Social and Economic Networks

Strategic Network Formation

Xiang Sun

October 25–November 1, 2017

Outline

Strategic network formation

- There are many settings where not only chance/randomness but also choice plays a central role in determining relationships (networks).
- Agents care about the relationships they form and maintain:
  - benefit,
  - cost: effort, time, or resources.
- Examples: trading relationships, political alliances, employer-employee relationships, marriages, professional collaborations, citations, emails, friendships, and so forth.

Modeling choices

How should we model incentives to form and sever links?

- Is consensus needed (undirected/directed)?
- Can they coordinate changes in the network?
- Is the process dynamic or static?
- How sophisticated are agents?
- What do they know when making a decision?
- Do they make errors?
- What happens on the network?
- Can they compensate each other for relationship?
- Are links ajustable in intensity?

Some questions

- Which networks are likely to form?
- Are some more stable than others to various perturbations?
- Are the networks that form efficient?
- How inefficient are they if they are not efficient?
- Can intervention help improve efficiency?
- Can such models provide insight into observed characteristics of networks?

Payoff of networks

- In order to model network formation in a way that accounts for individual incentives, we first need to model the utility that each agent receives as a function of networks.
- $$u_i \colon G(N) \to \mathbb{R},$$
  where $u_i(g)$ represents the payoff that i receives if the network g is in place.
- Depending on the setting, very different things can be covered.
- The agents are aware of changes in their own utility as they add or delete links, or at least react in terms of adding relationships that increase payoffs and delete relationships that decrease payoffs.

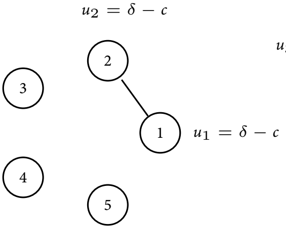
Connections model

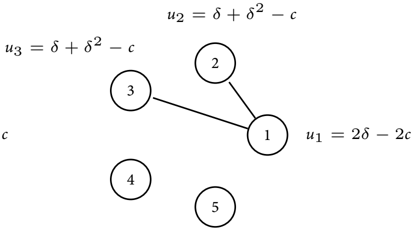
- $\delta \in [0, 1]$: benefit parameter for i from connection between i and j.
- $c_{ij} > 0$: cost to i of the link to j.
- $\ell(i, j)$: shortest path length between i and j.
- Payoff:

$$u_i(g) = \sum_j \delta^{\ell(i,j)} - \sum_{j \in N_i(g)} c_{ij}.$$

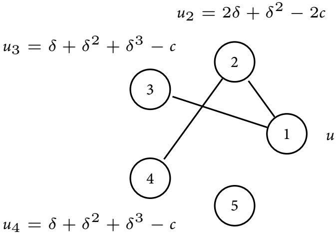
Symmetric version

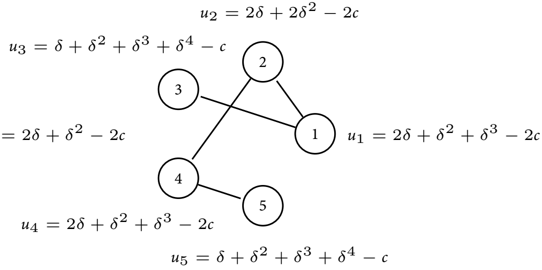
- Benefit from a friend is δ.
- Benefit from a friend of a friend is δ^2.
- Cost of a link is $c > 0$.





Symmetric version: Illustration





Questions

For each different network structure we can do different calculations. Once we have got those then we can talk about

- Which networks are best for society? (social incentive)
- Which networks are formed by the agents? (individual incentive)

Section 1

Pairwise stability

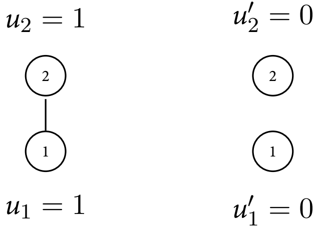
Modeling incentives

- Real world:
  - forming a relationship or link between two players usually involves mutual consent,
  - severing a relationship only involves the consent of one player.
- Modeling as a game: everybody just announces who they want to be friends with.
  - if two people both announce each other, then we form a friendship between them,
  - if they don't both announce each other, then we don't form a friendship.

