Tutorial: Economic Aspects of Social Networks (part II)

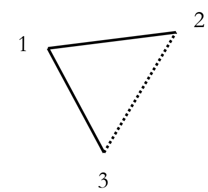
Brian Rogers

ACM EC08, Chicago

Outline

- Empirical regularities of social networks
- Capturing these features in models
- Strategic aspects of network formation
- Strategic interactions on networks

Illustration of Concepts



- Degree:
  - *With* link 23, degrees are (2, 2, 2)
  - *Without* link 23, degrees are (2, 1, 1)
- Clustering: What is the probability of link 23?
- Diameter:
  - *With* link 23, diameter = 1
  - *Without* link 23, diameter = 2

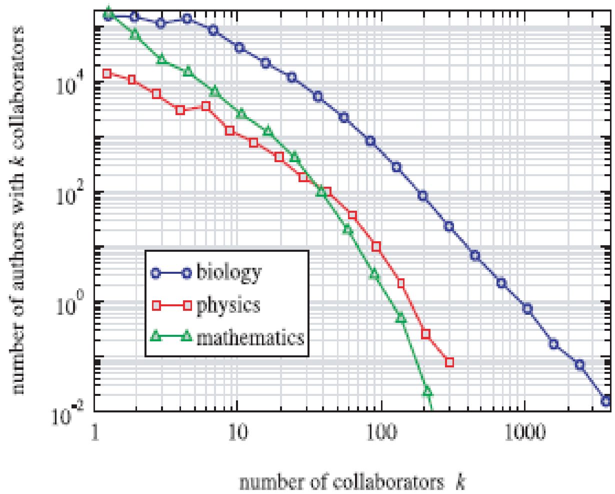
Social Networks Share Many Features

1. Degree distributions: Heavy tails, “scale-free”

   - Many more nodes with very high and very low degrees relative to what one would find in a completely random network

Degree Distributions

Co-Authorship Data (Newman and Grossman)



Features of Social Networks

1. Degree distributions: Heavy tails, "scale-free"
2. High clustering – "cliquishness"
   - Proportion of triads out of possible triads in a (sub)graph
   - Any two of a given node's neighbors are likely to themselves be neighbors

Movie Actors	Math Co-authors	Physics Co-authors	WWW
0.79	0.43	0.15	0.11

Features of Social Networks

1. Degree distributions: Heavy tails, “scale-free”
2. High clustering – “cliquishness”
3. Low diameter and average path length

	Movie Actors	Math Co-authors	Physics Co-authors	WWW
Average	3.7	7.6	5.9	3.1
Diameter		27	20	

Features of Social Networks

1. Degree distributions: “Scale-free’’– heavy tails
2. High clustering – “cliquishness”
3. Low diameter and average path length
4. Assortativity
   - Positive correlation in the degrees of linked nodes
   - Special to socially-generated networks

Assortativity

- Correlation in Degree (Newman 2003):
  - Socially generated networks:
    - .12 math co-authorship, .13 biology, .36 physics
    - .09 emails
    - .21 film actors
  - Technologically generated networks
    - -.19 internet
    - -.003 power grid
    - -.23 neural network

Many other regularities

- Negative relationship between degree and local clustering
- Homophily
- Strength of weak ties (Granovetter)
- Structural holes (Burt)
- Many more to be found. Which are the “right” ones?

Outline

- Empirical regularities of social networks
- Capturing these features in models
- Strategic aspects of network formation
- Strategic interactions on networks

Goals

- Develop models with several goals in mind
  - Consistent with observed regularities
  - Based on reasonable micro-level assumptions
  - Analytically tractable
  - Yield novel insights

A Network Formation Model
(Jackson, Rogers)

- Agents indexed by date of birth T={1,2,3,...}
- Upon birth, agent meets m_r others uniformly at random
- Search those neighborhoods to find m_n more nodes
  - Like entering at a random web page and following links
- Link to a given node if net utility is positive (probability p)

Key Intuitions

- Network-based meetings: likely to find nodes with many links
  » variation on scale-free: growth partly proportional to size
- Random aspect: not entirely scale-free
  » fit lower tail too
- Local search: may connect to two nodes that are already linked
  » high clustering

Key Intuitions (cont.)

