GRAPH AND SOCIAL NETWORK VISUALIZATIONS

Diego Alberto Iozzia

Dipartimento di Informatica

Università degli Studi statale di Milano

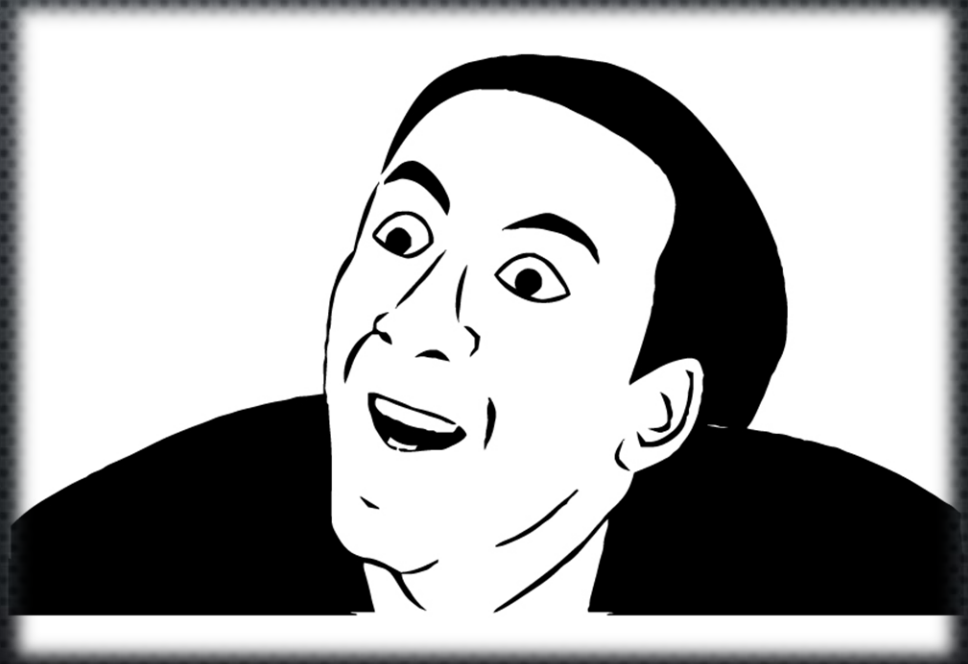
diegoalberto.iozzia@studenti.unimi.it

SOCIAL NETWORK ANALYSIS

- What is social network analysis?
- History of SNA
- What is a social network?
- Social Network Language
- Social Network representation
- Social Networks indexes
- Subgroups breakdowns
- applications of SNA

WHAT IS SOCIAL NETWORK ANALYSIS?

- Social Network Analysis is the study of social networks
- Its main idea is to study, measure and represent the relationships among individuals

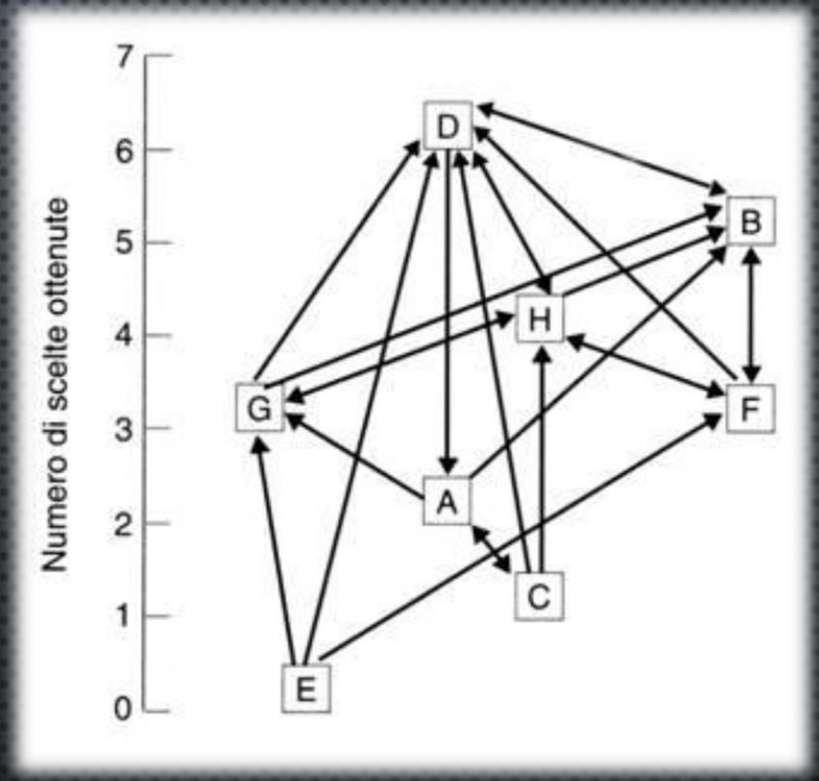


HISTORY OF SNA

- Gestalt theory
- Harvard researchers
- Manchester anthropologists
- Return of Harvard researchers

GESTALT THEORY

- Jacob L. Moreno: his book "Who Shall Survive?" contains some of the earliest graphical depictions of social networks (sociograms) (1934)
- Moreno sociograms represent formal properties of social configurations (1925)
- Individuals are dots and relationships are lines



Moreno sociogram

HARVARD RESEARCHERS

- They wanted to identify tecniques to reveal the subgroup structures in a social system
- Elton Mayo: studying the productivity of the Hawthorne power station in Chicago he shows the rilevance of the informal relationships structure in a work group (1920s/1930s)
- William Lloyd Warner: in his study about the New England community "Yankee City" he discovered that the community integration developed through informal relationships (1930s/1940s)

