# Social Capital: From Classics to Recent Trends

**Ramasuri Narayanam**

IBM Research, India
Email ID: ramasurn@in.ibm.com

24-July-2013

# Outline of the Presentation

1. **Introduction to Social Networks**
2. Introduction to Cooperative Game Theory
3. Social Capital: Classical Approach
4. Social Capital: Recent Trends
5. Summary of the Presentation

# Social Networks: Introduction

Recently there is a significant interest from research community to study social networks since:

- Such networks are fundamentally different from technological networks
- Networks are powerful primitives to model several real world scenarios such as interactions among individuals/objects

# Social Networks: Introduction (Cont.)

Social networks are ubiquitous and have many applications:

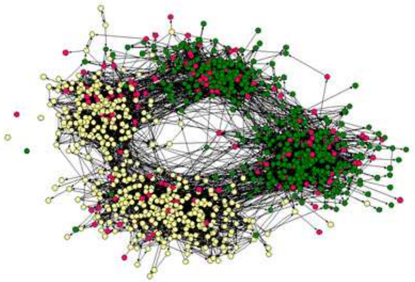
- For targeted advertising (or viral marketing)
- Monetizing user activities on on-line communities
- Job finding through personal contacts
- Predicting future events
- E-commerce and e-business
- ...

---

M.S. Granovetter. The Strength of Weak Ties. American Journal of Sociology, 1973.

# Example 1: Friendship Networks

## Friendship Network

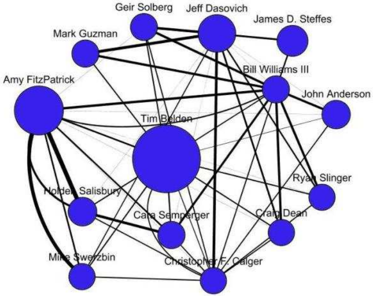


Nodes: Friends
Edges: Friendship

---

Reference: Moody 2001

## Email Network

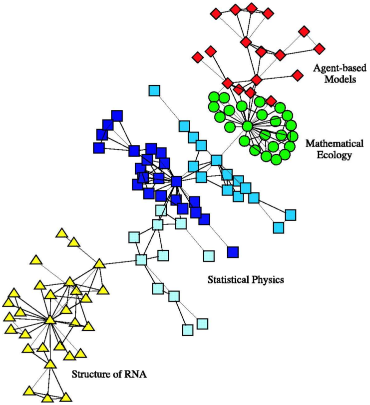



Nodes: Individuals
Edges: Email Communication

---

Reference: Schall 2009

# Example 2: Co-authorship Networks




Nodes: Scientists Edges: Co-authorship

Reference: M.E.J. Newman. Coauthorship networks and patterns of scientific collaboration. PNAS, 101(1):5200-5205, 2004

# Social Networks - Definition

- *Social Network:* A social system made up of individuals and interactions among these individuals
- Represented using graphs
  - Nodes - Friends, Publications, Authors, Organizations, Blogs, etc.
  - Edges - Friendship, Citation, Co-authorship, Collaboration, Links, etc.

---

S.Wasserman and K. Faust. Social Network Analysis. Cambridge University Press, Cambridge, 1994

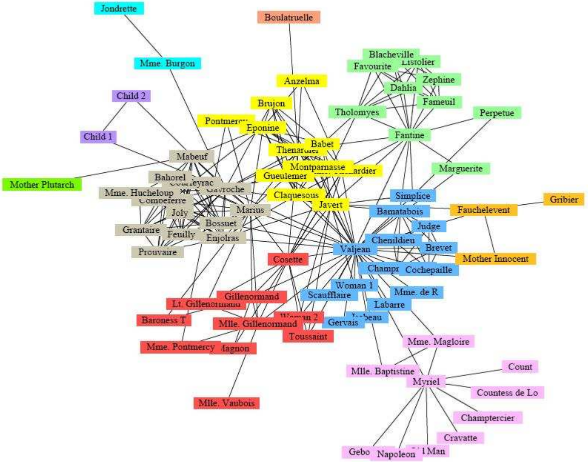
# Social Networks are Different from Computer Networks