- Search aspect generates hub-like nodes
  - » low diameter
- Randomness connects different neighborhoods
  - » even lower diameter: order ln(n)/ln(ln(n))
- Nodes enter sequentially and connect to existing nodes
  - » Assortativity: age is correlated with both degree and links

Mean-field approximation

- Network formation is stochastic and path-dependent
  - Difficult to analyze directly
- Use a deterministic continuous-time system
  - All changes occur at the mean rate of underlying stochastic system
- Apply, e.g., to expected changes in degrees of nodes
- Can analytically solve for steady state

Important Parameters

$r = m_r/m_n$

- Ratio of the number of links formed at random vs. by network-based meetings

$m = p(m_r+m_n)$

- Average number of out-links per node

(In-)Degree Distribution

Expected increase in the in-degree of a node i at time t is roughly

prob found at random

prob found through search

$$p\ (\ m_r\ /t\ +[d_i\ (m_r\ /t)]\ [m_n\ /(m_r\ m)]\ )$$

prob linked to given found

number of neighbors

prob a neighbor is entry point

prob found from search of neighborhoods of entry points

Theorem: Degree Distribution

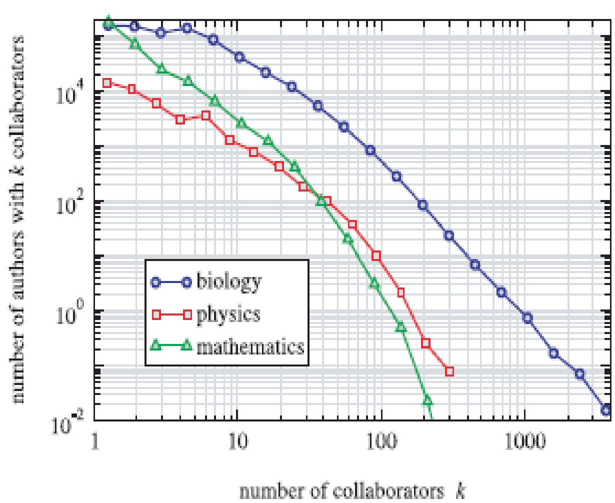
The in-degree distribution of the mean-field process has a degree distribution with cdf

$$F(d) = 1 - [rm/(d+rm)]^{1+r}$$

- Approximates a power distribution for large d
- Lower tail is thinner

Degree Distributions

Co-Authorship Data (Newman and Grossman)



Sketch of Proof

1. Mean-field:

   $\partial d_i(t)/\partial t = E[\text{change}(d_i(t))]; \quad d_i(i) = 0$

2. Solve the differential equation:

   $d_i(t) = rm(t/i)$^[1/(1+r)]-rm

3. $1-F_t(d)$ = proportion of nodes with degree > d at time t
   - Let i*(d) be such that $d_{i^*(d)}(t)=d$
   - => $1-F_t(d)=i^*(d)/t$
4. Solve for i*(d) from Step 2 and plug in

Degree distribution shifts

- Consider a distribution F with parameters (m,r) and a distribution F' with parameters (m',r'). Then:
  - If r' = r and m' > m, then F' strictly FOSD F
    - I.e., F'(d) < F(d) for all d
  - If m' = m and r' < r, then F' is a strict MPS of F

Simple implications for efficiency

- If utility to a node (information, employment, etc.) is increasing in degree, then average utility increases with m.
- If utility to a node is concave in degree, then average utility increases with r.