MANCHESTER ANTHROPOLOGISTS

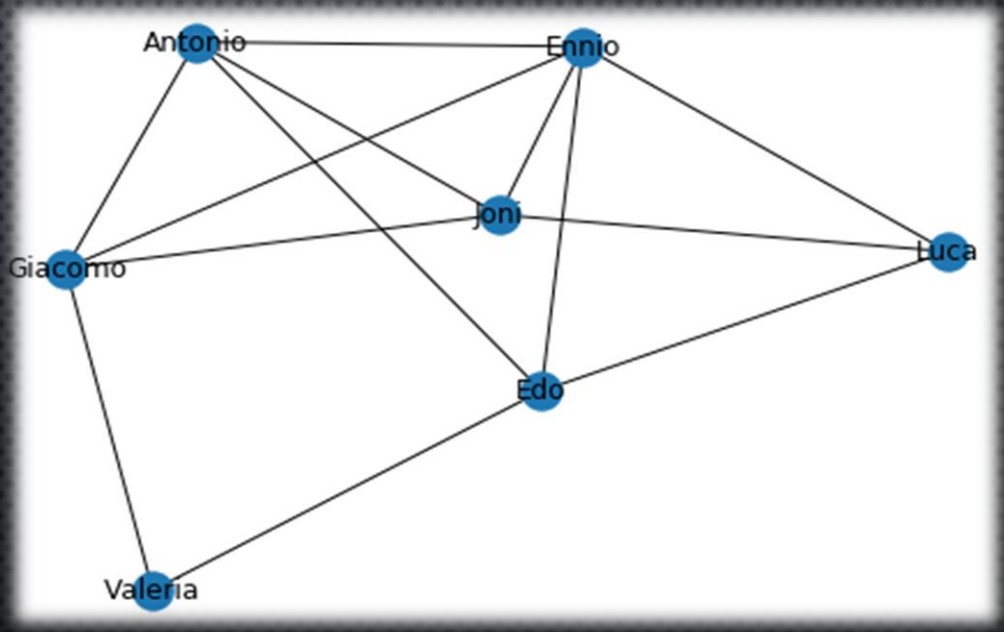
- Max Gluckman: he studied the changes in the pre and post colonial Africa, with particular interest in power and conflicts (1950s)
- He discovered to what extent power and conflict are integral elements in every social structure
- He and his students use social networks for the first time, initially as a metaphore and then as a mathematical concept in which to apply graph theory

RETURN OF HARVARD RESEARCHERS

- In the 60s and 70s the developement of algebric models to represent relationships as sets and multidimensional metodologies led to Social Network Analysis as we know it today
- Mark Granovetter: in "The strength of weak ties" he studied the processes by which people get information about employement (1974)
- He found out that It's more likely to get work information from acquaintances than from close friends.

WHAT IS A SOCIAL NETWORK?

- A social network is the totality of social actors and the relationships among them
- For instance a social network can represent a group of friends and the friendship relationships among them



