Game Theoretic Models for Social Network Analysis

Ramasuri Narayanam

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

20-Feb-2012

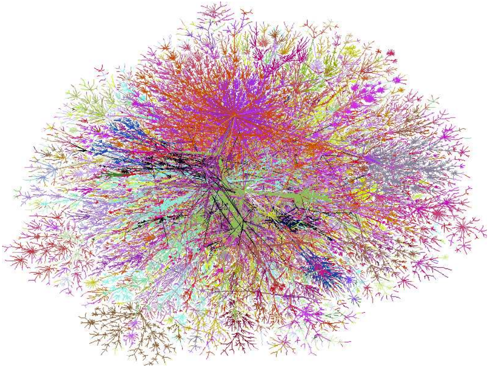
Acknowledgements

- Most of the content in this presentation is taken from the Ph.D. thesis of the speaker
- My sincere thanks to Prof. Y. Narahari for his valuable guidance
- Thanks to the Dept. of Computer Science and Automation, Indian Institute of Science, Bangalore, India
- Many thanks to IBM Research, India for supporting the collaborative work with IISc, Bangalore

Outline of the Presentation

1. **Social Network Analysis: A Quick Primer**
2. Foundational Concepts in Game Theory
3. Network Formation Problem
4. Summary and To Probe Further

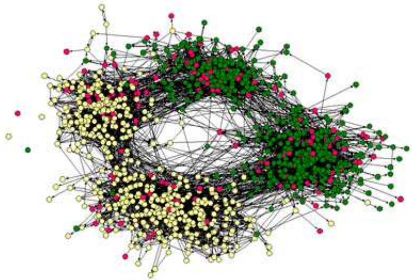
Example 1: Web Graph



Nodes: Static web pages
Edges: Hyper-links

Reference: Prabhakar Raghavan. Graph Structure of the Web: A Survey. In Proceedings of LATIN, pages 123-125, 2000.

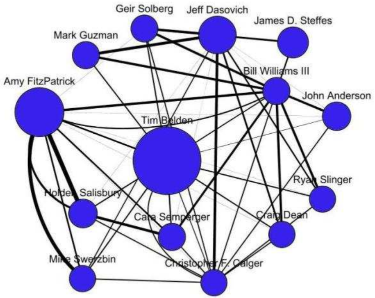
Example 2: Friendship Networks



Nodes: Friends
Edges: Friendship

Reference: Moody 2001

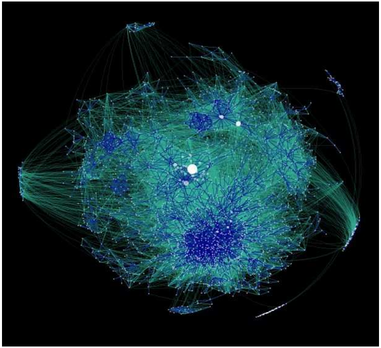
Example 3: Email Network



Nodes: Individuals
Edges: Email Communication

Reference: Schall 2009

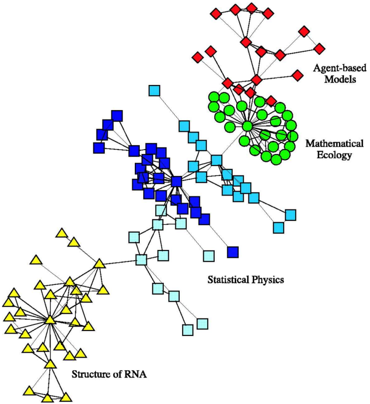
Example 4: Weblog Networks



Nodes: Blogs
Edges: Links

Reference: Hurst 2007

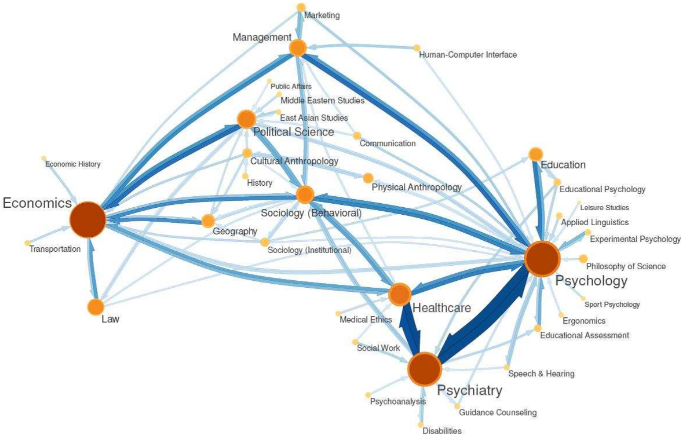
Example 5: Co-authorship Networks



Nodes: Scientists and Edges: Co-authorship

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

Example 6: Citation Networks



Nodes: Journals and Edges: Citation

Reference: http://eigenfactor.org/

Social Network Analysis (SNA)

- Study of structural and communication patterns
  - degree distribution, density of edges, diameter of the network

- Two principal categories:
  - **Node/Edge Centric Analysis:**
    - Centrality measures such as degree, betweeneness, stress, closeness
    - Anomaly detection
    - Link prediction, etc.

  - **Network Centric Analysis:**
    - Community detection
    - Graph visualization and summarization
    - Frequent subgraph discovery
    - Generative models, etc.

U. Brandes and T. Erlebach. Network Analysis: Methodological Foundations. Springer-Verlag Berlin Heidelberg, 2005.

Why is SNA Important?

- To understand complex connectivity and communication patterns among individuals in the network
- To determine the structure of networks
- To determine influential individuals in social networks
- To understand how social network evolve
- To determine outliers in social networks
- To design effective viral marketing campaigns for targeted advertising
- ...

A Few Key SNA Tasks: Measures to Rank Nodes

- **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

- **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

A Few Key SNA Tasks: Measures to Rank Nodes (Cont.)

- **Closeness Centrality:** The farness of a node is defined as the sum of its shortest distances to all other nodes, and its closeness is defined as the inverse of the farness. The more central a node is in the network, the lower its total distance to all other nodes.
- **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)).

