电子科技大学
University of Electronic Science and Technology of China

Strategic Network Formation

Zhongjing Yu

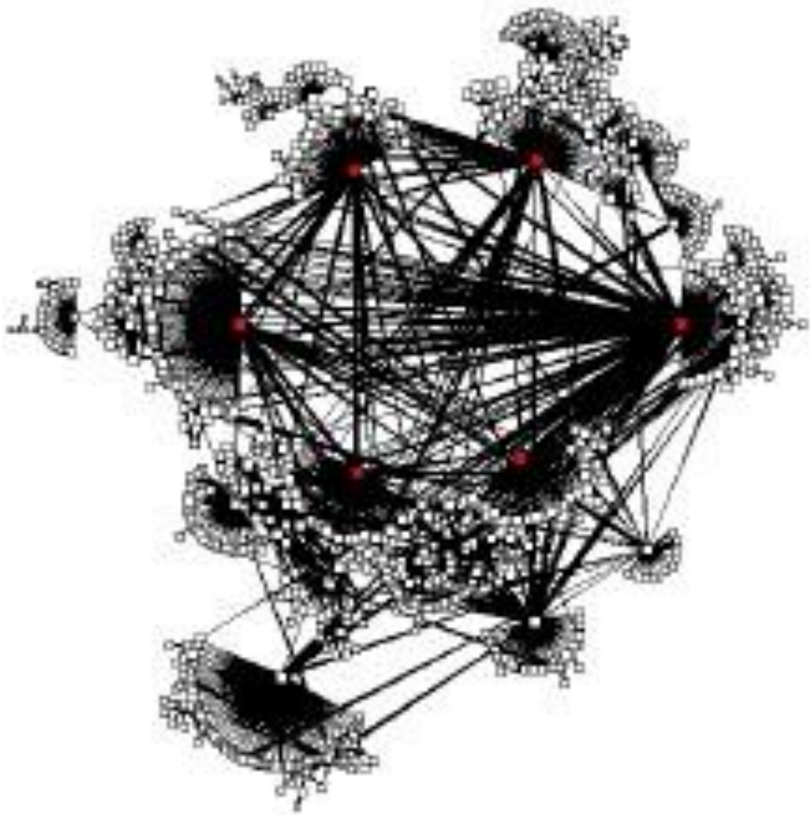
Data Mining Lab, Big Data Research Center, UESTC
Email：junmshao@uestc.edu.cn
http://staff.uestc.edu.cn/shaojunming

What's meaning of **Strategic Network Formation** ?

Node : a **individual**.

Edge : **strategy** that user choose.

Then user gets payoffs from his or her strategy. Each one **changes strategy to pursue** the **max payoffs** and it also is the reason **why** the network forms and **how** to form.



Outline

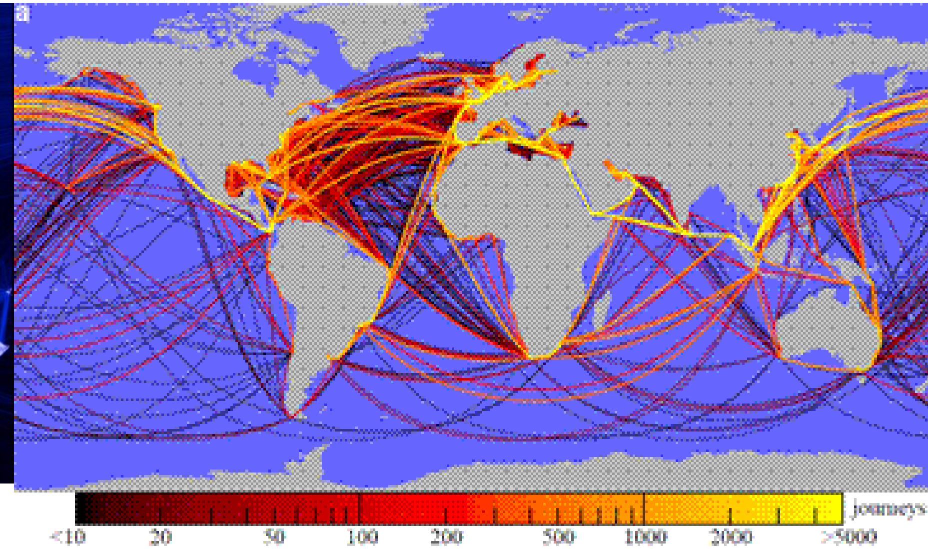
数据挖掘实验室
Data Mining Lab

- Pairwise Stability
- Efficient Networks
  - Efficient
  - Parato Efficient
- Distance-Based Utility
  - Externalities
  - Growing Networks and Inefficiency
  - The Price of Anarchy and the Price of Stability
- A Co-Author Model and Negative Externalities
- Small Worlds in an Islands-Connections Model

Introduction

数据挖掘实验室
Data Mining Lab

(Facebook network)



(Trade network)

How to **measure** those networks?