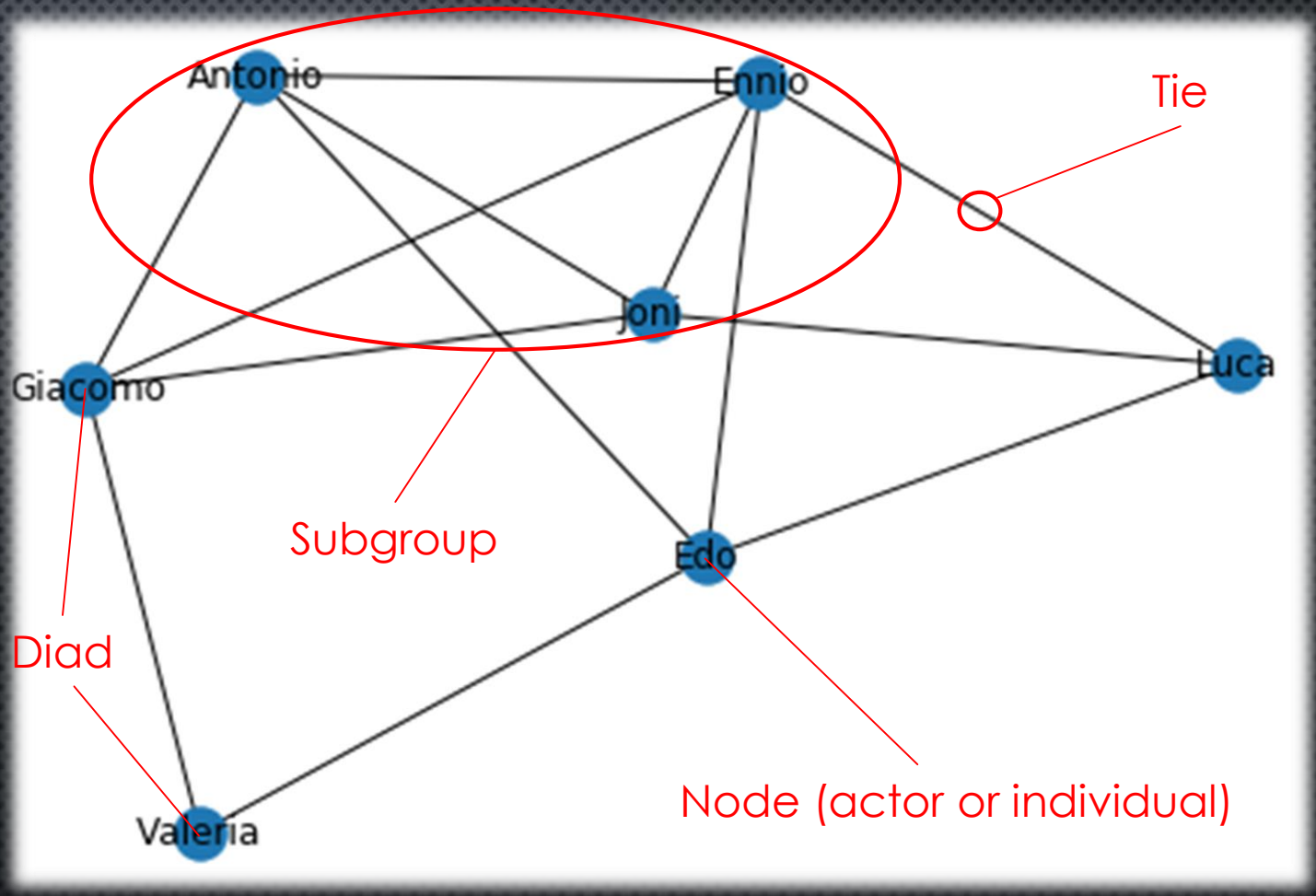
SOCIAL NETWORK LANGUAGE

- The language of Social Network Analysis is graph theory, especially:
- vertex → node (i.e. actor or individual)
- edge → social tie
- the edge set → relationships
- the vertex set → group
- subgraph → subgroup
- vertex pair → diad

SOCIAL NETWORK LANGUAGE

- group = {Antonio, Ennio, Joni, Luca, Edo, Giacomo, Valeria}

- relationships = {(Antonio, Ennio), (Joni, Ennio), (Edo, Valeria), …}



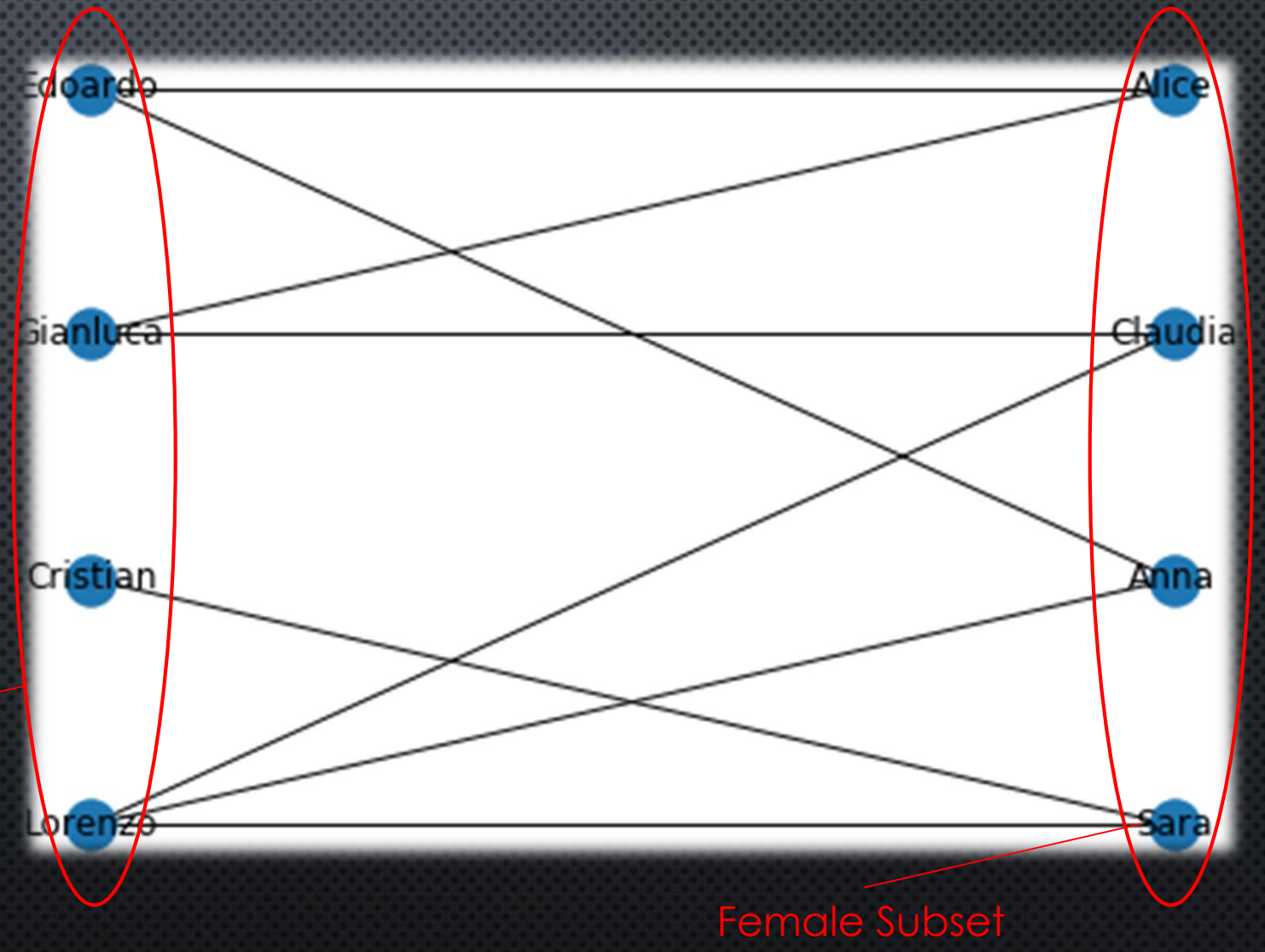
SOCIAL NETWORK LANGUAGE

- Another concept transposed from graph theory to social network analysis:

- The mode of a network is the number of subsets of entities in which the group is divided
- Partition → mode
- i.e. a two-mode network is represented with a bipartite graph

SOCIAL NETWORK LANGUAGE

- In this example this two-mode network can represent an old dating app



SOCIAL NETWORK LANGUAGE

- The maximum number of ties in a simple network is:
- Let n be number of actors:

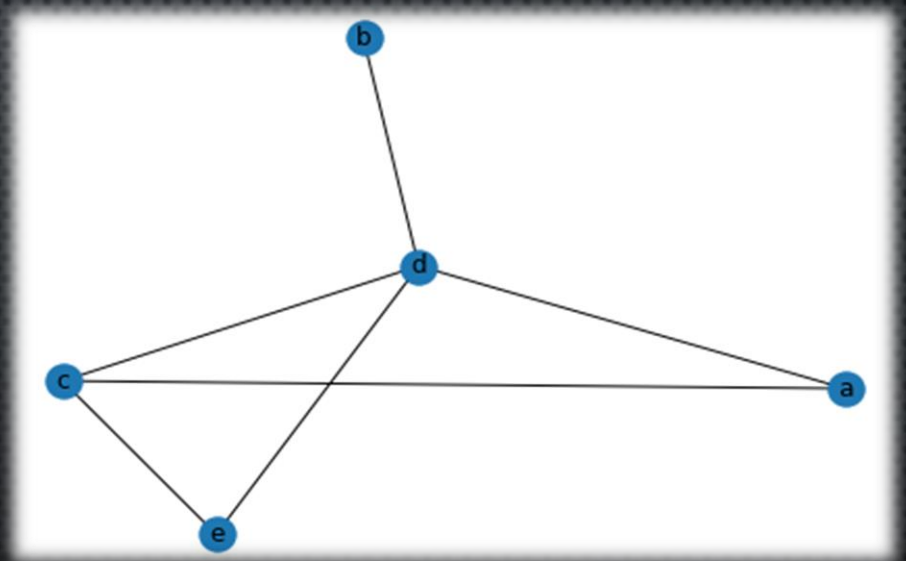
- n(n-1)/2 for an undirected relationship

- n(n-1) for a directed relationship

SOCIAL NETWORK REPRESENTATION

- Mathematically social networks are graphs, so they can be visualized like the way graphs do:

	a	b	c	d	e
a	0	0	1	1	0
b	0	0	0	1	0
c	1	0	0	1	1
d	1	1	1	0	1
e	0	0	1	1	0



SOCIAL NETWORKS INDEXES

The main network indexes are:

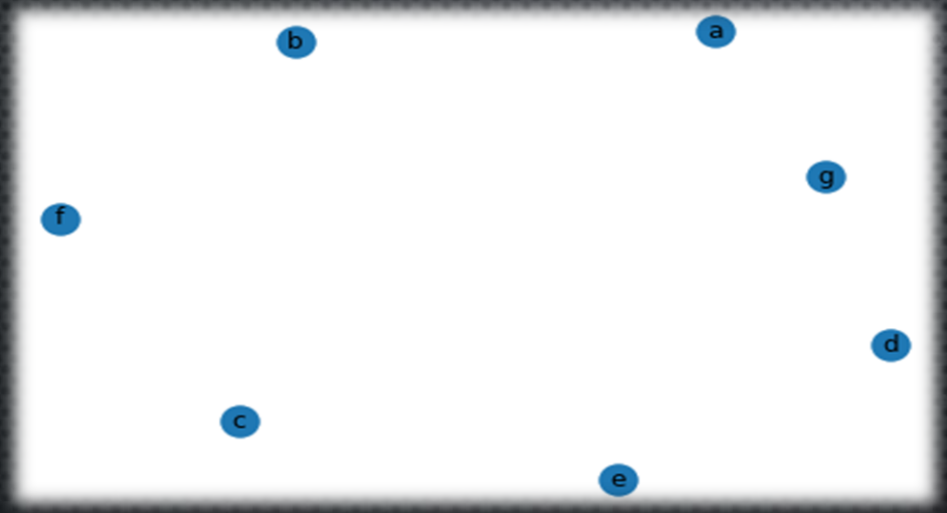
- Density
- Centrality (local and global)
- Cohesion
- Role

DENSITY

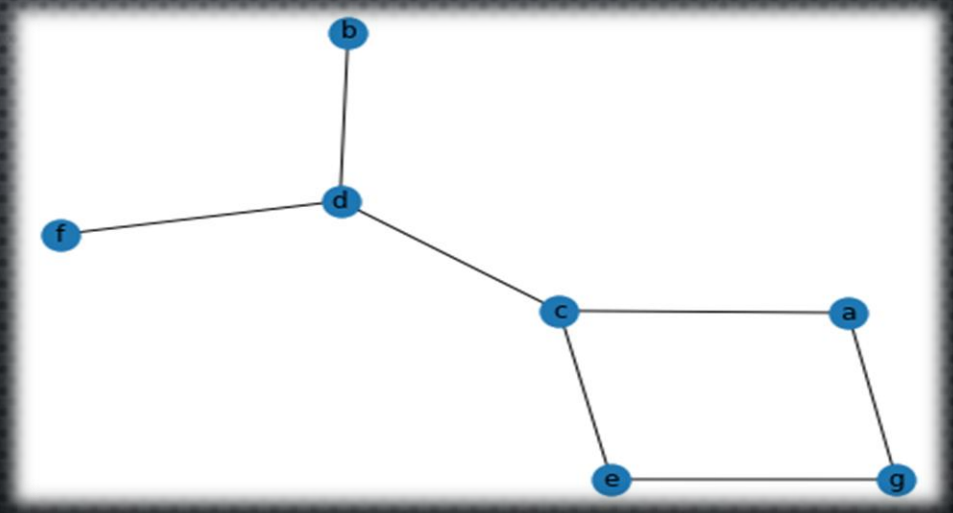
- The density of a graph is:
- 2m/n(n-1) for undirected graphs
- m/n(n-1) for directed graphs
- $\Sigma e_i/n(n-1)$ for weighted graphs
- Where n is the number of nodes, m is the number of ties and ei is the i-th ties with i = 1,...,m

