assortative and disassortative mixing (contd.)

Exemplary applications:

- Matchmaking algorithms
- International relations thinktanks
- Designing election campaigns (such as candidate nominations)

Strong and weak ties

We developed a strong foundation visualizing a social network as a **static** structure.

However, in most cases, the associations between social actors are constantly evolving with time. Hence, we will start thinking about a social network as an **evolving** structure where

- vertices are getting added and removed
- edges are getting added and removed

Strong and weak ties (contd.)

Mark Granovetter was doing PhD in Sociology at the Harvard University in late 1960s.

He was interviewing people who have recently changed jobs to figure out how they got their new jobs.

It turned out that many of the interviewees found the new jobs by leveraging their social networks. However, they found their jobs not through close friends (**strong ties**) but through acquaintances (**weak ties**).

Mark Granovetter. American Journal of Sociology, 1973

Triadic closure

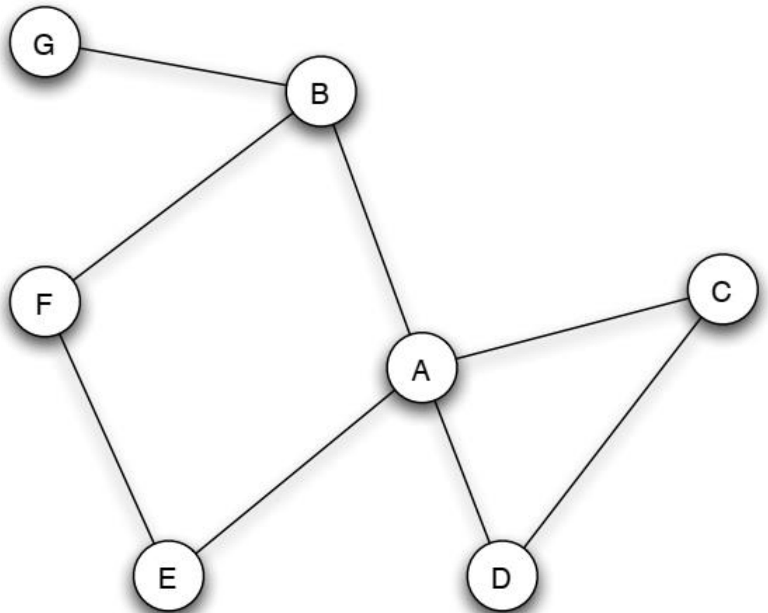
Granovetter’s observation in 1960s was in line with Anatole Rapoport’s observation in 1950’s. Rapoport observed:

If two people in a social network have a friend in common, then there is an increased likelihood that they will become friends themselves at some point in the future [1].

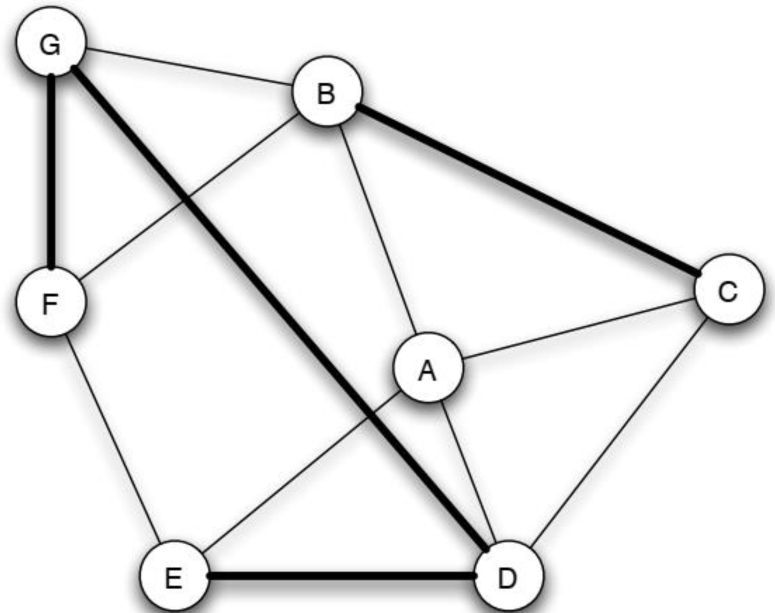
When they become friends, the triad is closed; in other words, the **triad closure** happens.