Nash equilibrium

- Nash equilibrium: A Nash equilibrium is a list of announcements by each player, such that no player would benefit by changing his or her announcement, given the announcements of the other player(s).
- Consider an example:
  - Two individuals.
  - They simultaneously announce whether whey are willing to form their relationship.
  - If they are separate, then they get a value of 0.
  - If they are connected, then they get a value of 1.

Nash equilibrium (Cont.)



There are two Nash equilibria:

- both players say they wish to form the link and it is formed,
- both players say they do not wish to form the link and it is not formed.

Nash equilibrium (Cont.)

- This second equilibrium does not make much sense in a social setting, where we would expect the players to talk to each other and form the link if it is in their mutual interest.
- Some standard game theoretic equilibrium notions are not well-suited for the study of network formation, as they do not properly account for the communication and coordination that is important in the formation of social relationships in networks.
- Two individuals should be able to coordinate on forming a link when it is in their mutual interest.

Pairwise stability

We are looking at a network:

- No agent gains from severing a link.
  - * relationships must be beneficial to be maintained.
- No two agents both gain from adding a link (at least one strictly).
  - * beneficial relationships are pursued when available.

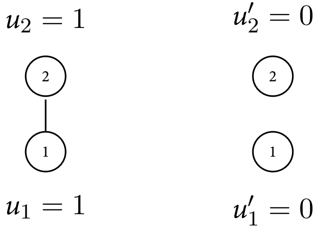
Pairwise stability (Cont.)

A network g is pairwise stable if

- for all $ij \in g$, $u_i(g) \geq u_i(g - ij)$ and $u_j(g) \geq u_j(g - ij)$.
  * no agent gains from severing a link.
- for all $ij \notin g$, if $u_i(g + ij) > u_i(g)$ then $u_j(g + ij) < u_j(g)$.
  * no two agents both gain from adding a link (at least one strictly).

It is sort of the minimal set of requirements for stability.

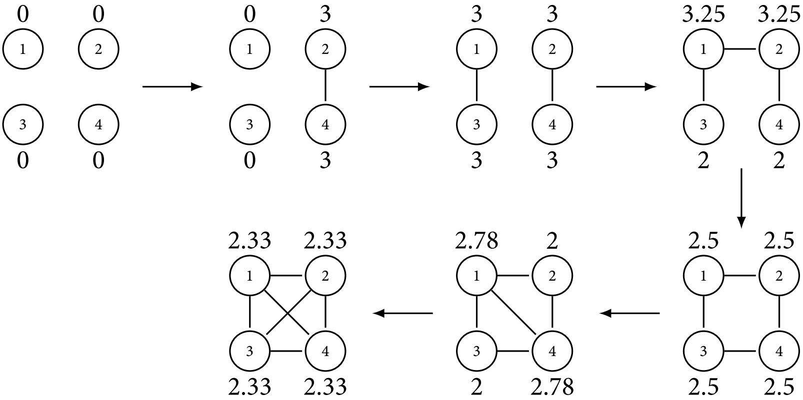
Pairwise stability: Illustration



Both are Nash equilibria, but only the left one is pairwise stable.

Pairwise stability: Illustration (Cont.)

Not pairwise stable



Pairwise stable

Pairwise stability: Limitations

- Pairwise stability is a weak notion in that it only considers deviations on a single link at a time.
  * For instance, it could be that a player would not benefit from severing any single link but would benefit from severing several links simultaneously, and yet the network could still be pairwise stable.
- Pairwise stability considers only deviations by at most a pair of players at a time.
  * It might be that some group of players could all be made better off by some more complicated reorganization of their links.

Pairwise stability might be thought of as a necessary but not sufficient requirement for a network to be stable over time.

Section 2

Efficient networks

Efficient networks

- Let us turn our attention to the evaluation of the overall benefits that society sees from a given network.
- Payoffs not only provide an individual's perspective on the network, but also enable us to at least partially order networks with regards to the overall societal benefits that they generate.

Efficient networks (Cont.)