Social networks differ from technological and biological networks in two important ways:

1. non-trivial clustering, and
2. the existence of dense groups or communities in the network

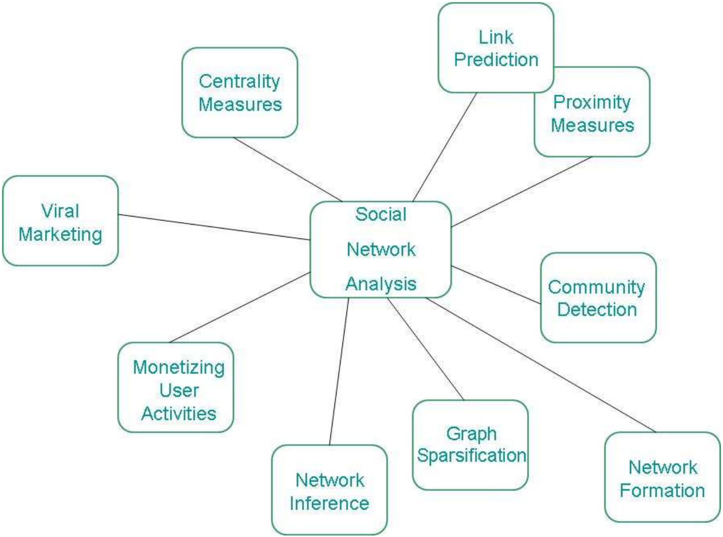
---


- M. E. J. Newman, Assortative mixing in networks. Phys. Rev. Lett. 89, 208701, 2002.
- M. E. J. Newman and Juyong Park. Why social networks are different from other types of networks. Physical Review E 68, 036122, 2003.

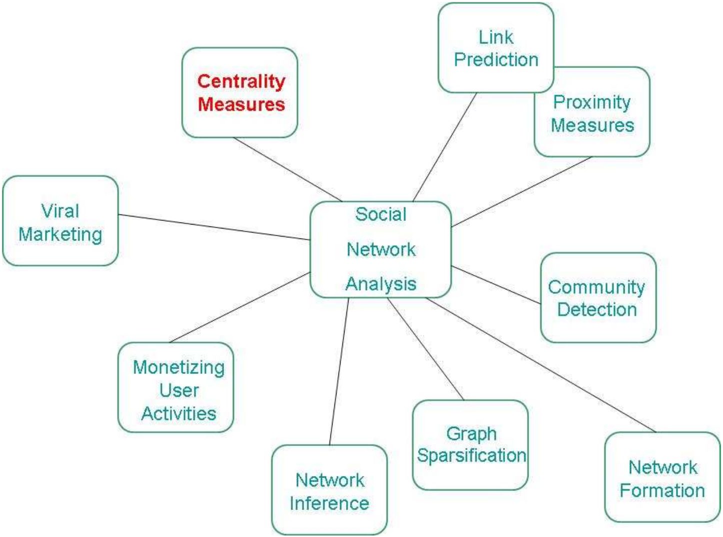



**Courtesy:** M. E. J. Newman and M. Girvan. *Finding and evaluating community structure in networks.* Phys. Rev. E 69, 026113, 2004.

# Social Networks: Some Key Topics



# Social Networks: Some Key Topics



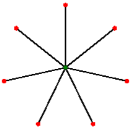
# Centrality Measures

- Significant amount of attention in the analysis of social networks is devoted to understand the centrality measures

- A centrality measure essentially ranks nodes/edges in a given network based on either their positional power or their influence over the network;

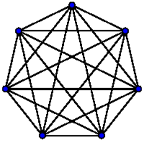
# Centrality Measures

- Significant amount of attention in the analysis of social networks is devoted to understand the centrality measures
- A centrality measure essentially ranks nodes/edges in a given network based on either their positional power or their influence over the network;



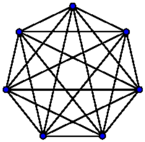
# Centrality Measures

- Significant amount of attention in the analysis of social networks is devoted to understand the centrality measures
- A centrality measure essentially ranks nodes/edges in a given network based on either their positional power or their influence over the network;