- The density is the social integration degree in a network

DENSITY

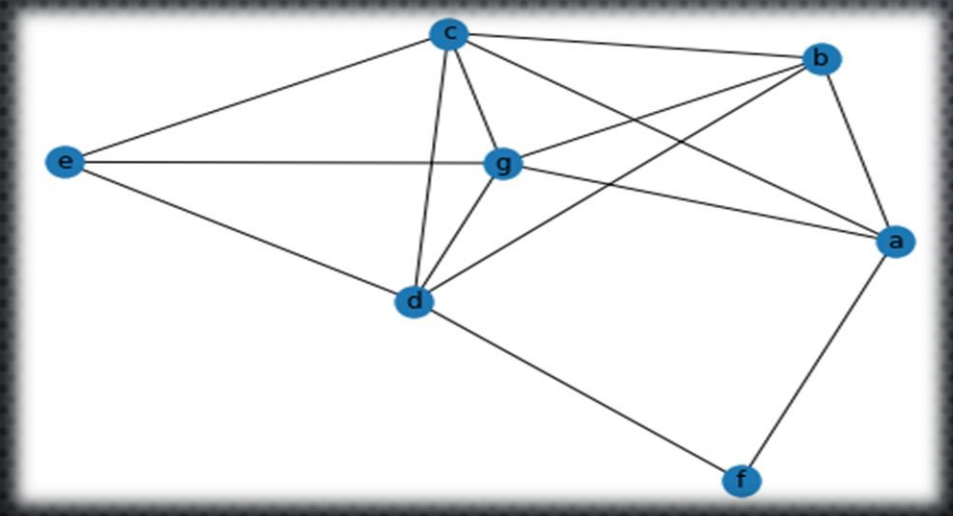


Set of disconnected components:
density = 0

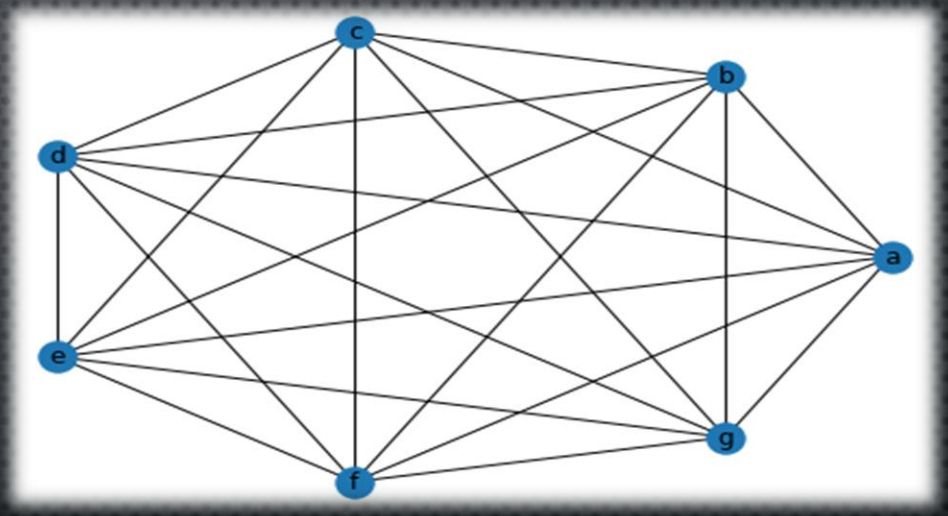


SPARSE GRAPH: DENSITY = 1/3

DENSITY



DENSE GRAPH: DENSITY = 2/3



COMPLETE GRAPH: DENSITY = 1

CENTRALITY

- There are two kinds of centrality:
- Local centrality
- the centrality indexes measure how much a node is connected to the others

- Global centrality
- the centralization indexes measure how strategic is the position of the node in the network

DEGREE CENTRALITY

- Local centrality index
- The actor with the highest number of ties is the place where things happen
- The higher the vertex degree, the greater the local centrality

DEGREE CENTRALITY

- Degree centrality index:
- $C_d(v_i) = d(v_i)$
- Where $d(v_i)$ is the degree of the i-th vertex

- Normalized centrality index:
- $C_d(v_i) = d(v_i)/(n-1)$
- Where n is the number of vertexes

CLOSENESS

- Global centrality index
- The closer the node is to the others, the more it is central
- The closeness degree increases as the distance between vertexes decreases

CLOSENESS

- Closeness centrality index:

- $C_c(v_i) = [\Sigma\, d(v_i,v_j)]\text{-}1$
- where $d(v_i,v_j)$ is the geodesic distance between v_i and v_j

- $0 \leq C_c(v_i) \leq 1$

BETWEENNESS

- Global centrality index
- The central node is located in most geodesic (shorter) paths in the network
- Actors in the mediator position have control on the information flow in the network