A Few Key SNA Tasks: Measures to Rank Nodes (Cont.)

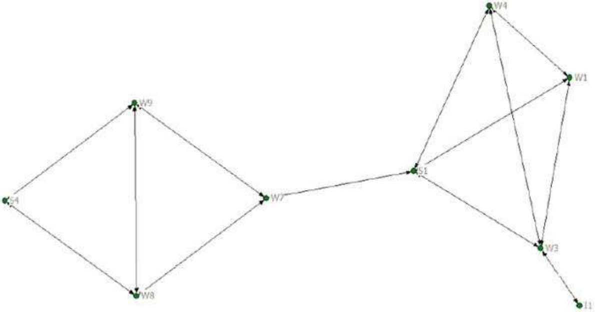
- **Eigenvector Centrality:** It assigns relative scores to all nodes in the network based on the principle that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes.
  - Formally, eigen-vector centrality (x_i) of a node i is given by

$$x_i = \frac{1}{\lambda} \sum_{j \in M(i)} x_j$$

    where $M(i)$ is the set of nodes directly connected to node i.
  - Google Page-Rank and Kats measure are variants of the Eigenvector centrality.
  - P. Bonacich and P. Lloyd. Eigenvector-like measures of centrality for asymmetric relations. Social Networks, 23(3):191-201, 2001.

A Few Key SNA Tasks: Measures to Rank Nodes (Cont.)

- *Example:*



Eigenvector Centrality	Degree Centrality	Closeness Centrality	Betweenness Centrality
S1 (.498)	W3,S1	S1	S1
W3 (.472)	W9,W8,W7,W1,W4	W7	W7
W1,W4 (.438)	S4	W3	W3
W7 (.254)	I1	W1,W4	W8,W9
W8,W9 (.159)		W8, W9	W1,I1,W4,S4 (0)
I1 (.147)		I1	
S4 (.099)		S4	

A Few Key SNA Tasks: Measures to Rank Nodes (Cont.)

- *Inadequacies of traditional ranking mechanisms for social networks:*
  - They are completely dependent on the structure of the underlying network. Often it is required to rank nodes/edges based on auxiliary information or data
  - Emergence of several applications wherein the ranking mechanisms should take into account not only the structure of the network but also other important aspects of the networks such as the value created by the nodes in the network
  - Several empirical evidences reveal that these ranking mechanisms are not scalable to deal with large scale network data
  - They are not tailored to take into account the strategic behavior of the nodes

A Few Key SNA Tasks: Diversity among Nodes

- Nodes in the network might be having various connectivity patterns
- Some nodes might be connected to high degree nodes, some others might be connected to bridge nodes, etc.
- Determining diversity among the connectivity patterns of nodes is an interesting problem
- L. Liu, F. Zhu, C. Chen, X. Yan, J. Han, P.S. Yu, and S. Yang. Mining Diversity on Networks. In DASFAA 2010.

A Few Key SNA Tasks: Link Prediction Problem

- Given a snapshot of a social network, can we infer which new interactions among its members are likely to occur in the near future?
- D. Liben-Nowell and J. Kleinberg. The link prediction problem for social networks. In CIKM 2003.

A Few Key SNA Tasks: Inferring Social Networks From Social Events

- In the traditional link prediction problem, a snapshot of a social network is used as a starting point to predict (by means of graph-theoretic measures) the links that are likely to appear in the future.
- Predicting the structure of a social network when the network itself is totally missing while some other information (such as interest group membership) regarding the nodes is available.
- V. Leroy, B. Barla Cambazoglu, F. Bonchi. Cold start link prediction. In SIGKDD 2010.

A Few Key SNA Tasks: Influence Maximization Problem

- With increasing popularity of online social networks, viral Marketing - the idea of exploiting social connectivity patterns of users to propagate awareness of products - has got significant attention
- In viral marketing, within certain budget, typically we give free samples of products (or sufficient discounts on products) to certain set of influential individuals and these individuals in turn possibly recommend the product to their friends and so on
- It is very challenging to determine a set of influential individuals, within certain budget, to maximize the volume of information cascade over the network
- P. Domingos and M. Richardson. Mining the network value of customers. In ACM SIGKDD, pages 5766, 2001.

A Few Key SNA Tasks: Community Detection

- *Based on Link Structure in the Social Network:*
  - Determining dense subgraphs in social graphs
  - Graph partitioning
  - Determining the best subgraph with maximum number of neighbors
  - Overlapping community detection
- Based on Activities over the Social Network
  - Determine action communities in social networks
  - Overlapping community detection
- J. Leskovec, K.J. Lang, and M.W. Mahoney. Empirical comparison of algorithms for network community detection. In WWW 2010.
- A.S. Maiya and T.Y. Berger-Wolf. Expansion and search in networks. In CIKM 2010.

A Few Key SNA Tasks: Design of Incentives in Networks

- With growing number of online social communities, users pose queries to the network itself, rather than posing queries to a centralized system.
- At present, the concept of incentive based queries is used in various question-answer networks such as Yahoo! Answers, Orkuts Ask Friends, etc.
- In the above contexts, only the person who answers the query is rewarded, with no reward for the intermediaries. Since individuals are often rational and intelligent, they may not participate in answering the queries unless some kind of incentives are provided.
- It is also important to consider the quality of the answer to the query, when incentives are involved.
- J. Kleinberg and P. Raghavan. Query incentive networks. In Proceedings of 46th IEEE FOCS, 2005.
- D. Dixit and Y. Narahari. Truthful and quality conscious query incentive networks. In WINE 2009.

A Few Key SNA Tasks: Determining Implicit Social Hierarchy

- Social stratification refers to the hierarchical classification of individuals based on power, position, and importance
- The popularity of online social networks presents an opportunity to study social hierarchy for different types of large scale networks
- M. Gupte, P. Shankar, J. Li, S. Muthukrishnan, and L. Iftode. Finding hierarchy in directed online social networks. In the Proceedings of World Wide Web (WWW) 2011.

Methods to Address SNA Tasks

- Traditional Approaches
  - Graph theoretic techniques
  - Spectral methods
  - Optimization techniques
  - ...