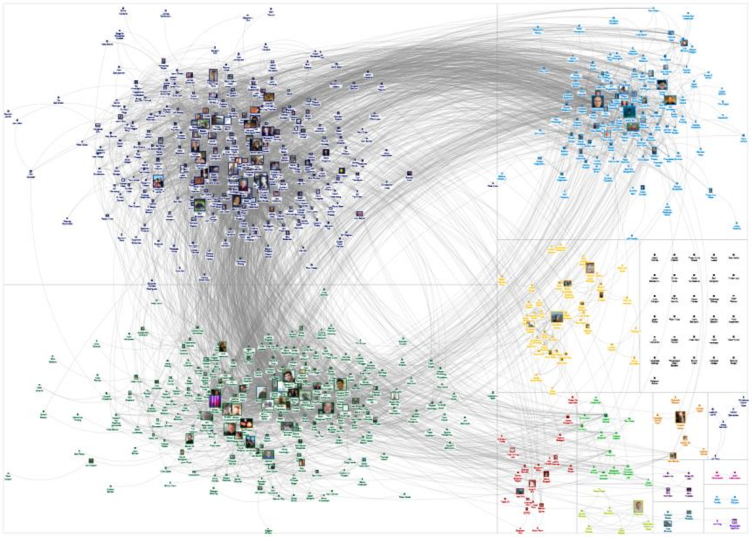
How to **model** those networks?

Modeling networks from “ ***Strategic***” point of view

Model with **benefit** and **cost**

- How to form network by **individual incentives**.
- How to provide well-defined measure by **social welfare**.

For example :



High clustering related to **low cost of connection**.

Low diameter related to **benefit of accessing the information**

Pairwise Stability

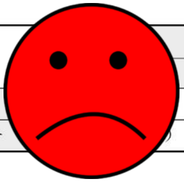
In order to capture the fact that forming a **relationship** or **link** between **two players**, we need to identify a state of **equilibrium or stability**.

Nash equilibrium can't capture the fact in real complex network.

Why ? …….let me explain.

Nash  **brium**

Real fact



		村民	
			不开发
地方政府	允许		妥协（1，2）
	不允许	)	遵守（4，1）

But Nash equilibrium could describe some info in other view (later)

Pairwise Stability

Directly define an equilibrium notion

Imposing individual *incentives*.

Players : $N = \{1,2,\dots,n\}$.

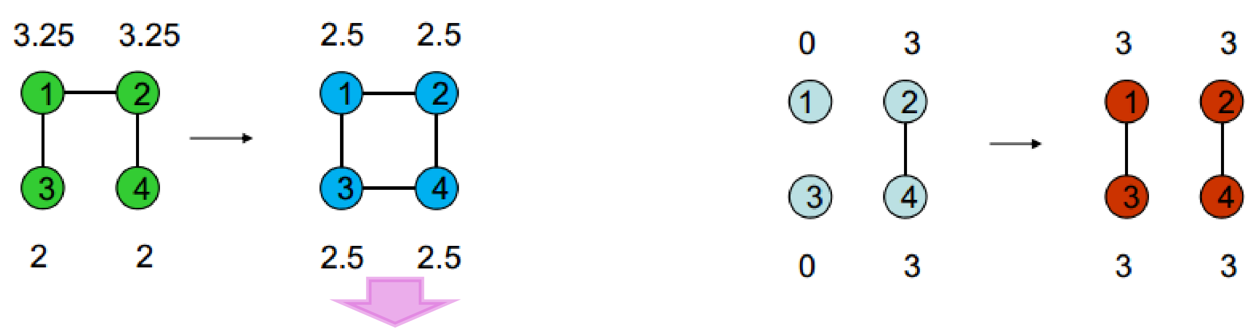
Utility function (payoff function) : u_i.

$u_i(g)$: benefit of node i in network g.

A network g is *pairwise stable* if

(i) for all $ij \in g$, $u_i(g) \geq u_i(g - ij)$ and $u_j(g) \geq u_j(g - ij)$, and

(ii) for all $ij \notin g$, if $u_i(g + ij) > u_i(g)$ then $u_j(g + ij) < u_j(g)$.

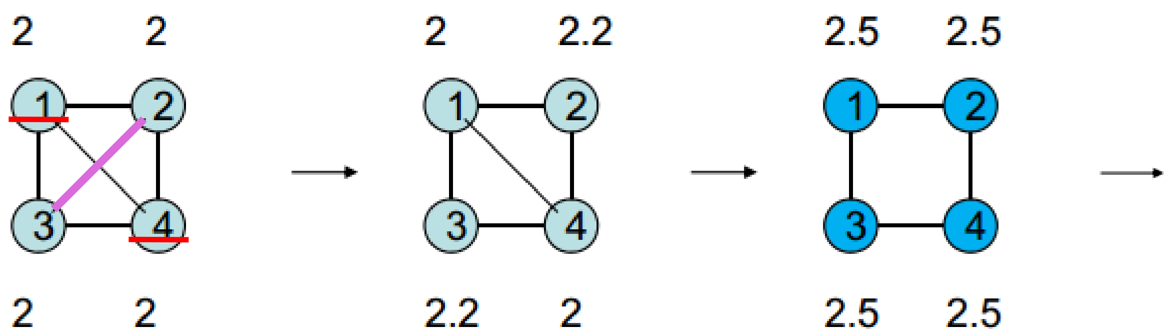


If node 3 changes strategy , the payoffs of node 3 will not increase. ----*Nash equilibrium*

Pairwise Stability

A network is **pairwise stable** if **no player** wants to **sever** a link and no two players neither want to **add** a link.

limitations: pairwise stability is a weak notion…….