Implications for Diffusion: SIS Model

- Stochastic process where behavior/state of agents change depending on behavior/states of neighbors
  - Change from "S" to "I" with probability proportional to # of neighbors in state "I"
  - Revert from "I" to "S" at random
- Originally used to model the spreading of disease (common cold, computer virus)
- Can also model behavior
  - Adopting a new behavior (buying new product, coordination game)
  - Myopic best-response dynamics, may revert to status quo

The model

- Fix the network. Diffusion occurs in discrete time.
- θ_i is proportion of i's infected neighbors
- "S" → "I": probability $\nu(d_i\theta_i)$
- "I" → "S": probability $0 < \delta < 1$
- Set $\lambda = \nu/\delta$

- $\rho(d)$ = average infection rate of degree-d nodes
- $\rho = \sum_d \rho(d)F(d)$ = average infection rate of society
- $\theta = \sum_d d\rho(d)F(d) / m$
  - Probability of linking to an infected node

Results for SIS diffusion (Jackson, Rogers)

Proposition: There is positive steady-state infection iff $\lambda > E[d]/E[d^2]$.

Proposition: Suppose F is such that $\theta^* > 0$.

- If F' strictly FOSD F, then $\theta^{*\prime} > \theta^*$ and $\rho^{*\prime} > \rho^*$.
- If F' is a strict MPS of F, then $\theta^{*\prime} > \theta^*$.
  - Note: overall infection rate ρ can go up or down

Corollaries for Network Formation Model

$$F(d) = 1 - [rm/(d + rm)]^{1+r}$$

- The minimal λ for positive steady-state infection is
  - 0 if $r < 1$
  - $(r - 1)/2rm$ if $r \geq 1$

Given F with (r,m) and F' with (r',m')

- If r' = r and m' > m, then $\theta^{*'} > \theta^*$ and $\rho^{*'} > \rho^*$
- If m' = m and r' < r, then $\theta^{*'} > \theta^*$

Outline

- Empirical regularities of social networks
- Capturing these features in models
- Strategic aspects of network formation
- Strategic interactions on networks

Main questions

- What kind of structures are likely to form when linking choices are made by self-interested individuals?
- What are the welfare properties of these likely outcomes?
- What are the implications for processes occurring through the network?

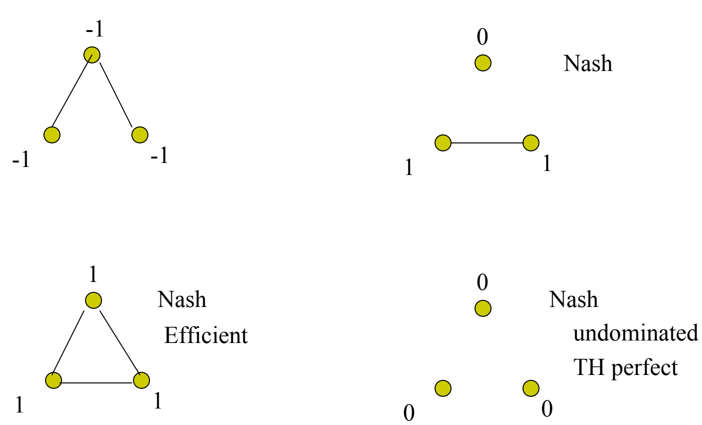
General setup

- Agents $N = \{1,2,\ldots,n\}$
- Network g, where $g_{ij} = 1$ if i and j are connected, 0 otherwise
- Utility functions $u_i(g) : G \rightarrow R$
- Welfare criterion: $U(g) = \sum_i u_i(g)$

Solution concepts

- Consider the following game
  - Agents simultaneously announce potential links
  - Links that are mutually agreed are formed
- Nash equilibrium behaves poorly!
  - Multiplicity
    - E.g., the empty network is always an equilibrium
  - Cannot be easily refined away
    - Undominated strategies
    - Trembling-hand perfection

Nash networks



Pairwise stability

- A network g is pairwise stable if
  - $g_{ij} = 1$ implies $u_i(g) \geq u_i(g - ij)$
    - Nobody wants to delete a link (unilateral)
  - $g_{ij} = 0$ implies if $u_i(g + ij) > u_i(g)$ then $u_j(g + ij) < u_j(g)$
    - No pair wants to add a link (bilateral)
- Game-form independent

A basic "connections" model (Jackson, Wolinsky)

- $u_i(g) = \sum_j \delta^{\wedge} l_{ij}(g) - cg_{ij}$
  - $c > 0$ link cost
  - $0 < \delta \leq 1$ decay parameter
  - $l_{ij}(g)$ length of shortest path between i and j
- What will the set of pairwise stable networks look like?
- Will they be efficient?