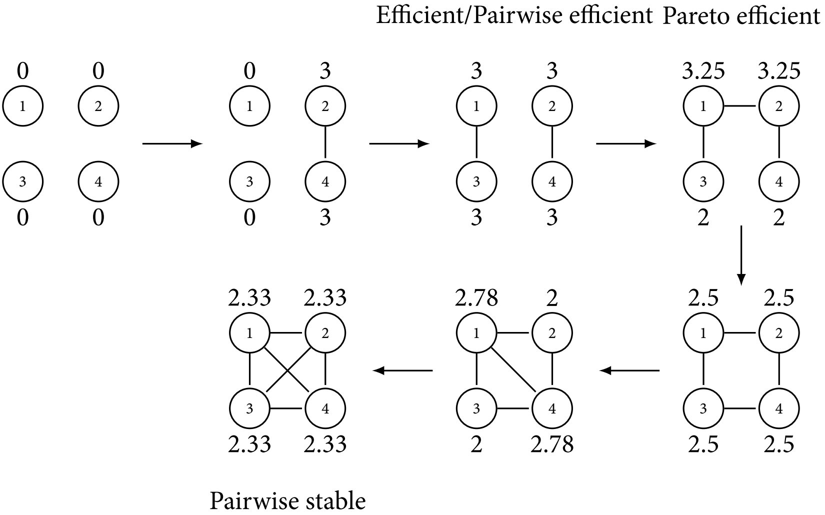
- Given agents' utility functions $(u_1, u_2, \ldots, u_n)$.
- A network g is Pareto efficient relative to $(u_1, u_2, \ldots, u_n)$ if there does not exist any $g' \in G(N)$ such that
  - $u_i(g') \geq u_i(g)$ for all i, and
  - $u_{i_0}(g') > u_{i_0}(g)$ for some i_0.
- A network g is efficient relative to $(u_1, u_2, \ldots, u_n)$ if

$$\sum_i u_i(g) \geq \sum_i u_i(g') \text{ for all } g' \in G(N),$$

or

$$g \in \arg\max_{g' \in G(N)} \sum_i u_i(g').$$

Efficiency and pairwise stability: Illustration



Efficiency and pairwise stability: Illustration (Cont.)

- Society would like to do in terms of picking something which maximizes over the total utility or even something which is Pareto efficient.
- The process can end up things, which are worse, in the sense that everybody is worse off than what would happen if the society could oppose the network.
- Part of it is due to the fact that individuals are not accounting for the harm that they can afflict on others when they make their decision.

Pareto efficiency vs. efficiency

- If g is efficient relative to $(u_1, u_2, \ldots, u_n)$, then it must also be Pareto efficient relative to $(u_1, u_2, \ldots, u_n)$.
- However, the converse is not true.

* Result: g is efficient relative to $(u_1, u_2, \ldots, u_n)$ if and only if it is Pareto efficient relative to all payoff functions $(\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n)$ such that $\sum_i \hat{u}_i = \sum_i u_i$.

Pareto efficiency vs. efficiency (Cont.)

- Efficiency is a more discriminating notion and is the more natural notion in situations where there is some freedom to change the way in which utility is allocated throughout the network, for instance by reallocating value through transfers.
- Pareto efficiency is more reasonable in contexts where the payoff functions are fixed and no transfers are possible.

Section 3

Connections model

Connections model

- $\delta \in [0, 1]$: benefit parameter for i from connection between i and j.
- $c > 0$: cost of a link.
- $\ell(i, j)$: shortest path length between i and j.
- Payoff:

$$u_i(g) = \sum_j \delta^{\ell(i,j)} - \sum_{j \in N_i(g)} c.$$

- To try and analyze what are the efficient networks, what are the Pareto efficient networks, and what are the pairwise stable networks.

Subsection 1

Efficiency in connections model

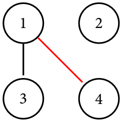
Efficient networks in symmetric connections model

- Low cost: $c < \delta - \delta^2$
  - complete network is uniquely efficient.
- Medium cost: $\delta - \delta^2 < c < \delta + \frac{n-2}{2}\delta^2$
  - star networks with all agents are uniquely efficient.
- High cost: $\delta + \frac{n-2}{2}\delta^2 < c$
  - empty network is uniquely efficient.
- Intuition: If links are so cheap you might as well just add them all. If links are so expensive, it does not make sense to add any.

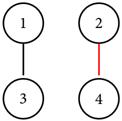
Why stars?

- We start with one relationship (between 1 and 3) that gives us $2\delta - 2c$, and we think about adding a second one.
- There are two different ways we can add this second relationship.