# Centrality Measures

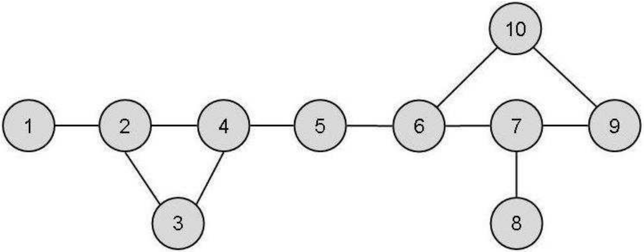
- Significant amount of attention in the analysis of social networks is devoted to understand the centrality measures
- A centrality measure essentially ranks nodes/edges in a given network based on either their positional power or their influence over the network;
- Some well known centrality measures:
  - Degree centrality
  - Closeness centrality
  - Clustering coefficient
  - Betweenness centrality
  - Eigenvector centrality, etc.



# Degree Centrality

- **Degree Centrality:** The degree of a node in a undirected and unweighted graph is the number of nodes in its immediate neighborhood.
  - Rank nodes based on the degree of the nodes in the network
  - Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. Social Networks, 1(3), 215-239
  - Degree centrality (and its variants) are used to determine influential seed sets in viral marketing through social networks

# Degree Centrality (Cont.)



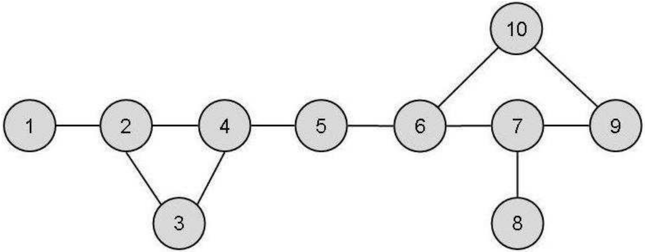

**Degree Centrality**

| **Node** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Value** | 1 | 3 | 2 | 3 | 2 | 3 | 3 | 1 | 2 | 2 |
| **Rank** | 9 | 1 | 5 | 1 | 5 | 1 | 1 | 9 | 5 | 5 |

# Closeness Centrality

- The farness of a node is defined as the sum of its shortest distances to all other nodes;
- The closeness centrality of a node is defined as the inverse of its farness;
- The more central a node is in the network, the lower its total distance to all other nodes.

# Closeness Centrality (Cont.)



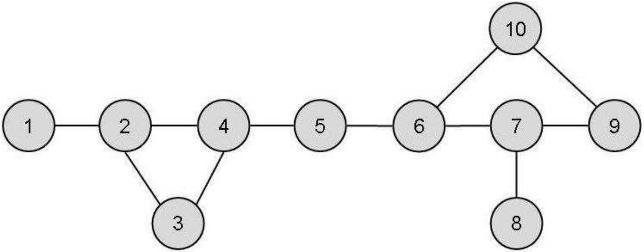

**Closeness Centrality**

| **Node** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Value** | $\frac{1}{34}$ | $\frac{1}{26}$ | $\frac{1}{27}$ | $\frac{1}{21}$ | $\frac{1}{19}$ | $\frac{1}{19}$ | $\frac{1}{23}$ | $\frac{1}{31}$ | $\frac{1}{29}$ | $\frac{1}{25}$ |
| **Rank** | 10 | 6 | 7 | 3 | 1 | 1 | 4 | 9 | 8 | 5 |

# Clustering Coefficient

- It measures how dense is the neighborhood of a node.
- The clustering coefficient of a node is the proportion of links between the vertices within its neighborhood divided by the number of links that could possibly exist between them.
- D. J. Watts and S. Strogatz. Collective dynamics of 'small-world' networks. Nature 393 (6684): 440442 , 1998.
- Clustering coefficient is used to design network formation models

# Clustering Coefficient (Cont.)




**Clustering Coefficient**

| **Node** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Value** | 0 | $\frac{1}{3}$ | 1 | $\frac{1}{3}$ | 0 | 0 | 0 | 0 | 0 | 0 |
| **Rank** | 3 | 2 | 1 | 2 | 3 | 3 | 3 | 3 | 3 | 3 |