After several steps, some edges that was stable ***to unstable***.

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

Efficient Networks

Definition

Efficiency :
The "**best**" network is the one which maximizes **the total utility** of the society.

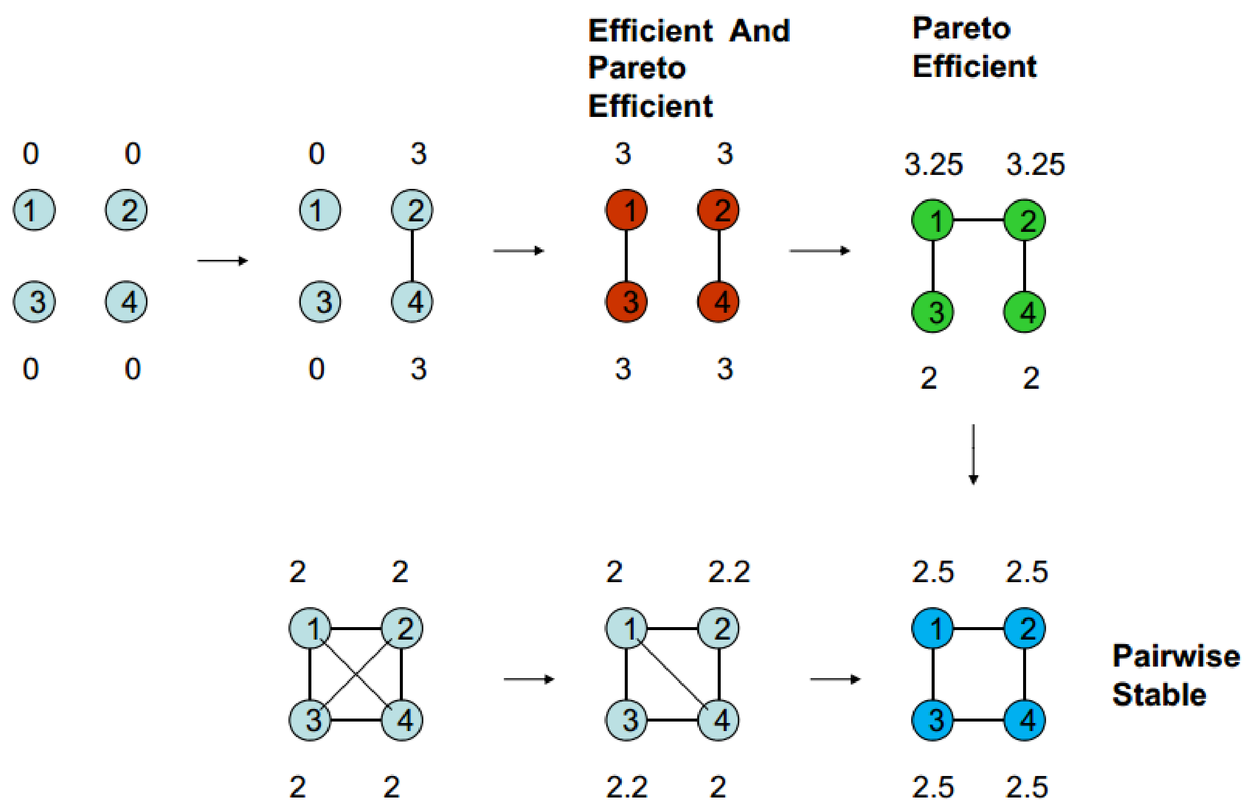
Formation

Utility functions $(u_1,\ldots,u_n)$ if $\sum_i u_i(g) \geq \sum_i u_i(g')$ for all $g' \in G(N)$