Methods to Address SNA Tasks

- Traditional Approaches
  - Graph theoretic techniques
  - Spectral methods
  - Optimization techniques
  - . . .
- Recent Advances
  - Data mining and machine learning techniques
  - **Game theoretic techniques**

Why Game Theoretic Models for SNA?

- Current metrics and measures in SNA are based on
  - Graph theoretic techniques
  - Optimization techniques
  - Spectral techniques, etc.
- Generative models can produce networks with similar structural properties
- In many network settings, the behavior of the system is driven by the actions of a large number of autonomous individuals (or agents)
  - Research collaborations among both organizations and researchers
  - Online social communities such as Orkut, Facebook, LinedIn
  - Telecommunication networks (Service Providers)

Why Game Theoretic Models for SNA? (Cont.)

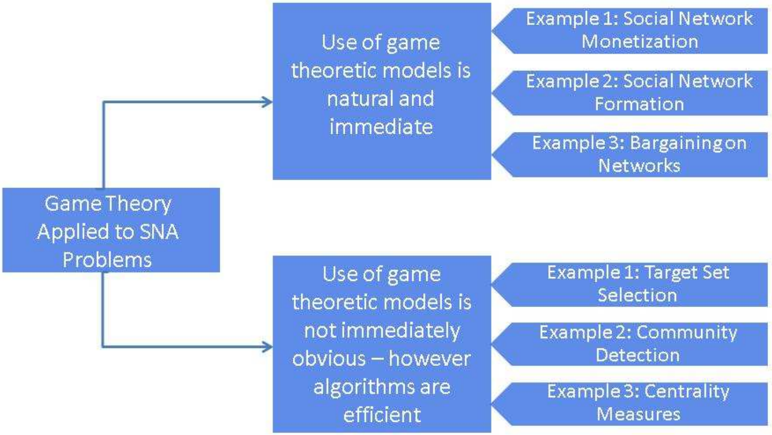
- Individuals are self-interested and optimize their respective objectives
- Inadequacies of current SNA approaches:
  - Social contacts (i.e. links) form more often by choice than by chance
  - There always exist social and economic incentives while forming links in the network
  - Do not satisfactorily capture the behavior of the individuals
  - Do not capture the dynamics of strategic interactions among the individuals in the network

Game theory helps to overcome this fundamental inadequacy

Initial Efforts in this Direction

- Siddharth Suri. The Effects of Network Topology on Strategic Behavior. PhD Thesis, Dept. of Computer and Information Science, University of Pennsylvania, USA, 2007.
- Sanjeev Goyal. Connections: An Introduction to the Economics of Networks. Princeton University Press, Princeton and Oxford, 2007.
- Eyal Even-Dar, Michael J. Kearns, Siddharth Suri. A network formation game for bipartite exchange economies. In SODA 2007.
- Jon M. Kleinberg, and Eva Tardos. Balanced outcomes in social exchange networks. In STOC, 2008.
- Jon M. Kleinberg, Siddharth Suri, Eva Tardos, and Tom Wexler. Strategic Network Formation with Structural Holes. In ACM EC, 2008.
- Matthew O. Jackson. Social and Economic Networks. Princeton University Press, Princeton and Oxford, 2008.

Game Theoretic Models for SNA: Two Viewpoints



Next Part of the Talk

1. Social Network Analysis: A Quick Primer
2. **Foundational Concepts in Game Theory**
3. Network Formation Problem
4. Summary and To Probe Further

Game Theory

- *Game Theory:* Mathematical framework for rigorous study of conflict and cooperation among rational, intelligent agents.
- *Applications:*
  - Microeconomics, Sociology, Evolutionary Biology
  - Auctions and Market Design
  - Computer Science: Algorithmic Game Theory, Internet and Network Economics, E-Commerce, etc.
- *Two classes of games:*
  - Non-cooperative Game Theory
  - Cooperative Game Theory

Non-Cooperative Game Theory

- *Representation:*
  - Extensive Form Games
  - Strategic Form Games (or Normal Form Games)
  - Graphical Games

Strategic Form Games

- $N = \{1, 2, \ldots, n\}$
- For each $i \in N$, $S_i = \{s_{i1}, s_{i2}, \ldots, s_{im}\}$ is a set of m pure strategies for player i
- $S = S_1 \times S_2 \times \ldots \times S_n$ is Cartesian product of n strategy sets
- $S = S_1 \times S_2 \times \ldots \times S_{i-1} \times S_{i+1} \times \ldots \times S_n$ is Cartesian product of $n - 1$ strategy sets
- Each element $s \in S$ is called a *strategy profile* and it is in the following form:

  $$s = (s_1, s_2, \ldots, s_n)$$

  where $s_i \in S_i$, $\forall i \in N$.
- $s_{-i} = \left(s_1 \times s_2 \times \ldots \times s_{i-1} \times s_{i+1} \times \ldots \times s_n\right) \in S_{-i}$ is a profile of $n - 1$ strategies
- For each $i \in N$, $u_i : S \to \mathbb{R}$ is a payoff function for player i
- $\Gamma = \left(N, (S_i)_{i \in N}, (u_i)_{i \in N}\right)$ is called the strategic form game

Pure Strategy Nash Equilibrium (PSNE)

- *Best Response Sets:* For each player $i \in N$,

$$B_i(s_{-i}) = \{s_i \in S_i \mid u_i(s_i, s_{-i}) \geq u_i(s_i^{'}, s_{-i}), \ \forall s_i^{'} \in S_i\}$$

  for every $s_{-i} \in S_{-i}$.
- A profile of strategies $(s_1^*, s_2^*, \ldots, s_n^*)$ is said to be a pure strategy Nash equilibrium (PSNE) if s_i^* is a best response strategy against s_{-i}^* $\forall i = 1, 2, \ldots, n$. That is,

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*), \qquad \forall i = 1, 2, \ldots, n.$$

