Networks, Economics, Communication Systems, Informatics and Multimedia Research Lab

http://medianetlab.ee.ucla.edu

Strategic Networks: Content Production, Dissemination and Link Formation among Self-interested Agents

Yu Zhang

yuzhang@ucla.edu

Ph.D. Defense

May 21, 2013

UCLA

Rating Protocol Design for Online Communities

- A new framework for incentive design in online communities
  - **Systematically designs indirect reciprocity-based incentive mechanisms (rating protocols)** to induce the voluntary contribution (resources, services etc.) of users and maximize the social welfare of the online communities (social networks, P2P networks, social computing systems, etc.)

Y. Zhang, J. Park and M. van der Schaar, "Rating Protocols for Online Communities," *ACM Trans. on Economics and Computation*, accepted and to appear.

Y. Zhang and M. van der Schaar, "Incentive Provision and Job Allocation in Social Cloud Systems," *IEEE J. on Sel. Areas in Commun.*, accepted and to appear.

Y. Zhang and M. van der Schaar, "Peer-to-Peer Multimedia Sharing based on Social Norms," *Elsevier J. on Signal Process.*, vol. 27, no. 5, pp. 383-400, 2012.

Y. Zhang and M. van der Schaar, "Strategic Learning and Robust Protocol Design for Online Communities with Selfish Users," *IEEE J. of Sel. Topics in Signal Process.*, accepted and to appear

Distributed Online Learning for Big Data

- Distributed online learning framework for large-scale Big Data mining
  - Designs efficient online learning algorithms for real-time, large-scale, distributed data mining tasks where
    - distributed learners have restricted data access, limited communication and computational capability
    - incoming data is time-varying, non-stationary
  - Coordinates among distributed learners to optimally trade-off mining accuracy against communication/computation costs

Y. Zhang, D. Sow, D. Turaga, and M. van der Schaar, "A Distributed Online Learning Framework for Vertically Distributed Big Data," *IEEE Trans. on Signal Process.*, submitted.

Y. Zhang, D. Sow, D. Turaga, M. van der Schaar, "A Fast Online Learning Algorithm for Distributed Mining of BigData," *Big Data Analytics Workshop at SIGMETRICS*, 2013.

Stochastic Control in Multimedia Streaming

- A new systematic cross-layer optimization framework for real-time multimedia streaming in unknown, dynamic environments (time-varying networks, time-varying delay requirements, time-varying source characteristics)
  - Rigorously formalizes the cross-layer optimization problem as stochastic control problems
  - Online learns to make fast and efficient decisions in unknown, dynamic environments (structure-dependent reinforcement-learning solutions)
  - Challenges addressed:
    - Large state space
    - Delay-sensitive applications => Fast learning required

Y. Zhang, F. Fu, and M. van der Schaar, "Online Learning and Optimization for Wireless Video Transmission," *IEEE Trans. on Signal Process.*, vol. 58, no. 6, pp. 3108-3124, 2010.

Outline

- Motivation
- Network Formation with Strategic Content Acquisition
- Network Formation with Strategic Content Dissemination
- Conclusion

Socio-technical Networks - Emergence

– *Social-technical networks* enable individuals to share content, contribute expertise, collectively solve tasks, disseminate information at a low cost.

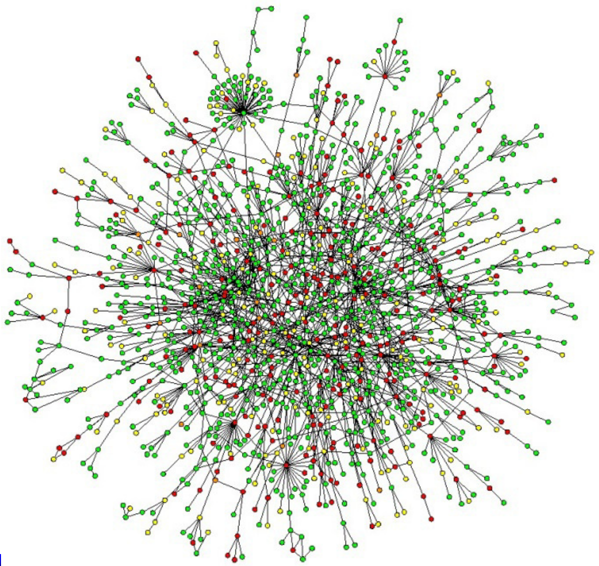
Challenges

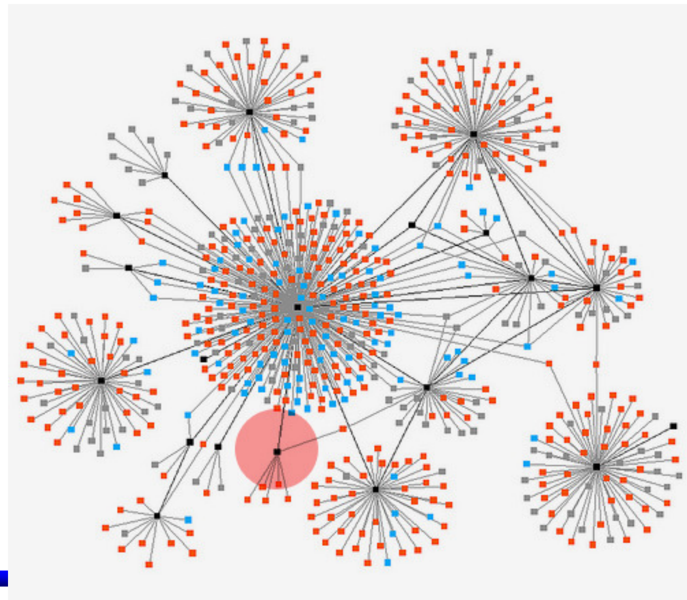
Understand and influence how strategic agents proactively make decisions on:

- **Content production** - *whether* to personally produce content and *how much*
  - "Content" – any knowledge, data, file, service etc.
    - Tweets/posts (Twitter/Facebook)
    - Self-made video (YouTube)
    - Data file (P2P)
- **Link formation** - *whether* to form/severe links
  - "Link" – any social/physical connection for content exchange
    - Friendship connection (Twitter/Facebook)
    - Peer-to-peer connection (P2P)

Network Formation Games

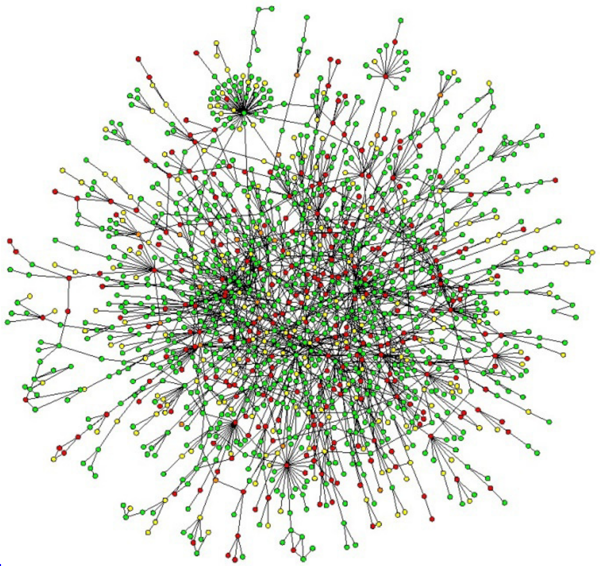
– Key questions - Given the strategic content production and link formation
  - **What network topologies arise at equilibrium?**
    - Small-world
    - Scale-free
    - Short-diameter (6-degree separation)