Efficient networks in the connections model

Proposition: The efficient network is unique, and given by

- if $c < \delta - \delta^2$ — Complete
- if $\delta - \delta^2 < c < \delta + (n-2)/2*\delta^2$ — Star
- if $\delta + (n-2)/2*\delta^2 < c$ — Empty

Pairwise stability in the connections model

PS networks have the following properties

- There is at most one component
- If $c < \delta - \delta^2$ the only PS network is Complete
- If $\delta - \delta^2 < c < \delta$ then Star, and maybe others, are PS
- If $c > \delta$ then no agent has exactly one link

- Implication: PS networks are not necessarily efficient
  - When c is slightly bigger than δ then no PS can be efficient

General tension between PS and Efficiency

- PS is a mild requirement for stability
  - Not necessarily robust to many kinds of deviations
- What if we also consider transfers among players with the hope that we can induce efficient networks to be PS?
- This turns out to not be generally possible!

General framework

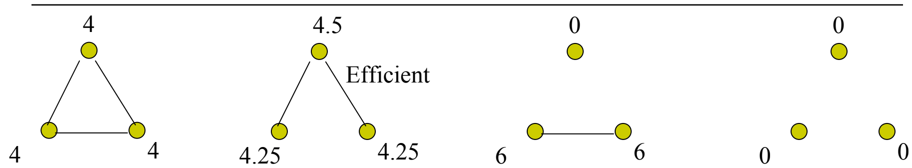
Transfers $t{:}G \rightarrow R^n$ such that $\sum_i t_i(g) = 0$

- Require transfers to be
  - "Component Balanced" whenever utilities are component-based
    - Cannot make net transfers across components
  - Satisfy "Equal Treatment of Equals"
    - If two agents have the same set of neighbors, and always generate the same value to other agents, they must receive the same transfer
- Augment pairwise stability to respect transfers:
  - $g_{ij} = 1$ implies $u_i(g) + t_i(g) \geq u_i(g - ij) + t_i(g - ij)$
  - $g_{ij} = 0$ implies if $u_i(g + ij) + t_i(g + ij) > u_i(g) + t_i(g)$ then $u_j(g + ij) + t_j(g + ij) < u_j(g) + t_j(g)$

Main result (Jackson, Wolinsky)

Proposition: There exist component-based utilities such that every network that is stable relative to a given transfer rule satisfying CB and ETE is not efficient.

Sketch of proof



- Try to design t so that Eff network is PS
- ETE implies all $t_i = 0$ for Complete & Empty
- ETE+CB imply $t_i = 0$ for one-link network
- Take Eff network with agent 2 in center
  - ETE implies $t_1 = t_3$
  - Need $t_1 = t_3 \geq -.25$ (else add link 13)
  - Need $t_2 \geq 1.5$ (else he severs a link)
  - But then $t_1 + t_2 + t_3 > 0$ – a contradiction

An "islands-connections" model (Jackson, Rogers)

- Agents are grouped on "islands"
  - K islands
  - J agents per island
- Linking costs:
  - C for inter-island link, c for intra-island link
  - $C > c > 0$
- Benefit truncation
  - No benefit when distance is greater than D

Small-world features

Proposition: Let $c < \delta - \delta^2$ and $C < \delta + (J - 1)\delta^2$. Any network that is PS and/or Eff is such that:

- Islands are fully intra-connected
- Diameter is no greater than D+1
- Clustering is greater than $(J-1)(J-2)/(J^2K^2)$

Many other aspects of strategic network formation

- Strong stability
- Extensive form games
- Dynamic processes of link addition/deletion
- Directed networks, weighted networks
- A huge variety of specific utility models