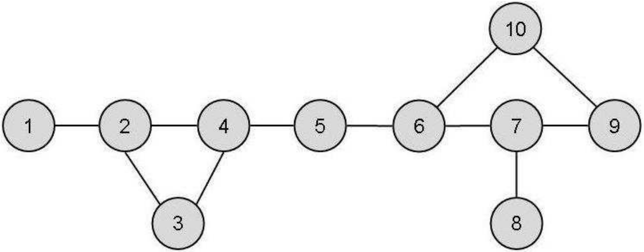
# Betweeness Centrality

- **Between Centrality:** Vertices that have a high probability to occur on a randomly chosen shortest path between two randomly chosen nodes have a high betweenness.
  - Formally, betweenness of a node $v$ is given by

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{s,t}(v)}{\sigma_{s,t}}$$

    where $\sigma_{s,t}(v)$ is the number of shortest paths from $s$ to $t$ that pass through $v$ and $\sigma_{s,t}$ is the number of shortest paths from $s$ to $t$.
  - L. Freeman. A set of measures of centrality based upon betweenness. Sociometry, 1977.
  - Betweenness centrality is used to determine communities in social netwoks (Reference: Girvan and Newman (2002)).

# Betweenness Centrality (Cont.)




| Betweenness Centrality | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Node** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| **Value** | 0 | 8 | 0 | 18 | 20 | 21 | 11 | 0 | 1 | 6 |
| **Rank** | 8 | 5 | 8 | 3 | 2 | 1 | 4 | 8 | 7 | 6 |

# A Simple Observation

| ID | Degree Centrality | Closeness Centrality | Clustering Centrality | Betweenness Centrality | Eigenvector Centrality |
|---|---|---|---|---|---|
| 1 | 9 | 10 | 3 | 8 | 9 |
| 2 | 1 | 6 | 2 | 5 | 2 |
| 3 | 5 | 7 | 1 | 8 | 3 |
| 4 | 1 | 3 | 2 | 3 | 1 |
| 5 | 5 | 1 | 3 | 2 | 5 |
| 6 | 1 | 1 | 3 | 1 | 3 |
| 7 | 1 | 4 | 3 | 4 | 6 |
| 8 | 9 | 9 | 3 | 8 | 10 |
| 9 | 5 | 8 | 3 | 7 | 8 |
| 10 | 5 | 5 | 3 | 6 | 7 |

# Outline of the Presentation

1. Introduction to Social Networks
2. **Introduction to Cooperative Game Theory**
3. Social Capital: Classical Approach
4. Social Capital: Recent Trends
5. Summary of the Presentation

# Cooperative Game Theory

- **Definition:** A cooperative game with transferable utility is defined as the pair $(N, v)$ where $N = \{1, 2, \ldots, n\}$ is a set of players and $v : 2^N \rightarrow \mathbb{R}$ is a characteristic function, with $v(.) = 0$. We call such a game also as a game in coalition form, game in characteristic form, or coalitional game or TU game.

- **Example:** There is a seller $s$ and two buyers $b_1$ and $b_2$. The seller has a single unit to sell and his willingness to sell the item is 10. Similarly, the valuations for $b_1$ and $b_2$ are 15 and 20 respectively. The corresponding cooperative game is:
  - $N = \{s, b_1, b_2\}$
  - $v(\{s\}) = 0 \quad , v(\{b_1\}) = 0 \quad , v(\{b_2\}) = 0 \quad , v(\{b_1, b_2\}) = 0$
    $v(\{s, b_1\}) = 5 \quad , v(\{s, b_2\}) = 10 \quad , v(\{s, b_1, b_2\}) = 10$

# Cooperative Game Theory (Cont.)

- **Key Question:** How should the grand coalition that forms divides its value among its members?
- Certain well known solution concepts
  - Core,
  - Shapley Value,
  - Bargaining sets,
  - Nucleolus, etc.