BETWEENNESS

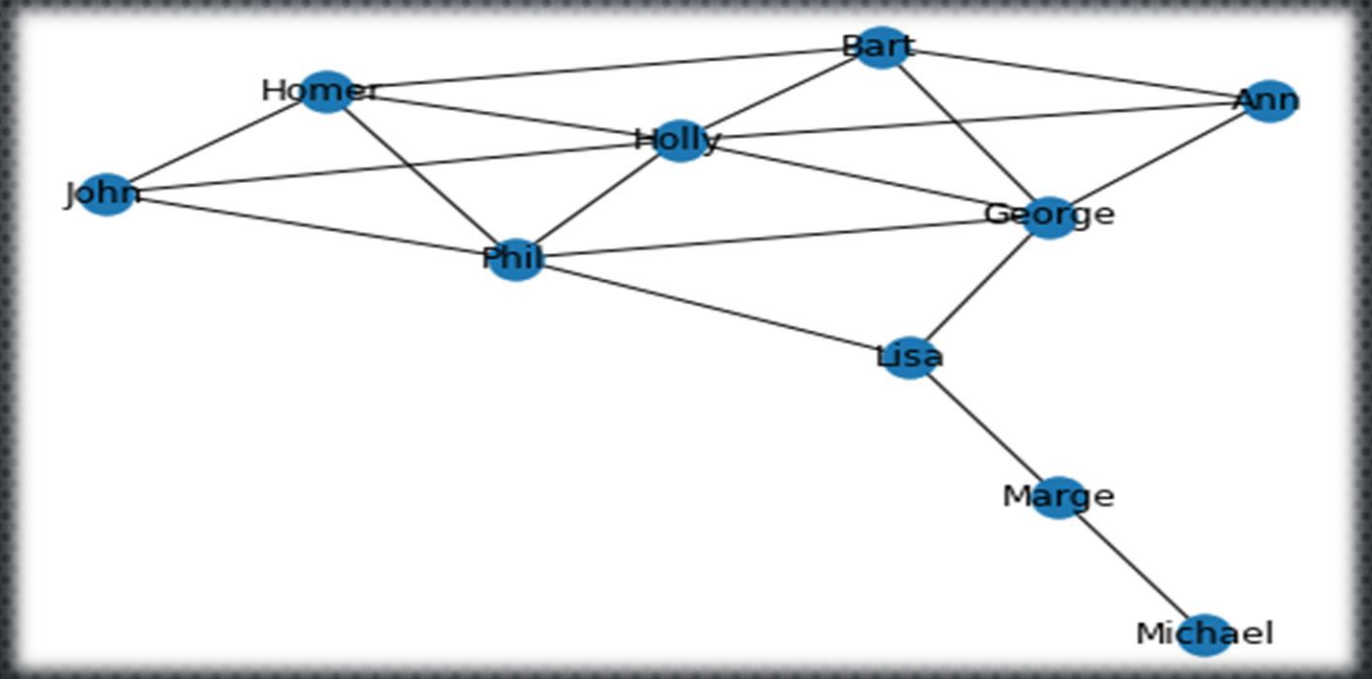
- betweenness centrality index

- $C_b(v_i) = \Sigma\, g_{jk}(v_i)/g_{jk}$
- where $g_{jk}(v_i)$ is the number of geodesic paths between v_j and v_k that contains v_i

- The betweenness of a node is the sum of all the partial betweennesses between every pair of nodes

	Ann	Bart	Holly	Homer	John	Phil	George	Lisa	Marge	Michael
Ann	0	1	1	0	0	0	1	0	0	0
Bart	1	0	1	1	0	0	1	0	0	0
Holly	1	1	0	1	1	1	1	0	0	0
Homer	0	1	1	0	1	1	0	0	0	0
John	0	0	1	1	0	1	0	0	0	0
Phil	0	0	1	1	1	0	1	1	0	0
George	1	1	1	0	0	1	0	1	0	0
Lisa	0	0	0	0	0	1	1	0	1	0
Marge	0	0	0	0	0	0	0	1	0	1
Michael	0	0	0	0	0	0	0	0	1	0

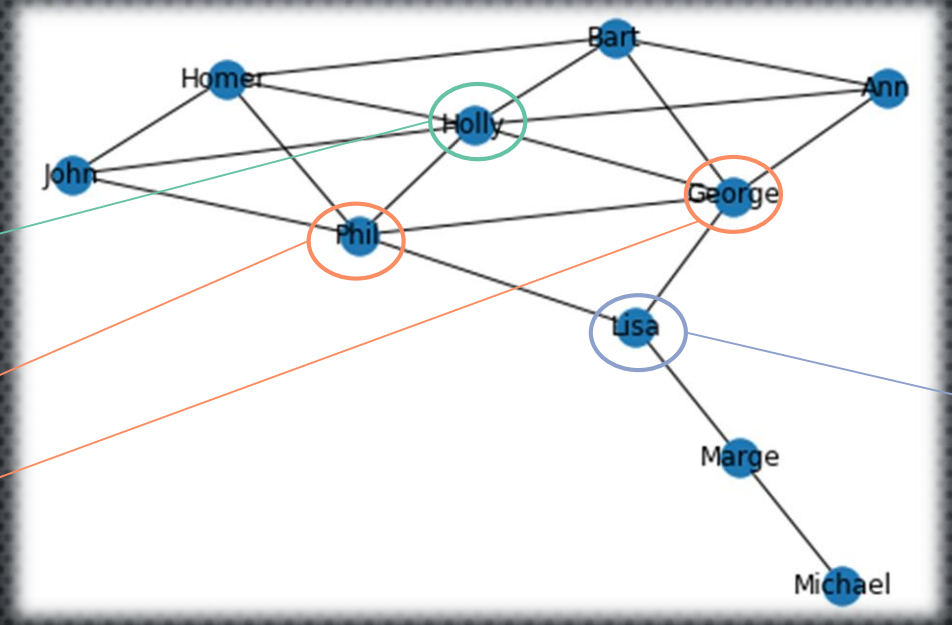
EXAMPLE: ADJACENCY MATRIX



EXAMPLE: GRAPH VISUALIZATION

Actors	Degree	Closeness	Betweenness
Ann	3	0.500	0.0
Bart	4	0.529	0.023
Holly	6	0.600	0.101
Homer	4	0.529	0.023
John	3	0.500	0.0
Phil	5	0.642	0.231
George	5	0.642	0.231
Lisa	3	0.600	0.388
Marge	2	0.428	0.22
Michael	2	0.310	0.0