$4\delta + 2\delta^2 - 4c$



$4\delta - 4c$

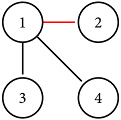


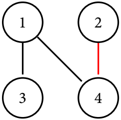
- The indirect benefits that flow through the network generate extra value. And so connecting in this way it gives us a higher value than connecting in this separate way.

Why stars? (Cont.)

- Consider the fourth person.

$$6\delta + 6\delta^2 - 6c \qquad\qquad 6\delta + 4\delta^2 + 2\delta^3 - 6c$$



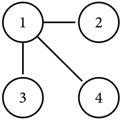


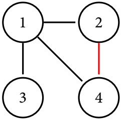
- In a star form, all these indirect connections now are at a distance two. Whereas in the right network one some of the indirect connections is at a distance three.
- In a star form, we end up with a higher value for all the indirect connections.
- The stars are coming out because they are the most efficient way to connect people with a given number of links with the least distance between them.

Star vs. complete

- When is it that you want to keep connecting?

$6\delta + 6\delta^2 - 6c$ $\qquad$ $8\delta + 4\delta^2 - 8c$





- When $2\delta - 2\delta^2 - 2c > 0$, adding the link 2-4 is better off.

Proof: The case $c < \delta - \delta^2$

- Intuition: It is more beneficial to have a direct relationship than to have an indirect relationship of distance 2 (and others).
- Suppose that $ij \notin g$.
- The value that i and j getting from their relationship is going to be $\leq \delta^2$. And if they add a direct link, they are going to get $\delta - c$ for that relationship.
  - $u_i(g + ij) > u_i(g)$.
  - $u_j(g + ij) > u_j(g)$.
- Everybody else benefits: $u_k(g + ij) \geq u_k(g)$ for every k.
- Thus,

$$\sum_{\ell} u_\ell(g + ij) > \sum_{\ell} u_\ell(g).$$

- Therefore, the complete network is uniquely efficient.

Proof: The case $c > \delta - \delta^2$

Idea:

1. To show that the value of a component is highest when a component is a star.

   If you are going to arrange people, you are best off doing it in a star.

2. To show that you do not want to have multiple stars, you would be better off having one star.

3. Compare whether it is better to have a big star with everybody in it, or no star at all.

   It is the difference between the medium cost and the really high cost.

Proof: The case $c > \delta - \delta^2$: Step 1

- The value of a star with k agents is

$$2(k-1)[\delta - c] + (k-1)(k-2)\delta^2.$$

- The value of a network with k agents and m links ($m \geq k-1$) is at most

$$2m[\delta - c] + [k(k-1) - 2m]\delta^2.$$

- The difference is

$$2[m-(k-1)][\delta^2 - (\delta - c)],$$

which is positive when $m > k-1$.

Proof: The case $c > \delta - \delta^2$: Step 1 (Cont.)

If $m = k - 1$ and not a star, then some pair is at a distance of more than 2, so less value than a star.

- The value of a star with k agents is

$$2(k-1)[\delta - c] + (k-1)(k-2)\delta^2.$$

- The value of a component with k agents and $k-1$ links that is not a star is at most

$$2(k-1)[\delta - c] + [(k-1)(k-2) - 1]\delta^2 + \delta^3.$$

- Star is better.

Proof: The case $c > \delta - \delta^2$: Step 2

If each of two separate star components has nonnegative total utility, then one star with all those agents generates higher total utility.

- Separate:

$$
\begin{aligned}
& 2(k-1)[\delta - c] + (k-1)(k-2)\delta^2 \\
& + 2(k'-1)[\delta - c] + (k'-1)(k'-2)\delta^2 \\
= \; & 2(k+k'-2)[\delta - c] + [(k-1)(k-2) + (k'-1)(k'-2)]\delta^2
\end{aligned}
$$

- As one star:

$$
2(k+k'-1)[\delta - c] + (k+k'-1)(k+k'-2)\delta^2.
$$

- The second expression is bigger.

Proof: The case $c > \delta - \delta^2$: Step 3

- When $c > \delta - \delta^2$, the efficient networks are collections of stars and empty networks.

$\Rightarrow$ Either a star with all agents or empty.