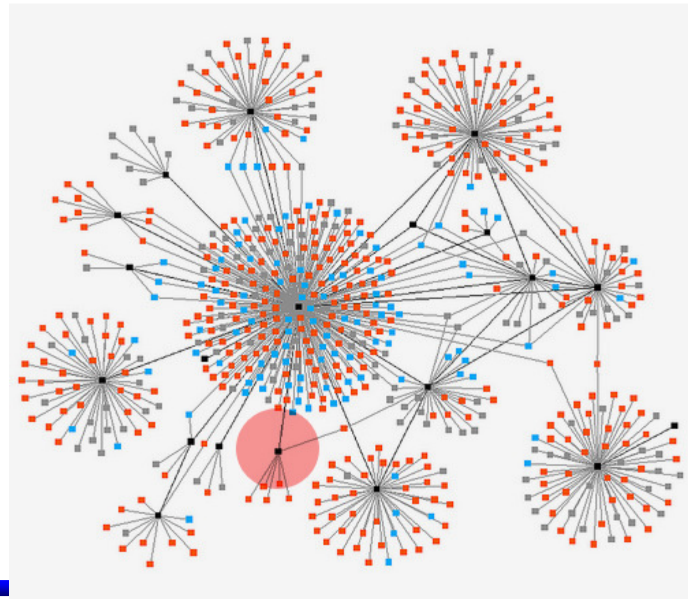




Network Formation Games

- Key questions - Given the strategic content production and link formation
  - What network topologies arise at equilibrium?
  - **How efficient the (equilibrium) topologies are?**
    - Content production efficiency
    - Content sharing efficiency
    - Fairness





Related Works – Network Formation

CS literature

- A. Fabrikant, A. Luthra, E. Maneva, C. Papadimitriou, and S. Shenker, "On a network creation game", 2003.

  - network formation and price of anarchy in networks with indirect information transmission (agents can access not only information from "neighbors", but also from neighbors of neighbors)

- J. Corbo and D. C. Parkes, "The price of selfish behavior in bilateral network formation", 2005.

  - bilateral network formation and price of anarchy in networks where link creation requires mutual consent and cost is two-sided

Related Works – Network Formation

CS literature

- E. Anshelevich, A. Dasgupta, E. Tardos, and T. Wexler, Near-optimal network design with selfish agents, 2003.

  - proposes efficient (polynomial time) algorithms to find Nash equilibria that are near-optimal given that agents have specific connectivity requirements

- L. Blume, D. Easley, J. Kleinberg, R. Kleinberg, and E. Tardos, Network formation in the presence of contagious risk, 2011.

  - network formation game with contagious risk, where an agent is exposed to the risk of being hit by a cascading failure based on its connectivity

Related Works – Network Formation

Game-theoretic/Economics literature

- Network games: many interesting works (Jackson, Goyal, etc.)
- Homogeneous agents (Bala and Goyal, 2000)
  - Equilibrium topologies are symmetric: circles, stars, variants of stars.
- Heterogeneous agents (Galeotti and Goyal, 2006)
  - A strict equilibrium is a minimal network, and every minimal network could be a strict equilibrium for some benefits and costs.
- Indirect information flow (Hojman and Szeidl, 2005)
  - When value of information is decaying over the distance from which is acquired, the equilibrium topologies usually have small diameters.

Limitations

- **Agents are non-strategic on content production**
  - The amount of each agent's possessed content is exogenously determined
  - Neglect the interaction between strategic content production and link formation
- **No variety in content**
  - Content is perfectly substitutable has the same value in consumption
  - Neglect the agents' preference on content variety
- **No model for content dissemination**
  - Agents benefit solely from consuming acquired content from others
  - Neglect the benefit from dissemination (e.g. advertisement)

Our Contribution

– **Network Formation Game with strategic content production [1]**
  - Captures agents' strategic behavior on both *content production* and *link formation*
  - Explicitly considers *agents' preference on content variety* (Dixit-Stiglitz model)
    - An agent's benefit does not only depend on the total amount of its consumed content, but also on its variety

[1] **Y. Zhang**, M. van der Schaar, "Information Production and Link Formation in Social Computing Systems," *IEEE J. on Sel. Areas in Commun.*, vol. 30, no. 11, pp. 2136-2145, 2012.

Our Contribution

- **Network Formation Game with strategic content production [1]**
  - Captures agents' strategic behavior on both *content production* and *link formation*
  - Explicitly considers *agents' preference on content variety* (Dixit-Stiglitz model)
    - An agent's benefit does not only depend on the total amount of its consumed content, but also on its variety
- **Network Formation Game with strategic content dissemination [2]**
  - Considers the scenario in which agents benefit from *content dissemination* (instead of content consumption)

[1] **Y. Zhang**, M. van der Schaar, "Information Production and Link Formation in Social Computing Systems," *IEEE J. on Sel. Areas in Commun.*, vol. 30, no. 11, pp. 2136-2145, 2012.

[2] **Y. Zhang**, M. van der Schaar, "Strategic Networks: Information Dissemination and Link Formation Among Self-interested Agents," *IEEE J. on Sel. Areas in Commun.*, vol. 31, no. 6, to be published in June 2013.

Outline

- ❑ Motivation
- ❑ **Network Formation with Strategic Content Acquisition**
- ❑ Network Formation with Strategic Content Dissemination
- ❑ Conclusion

Model

- We consider a network consisting of *n* agents
- Each individual agent can share its produced content with other agents
- Consider unilateral link formation and undirected link
  - Links are created by the unilateral actions of agents, and link costs are one-sided (paid by the creator of a link)
  - Content can be transmitted in both directions over an established link
- Content propagation
  - Local: Agents only exchange content with the one-hop neighbors
  - Indirect: Agents exchange content with the multi-hop neighbors

Settings

- $N = \{1, \ldots, n\}$: set of agents
- Action
  - $x_i \in \mathbb{R}^+$: production level
    - the amount of content produced by agent *i*
  - $\boldsymbol{g}_i \triangleq (g_{i1}, \ldots, g_{in}) \in \{0,1\}^n$: link formation decision
    - $g_{ij} = 1$: agent *i* forms a link to agent *j*
    - $\overline{g}_{ij} = \max\{g_{ij}, g_{ji}\} = 1$: agents *i* and *j* are connected (neighbors)
- $N_i(\boldsymbol{g}) \triangleq \{j \mid g_{ij} = 1\}$: the set of agents who agent *i* forms links to
  - Determines the link formation cost
- $N_i(\overline{\boldsymbol{g}}) \triangleq \{j \mid \overline{g}_{ij} = 1\}$: the set of agent *i*'s neighbors
  - Determines the total content consumption benefit of agent *i*