EXAMPLE: CENTRALITY INDEXES



degree

Closeness

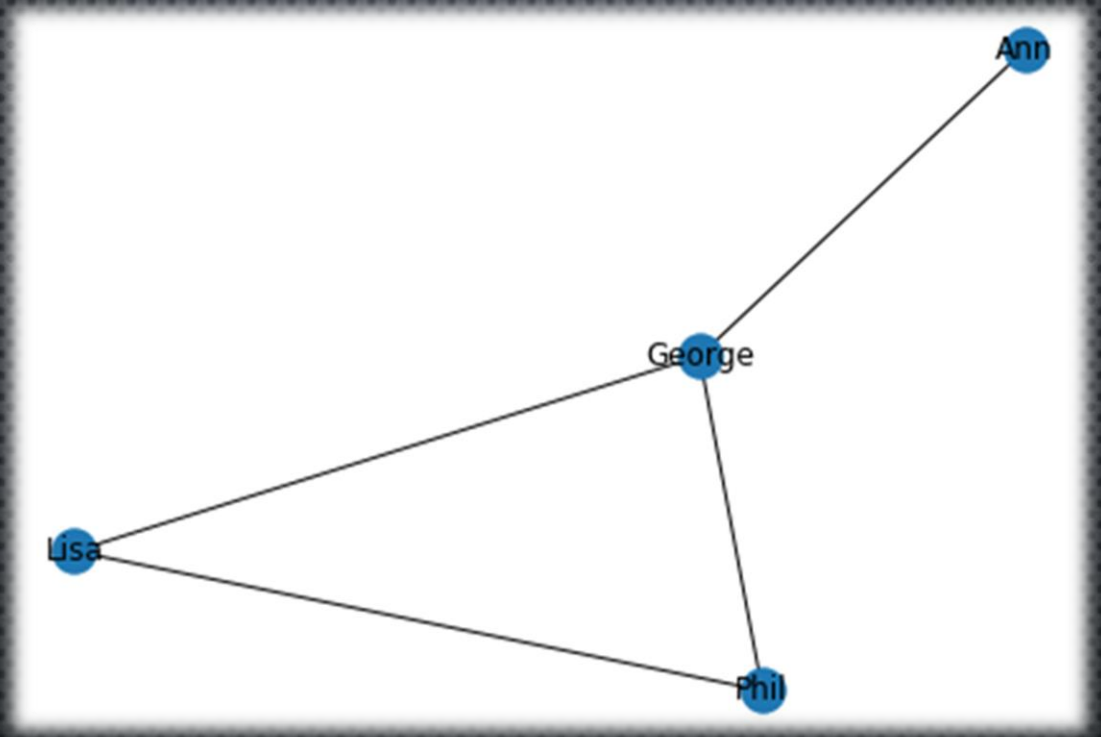
Betweenness

EXAMPLE: INDEXES VISUALIZATION

SUBGROUPS BREAKDOWNS

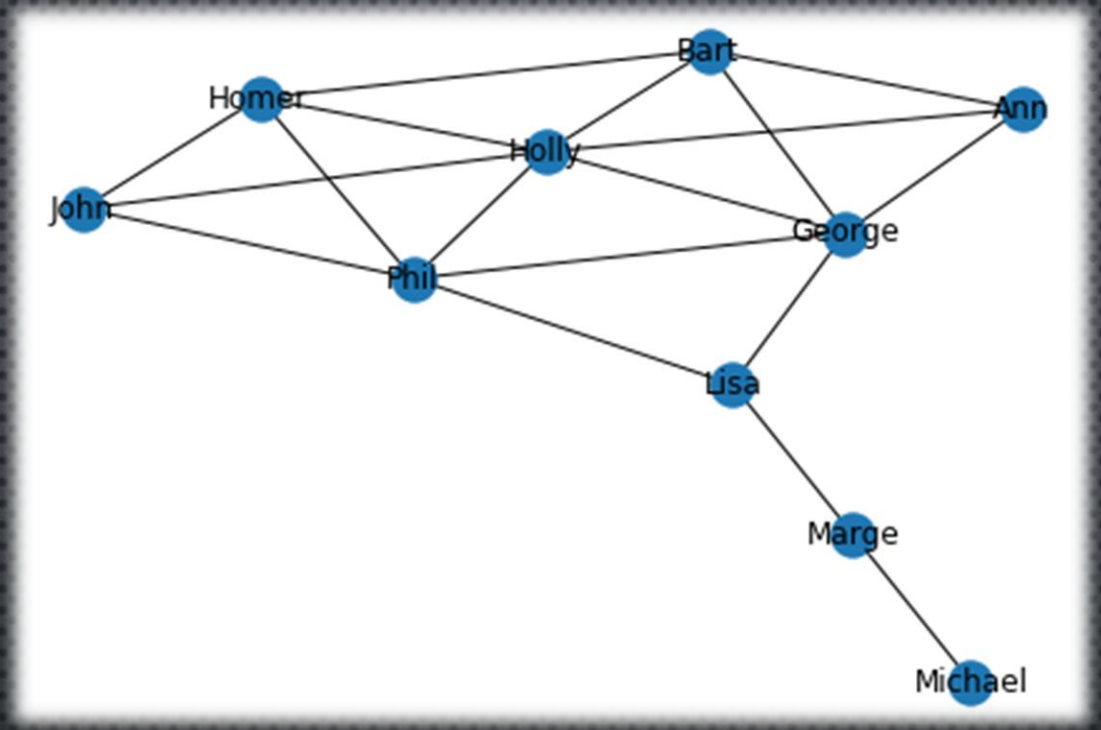
- Identification of associated subgroups representing places of exchange and interactions among members