Outline

- Empirical regularities of social networks
- Capturing these features in models
- Strategic aspects of network formation
- Strategic interactions on networks

Semi-anonymous games on graphs

- Agents and graph given exogenously
- Agents choose binary action $x_i = 0$ or 1
- Utility has the form $u_d(x_i,m)$
  - Depends on degree d, own action, and
  - m = number of neighbors choosing 1
- Goal: analyze equilibrium properties

Two examples

- Coordination game
  - $u_d(1,m) = a*m/d$
  - $u_d(0,m) = b*(d-m)/d$
    - Action 1 is optimal iff $m/d \geq b/(a+b)$
- Best-shot game
  - $u_d(1,m) = 1- c$ $(0 < c < 1)$
  - $u_d(0,m) = 1$ if $m \geq 1$, 0 otherwise
    - Action 1 is optimal iff $m = 0$

General strategic features

- A game has Strategic Complements if for all d and for all m ≥ m',
  - $u_d(1,m) - u_d(0,m) \geq u_d(1,m') - u_d(0,m')$
  - Ex: coordination game
- Strategic Substitutes if "≤"
  - Ex: best-shot game

Equilibrium properties under SC and SS

- In both cases optimal behavior is characterized by "threshold" strategies
  - There is a function t(d) such that
    - SC: $x_i = 1$ iff $m \geq t(d)$
    - SS: $x_i = 0$ iff $m \geq t(d)$
- Games of SC: There always exists a pure strategy equilibrium, and the set of equilibria form a lattice
- Games of SS: There may not exist a pure strategy equilibrium

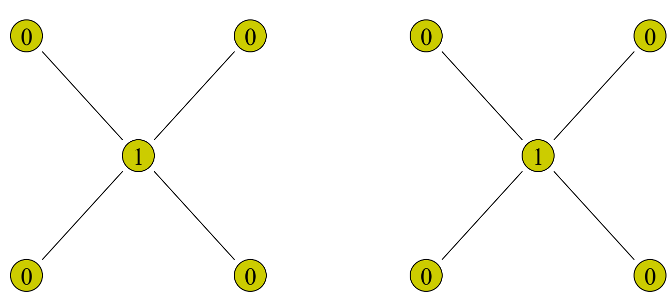
Comparative statics w.r.t. graph

Proposition: Take a game of SC with a non-increasing threshold t(d).

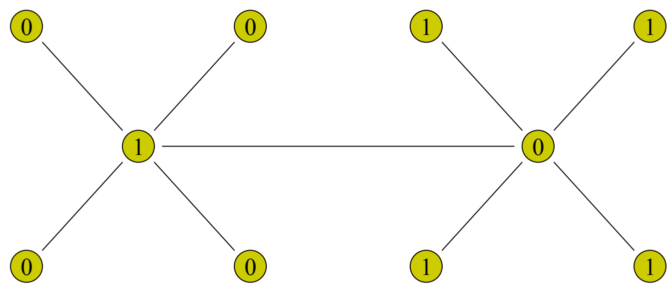
- For every equilibrium x of g and for every g' larger than g, there exists an equilibrium x' of g' where x' ≥ x.

- With SS, adding links can have hard-to-predict consequences
  - The "opposite" result of getting lower actions does not hold

An equilibrium of the best-shot game



Adding link causes many agents to switch



Result for best-shot game
(Galeotti et al)

Proposition: Consider the best-shot game on g and an equilibrium x of g+ij.

- Either x is an equilibrium of g, or
- There exists an equilibrium of g in which a strict superset of agents chooses 1

- "Adding links decreases actions"

Other processes on networks

- Semi-anonymous games in the configuration model
- Diffusion
- Information transmission (job openings)
- Opinion formation
- Learning (both Bayesian and non-Bayesian)
- Evolutionary dynamics
- Simultaneous evolution of links and actions

Concluding remarks

- Incorporating network elements is essential for many modeling applications
- The range of settings and analyses is very broad
- Many open questions
- Current research is becoming increasingly interdisciplinary
  - E.g., designing incentives in peer-to-peer applications