The Law of the Few (Galeotti & Goyal, 2010)

- Payoff function

$$u_i\left(\boldsymbol{x},\boldsymbol{g}\right) = v\left(x_i + \sum_{j \in N_i\left(\bar{\boldsymbol{g}}\right)} x_j\right) - cx_i - k\left|N_i\left(\boldsymbol{g}\right)\right|$$

benefit from content consumption

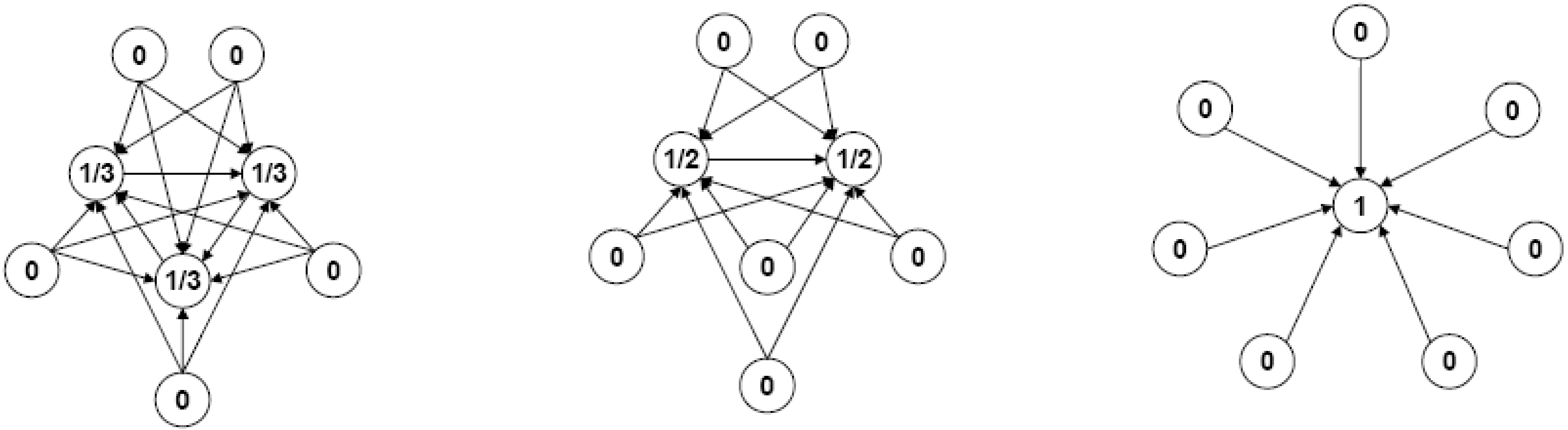
cost for content production

cost for link formation

  - c: unit content production cost
  - k: link formation cost
  - Assumption
    - Content is perfectly substitutable – The same amount of content produced by any agent has the same value in consumption
    - Agents have no preference on the content variety!

The Law of the Few (Galeotti & Goyal, 2010)

- Result: The law of the few
  - The number of content producers and the total amount of produced content are upper-bounded and independent of the network size
  - The network is dominated by content consumers when its size goes to infinity



Our Model (Content Variety)

- Payoff function

$$u_i(\boldsymbol{x}, \boldsymbol{g}) = v\left(\left[x_i^{\rho} + \sum_{j \in N_i(\overline{\boldsymbol{g}})} x_j^{\rho}\right]^{1/\rho}\right) - c x_i - k \left| N_i(\boldsymbol{g}) \right|$$

  - Assumption
    - Content is not perfectly substitutable (Dixit-Stiglitz model)

Our Model (Content Variety)

- Payoff function

$$u_i(\boldsymbol{x}, \boldsymbol{g}) = v\left(\left[x_i^{\rho} + \sum_{j \in N_i(\overline{\boldsymbol{g}})} x_j^{\rho}\right]^{1/\rho}\right) - cx_i - k\left|N_i(\boldsymbol{g})\right|$$

  - $\rho \in (0,1)$ measures the influence of content variety
    - When $\rho = 1$, content is perfectly substitutable and variety does not matter
    - When ρ is small, variety matters a lot
  - $X_i \triangleq \left[x_i^{\rho} + \sum_{j \in N_i(\overline{\boldsymbol{g}})} x_j^{\rho}\right]^{1/\rho}$: the amount of agent *i*'s effective consumption
    - Negative externality on the production level

Assumptions on the Payoff Function

- $v(\bullet)$ is a twice continuously differentiable, increasing and strict concave function
- $v(\bullet)$ satisfies
  - $v(0) = 0$
  - $\lim_{x \to \infty} v'(x) = 0$
  - $v'(0) \triangleq \lim_{x \to \infty} v'(x) < \infty$
  - $v'(0) > \alpha$

- An agent's benefit increases with the amount of effective content, but saturates with the rate of increase approaching to 0
- There always exists an upper bound on the amount of content that an individual agent is willing to consume

Equilibrium Concept

- One-stage simultaneous move game with pure strategies

- Strategy profile

$$\boldsymbol{s} = ((x_i, \boldsymbol{g}_i)_{i=1}^{N})$$

- Consider equilibrium: Strict Nash equilibrium

$$u_i\left(\boldsymbol{s}_i^*, \boldsymbol{s}_{-i}^*\right) > u_i\left(\boldsymbol{s}_i, \boldsymbol{s}_{-i}^*\right),\ \forall \boldsymbol{s}_i \in \mathbb{R}^+ \times \left\{0,1\right\}^{n-1}, \forall i \in N$$

Any agent's utility strictly decreases if it unilaterally deviates from $(x_i^*, \boldsymbol{g}_i^*)$

Symmetric/Asymmetric Strategy Profile

- For a strategy profile $\boldsymbol{s}$, agents are ordered by their production levels $x_1 \geq x_2 \geq \cdots \geq x_n$.
  - high producer: $\{i \mid x_i = x_1\}$
  - low producer: $\{i \mid x_i < x_1\}$
  - The number of high producers: $n_h(\boldsymbol{s})$
  - The number of low producers: $n_l(\boldsymbol{s}) = n - n_h(\boldsymbol{s})$

- Symmetric profile: $n_h(\boldsymbol{s}) = n$
  - Agents produce the same amount of content
- Asymmetric profile: $n_h(\boldsymbol{s}) < n$
  - Agents produce different amounts of content

Equilibrium Properties

- **Lemma 1.** In any equilibrium $\boldsymbol{s}^* = \left(\boldsymbol{x}^*, \boldsymbol{g}^*\right)$,
  - $g_{ij}^* g_{ji}^* = 0, \forall i, j \in N$
    - No redundant link formation
    - Due to the undirected information flow
  - $x_i^* > 0, \forall i \in N$
    - Each agent produces a positive amount at equilibrium
    - Its self-produced content cannot be fully replaced by acquired content
  - $x_i^* \leq \bar{x}, \forall i \in N$, where $\bar{x}$ is the unique solution of $v'\left(\bar{x}\right) = c$
    - The production level is upper-bounded
    - The concave benefit function and the linear production cost