- The star is valuable if and only if

$$2(n-1)[\delta - c] + (n-1)(n-2)\delta^2 > 0,$$

or

$$c < \delta + \frac{n-2}{2}\delta^2.$$

Subsection 2

Pairwise stability in connections model

Pairwise stability in connections model

- Low cost: $c < \delta - \delta^2$
  - complete network is uniquely pairwise stable.
- Medium/low cost: $\delta - \delta^2 < c < \delta$
  - star network is pairwise stable.
  - others are also pairwise stable.
- Medium/high cost: $\delta < c < \delta + \frac{n-2}{2}\delta^2$
  - star network is not pairwise stable (no loose ends).
  - nonempty pairwise stable networks are over-connected and may include too few agents.
- High cost: $\delta + \frac{n-2}{2}\delta^2 < c$
  - empty network is pairwise stable.

Case 2: $\delta - \delta^2 < c < \delta$

- When $c < \delta$, it is still valuable to have a connection.
- When $\delta - \delta^2 < c < \delta$, we are in a situation where it is valuable to have connections but it is not worth it to shorten indirect connections necessarily to direct ones.
- The star network turns out to be pairwise stable.
- There can also be other pairwise stable networks (inefficient), so it is not the only pairwise stable.

Case 3: $\delta < c < \delta + \frac{n-2}{2}\delta^2$

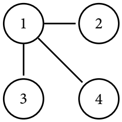
- In this case, star is efficient.
- Since $\delta < c$, it is not worth to have a relationship with somebody that only brings that one person.

$\Rightarrow$ The only reason you want to have a relationship is if it is bringing also some indirect benefits with it.

$\Rightarrow$ Star is not worthwhile: The center agent is not willing to have connections with other individuals.

- No loose ends: there is no individual that is going to want to connect to some other individual that does not bring them any indirect benefits.

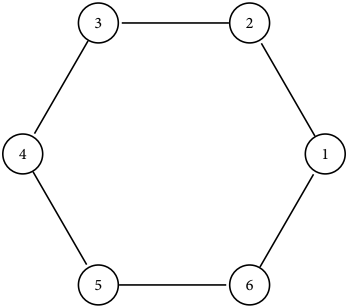
Case 3: $\delta < c < \delta + \frac{n-2}{2}\delta^2$: Illustration



- Payoff to the center: $3\delta - 3c$.
- Overall payoff: $6\delta + 6\delta^2 - 6c$.
- It is efficient, but not pairwise stable:
  - The peripheral players are actually getting indirect benefits and the center does not get those.
  - So the center is willing to sever the links even though the peripheral players would rather have the center maintain the star.

Exercise

Prove: When $n = 6$ and $\delta < c < (\delta + \delta^2 + \delta^3)(1 - \delta^2)$, the following is the unique nonempty pairwise stable network.



Section 4

The co-author model

Subsection 1

Externality

Externality

- There are nonnegative externalities under $u = (u_1, \dots, u_n)$ is

$$u_k(g + ij) \geq u_k(g)$$

  for all $k, g \in G(N)$ and ij such that $k \neq i, j$.
- There are positive externalities under $u = (u_1, \dots, u_n)$ if there are nonnegative externalities under $u = (u_1, \dots, u_n)$ and the inequality above is strict in some instances.

Inefficiency in connections model

- Inefficiency in the connections model is due to the fact that there are positive externalities.
- The star is not willing to maintain these external relationships is coming from the fact that those are not giving the center of the star any value.
- However, there are positive externalities to the other players that the center is not taking into account.

Subsection 2

The co-author model

The co-author model

- People are going to be involved in research collaborations.
- The value from each relationship depends on:
  - how much time people put into those relationships,
  - an interaction term which is going to capture the some sort of synergies.
  - * if I spend more time collaborating with somebody, we have more time to get better ideas, and that is going to be valuable.
- Utility:

$$u_i(g) = \sum_{j:\ ij \in g} \left[\frac{1}{d_i} + \frac{1}{d_j} + \frac{1}{d_i d_j}\right].$$

⇒ Negative externalities.