# The Shapley Value

- Shapley value is a solution concept which is motivated by the need to have a theory that would predict a unique expected payoff allocation for every given coalitional game
- The Shapley value concept was proposed by Shapley in 1953
- It was part of his doctoral dissertation at the Princeton University
- Given a cooperative game $(N, v)$, the Shapley value is denoted by $\phi(v) = \{\phi_i(v), \phi_2(v), \ldots, \phi_n(v)\}$
- **Theorem:** There is exactly one mapping $\phi : \mathbb{R}^{2^N - 1} \rightarrow \mathbb{R}^N$ that satisfies Symmetry, Linearity, and Carrier axioms. This function satisfies: $\forall i \in N$, $\forall v \in \mathbb{R}^{2^N - 1}$,

$$\phi_i(v) = \sum_{C \subseteq N \setminus \{i\}} \frac{|C|!(n - |C| - 1)!}{n!} \{v(C \cup \{i\}) - v(C)\}$$

# Shapley Value: An Example

- **Example:** Consider the following cooperative game: $N = \{1, 2, 3\}$, $v(1) = v(2) = v(3) = v(23) = 0$, $v(12) = v(13) = v(123) = 300$.

- Then we have that

$$\phi_1(v) = \frac{2}{6}v(1) + \frac{1}{6}(v(12) - v(2)) + \frac{1}{6}(v(13) - v(3)) + \frac{2}{6}(v(123) - v(23))$$

- It can be easily computed that $\phi_1(v) = 200$, $\phi_2(v) = 50$, $\phi_3(v) = 50$.

# Outline of the Presentation

1. Introduction to Social Networks
2. Introduction to Cooperative Game Theory
3. **Social Capital: Classical Approach**
4. Social Capital: Recent Trends
5. Summary of the Presentation

# Social Capital

- Social capital a fundamental concept in sociology literature
- Social capital can be thought of as the links, shared values and understandings in society that enable individuals and groups to trust each other and work together
- The first approach conceives of social capital as a value/quality of groups
- The second approach conceives of social capital as the value/quality of an individuals social connections

---

S.P. Borgatti, C. Jones, and M.G. Everett. Network measures of social capital. CONNECTIONS 21(2), 1998.

# Social Capital (Cont.)

- The value of social capital (for both groups and individuals) can be determined either internally or externally
- This immediately leads to three different forms of social capital:
  - The value of each individual is determined using the connections with others (First Form of Social Capital)
  - The value of each group is determined using the connections among themselves only (Second Form of Social Capital)
  - The value of each group is determined using the connections that the group members have outside of it (Third Form of Social Capital)

# Classical Measures of Social Capital

- Measures for First Form of Social Capital: Degree centrality, Closeness centrality, Clustering Coefficient, Betweenness centrality, etc.
- Measures for Second Form of Social Capital: Average distance, Maximum distance, etc.
- Measures for Third Form of Social Capital: Group degree, Group closeness, Group betweenness, etc.

# Outline of the Presentation

# Limitations of Classical Approach

- The common phenomenon of these standard centrality measures is that they assess the importance of each node by focusing only on the role played by that node itself.
- Such an approach is inadequate to capture the synergies that may occur if the functioning of nodes as groups is considered.

# Modern Application 1: Virus Contamination

- Consider a computer network (example, intranet of a company) where nodes represent workstations and edges represent connections between them;
- Let us assume that every workstation can be potentially attacked by a virus which then propagates over the network;
- Also, let us consider a simple virus propagation model where an infected node infects all the unprotected nodes (i.e. those without anti-virus software) that are reachable from it;
- Assume that the network administrator has a limited budget to install anti-virus software
- If the virus spreads from some initial node chosen uniformly at random, on which machines does it make sense to install anti-virus software to minimize the expected number of infected nodes?

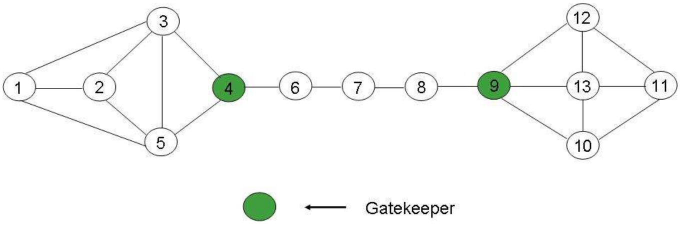
# Modern Application 2: Limiting the Spread of Misinformation through Social Networks

- More recently, companies often rely on viral marketing of products to maximize their revenue;
- At times, not only positive opinions, but also negative opinions may emerge and spread over the network of potential buyers;
- The company who owns this product wants to minimize the loss incurred due to the negative opinions;
- The question is which individual buyers the company should target (for convincing) in order to minimize the number of individuals that receive the negative opinion.