Equilibrium Properties

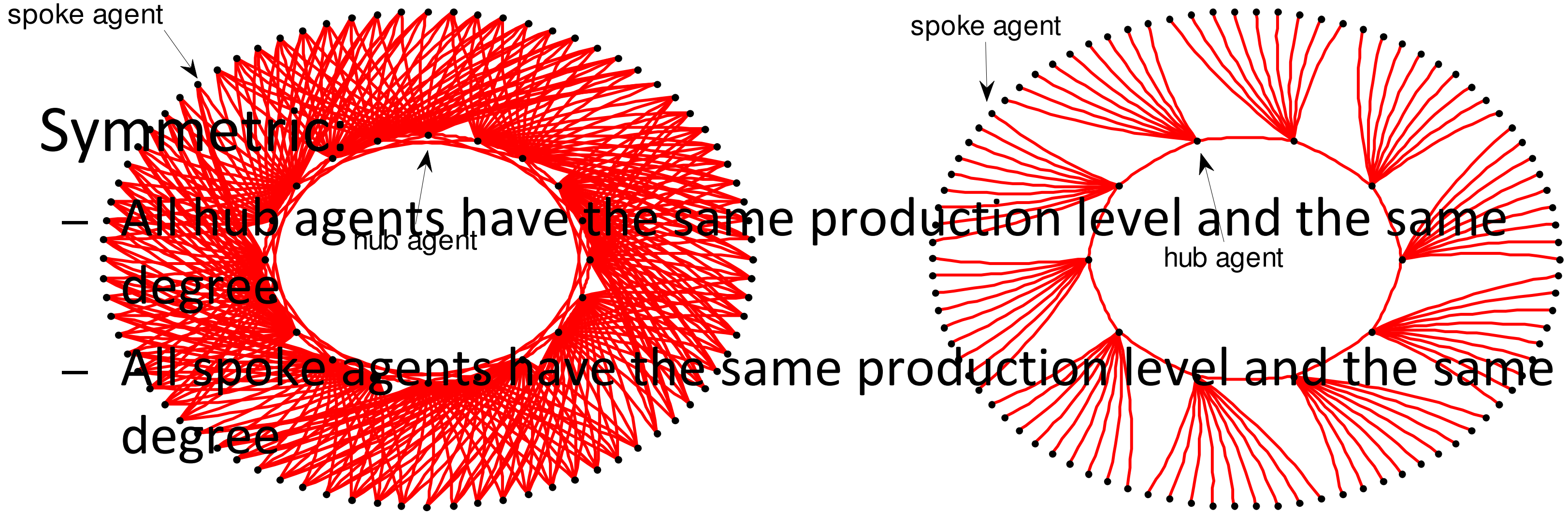
- **Lemma 2.** In any equilibrium $\boldsymbol{s}^* = (\boldsymbol{x}^*, \boldsymbol{g}^*)$
  - $\overline{g}_{ij}^* = 0$, for some $i, j \in N$ and $i \neq j$
    - The network is never complete at equilibrium
    - Only symmetric profiles form complete networks at equilibrium
  - For each $i \leq n_h(\boldsymbol{s}^*)$, $g_{ij}^* = 0, \ \forall j > n_h(\boldsymbol{s}^*)$
    - No high producer forms links to low producers
    - An high producer always gets sufficient content from other high producers

Equilibrium properties

- **Lemma 3.** In any equilibrium $s^* = (x^*, g^*)$
  - For each $i \le n_h(s^*)$, $\bar{g}^*_{ii'} = 0$, for some $i' \le n_h(s^*)$ and $i' \ne i$
    - High producers are not all linked to each other
    - Suppose *i* is connected with all high producers
    - There is at least one low producer *j* who is not connected with *i.*
    - Each neighbor of *j* should also be a neighbor of *i.*
    - *i* should produce less information than *j* -> Contradiction
  - For each $j > n_h(s^*)$, $g^*_{ji} = 1$, for some $i \le n_h(s^*)$.
    - Each low producer forms links to at least one high producer.
  - $g^*_{jj'} = 1,\ for\ \ some\ \ j, j' > n_h(s^*)$ only if $g^*_{ji} = 1,\ \forall i \le n_h(s^*)$
    - A low producer *j* will only form a link to another low producer *j'* only if it has already connected to all high producers.

Symmetric Core-periphery Network

- A strategy profile s forms a symmetric core-periphery network if only two types of agents exist in s:
  - Each high producer (hub agent) produces an amount $\tilde{x}(\boldsymbol{x})$
  - Each low producer (spoke agent) produces an amount $\underset{\sim}{x}(\boldsymbol{x}^*)$ and forms links with $q(\boldsymbol{g}^*)$ high producers



- Symmetric:
  - All hub agents have the same production level and the same degree
  - All spoke agents have the same production level and the same degree

Asymptotic Equilibrium

- **Theorem 2.** Given *c*, *k*, ρ, and *v*,
  - $\lim_{n\to\infty} S_n^{SCP} = S_n^*$
    - S_n^* : the set of (asymmetric) equilibrium strategy profiles when the network size is n
    - S_n^{SCP}: the set of (asymmetric) equilibrium strategy profiles that form symmetric core-periphery networks when the network size is n
  - $\lim_{n\to\infty} \inf_{\boldsymbol{s}^* \in S_n^*} \left\{ n_h\left(\boldsymbol{s}^*\right) \right\} = \infty$ and $\lim_{n\to\infty} \inf_{\boldsymbol{s}^* \in S_n^*} \left\{ n_l\left(\boldsymbol{s}^*\right) \right\} = \infty$
- When the network size goes to infinity,
  - Every equilibrium forms a symmetric core-periphery network
  - The numbers of both hub agents and spoke agents goes to infinity