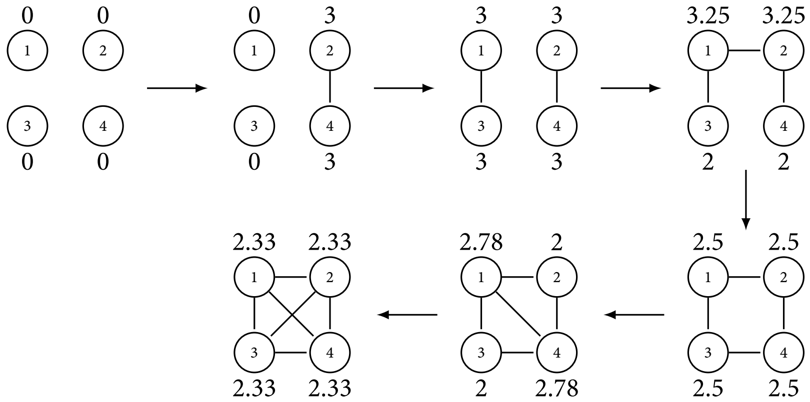
The co-author model (Cont.)

We are not going to put in explicit costs to links:

- The costs from adding extra links come from the fact that you are diluting your synergies with different collaborations.
- You are just spreading your time out and the more thinly you spread your time the lower the value from any relationship you get.

The co-author model: Illustration

Efficient



Pairwise stable

The co-author model: Efficiency and stability

Suppose that n is even.

- Efficient networks: pairs.
- Pairwise stable networks consist of completely connected components, each of a different size, one has more than the square of the number of nodes in the other.
- By adding a link, you would dilute existing synergies and so you only want to add a new coauthor if they bring in sort of comparable worth to your own values.

* It gives these the fact that pairwise stable networks, if they have separate components, have to have very different sizes, so that one is not going to group with another.

Section 5

Network formation and transfers

Transfers

- Stable and efficient networks are only going to coincide in special cases.
- Can transfer help in other cases?
- What can we say about when transfers improve efficiency?
- Are transfers in players' interests?

What are transfers

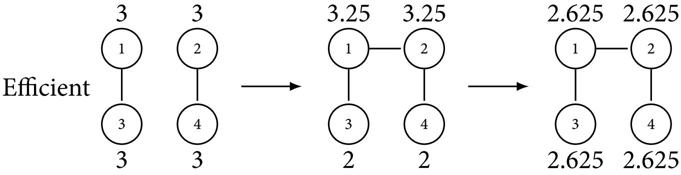
- Utility could be moved from one node to another.
- Outside intervention, taxing or subsidizing relationships.
- Bargaining among the individuals involved.

Modeling transfers

- Change utility from $u_i(g)$ to $u_i(g) + t_i(g)$.
- $t_i(g)$ could be either a positive or negative number depending on whether somebody is making net payments or getting that receipts as a function of the network.

Transfers in co-author model

- Problem: people want to over connect. (individual incentive)



- Consider: government says that we are going to tax people who form extra links and then move that to to the other players.
  * Charge 1 and 2 a 0.625 each, and then pay that to 3 and 4.
- Individuals no longer have an incentive to form this extra link.
- The left network turns out to be pairwise stable.

Egalitarian transfer

- Set $t_i(g) = \frac{1}{n}\sum_j u_j(g) - u_i(g)$.
- Then $u_i(g) + t_i(g) = \frac{1}{n}\sum_j u_j(g)$.
- We are just going to adjust the transfers to move everybody back to the average.
- Now the utility anybody gets is exactly proportional to the efficiency of the network.

$\Rightarrow$ Now everybody in the society has exactly the same incentives as a utilitarian planner would have.

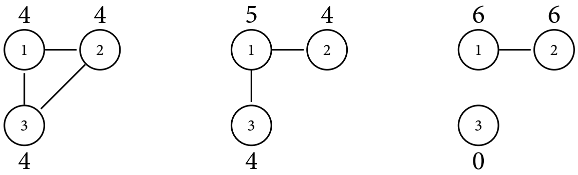
$\Rightarrow$ Efficient network is going to be pairwise stable.

Requirements on transfer

- Making transfers are going to violate some fairly basic conditions.
- Some very basic requirements on transfers:
  - Completely isolated nodes that generate no value get 0.
  - Nodes that are completely interchangeable get the same transfers.

Transfers cannot always help

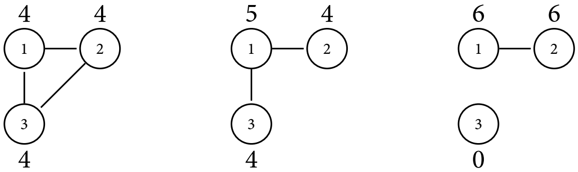
Efficient



- The middle network is the efficient one. And it is not pairwise stable.

  ⋆ 1 benefits from deleting the link 1-3.