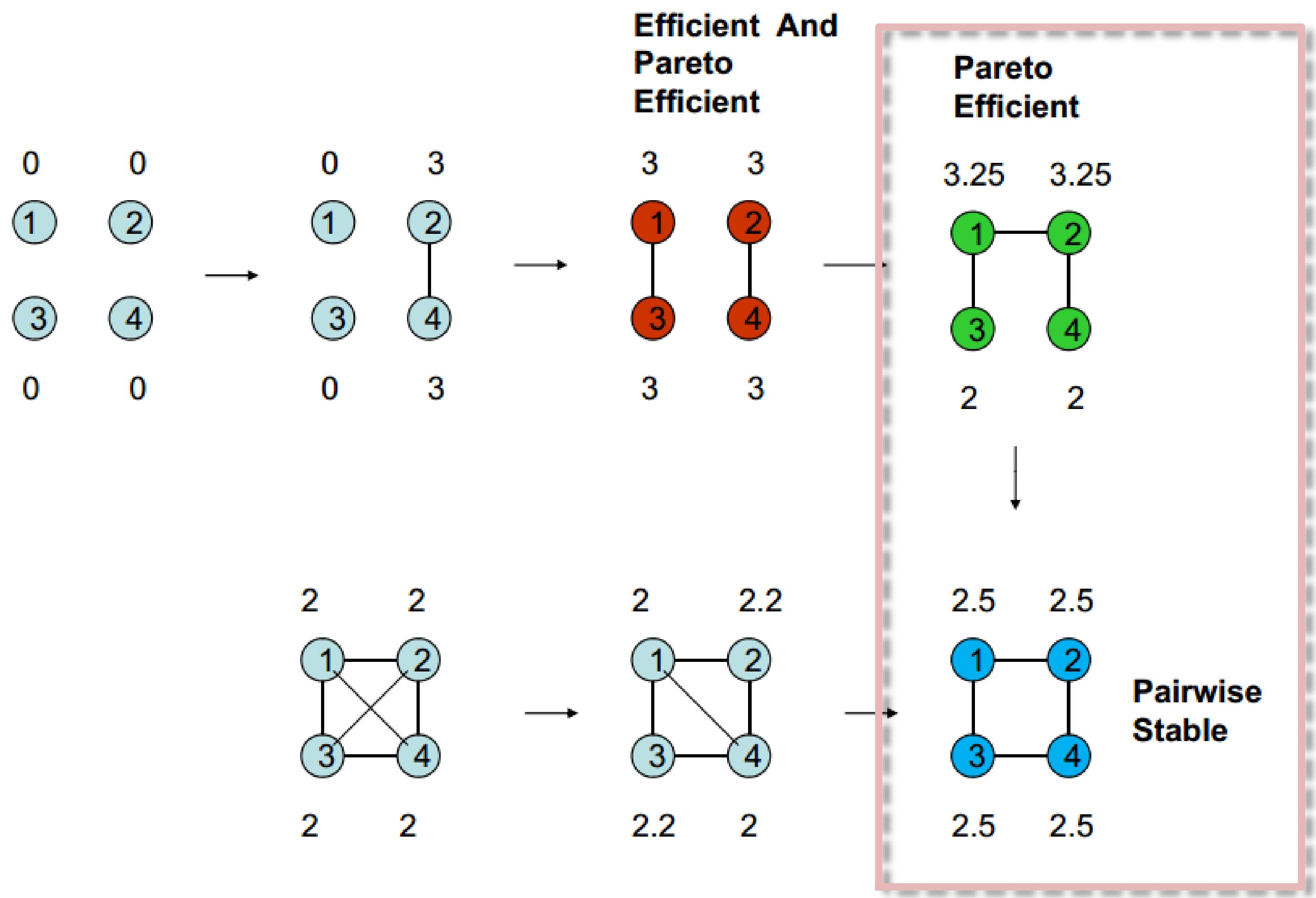
Pareto Efficiency

➢An outcome of a game is **Pareto dominated** if some other outcome would make **at least one** player better off **without hurting any other player**。

➢ If an outcome is not Pareto dominated by any other, then it is **Pareto optimal**, named after Vilfredo Pareto.



Remind: *inconformity* between stability and efficiency



How to model ?

(How to define utility function ?)

Distance-Based Utility

The basic idea : players get utility from their ***direct connections*** and also from their ***indirect connections.***

Intuitive : the ***closer*** two individuals are, the ***higher benefit*** they get.

Definition : ***distance-based utility***

$$u_i(g) = \sum_{j \neq i : j \in N^{n-1}(g)} b(\ell_{ij}(g)) - d_i(g)c$$

$\ell_{ij}(g)$ is the ***shortest path length*** between *i* and *j*.

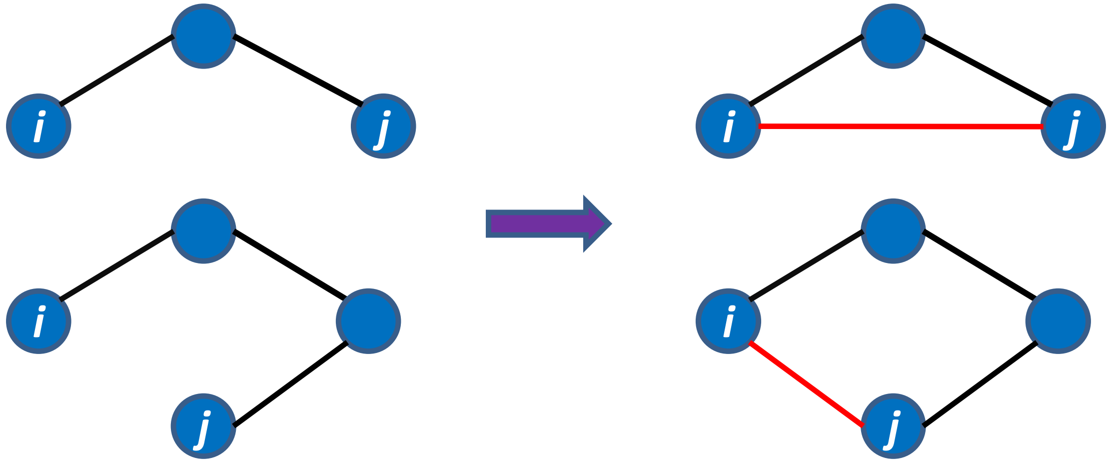
$b(k) > b(k+1) > 0$: the closer two individuals are, the higher benefit they get. usually ***b(k)*** = δ^k (***decay rate***).

$d_i(g)$: ***degree*** of node ***i***; ***c*** : the ***cost*** to maintain one link.