Asymptotic Equilibrium

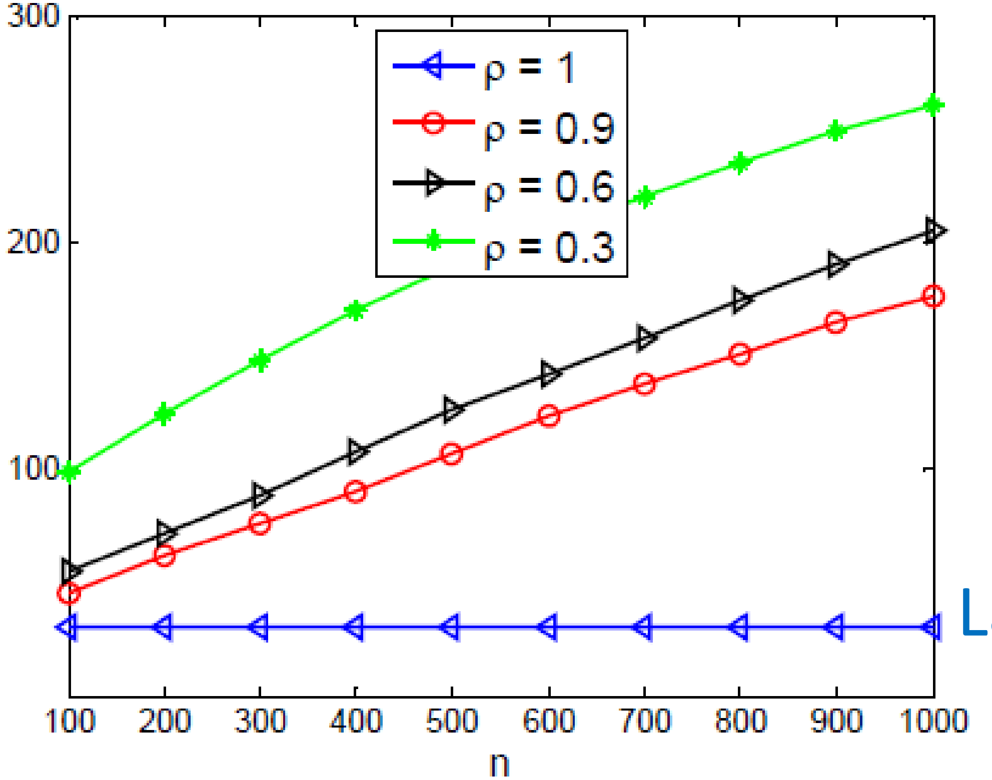
- **Theorem 3.** Given *c*, *k*, ρ, and *v*, the number of hub agents grows at the same order as the entire population, i.e. $\inf_{\boldsymbol{s}^* \in S_n^*} \left\{ n_h \left(\boldsymbol{s}^* \right) \right\}$ is $\Omega\left(n\right)$
  - There are two constants δ_1 and δ_2 such that

$$\delta_1 n \le \inf_{\boldsymbol{s}^* \in S_n^*} \left\{ n_h \left(\boldsymbol{s}^* \right) \right\} \le \delta_2 n, \ \forall n$$

The law of the few no longer holds!

Illustrative Result

Total amount of produced content at equilibrium



Law of the few

Efficiency in the Network Formation Game

- The *social welfare* under a strategy profile

$$W\left(\boldsymbol{x},\boldsymbol{g}\right)=\sum_{i\in N}u_i\left(\boldsymbol{x},\boldsymbol{g}\right)$$

- A strategy profile $\boldsymbol{s}^{\#}=\left(\boldsymbol{x}^{\#},\boldsymbol{g}^{\#}\right)$ achieves the *social optimum* $W^{\#}$ if and only if

$$W^{\#}=W\left(\boldsymbol{x}^{\#},\boldsymbol{g}^{\#}\right)\geq W\left(\boldsymbol{x},\boldsymbol{g}\right),\ \forall\left(\boldsymbol{x},\boldsymbol{g}\right)$$

- **Lemma 4.** Given *c*, *k*, ρ, and *v*, the social optimum is upper-bounded as follows

$$W^{\#}\leq\max\left\{n\left[v\left(n^{1/\rho}\widehat{x}_n\right)-c\widehat{x}_n\right]-nk/2, n\left[v\left(\overline{x}\right)-c\overline{x}\right]\right\}$$

  - $\widehat{x}_n$ is the solution of $v'\left(n^{1/\rho}\widehat{x}_n\right)=c/n^{\left(1-\rho\right)/\rho}$

the upper bound when the optimal network is non-empty

the upper bound when the optimal network is empty

Efficiency Loss at Equilibrium

- **Price of Stability (PoS)**
  - The ratio between the ***social optimum*** and the ***"best" equilibrium***

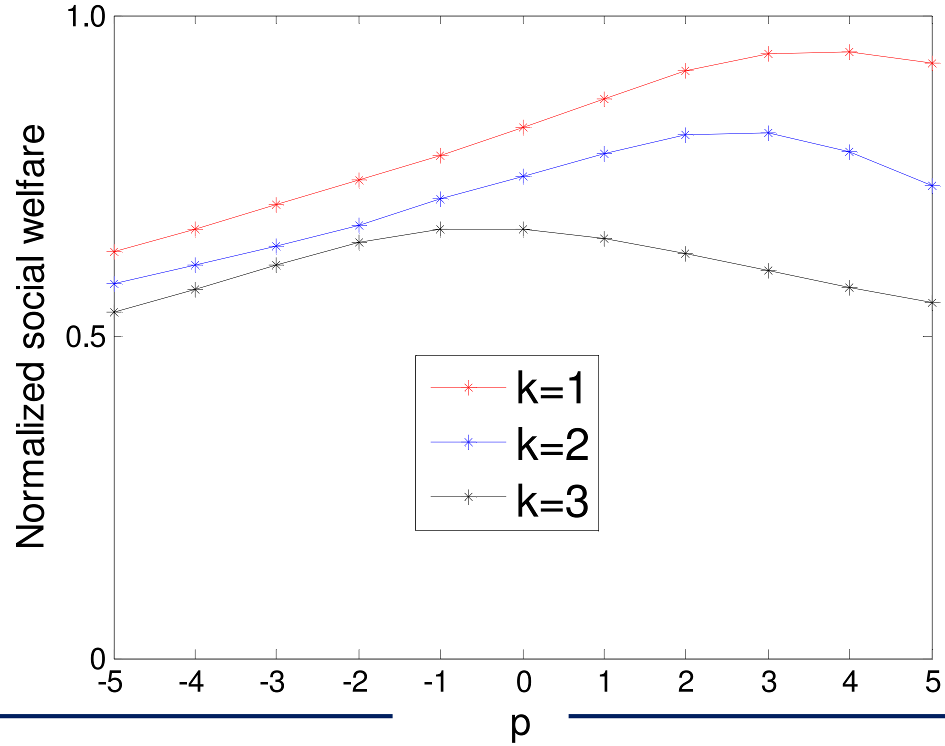
- **Theorem 5.** Given *c*, *k*, ρ, and *v*, the PoS is upper-bounded as

$$PoS \leq \max\left\{\left(\left[v\left(n^{1/\rho}\widehat{x}_n\right) - c\widehat{x}_n\right] - k/2\right) / \left[v\left(\overline{x}\right) - c\overline{x}\right], 1\right\}$$

- Idea of proof:
  - The social welfare $W\left(\boldsymbol{x}^*, \boldsymbol{g}^*\right)$ of any equilibrium strategy is lower bounded by $n\left[v\left(\overline{x}\right) - c\overline{x}\right]$

Can We Reduce the PoS?