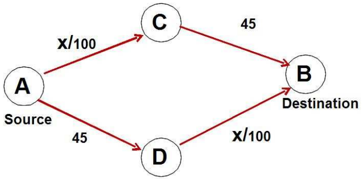
- A Nash equilibrium profile is:
  - Robust to unilateral deviations
  - Captures a stable and self-enforcing agreement among the players
  - A principled way of predicting a steady-state outcome of a dynamic adjustment process

PSNE Example: Presoner's Dilemma

	No Confess NC	Confess C
No Confess NC	- 2, - 2	- 10, - 1
Confess C	-1, - 10	- 5, - 5

- $N = \{1, 2\}$
- $S_1 = \{NC, C\}$ and $S_2 = \{NC, C\}$
- $S = \{(NC, NC), (NC, C), (C, NC), (C, C)\}$
- $u_1(NC, NC) = -2$, $u_1(NC, C) = -10$, $u_1(C, NC) = -1$, and $u_1(C, C) = -5$
- $u_2(NC, NC) = -2$, $u_2(NC, C) = -1$, $u_2(C, NC) = -10$, and $u_2(C, C) = -5$
- (C, C) is a PSNE

PSNE Example: Traffic Routing Game



- $N = \{1, 2, \ldots, 4000\}$
- $S_1 = S_2 = \ldots = S_{4000} = \{C, D\}$
- For each $i \in N$ and $s \in S$, $u_i(s)$ is the sum of the latencies of the path chosen from A to B
- Any strategy profile with 2000 C's and 2000 D's is a PSNE
    - $u_i\Big(C, \ldots, C, D, \ldots, D\Big) = -(20 + 45) = -65$

Mixed Strategy Nash Equilibrium

- **Nash's Result:** Every finite strategic form game has at least one mixed strategy Nash equilibrium.

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:
  - $N = \{s, b_1, b_2\}$
  - $v(\{s\}) = 0$, $v(\{b_1\}) = 0$, $v(\{b_2\}) = 0$, $v(\{b_1, b_2\}) = 0$
    $v(\{s, b_1\}) = 5$, $v(\{s, b_2\}) = 10$, $v(\{s, b_1, b_2\}) = 10$

Cooperative Game Theory (Cont.)

- **Key Question:** How should a coalition that forms divide its winnings among its members?
- A payoff allocation $x = (x_1, x_2, \ldots, x_n)$ is any vector in $\mathbb{R}^n$ where x_i is the utility payoff to player i
- Any allocation $x = (x_1, x_2, \ldots, x_n)$ is said to be *feasible* for a coalition C if and only if

$$\sum_{i \in C} x_i \leq v(C)$$

- Any allocation $x = (x_1, x_2, \ldots, x_n)$ is said to be *individually rational*, if $x_i \geq v(\{i\}), \forall i \in N$
- Any allocation $x = (x_1, x_2, \ldots, x_n)$ is said to be *collectively rational*, if $\sum_{i \in N} x_i = v(N), \forall i \in N$
- Any allocation $x = (x_1, x_2, \ldots, x_n)$ is said to be *coalitionally rational*, if $\sum_{i \in C} x_i \geq v(C), \forall C \subseteq N$

The Core

- The core of a TU game (N, v) is the set of all payoff allocations that are individually rational, coalitionally rational, and collectively rational. That is,

$$Core = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : \sum_{i \in N} x_i = v(N); \sum_{i \in C} x_i \geq v(C), \forall C \subseteq N\}$$

- **Example:** Consider a glove market. Let $N = \{1, 2, 3, 4, 5\}$. The first two players can supply left gloves and the other three players can supply right gloves; $N_L = \{1, 2\}$ and $N_R = \{3, 4, 5\}$. Suppose the worth of each coalition is the number of matched pairs that it can assemble. That is,

$$v(C) = min\{|C \cap N_L|, |C \cap N_R|\}$$

The core for this game is a singleton set $\{(1, 1, 0, 0, 0)\}$.

The Core (Cont.)

- If $N_L = \{1, 2, 3\}$ and $N_R = \{4, 5\}$, then the core of this modified game would be a singleton set $\{(0, 0, 0, 1, 1)\}$
- If $N_L = \{1, 2\}$ and $N_R = \{3, 4\}$, then the core of this modified game would be a singleton set $\{(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})\}$
- The core of a cooperative game can be a singleton set or very large or empty

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, following an axiomatic approach. This was part of his doctoral dissertation at the Princeton University. Given a cooperative game (N, v), the Shapley value is denoted by $\phi(v)$:

$$\phi(v) = \{\phi_i(v), \phi_2(v), \ldots, \phi_n(v)\}$$

  where $\phi_i(v)$ is the expected payoff to player i
- Shapley proposed three axioms: Symmetry, Linearity, and Carrier

The Shapley's Theorem

- **Theorem:** There is exactly one mapping $\phi : \mathbb{R}^{2^N-1} \to \mathbb{R}^N$ that satisfies Symmetry, Linearity, and Carrier axioms. This function satisfies: $\forall i \in N$, $\forall v \in \mathbb{R}^{2^N-1}$,

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

- **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 $\phi_1(v) = 200$, $\phi_2(v) = 50$, $\phi_3(v) = 50$

Next Part of the Talk

Network Formation

- In the process of information dissemination or in general value creation, nodes receive not only various kinds of benefits but also incur costs in terms of time, money, and effort.

Network Formation

- In the process of information dissemination or in general value creation, nodes receive not only various kinds of benefits but also incur costs in terms of time, money, and effort.
- Hence individual nodes do act strategically.

Network Formation

- In the process of information dissemination or in general value creation, nodes receive not only various kinds of benefits but also incur costs in terms of time, money, and effort.
- Hence individual nodes do act strategically.
- It is essential to study:
  - How to model the formation of social networks in the presence of strategic nodes that are interested in maximizing their payoffs from the social interactions?
  - What are the networks that will emerge due to the dynamics of network formation and what their characteristics are likely to be?

Models of Social Network Formation

Two popular models:

1. **Random graph models:**
    - Links form by chance and simply governed by probabilistic rules.
2. **Game theoretic models:**
    - Links form by choice more often.
    - Capture social and economic incentives while forming links.
    - Nodes often act strategically as link formation incurs costs in terms of cost, money and effort.