Distance-Based Utility

Indentify some properties :

i): $b(1) - c > b(2)$:

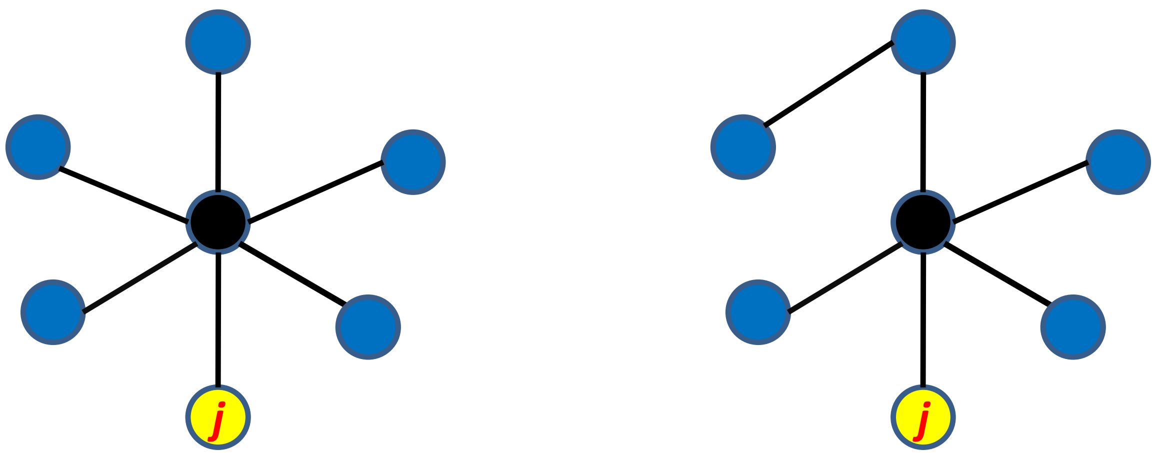


The best strategy is that nodes link each other, then networks become a complete network.

Distance-Based Utility

ii): $b(1)-b(2) < c < \underbrace{b(1)+\frac{(n-2)}{2}b(2)}$

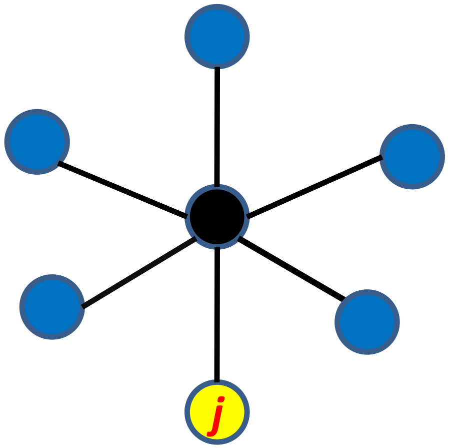
Star networks get the highest benefit.



$$U_{star}(j) > U_{normal}(j)$$

Distance-Based Utility

For a star network with n-1 links



Total utility :

$$U_{star} = 2(n\text{-}1)(b(1)\text{-}c) + (n\text{-}1)(n\text{-}2)b(2)$$

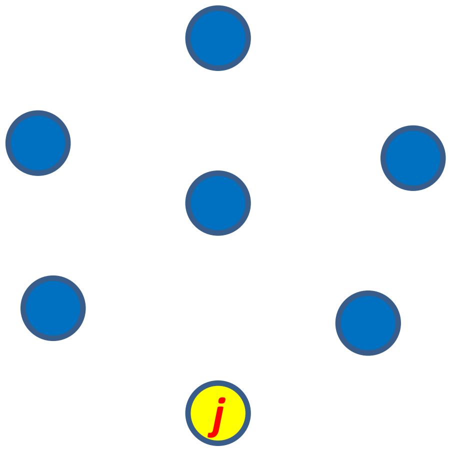
$$U_{star} = 0$$

$$c = b(1) + \frac{(n-2)}{2} b(2)$$

Distance-Based Utility

数据挖掘实验室
Data Mining Lab

$$iii):\ b(1)+\frac{(n-2)}{2}b(2)<\frac{c}{2}$$



The best strategy is no link with other individuals.

➔ *empty networks.*

Externalities

Externalities refer to situations where the **utility or payoffs** to one individual are **affected** by the **actions of others**, where those actions do not directly involve the individual in question.

- **Positive** : increase in payoffs as its neighbors form more links or even if indirectly connected players form more links.

  nonnegative externalities under $(u_1, \ldots, u_n)$

  $$u_i(g + jk) \geq u_i(g)$$

- Negative : increase in payoffs as its neighbors form less links ,…..