# The Problem Scenario

- We consider a network of individuals (such as social network of the buyers) or a network of objects (such as intranet of a company);
- Assume that certain unwanted process may attack a node uniformly at random and then starts spreading over the network effecting the function of all reachable nodes/individuals;
- We have some limited budget to reach out at most $k$ nodes;
- The problem is which $k$ nodes that we should target to minimize the expected number of the nodes that receive the misinformation.

# The Problem Scenario - Example when k = 2




| **Centrality Measure** | **Rank** 1 | **Rank** 2 |
|---|---|---|
| Degree | 9 | 3,5,10,11,12,13 |
| Closeness | 7 | 6,8 |
| Betweenness | 7 | 6,8 |
| Clustering Coefficient | 10,11,12,13 | 9 |
| EigenVector | 1,2,10,11,12,13 | 3,5 |
| PageRank | 9 | 3,5 |

# A New Approach to Social Capital: Game Theoretic Approach

- Motivated by the above, consider the following two step approach:
  - Using any standard measure for group level social capital, derive the value of social capital for each group in the network, and
  - Then compute the social capital for each individual actor in the network from the values derived already for each group.
- Cooperative game theory is a natural tool to model the above framework!

# Outline of the Presentation

# Conclusions

- A quick overview of social networks and then centrality measures is presented;
- A brief introduction to cooperative game theory is given
- Classical methods to social capital
- Recent advances in the space of social capital where the role of game theory is prominent

# Some Important Text Books

- D. Easley and J. Kleinberg. Networks, Crowds, and Markets. Cambridge University Press, 2010.
- M.E.J. Newman. Networks: An Introduction. Oxford University Press, 2010.
- M.O. Jackson. Social and Economic Networks. Princeton University Press, 2008.
- U. Brandes and T. Erlebach. Network Analysis: Methodological Foundations. Springer-Verlag Berlin Heidelberg, 2005.

# Some Leading Researchers

- Jon M. Kleinberg
- Christos Faloutsos
- Matthew O. Jackson
- Sanjeev Goyal
- Eva Tardos
- Jure Leskovec
- Nicole Immorlica
- David Kempe
- Krishna P. Gummadi
- Tanya Berger-Wolf
- ...

# Network Dataset Repositories

- Jure Leskovec: http://snap.stanford.edu/data/index.html
- MEJ Newman: http://www-personal.umich.edu/m̃ejn/netdata
- Albert L. Barabasi: http://www.nd.edu/ñetworks/resources.htm
- NIST Data Sets: http://math.nist.gov/R̃Pozo/complex_datasets.html
- MPI Data Sets: http://socialnetworks.mpi-sws.org/
- ...

# Software Tools for Network Analysis

- Gephi (Graph exploration and manipulation software)
- Pajek (Analysis and Visualization of Large Scale Networks)
- UCINET (Social Network Analysis tool)
- CFinder (Finding and visualizing communities)
- GraphStream (Dynamic graph library)
- Graphviz (Graph vizualisation software)
- Refer to Wikipedia for more information (http://en.wikipedia.org/wiki/Social_network_analysis_software)

# A List of Important Conferences

- ACM Conference on Electronic Commerce (ACM EC)
- Workshop on Internet and Network Economics (WINE)
- ACM SIGKDD
- WSDM
- ACM Internet Measurement Conference (ACM IMC)
- CIKM
- ACM SIGCOMM
- Innovations in Computer Science (ICS)
- AAMAS
- AAAI
- IJCAI
- ...

# A List of Important Journals

- American Journal of Sociology
- Social Networks
- Physical Review E
- Data Mining and Knowledge Discovery
- ACM Transactions on Internet Technology
- IEEE Transactions on Knowledge and Data Engineering
- Games and Economic Behavior
- ...

**Thank You**