Game Theoretic Models of Social Network Formation

- How does networks form?
  1. **Defining a Model of Social Network Formation:** Need to capture major key determinants of network formation process
- What are the networks that will emerge finally that satisfy certain desirable properties?
  1. **Analysis of the Model:** We analyze the topologies of the networks that satisfy two key features namely stability and efficiency

Network Formation Game

We represent the corresponding strategic network formation game with 3-tuple $\langle N, (S_i)_{i \in N}, (u_i)_{i \in N} \rangle$ where

1. N is the set of individuals in the network and we call them players.
2. For each $i \in N$, S_i is the set of strategies of player i. A strategy $s_i \in S_i$ of player i is the set of individuals with which player i wants to form a link.
3. For each $i \in N$, u_i is the utility of individual i and this utility depends on its neighborhood and the structure of the network.

Network Formation Game

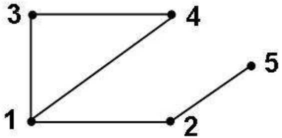
Two fundamentally ways of modeling the formation of social contacts:

1. *Two-sided Link Formation:* A link is formed under mutual consent; and
2. *One-sided Link Formation:* A link if formed under the consent of either of the individuals involved in the link formation.

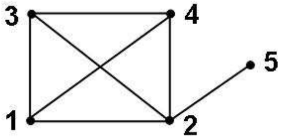
Network Formation Game: An Example

Consider $N = \{1,2,3,4,5\}$ be the set of players. Assume that the strategies of the players are as follows:

$$\begin{aligned} S_1 &= \{2,3,4\}, \\ S_2 &= \{1,3,4,5\}, \\ S_3 &= \{1,4\}, \\ S_4 &= \{1,3\}, \\ S_5 &= \{2\}. \end{aligned}$$



(i)



(ii)

A Glimpse of State-of-the-Art

- M. O. Jackson. Social and Economic Networks. Princeton University Press, Princeton and Oxford, 2008.
- S. Goyal. Connections: An Introduction to the Economics of Networks. Princeton University Press, Princeton and Oxford, 2007.
- G. Demange and M. Wooders. Group Formation in Economics: Networks, Clubs, and Coalitions. Cambridge University Press, Cambridge and New York, 2005.
- M. Slikker and A. van den Nouweland. Social and Economic Networks in Cooperative Game Theory. Kluwer Academic Publishers, Massachusetts, USA and The Netherlands, 2001.

A Glimpse of State-of-the-Art (Cont.)

- M. O. Jackson and A. Wolinsky. A strategic model of social and economic networks. Journal of Economic Theory, 71(1):44-74, 1996.
- S. Goyal and F. Vega-Redondo. Structural holes in social networks. Journal of Economic Thoery, 137(1):460-492, 2007.
- V. Buskens and A. van de Rijt. Dynamics of networks if everyone strives for structural holes. American Journal of Sociology, 114(2):371-407, 2008.
- J. Kleinberg, S. Suri, E. Tardos, and T. Wexler. Strategic network formation with structural holes. In Proceedings of the 9th ACM Conference on Electronic Commerce (EC), pages 284-293, 2008.

Key Observations

- Several game theoretic models are proposed to capture the rational behavior of nodes.
- Some of these studies are able to yield sharp predictions on the network topologies that emerge, if stability and efficiency are to be satisfied.
- Various notions of stability and efficiency have been employed.
- Significant emphasis on the tradeoffs between stability and efficiency.

Stability

Stability: A network is stable if it is in a strategic equilibrium.
Examples: Nash equilibrium, Pairwise stability.

Nash Equilibrium: A network is said to be in Nash equilibrium if no node unilaterally forms or deletes a link to any other node.

Pairwise Stability: A network is said to be pairwise stable if

- if no node gains by deleting a link to any other node
- no pair of nodes wants to add a link between them

M.O. Jackson and A. Wolinsky. A strategic model of social and economic networks. In Journal of Economic Theory, 71:44–74, 1996.

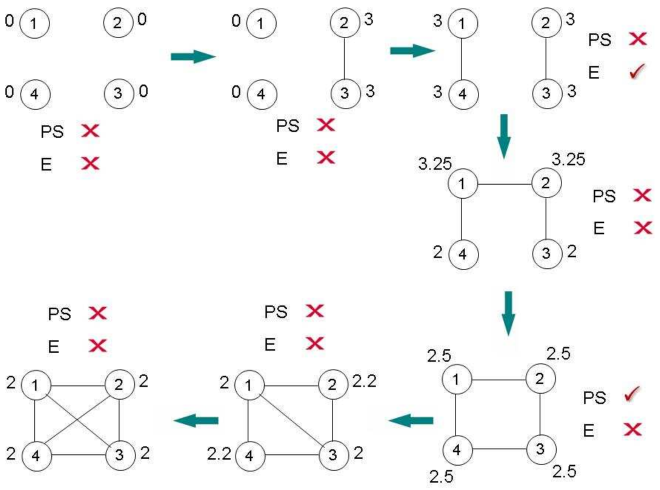
Efficiency

Efficiency: A network is efficient if a function of the utilities of the nodes is maximized.
Example: Pareto efficiency, Maximize sum of utilities.

Pareto Efficiency: A network g is said to be Pareto efficient if there is no network g' in which the utility of at least one node is strictly greater than that of in g and the utilities of the rest of the nodes are greater than or equal to that of in g.

Sum of Utilities: A network g is said to be efficient if the sum of utilities of the nodes in g is greater than or equal to that of any other network.