- The cost of forming a link can be adjusted and more efficient equilibrium can be induced -> pricing on links



subsidize the link creator to increase agents' incentive to form links and decrease self-production

charge the link creator to reduce agents' incentive to form links and increase self-production

Indirect Content Sharing

– Content propagation is indirect: an agent can also consume content which its neighbors acquired from other agents

- Payoff function

$$w_i\left(\boldsymbol{x},\boldsymbol{g}\right) = v\left(\left[x_i^{\rho} + \sum_{l=1}^{n-1} \sum_{j\in N_i^l\left(\overline{\boldsymbol{g}}\right)} x_j^{\rho}\right]^{1/\rho}\right) - cx_i - k\left|N_i\left(\boldsymbol{g}\right)\right|$$

An agent acquires content from all neighbors with whom it connects via a path

Indirect Content Sharing

- **Theorem 6.** There is a value $\bar{k}_n$ such that
  - When $k > \bar{k}_n$, there exists a unique equilibrium where each agent personally acquires an amount $\bar{x}$ of content and no agent forms links (empty network)
  - When $k < \bar{k}_n$, then each equilibrium is minimally connected
    - Minimally connected network: There is a unique path between every pair of agents

Key difference to local content sharing:

here the agents with most connections become the network's hubs and share the content unlike in the local content sharing (high producers are the hubs)

UCLA

Outline

- ❑ Motivation
- ❑ Network Formation with Strategic Content Acquisition
- ❑ **Network Formation with Strategic Content Dissemination**
- ❑ Conclusion

Model

- We consider a network consisting of *n* agents
- Two scenarios
  - Content is exogenously endowed
  - Content is endogenously produced
- Consider unilateral link formation and undirected link
- Consider indirect content sharing

Settings

- $N = \{1, \ldots, n\}$: set of agents

- Action
  - $\boldsymbol{g}_i \triangleq (g_{i1}, \ldots, g_{in}) \in \{0,1\}^n$: link formation decision of agent i

- Connectivity
  - $cl(\overline{\boldsymbol{g}}) = \{(i,j) \in N \times N \mid i \neq j,\ \overline{g}_{ij} = 1\}$: the topology
  - $path_{ij} = ((i,j_1),(j_1,j_2),\ldots,(j_m,j)) \subseteq cl(\overline{\boldsymbol{g}})$: a path between agents i and j
  - $N_i(\overline{\boldsymbol{g}}) \triangleq \{j \mid i \to j\}$: the set of agents who agent i connects with (can reach via a path)

Content Dissemination Model

- Payoff function

$$u_i(\boldsymbol{g}) = f(x_i \mid N_i(\overline{\boldsymbol{g}}) \mid) - \sum_{j \in N_i(\boldsymbol{g})} k_{ij}$$

Link formation costs

Benefit from content dissemination

Possessed content (exogenously determined)

  - The benefit monotonically increases with the total amount of *disseminated content*
    - Proportional to the production level x_i
    - Proportional to the number of reachable agents $\mid N_i(\overline{\boldsymbol{g}}) \mid$
  - Agents are heterogeneous in terms of benefit and cost
    - Significantly impacts the shape of the equilibrium topologies!

Difference to Content Acquisition

Content dissemination

$$u_i(\boldsymbol{g}) = f(x_i \mid N_i(\overline{\boldsymbol{g}}) \mid) - \sum_{j \in N_i(\boldsymbol{g})} k_{ij}$$

Content acquisition

$$u_i(\boldsymbol{g}) = f(x_i + \sum_{j \in N_i(\overline{\boldsymbol{g}})} x_j) - \sum_{j \in N_i(\boldsymbol{g})} k_{ij}$$

- *Source of benefit*
  - Disseminating content: reach as many agents as possible (also values the variety of the connections!)
  - Acquiring content: collect as much content as possible (the number of reachable agents does not matter!)
- *Incentives on link formation*
  - Increases with x_i in content dissemination
  - Decreases with x_i in content acquisition
- Different *equilibrium behavior and topology*

Main Results

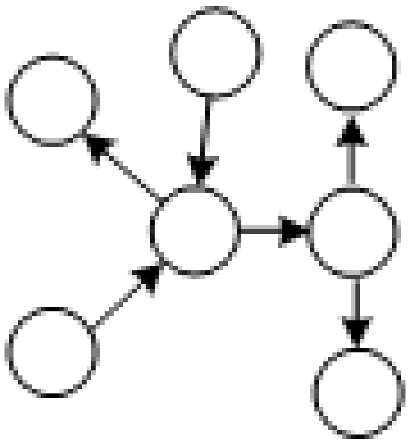
- Minimal connectedness
  - Agents do not form redundant links
  - Equilibrium topologies usually form hierarchical tree structures with rare occurrence of loops
- Short distance (Small-world phenomenon)
  - Agents tend to get closer to make efficient dissemination
  - The diameter of the equilibrium topology is upper-bounded, which is usually small
- Centrality
  - Equilibrium topologies often have core-periphery structure
  - The connections in the network is supported by several core agents (at the center of the topology)

Main Results

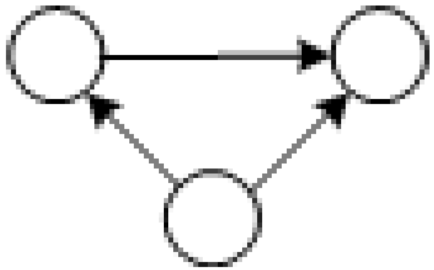
- High efficiency
  - Little efficiency loss compared to the social optimum
  - Due to the minimal connectedness
- Scale-free
  - Agent degree distribution is upper-bounded by the power-law distribution
  - The fraction of agents with high degrees is small

Minimal Connectedness

- **Component**
  - A *component* C is a set of agents such that $i \to j,\ \forall i, j \in C$ and $i \not\to j',\ \forall i \in C\ \ and\ \forall j' \notin C$.
- **Minimal component**
  - A component C is *minimal* if there is only one path in $cl(\overline{\boldsymbol{g}})$ from any agent $i \in C$ to any agent $j \in C$



A minimal component



A non-minimal component

Minimal Connectedness

- **Proposition 1.** In a NE g^*, each component is minimal.

- **Proposition 2.** The network in each NE is always minimally connected if $f(x_i \mid N \mid) - f(x_i(\mid N \mid -1)) > \max_{i,j}\{k_{ij}\}$.

The network is fully connected if the link formation cost is sufficiently large

Considered Network

- The exact shape of the equilibrium topology depends on the characteristics of link formation costs $\{k_{ij}\}$
- Consider two types of networks
  - **Networks with recipient-dependent costs**
    - $k_{ij} = k_j, \forall i \in N / \{j\}$: The cost of forming a link is recipient-specific
    - $k_i \in \{k^1, ..., k^L\}, \forall i \in N$: *L* different types of link formation costs
  - **Networks with groups**
    - Agents belong to Z different groups
    - Each agent only belongs to one group

$$N = \cup \{N_z\}_{z=1}^{Z} \qquad N_z \cap N_{z'} = \phi$$

    - The cost of intra-group links $\underline{k}$ is lower than the cost of inter-group links $\overline{k}$

Agents within the same group are more often interconnected

Networks with Recipient-dependent Costs

- **Theorem 1 (Short Distance).** Under each strict NE, the diameter of the network is at most $2L$.
  - The *diameter* of a network is the length of the longest path in the network

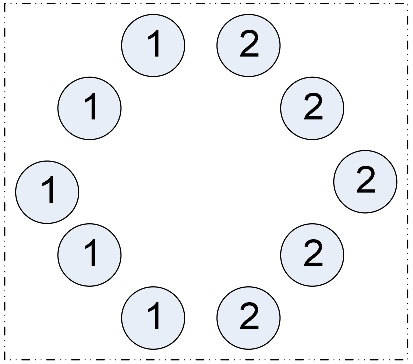
  Agents cannot be separated by more than $2L$ hops at equilibrium

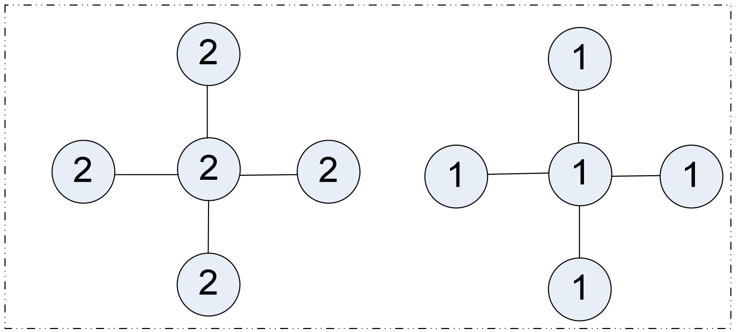
- **Corollary 1 .** When $k_{ij} = k, \forall i, j \in N$, each component forms a star topology at equilibrium.