Transfers cannot always help (Cont.)

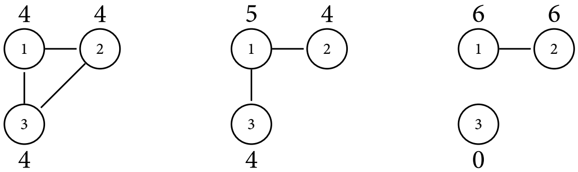


We want see if we can do some transfers to try and help this.

- Consider the right network: 3 is completely disconnected, not generating any value.

⇒ Value should be split between 1 and 2. They're completely symmetric, doing the same things so each one of them has to be 6.

Transfers cannot always help (Cont.)



- Consider the middle network: in order to be pairwise stable, 1 is going to have to get a transfer at least 1.
- In order for 2 and 3 not to want to form a new link, they have to stay at least 4.

$\Rightarrow$ You cannot take anything away from them.

The only way to make the efficient thing stable is by somehow infusing extra value into this.

Transfer

- Transfers can be helpful sometimes but not necessarily always.
- It is not necessarily entirely correctable with bargaining or transfers.
- It is going to depend on exactly what kinds of transfers we allow, and what situations.

Section 6

Small worlds in an islands-connections model

Can economic models match observables?

Can small worlds be derived from costs/benefits?

- Low costs to local links—high clustering
- High value to distant connections—low diameter
- * if there were no short enough paths between two given nodes, then even if there were a high cost to adding a link, that link would bridge distant parts of the network and bring high benefits to that pair of nodes.
- High cost of distant connections—few distant links

Strategic model vs. random model

- The random models can identify processes which generate certain features, but do not explain why those processes might arise.
- In a strategic model, the explanation for a specific characteristic of a network is instead traced back to more primitive elements such as costs and benefits from social relationships.
- The strategic model can be thought of as explaining why, whereas the random-graph models can be thought of as explaining how. This is not to say that strategic models are better.

Islands connections model

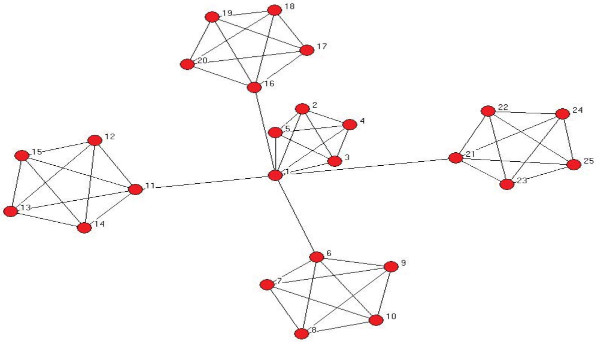
- J players live on an island, K islands.
- cost c of link to player on this island.
- cost $C > c$ of link to player on another island.
- Result:
  - High clustering with islands, few links across.
  - Small distances.

Island

- It could be geography.
- It also could be characteristics so people with very similar characteristics find it very easy to link to each other.
  * People with different characteristics find it more costly so the islands are.

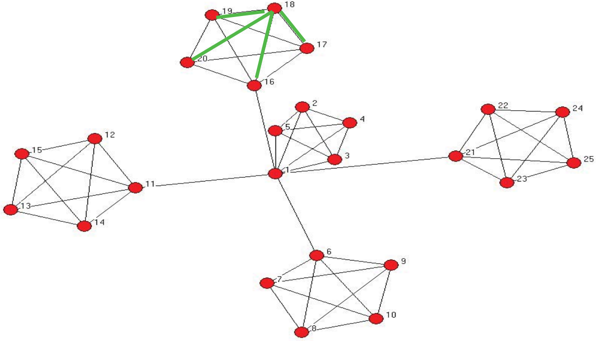
Islands connections model: Illustration

$$4\delta \ - 4c + \delta^2 + 7\delta^3 + 12\,\delta^4$$



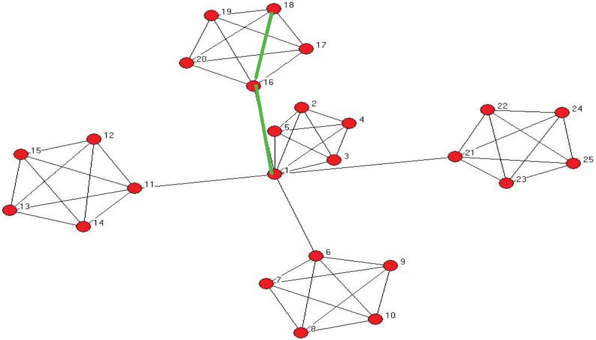
Islands connections model: Illustration (Cont.)