  non-positive externalities under $(u_1, \ldots, u_n)$ if

  $$u_i(g + jk) \leq u_i(g)$$

where $i \in N, g \in G(N)$ and jk such that $j \neq k \neq i$

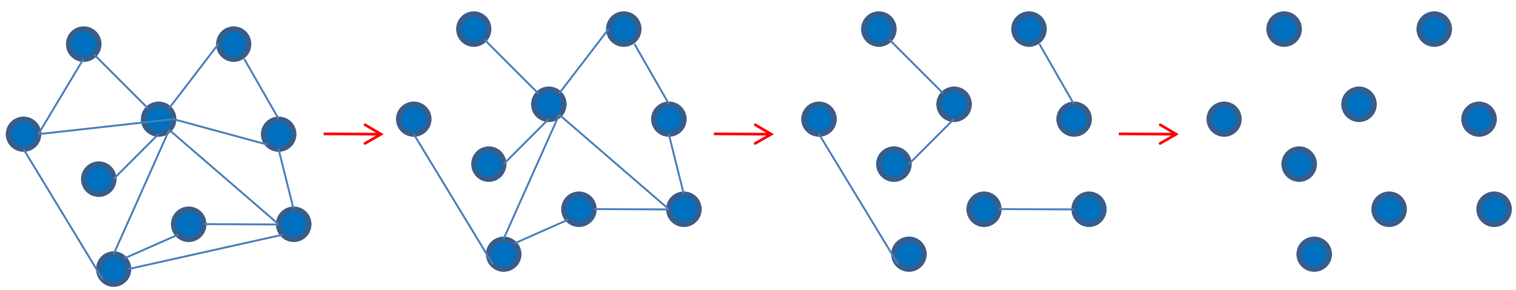
Distance-Based Utility

Growing Networks and Inefficiency

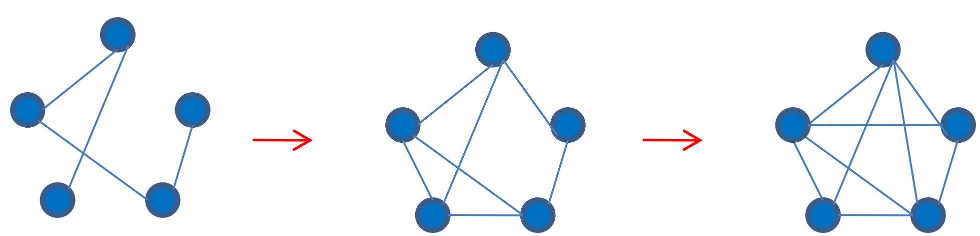
How can we predict which networks are likely to emerge?

Situation : a **random ordering** over links, where at **any point in time any link** is as likely as any other to be identified. And **at least one player benefit** from adding it or severing.

- *c>b(1)* : ***nonempty networks*** are strictly preferred by all players to the empty network.



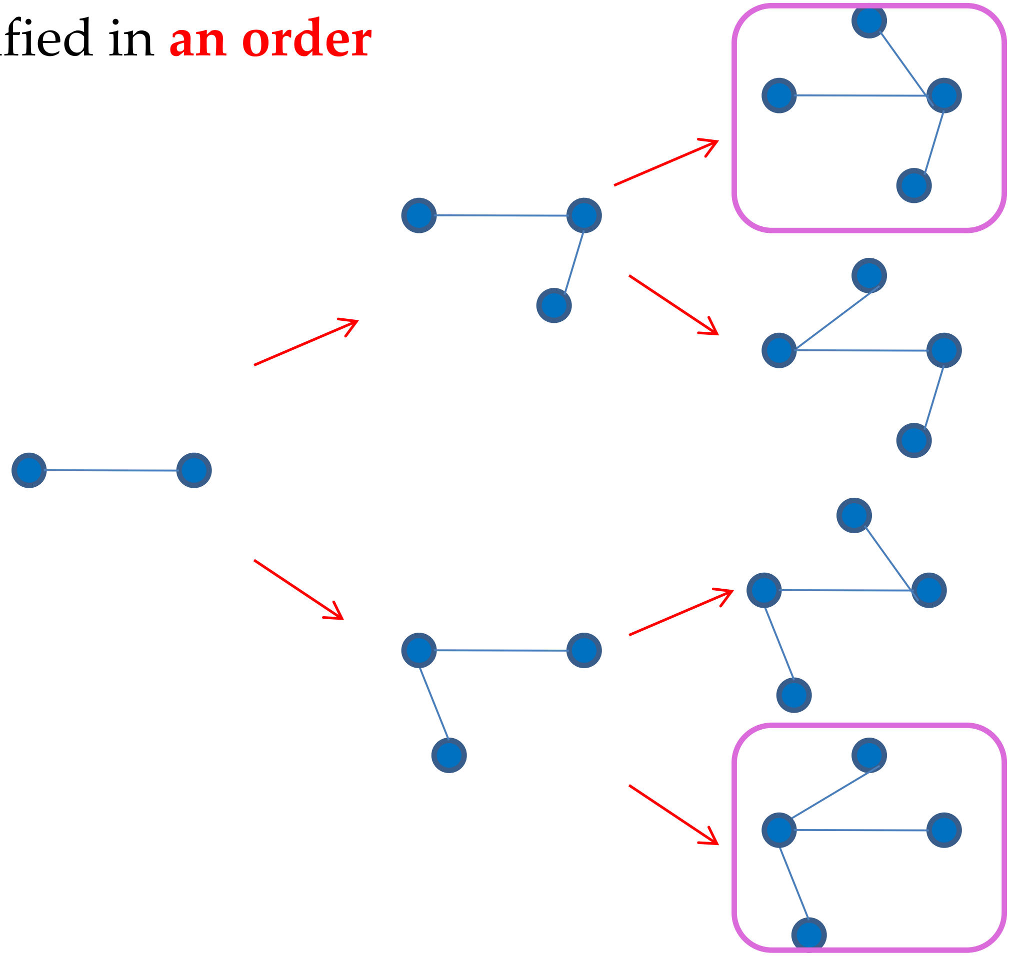
- b(1)-c>b(2) : all links will from and the efficient complete network will be reached.