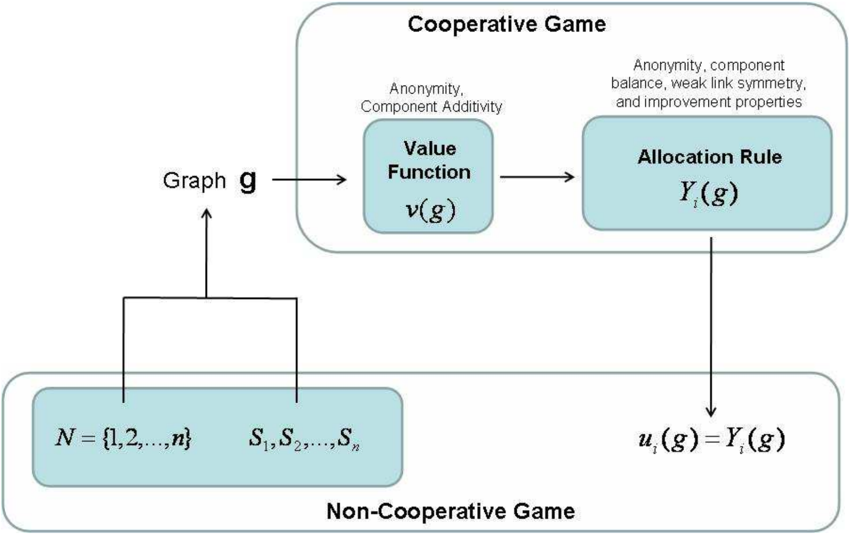
Stability and Efficiency Tradeoffs



Four Key Network Determinants

1. Benefits from Direct Links
2. Cost of the Direct Links
3. Decaying Benefits from Non-neighbor Nodes
4. Bridging Benefits

The Model


Cooperative Game
Anonymity, Component Additivity
Value Function
$v(g)$
Anonymity, component balance, weak link symmetry, and improvement properties
Allocation Rule
$Y_i(g)$
Graph g
$N = \{1, 2, \ldots, n\}$
$S_1, S_2, \ldots, S_n$
$u_i(g) = Y_i(g)$
Non-Cooperative Game

The Model

- Let $N = \{1, 2, \ldots, n\}$ be the set of n (≥ 3) nodes.
- A strategy s_i of a node i is any subset of nodes with which it establishes links.
- Links are formed under mutual consent.
- S_i is the set of strategies of node i.
- Each $s = (s_1, s_2, \ldots, s_n)$ leads to an undirected graph and we represent it by $g(s)$.
- Let $\Psi(S)$ be the set of all such undirected graphs.
- When the context is clear, we use g and Ψ instead of $g(s)$ and $\Psi(S)$ respectively.

The Model (Cont.)

- $\forall i, j \in N$, $d_g(i,j)$ = length of shortest path between i and j.
- *Costs:* If nodes i and j are connected by a link in g, then we assume that the link incurs a cost $c > 0$.
- *Benefits:* The communication between i and j leads to a benefit of $b(d_g(i,j))$.
- We assume that $b(.)$ is a non-increasing function, implying that the benefit of communication decays as the length of shortest path increases.
- A value function $v : \Psi \to \mathbb{R}$ for a given graph $g \in \Psi$ is as follows:

$$v(g) = \sum_{\substack{x,y \in N, \\ (x,y) \in g}} [b(1) - c] + \sum_{i \in N} \sum_{\substack{j \in N, \\ j > i, \\ (i,j) \notin g}} b(d_g(i,j)) \qquad (1)$$

The Model (Cont.)

- **Lemma:** The proposed value function $v(.)$ satisfies anonymity and component additivity.
- The network value $v(g)$ is divided among the nodes in g as utilities using an allocation rule. Allocation rule $Y : \Psi \to \mathbb{R}^n$ distributes the network value $v(g)$ among nodes as utilities such that

$$\sum_{i \in N} Y_i(g) = v(g), \ \forall g \in \Psi.$$

- We define $Y_i(g)$ to be the utility $(u_i(g))$ of node i.

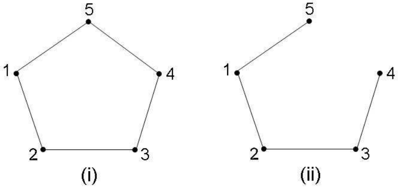
$$u_i(g) = Y_i(g), \ \forall i \in N.$$

- This framework clearly defines a strategic form game: $\Gamma = (N, (S_i)_{i \in N}, (u_i)_{i \in N})$.

Axiomatic Allocation Rule

- **Anonymity:** Allocation rule Y is anonymous if for any v, $g \in \Psi$, and any permutation of the players π, $Y_{\pi(i)}(g^{\pi}) = Y_i(g)$.
- **Component Balance:** Allocation rule Y is *component balanced* if $\sum_{i \in C} Y_i(g) = v(C)$ for each component additive v, $g \in \Psi$, and for each component C in $\Omega(g)$ (the set of all components of the graph g).
- **Weak Link Symmetry:** Allocation rule Y satisfies *weak link symmetry* if for each link $e = (i,j) \notin g$, it holds that if $Y_i(g \cup \{e\}) > Y_i(g)$, then $Y_j(g \cup \{e\}) > Y_j(g)$.
- **Improvement Property:** Allocation rule Y satisfies *improvement property* if for each link $e = (i,j) \notin g$, whenever there exists a node $z \in N \setminus \{i,j\}$ such that $Y_z(g \cup \{e\}) > Y_z(g)$, then $Y_i(g \cup \{e\}) > Y_i(g)$ or $Y_j(g \cup \{e\}) > Y_j(g)$.