$4\delta - 4c + \delta^2 + 7\delta^3 + 12\delta^4$



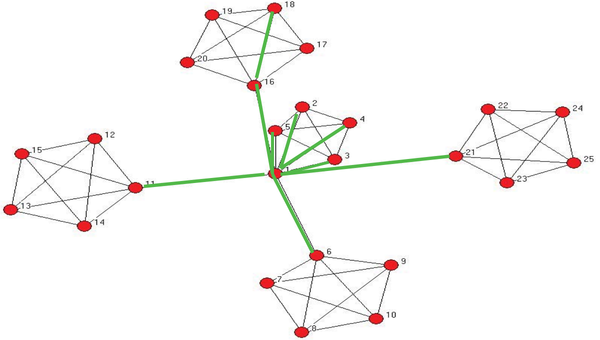
Islands connections model: Illustration (Cont.)

$$4\delta - 4c + \delta^2 + 7\delta^3 + 12\delta^4$$



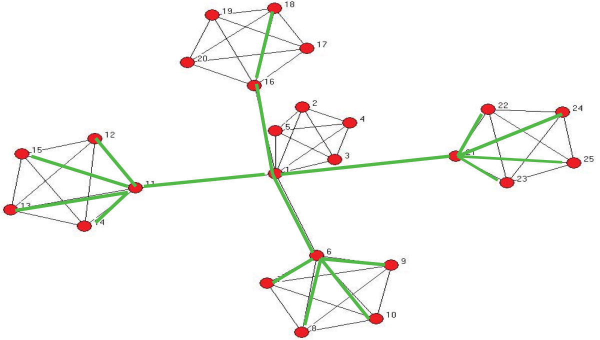
Islands connections model: Illustration (Cont.)

$4\delta - 4c + \delta^2 + 7\delta^3 + 12\delta^4$

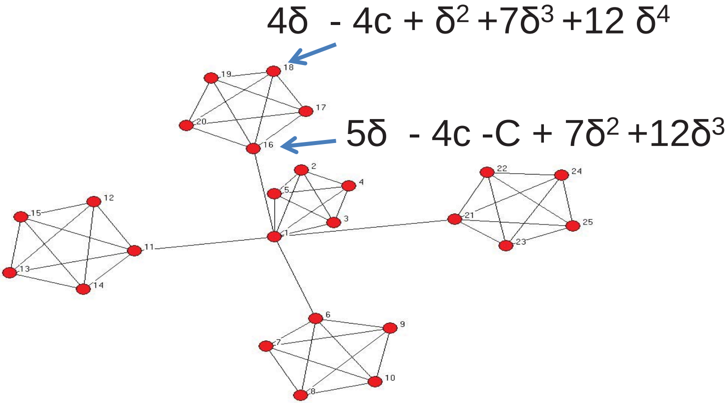


Islands connections model: Illustration (Cont.)

$4\delta - 4c + \delta^2 + 7\delta^3 + 12\delta^4$



Islands connections model: Illustration (Cont.)



Islands connections model: Result

- Low cost to an island: you want to connect within your island.
- High cost across islands: you only want to have limited number of connections across islands.
- If $c < 0.04$, $1 < C < 4.5$ and $\delta = 0.95$, then the following network is pairwise stable.
  - High clustering.
  - Low diameter.

Islands connections model: Result (Cont.)

- It gives us a different explanation and reasoning behind why you might see small worlds.
- We can begin to enrich this kind of model with some random formation to begin to try and fit things to data.

Islands connections model: Result (Cont.)

General result:

- Truncate connections:

$$u_i(g) = \sum_{j:\ \ell(i,j) \leq D} \delta^{\ell(i,j)} - d_i(g)c.$$

- If $c < \delta - \delta^2$ and $C < \delta + (J-1)\delta^2$, then
  - players on each island form a clique.
  - diameter is bounded by $D + 1$.
  - $\delta - \delta^3 < C$ implies a lower bound on individual clustering is $\frac{(J-1)(J-2)}{J^2K^2}$.