A network with homogeneous agents (in terms of the link formation cost) has a maximum distance of 2!

Networks with Groups

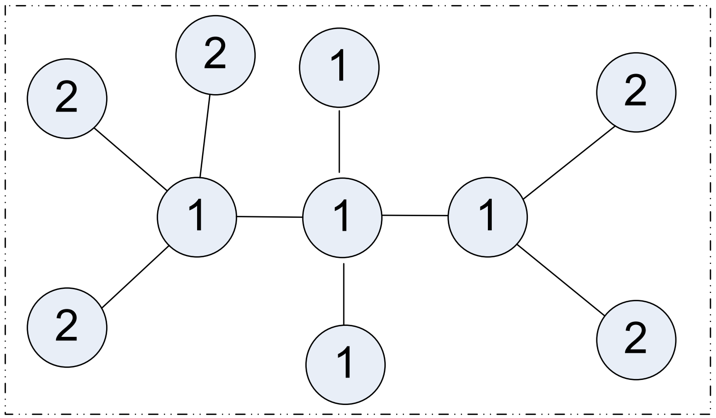
- **Theorem 2 (Centrality and Clustering).**
  - When $f(x) - f(0) < \underline{k}$, the unique strict NE $\boldsymbol{g}^*$ satisfies $g_{ij}^* = 0, \forall i, j$
    - The network is empty when the benefit of disseminating content is small
  - When $f(x) - f(0) \in (\underline{k}, \overline{k})$, the unique strict NE consists of Z components, where each component only contains agents from the same group and forms a star topology
    - Agents from the same group are *clustered* at equilibrium due to the small inter-group link formation cost





Networks with Groups

- **Theorem 2.** (cont'd)
  - When $f(x) - f(0) > \overline{k}$, in each strict NE there is a group N_z and an agent $i \in N_z$ such that $g^*_{ij} = 1, \forall j \in N_z / \{i\}$. Also for each agent $j' \notin N_z$, there is an agent $j \in N_z$ such that $g^*_{jj'} = 1$.
    - The network is connected when the benefit of disseminating content is sufficiently large
    - There is one group in the ***center*** of the network that connects all the other groups



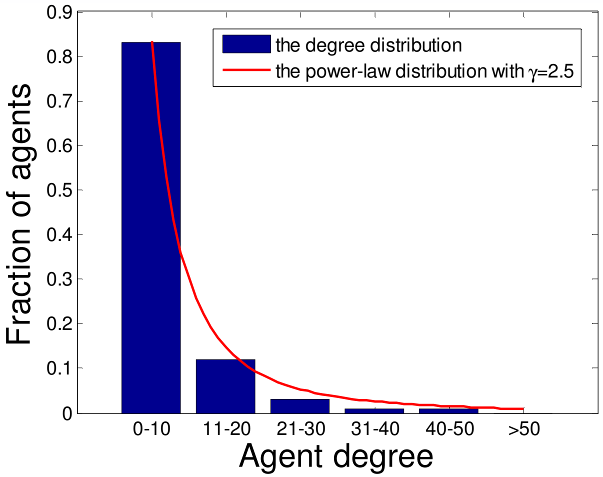
Equilibrium Efficiency

- **Metric**
  - Price of Stability (PoS): $U^{\#} / \max_{\boldsymbol{g}^*} U(\boldsymbol{g}^*)$
  - Price of Anarchy (PoA): $U^{\#} / \min_{\boldsymbol{g}^*} U(\boldsymbol{g}^*)$
    - The ratio between the ***social optimum*** and the ***"worst" equilibrium***
- **Theorem 3.** If $k_{ij} = k_j, \forall i \in N / \{j\}$, the PoS is always 1.
  - There is always an equilibrium that achieves social optimum
- **Theorem 4.** If $k_{ij} = k_j, \forall i \in N / \{j\}$, the PoA is upper-bounded by $\max_{i,j \in N} k_i / k_j$.
  - The PoA is small when agents have close link formation costs
  - Each equilibrium is efficient when agents are homogeneous

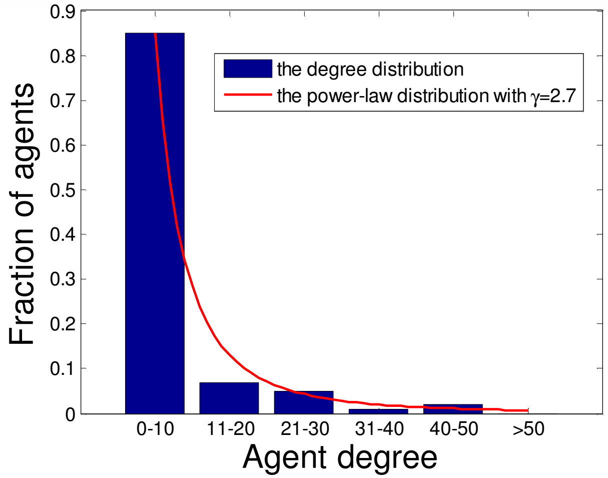
The "minimally connectedness" property ensures the efficient dissemination of content with the fewest possible links and the efficiency at equilibrium

Degree Distribution at Equilibrium

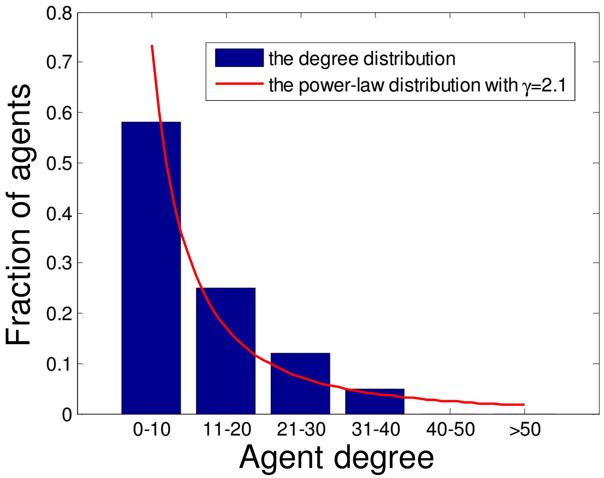
Network with homogeneous agents



Network with groups



Network with recipient-dependent cost



The degree distribution is always upper-bounded by a power-law distribution

Content Dissemination with Strategic Production

- What happens if agents strategically determine their production levels?
- Each agent determines its production level x_i and the link formation strategy $\boldsymbol{g}_i$ to maximize its payoff
- **Equilibrium**

$$u_i(x_i^*, \boldsymbol{g}_i^*, \boldsymbol{x}_{-i}^* \boldsymbol{g}_{-i}^*) \geq u_i(x_i', \boldsymbol{g}_i', \boldsymbol{x}_{-i}^* \boldsymbol{g}_{-i}^*), \forall i \in N$$

Content Dissemination with Strategic Production

- **Theorem 5.** When agents strategically determine their production levels and benefit from content dissemination, $\inf_{\boldsymbol{s}^* \in S_N^*} \{\sum_{i \in N} x_i^*\} \sim \Theta(|N|)$
  - The total amount of content produced in the network at equilibrium grows proportionally to the network size (In sharp contrast to the law of the few!)