- Types:
  - Subgraph
  - Component
  - Clique



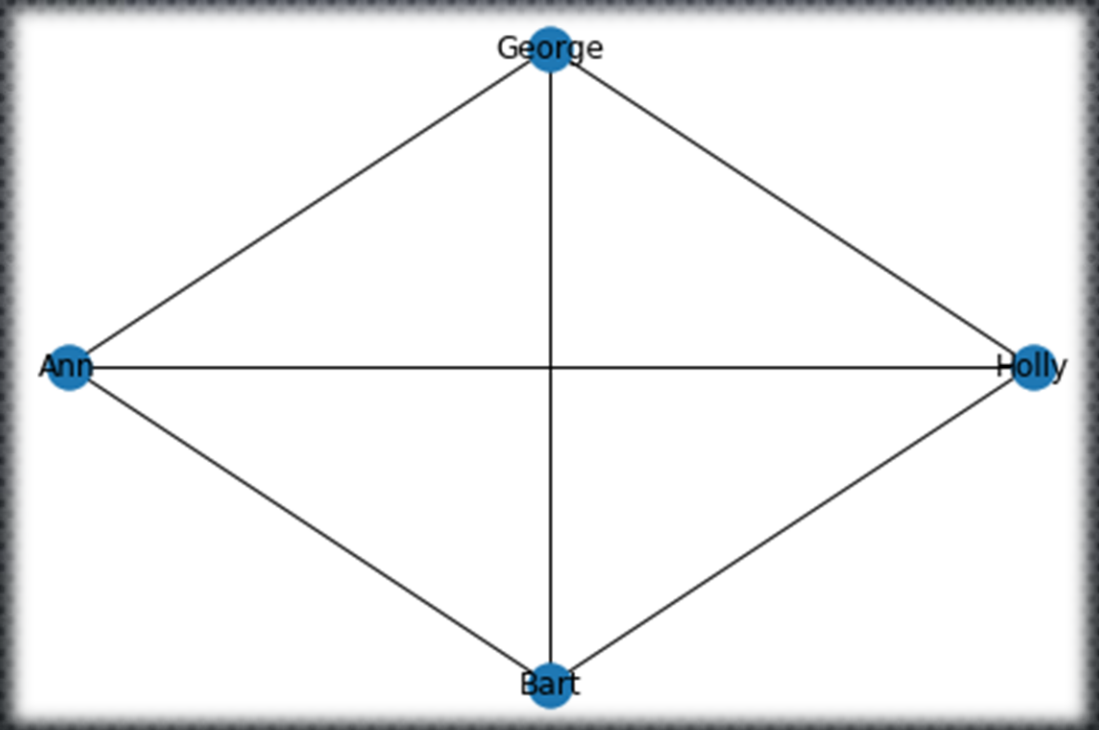
SUBGRAPH:

selection of nodes and the ties among them



COMPONENT:

internally connected subgraph disconnected with the others

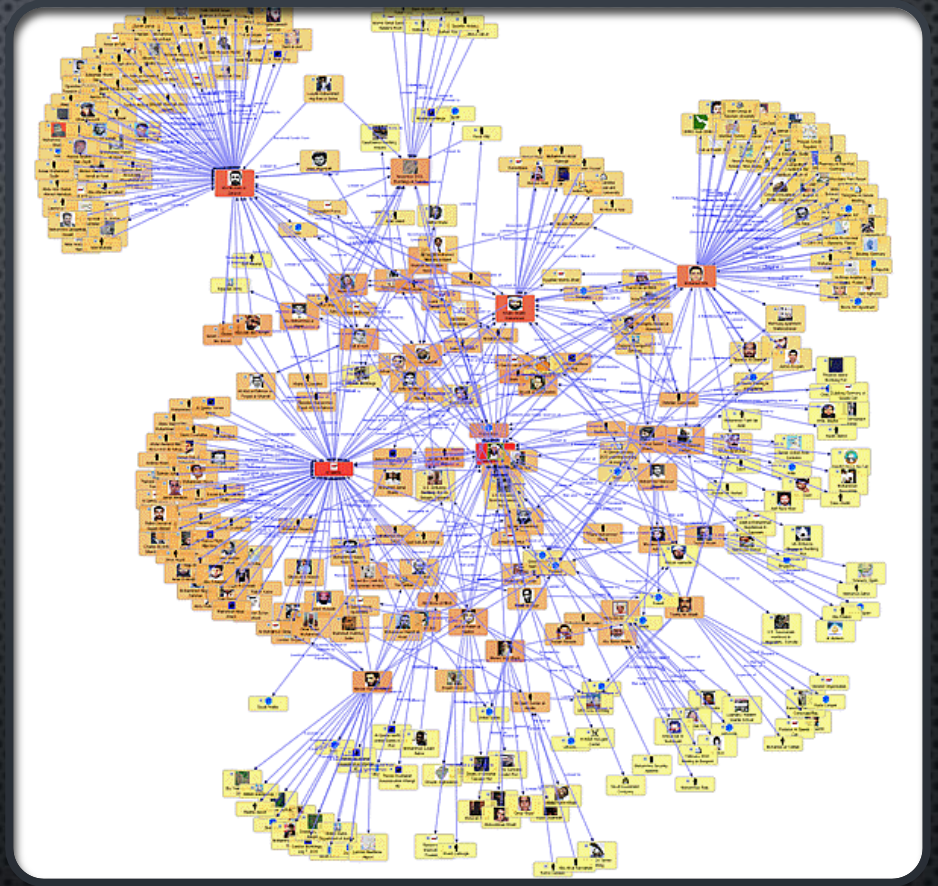


CLIQUE:

a complete subgraph with 3 or more nodes in which every possible pair is connected

APPLICATIONS OF SNA

- Security
- Sociology
- Textual analysis
- Recommender system



Al Qaeda terrorist network
(by Sentinel Visualizer)