[1] Anatole Rapoport, Bulletin of Mathematical Biophysics, 1953.

Triadic closure (contd.)



(a) *Before new edges form.*



(b) *After new edges form.*

If we observe a social network for a long period of time, many of the new edges (such as edges B-C and G-F) are added due to the triad closure principle.
At the same time, some new edges (e.g., D-G) are added due to other reasons as well.

Strong triadic closure

Granovetter proposed **the strong triadic closure property**.

Any vertex A violates the strong triadic closure property if it has strong ties to two other vertices B and C, and there is no edge at all (either a strong or weak tie) between B and C.

We say that a vertex A satisfies the strong triadic closure property if it does not violate it.

The Strength of Weak Ties

Mark S. Granovetter

American Journal of Sociology

Vol. 78, No. 6 (May, 1973), pp. 1360-1380 (21 pages)

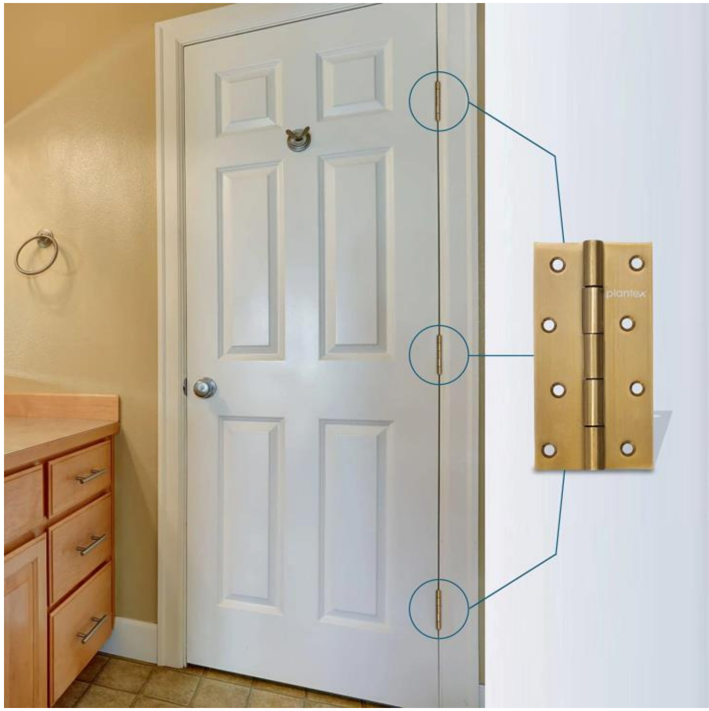
Published By: The University of Chicago Press

https://www.jstor.org/stable/2776392

Why does triadic closure work?

- Opportunity
- Trust
- **Incentive**

Three hinges to hold on to just a single door.
Is it unnecessary?