An Illustrative Example



$$v(g) = 5(b(1) - c) + 5b(2),$$
$$u_i(g) = (b(1) - c) + b(2) \quad \forall i \in \{1, 2, 3, 4, 5\}.$$

$$v(g') = 4(b(1) - c) + 3b(2) + 2b(3) + b(4),$$
$$u_1(g') = u_3(g') = (b(1) - c) + \frac{2}{3}b(2) + \frac{1}{2}b(3) + \frac{1}{5}b(4),$$
$$u_2(g') = (b(1) - c) + b(2) + \frac{1}{2}b(3) + \frac{1}{5}b(4),$$
$$u_4(g') = u_5(g') = \frac{1}{2}(b(1) - c) + \frac{1}{3}b(2) + \frac{1}{4}b(3) + \frac{1}{5}b(4).$$

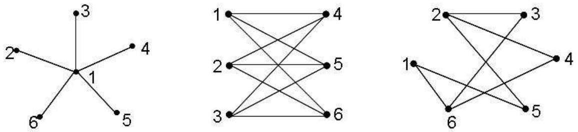
Pairwise Stability and Efficiency

- **Pairwise Stability:** A network g is said to be *pairwise stable* with respect to the value function v and the allocation rule Y if (i) for each edge $e = (i,j) \in g$, $Y_i(g) \geq Y_i(g \setminus \{e\})$ and $Y_j(g) \geq Y_j(g \setminus \{e\})$, and (ii) for each edge $e' = (i,j) \notin g$, if $Y_i(g) < Y_i(g \cup \{e'\})$ then $Y_j(g) > Y_j(g \cup \{e'\})$.

- **Efficiency:** A network $g \in \Psi$ is said to be efficient if $v(g) \geq v(g') \;\; \forall g' \in \Psi$.

Minimal Edge Graphs with Diameter p $(1 < p < n)$

- **Definition:** The diameter of a graph is the length of a longest shortest path between any two vertices of the graph.
- **Definition:** A graph with diameter p is said to be a *minimal edge graph with diameter p* if the deletion of any edge in the graph results in a graph with diameter greater than p.
- Given a set of n nodes, there may be multiple minimal edge graphs with diameter p for $1 < p < n$. For example, the following figure shows three different minimal edge graphs with diameter 2.



(i) (ii) (iii)

Analysis of Efficient Networks

The following two results useful in characterizing the topologies of efficient networks.

Lemma: Given a graph g, if $(b(1) - b(2)) < c < (b(1) - b(3))$ and there exists a pair of nodes x and y such that $d_g(x, y) > 2$, then forming a link between x and y strictly increases the value of g.

Proof

Lemma: If $(b(1) - b(2)) < c < (b(1) - b(3))$, then every efficient network is a minimal edge graph with diameter 2.

Proof

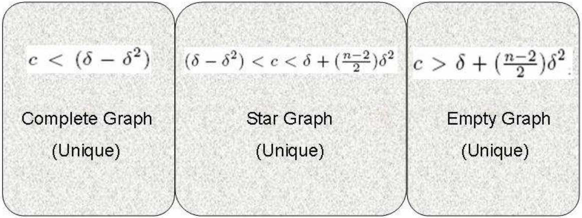
Analysis of Efficient Networks (Cont.)

Theorem: Following our proposed model,

(i) if $c < (b(1) - b(2))$, then the complete graph is the unique topology possible for an efficient network

(ii) if $(b(1) - b(2)) < c \leq b(1) + (\frac{n-2}{2})b(2)$, then the star network is the unique topology possible for an efficient network

(iii) if $c > b(1) + (\frac{n-2}{2})b(2)$, then the only efficient network is the empty graph.



Analysis of Pairwise Stable Networks

A few useful results are as follows.

- **Lemma:** For any graph g, if a pair of non-neighbor nodes i and j form a link (i,j) such that $v(g \cup \{(i,j)\}) > v(g)$, then it holds that both $Y_i(g \cup \{(i,j)\}) > Y_i(g)$ and $Y_j(g \cup \{(i,j)\}) > Y_j(g)$.
- **Lemma:** For any graph g, if a node i severs a link $e = (i,j) \in g$ with a node j such that $v(g \setminus \{(i,j)\}) \leq v(g)$, then it holds that $Y_i(g \setminus \{(i,j)\}) \leq Y_i(g)$.
- **Corollary:** For any graph g, under our model, if a node i severs a link $e = (i,j) \in g$ with a node j such that $v(g \setminus \{(i,j)\}) < v(g)$, then it must hold that $Y_i(g \setminus \{(i,j)\}) < Y_i(g)$.

Analysis of Pairwise Stable Networks (Cont.)

- **Lemma:** If $c < (b(1) - b(2))$, then the complete graph is the unique topology possible for a pairwise stable graph.
- **Regularity Condition (RC):** This involves a couple of conditions:
(a) If a pair of nodes i and j in a graph g are not neighbors and form a link (i, j) such that $v(g \cup \{(i, j)\}) \leq v(g)$, then it implies that either $Y_i(g \cup \{(i, j)\}) \leq Y_i(g)$ or $Y_j(g \cup \{(i, j)\}) \leq Y_j(g)$.
(b) $Y_i(g) \geq 0,\ \forall i \in N$.
- **Lemma:** If $c \in (b(1) - b(2), b(1)]$ and RC is satisfied, then any minimal edge graph with diameter 2 is pairwise stable.

Analysis of Pairwise Stable Networks (Cont.)

- **Corollary:** If $c \in (b(1) - b(2), b(1)]$ and RC is satisfied, then the star graph and the completely connected bi-partite graph are pairwise stable.
- **Lemma:** If $(b(1) - b(p)) < c < (b(1) - b(p+1))$ for any integer $p > 1$ and if g is a pairwise stable graph, then g is a graph with diameter p.
- **Lemma:** If $c > b(1) + b(2)$, then the empty graph is pairwise stable.

Efficiency versus Pairwise Stability

Theorem: Consider an anonymous and component additive value function v; and an anonymous, component balanced allocation rule $Y(.)$ satisfying weak link symmetry and improvement properties. Suppose g is an efficient graph relative to v. Then g is pairwise stable if and only if v, Y, and g satisfy the *regularity condition (RC)*.

▸ Proof

Next Part of the Talk

To Probe Further: Important Research Directions

- **Time Varying Graphs:** Typically, the structure of networks change over time. Designing game theoretic models for such time varying graphs is a challenging and interesting research direction
- **Probabilistic Graphs:**
    - Complex networks often entail uncertainty and thus can be modeled as probabilistic graphs
    - M. Potamias, F. Bonchi, A. Gionis, and G. Kollios. *k*-nearest neighbors in uncertain graphs In VLDB Endowment, Vol. 3, No. 1, 2010

To Probe Further: Important Research Directions (Cont.)