Growing Networks and Inefficiency

- b(1)- b(2)<c<b(1): the **star network** is the **efficient network**, but players are **willing to add** a link to players with whom they do not have any indirect connection.

In order to form a star network, it must be that the links are identified in **an order**



……

Distance-Based Utility

Growing Networks and Inefficiency

The symmetric **distance-based utility** model in the case where $b(1)-b(2)<c<b(1)$. As the number of players **grows**, the probability that the above described **dynamic process** leads to an efficient network(star) **converge to 0**.

And the growing network always trends to be **inefficient** networks.

Distance-Based Utility

Data Mining Lab

The Price of Anarchy and the Price of Stability

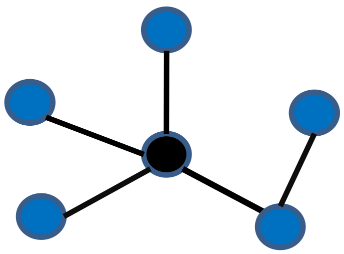
Situation is somehow worse if the stable networks are **"very" inefficient** compared to if they are **"nearly" efficient**.

This issue of quantifying the social inefficiency that results from selfish individuals acting, it also is critical to variety of settings and this has become known as the "price of anarchy".

Definition of utility function:

$$u_i(g) = \sum_{i \neq j} -l_{ij}(g) - d_i(g)c \qquad u(g) = \sum_{i \in N(g)} u_i(g)$$

where $l_{ij}(g)$ is distance between node i and node j in network g.
$d_i(g)$:degree of node i.



$$u_{black} = -(1+1+1+1+2)-4*c$$

Distance-Based Utility

The Price of Anarchy(POA)

$$R_{anarchy} = \frac{\max\{|u(g)_1|,\ldots,|u(g)_k|\}}{|u(g)_{efficient}|}$$

Where $|u(g)_i|, i = 1,\ldots,k$, is the cost and different networks with various order sequence to form networks.

$R_{anarchy} = 1$: all pairwise stable networks are efficient.

$R_{anarchy} > 1$: there are higher costs(lower payoffs) with some pairwise stable networks than the efficient network.

Distance-Based Utility

The Price of stability(POS)

$$R_{stability} = \frac{\min\{|u(g)_1|,\ldots,|u(g)_k|\}}{|u(g)_{efficient}|}$$

$R_{stability} = 1$: the efficient network will be stable

$R_{anarchy} = 1$: all pairwise stable networks are efficient.

$R_{stability} > 1$: all stable networks are inefficient

$R_{stability} = 1, R_{anarchy} > 1$: some stable networks are efficient while others are not.

$$R_{anarchy} \geq R_{stability} \geq 1$$

Give two model.

A Co-Author Model and Negative Externalities

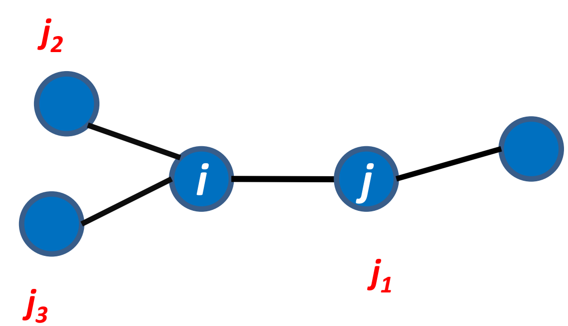
Situation : a given individual would rather that his or her neighbors have fewer connections rather than more.

Story

In **collaborating** on **a research project**. If **an individual's collaborator** **increases** the **time** spending on other projects, then the **individual sees less synergy** with that collaborator. Effectively, each player has a **fixed** amount of time to spend on projects and the time that researcher *i* spends on a given project is inversely related to the number of projects.

$$u_i(g) = \sum_{j:e_{ij}\in g} \Big(\underbrace{\frac{1}{d_i(g)}}_{\text{fixed time } (i)} + \underbrace{\frac{1}{d_j(g)}}_{\text{fixed time } (j)} + \underbrace{\frac{1}{d_i(g)d_j(g)}}_{\text{Synergy time } (i,j)}\Big)$$

Identify all neighbors of node *i*.



Example:

node 2 : $\frac{1}{1}+\frac{1}{1}+\frac{1}{1}=3$

node 1: $\frac{1}{2}+\frac{1}{2}+\frac{1}{4}=1.25$ (1-2)

$\frac{1}{2}+\frac{1}{1}+\frac{1}{2}=2$ (1-3)

3.25

Small Worlds in an Islands-Connections Model

It belongs to **distance-based Model**, and most features stem from a distance-based cost structure.

e.g high clustering : find a **cheaper** to maintain links to each other.

smaller 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 benefit*** to that pair of nodes