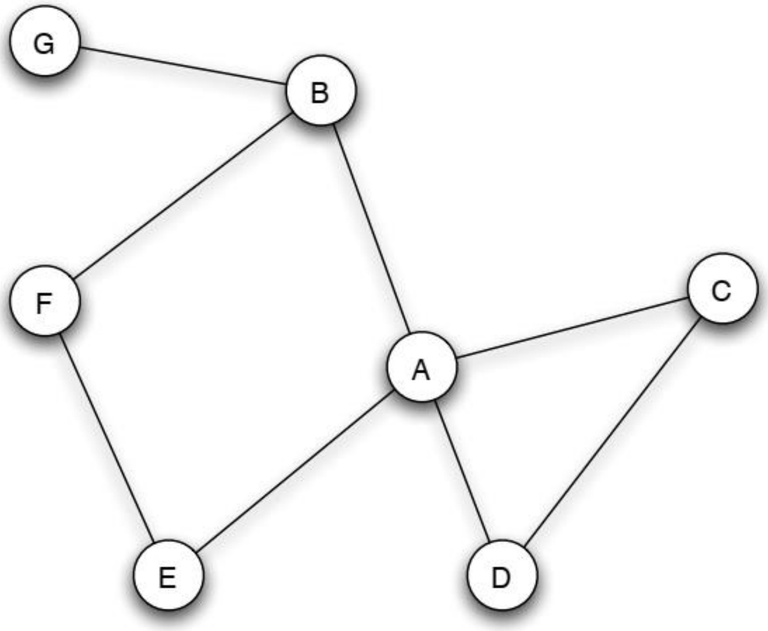
Incentive (in mathematical terms)

Incentive increases the **clustering coefficient** of a vertex.

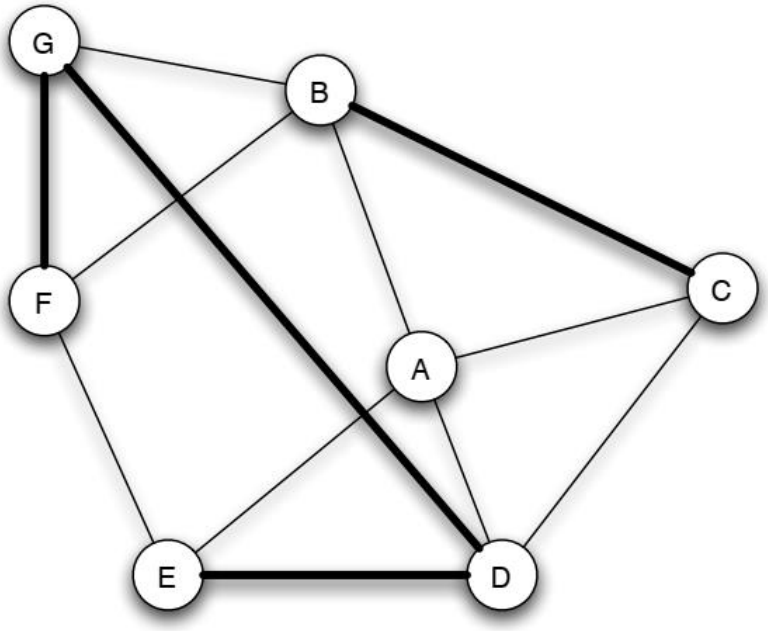
Clustering coefficient of a vertex (in an undi graph)

= (How many ties are there between its friends / How many ties are possible among its friends)

Clustering coefficient



(a) *Before new edges form.*



(b) *After new edges form.*

Vertex A's clustering coefficient
= (1 / 4C2) = 1/6

Vertex A's updated clustering coefficient
= (3 / 4C2) = 3/6 = 1/2

High clustering coefficient reduces internal stress

In social psychology, it has been **hypothesized** that social actors with higher clustering efficients (i.e. whose friends are also friends among each other) encounter less internal stress.

Example:

Am J Public Health. 2004 January; 94(1): 89–95.
doi: 10.2105/ajph.94.1.89

Suicide and Friendships Among American Adolescents

Peter S. Bearman, PhD and James Moody, PhD

In 2004, a study on American adolescents found that teenage girls who have a high clustering coefficient in their friendship networks are significantly less likely to contemplate suicide than those whose clustering coefficient is low.

References

- Sections 7.8, 7.11, 7.13, ‘Networks’ by Mark Newman, Oxford University Press, 1st edition, 2010.

- Strong and weak ties:
  Pages 47-56,
  https://www.cs.cornell.edu/home/kleinber/networks-book/networks-book-ch03.pdf

Thank you