- Exploit games with special structure such as convex games, potential games, matrix games, etc. to problems in SNA
- Designing scalable approximation algorithms with worst case guarantees
- Problems such as incentive compatible learning and social network monetization are at the cutting edge
- Explore numerous solution concepts available in the ocean of game theory literature

To Probe Further: 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.

To Probe Further: Important References

- Ramasuri Narayanam and Y. Narahari. A Shapley Value based Approach to Discover Influential Nodes in Social Networks. In IEEE Transactions on Automation Science and Engineering (IEEE TASE), 2011.
- Ramasuri Narayanam and Y. Narahari. Topologies of Strategically Formed Social Networks Based on a Generic Value Function - Allocation Rule Model. Social Networks, 33(1), 2011.
- Ramasuri Narayanam and Y. Narahari. Determining Top-k Nodes in Social Networks using the Shapley Value. In AAMAS, pages 1509-1512, Portugal, 2008.
- Ramasuri Narayanam and Y. Narahari. Nash Stable Partitioning of Graphs with Application to Community Detection in Social Networks. Under Review, 2010.
- D. Dikshit and Y. Narahari. Truthful and Quality Conscious Query Incentive Networks. In Workshop on Internet and Network Economics (WINE), 2009.
- Mayur Mohite and Y. Narahari. Incentive Compatible Influence Maximization in Social Networks with Application to Viral Marketing. AAMAS 2011.

To Probe Further: Useful Resources (Cont.)

Network Data Sets:

- 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
- ...

To Probe Further: Useful Resources (Cont.)

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
- ...

To Probe Further: Useful Resources (Cont.)

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
- . . .

To Probe Further: Useful Resources (Cont.)

- Y. Narahari, Dinesh Garg, Ramasuri Narayanam, Hastagiri Prakash. Game Theoretic Problems in Network Economics and Mechanism Design Solutions. In Series: Advance Information & Knowledge Processing (AIKP), Springer Verlag, London, 2009.
- Home page of Y. Narahari: http://lcm.csa.iisc.ernet.in/hari/
- Home page of Ramasuri Narayanam: http://lcm.csa.iisc.ernet.in/nrsuri/
- Blog on Social Networks: http://cs2socialnetworks.wordpress.com/

Summary of the Tutorial

- We first presented the important fundamental concepts in social network analysis and game theory
- We then presented game theoretic models for four important problems in social network analysis
  - Social network formation
- Game theory imparts more power, more efficiency, more naturalness, and more glamour to SNA problem solving
- Sensational new algorithms for SNA problems? Still a long way to to but the potential is good. Calls for a much deeper study

Thank You

Proof:

- Given that $(b(1) - b(2)) < c < (b(1) - b(3))$.
- Consider any network g and assume that there exists a pair of nodes x and y such that $d_g(x, y) > 2$.
- Recall that the communication between nodes x and y in g leads to a benefit of $b(d_g(x, y))$.
- Assume that x and y form a link and call the link $e = (x, y)$ and also call the new graph $g' = g \cup \{e\}$.

Go Back

- Link (x, y) leads to a net benefit of $(b(1) - c)$.
- Note that the length of a shortest path between any pair of nodes in g' either remains same or decreases when compared to that in g.
- From the above observations, we get that $v(g') - v(g) \geq [(b(1) - c) - b(d_g(x, y))] > 0$, since $d_g(x, y) > 2$ and $(b(1) - c) > b(3) \geq b(d_g(x, y))$.
- That is, $v(g') > v(g)$.

Go Back

Proof:

- Consider that g is an efficient graph.
- Due to previous lemma, the shortest distance between any pair of nodes is at most 2 in g.
- That is, g is a graph with diameter 2.
- Suppose that g is not a minimal edge graph with diameter 2.

Go Back

- Then g contains a link (x, y) such that severing the link (x, y) does not lead the diameter to exceed 2.
- Thus, if we remove the link (x, y), only the shortest distance between the nodes x and y increases to 2.
- Since $(b(1) - b(2)) < c$, the value of g strictly increases if the link (x, y) is severed.
- Contradiction to the fact that g is efficient.

Go Back

Proof Part 1:

- Given that g is efficient. Assume that g is pairwise stable.
- *Claim:* Regularity condition (RC) is holds.
- Let i and j be a pair of non-neighbor nodes in g and form a link, (i, j). Call the new graph $g' = g \cup \{(i, j)\}$.
- Since g is efficient and v is component additive, we get that $v(g') \leq v(g)$.
- If $Y_i(g') > Y_i(g)$ (or $Y_j(g') > Y_j(g)$), then due to weak link symmetry, we get that $Y_j(g') > Y_j(g)$ (or $Y_i(g') > Y_i(g)$).
- Contradicts the fact that g is pairwise stable.
- Also, since g is pairwise stable, we have that $Y_i(g) \geq 0, \forall i \in N$.

Go Back

Proof Part 2:

- Given that g is efficient. Assume that the regularity condition is satisfied.
- *Claim:* g is pairwise stable for Y relative to v.
- *Severing a Link:* Suppose a node x_1 severs a link (x_1, y_1) with node y_1 in g. Call the new graph $g_1 = g \setminus \{(x_1, y_1)\}$.
- Since g is efficient and v is component additive, we get that $v(g_1) \leq v(g)$.
- From previous results, it is clear that node x_1 is not strictly better off by severing the link.

Go Back

- *Adding a Link:* Suppose two non-neighbor nodes i and j form a link (i, j) in g and call the new graph $g' = g \cup \{(i, j)\}$.
- Since g is efficient and v is component additive, it holds that $v(g') \leq v(g)$.
- If $Y_i(g') > Y_i(g)$ (or $Y_j(g') > Y_j(g)$), then due to weak link symmetry, we get that $Y_j(g') > Y_j(g)$ (or $Y_i(g') > Y_i(g)$). Contradiction to *RC*!
- Implies that neither i nor j is strictly better off. Again since *RC* is satisfied, we get that $Y_i(g) \geq 0$, $\forall i \in N$.
- Hence, g is pairwise stable.

Go Back