Islands-Connections Model

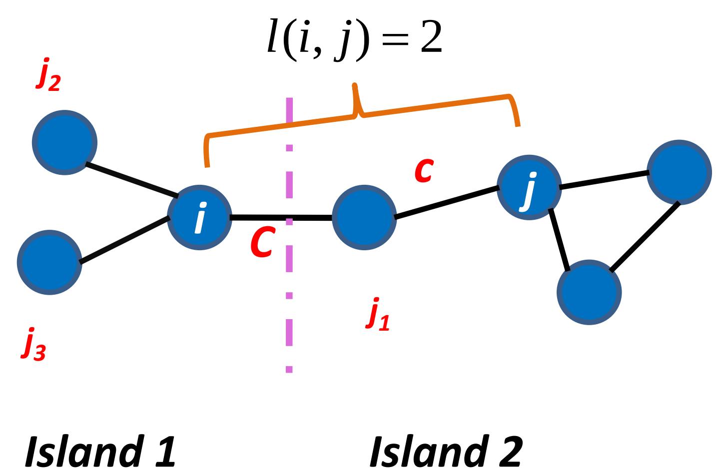
assumption

1. If the min path length between two players is more than ***D* links**, **no information** could **received** from each other.
2. For simply, there are **K islands (clusters)**, each of which has ***J* players** on it.
3. Two players connected by a link **cost *c*** for each other if they are on same island, and *C* otherwise, where ***C > c > 0*** .

Utility function :

$$u_i(g) = \sum_{k \neq i: l(i,k) \leq D} \delta^{l(i,j)} - \sum_{j:ij \in g} c_{ij}$$

Where $c_{ij} = c$ if *i* and *j* are on the same island and *C* otherwise. δ: decay rate.

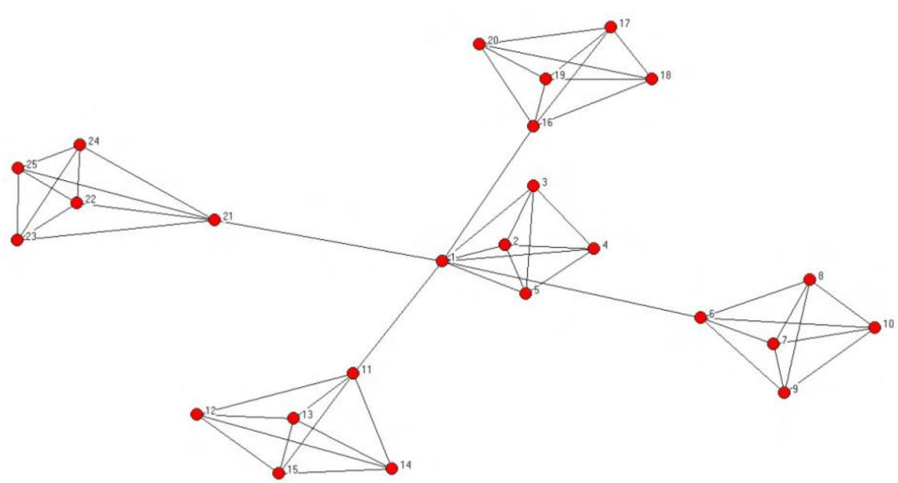


Islands-Connections Model

proposition

If $c < \delta - \delta^2$ and $C < \delta + (J-1)\delta^2$, then any network that is pairwise stable or efficient is such that

- The players o any given island are completely connected to each other.
- And if $\delta - \delta^3 < C$, then a lower bound on individual, average, and overall clustering is $\frac{(J-1)(J-2)}{J^2K^2}$

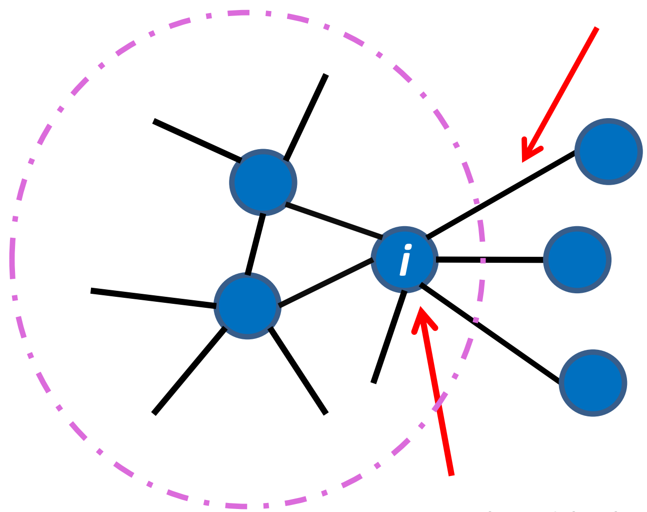


$c < .04,\ 1 < C < 4.5,\ \delta = .95$

数据挖掘实验室

Data Mining Lab

Islands-Connections Model



At least C_{J-1}^{2} pairs of *i*'s neighbors.

Max total of C_{J-L-1}^{2} pairs of neighbors.

→ Low bound:

$$\frac{(J-1)(J-2)}{[(J-1-L)(J-2-L)]} \approx \frac{(J-1)(J-2)}{J^2K^2}$$

Conclusion

数据挖掘实验室
Data Mining Lab

Propose **fundamental concepts** (measures) to capture stability of edges and efficient of networks (Efficient, Parato Efficient). In addition, we define distance-based utility to explore the strategy choose by users. Moreover, Co-author and Island model are proposed as instances.

Thanks

Zhongjing Yu
yuzhongjing@foxmail.com