- Reason for the "law of the few" to disappear
  - If benefited from content dissemination, an agent's production decision is not affected by the production levels of other agents
  - Its production level always increases when the network size grows

UCLA

Outline

- ❑ Motivation
- ❑ Network Formation with Strategic Content Acquisition
- ❑ Network Formation with Strategic Content Dissemination
- ❑ **Conclusion**

Conclusion

- If content variety matters, and content production and link formation are choices
  - Many hubs and many spokes
  - "Law of the few" disappears
  - Total content production grows as the same rate as size of the network
- If agents are interested in disseminating information
  - Minimally connected network
  - Short network diameter
  - Clustering and centrality
  - Disappearance of the "law of the few"
- **Strategic design: new research field where network agents are strategic**
- **Future work: BIG RESEARCH AGENDA**

Selected Journal Publications

- [9] **Y. Zhang**, J. Park and M. van der Schaar, "Rating Protocols for Online Communities," *ACM Trans. on Economics and Computation*, accepted and to appear.
- [8] **Y. Zhang** and M. van der Schaar, "Incentive Provision and Job Allocation in Social Cloud Systems," *IEEE J. on Sel. Areas in Commun.*, accepted and to appear.
- [7] **Y. Zhang** and M. van der Schaar, "Information Production and Link Formation in Social Computing Systems," *IEEE J. on Sel. Areas in Commun.*, vol. 30, no. 11, pp. 2136-2145, 2012.
- [6] **Y. Zhang** and M. van der Schaar, "Peer-to-Peer Multimedia Sharing based on Social Norms," *Elsevier J. on Signal Process.*, vol. 27, no. 5, pp. 383-400, 2012.
- [5] **Y. Zhang**, F. Fu, and M. van der Schaar, "Online Learning and Optimization for Wireless Video Transmission," *IEEE Trans. on Signal Process.*, vol. 58, no. 6, pp. 3108-3124, 2010.
- [4] **Y. Zhang** and M. van der Schaar, "Strategic Learning and Robust Protocol Design for Online Communities with Selfish Users," *IEEE J. of Sel. Topics in Signal Process.*, accepted and to appear.
- [3] **Y. Zhang** and M. van der Schaar, "Strategic Networks: Information Dissemination and Link Formation Among Self-interested Agents," *IEEE J. on Sel. Areas in Commun.*, accepted and to appear.
- [2] **Y. Zhang** and M. van der Schaar, "Collective Rating Design for Secure Internet Peering," *IEEE/ACM Trans. on Networking*, submitted.
- [1] **Y. Zhang**, D. Sow, D. Turaga, and M. van der Schaar, "A Distributed Online Learning Framework for Vertically Distributed Big Data," *IEEE Trans. on Signal Process.*, submitted.

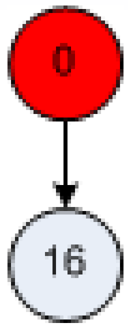
Selected Conference Publications

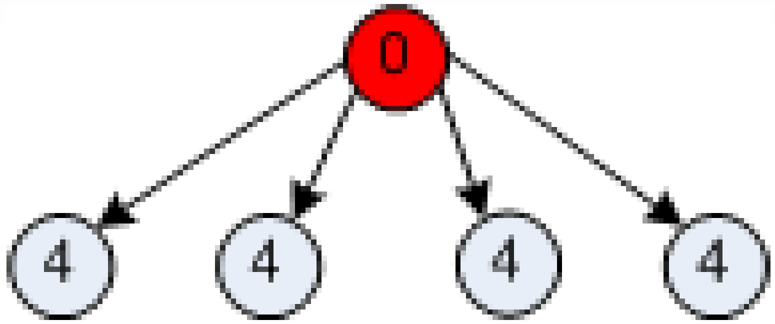
- [8] Y. Xiao, **Y. Zhang**, and M. van der Schaar, "Socially Optimal Design of Crowdsourcing Platforms with Reputation Update Errors," *IEEE ICASSP*, 2013.
- [7] **Y. Zhang** and M. van der Schaar, "Strategic Information Dissemination and Link Formation in Social Networks," *IEEE ICASSP*, 2013.
- [6] J. Xu, **Y. Zhang**, and M. van der Schaar, "Rating System for Enhanced Cyber-security Investments," *IEEE ICASSP* 2013.
- [5] **Y. Zhang** and M. van der Schaar, "Reputation-based Protocols in Crowdsourcing Applications," *IEEE INFOCOM* 2012.
- [4] **Y. Zhang** and M. van der Schaar, "Collective Rating for Online Labor Markets," *Allerton Conf. on Commun., Control and Computing*, 2012.
- [3] **Y. Zhang** and M. van der Schaar, "User Adaptation and Long-run Evolution in Online Communities," *IEEE CDC*, 2011.
- [2] **Y. Zhang**, J. Park, and M. van der Schaar, "Social Norm based Incentive Mechanisms for Peer-to-Peer Networks," *IEEE ICASSP*, 2011.
- [1] **Y. Zhang**, F. Fu, and M. van der Schaar, "Online Learning for Wireless Video Transmission with Limited Information," *Int'l Packet Video Workshop*, 2009.

References

- [1] A. Fabrikant, A. Luthra, E. Maneva, C. Papadimitriou, and S. Shenker, “On a network creation game”, 2003.
- [2] J. Corbo and D. C. Parkes, “The price of selfish behavior in bilateral network formation”, 2005.
- [3] E. Anshelevich, A. Dasgupta, E. Tardos, and T. Wexler, “Near-optimal network design with selfish agents”, 2003.
- [4] L. Blume, D. Easley, J. Kleinberg, R. Kleinberg, and E. Tardos, “Network formation in the presence of contagious risk”, 2011.
- [5] V. Bala and S. Goyal, “A non-cooperative model of network formation”, 2000.
- [6] A. Galleotti, S. Goyal, and J. Kamphorst, “Network formation with heterogeneous players”, 2006.
- [7] D. Hojman and A. Szeidl, “Core and periphery in endogenous networks”, 2006.
- [8] A. Galleotti and S. Goyal, “The law of the few”, 2010.
- [9] M. Jackson and A. Watts, “The evolution of social and economic networks”, 2002.
- [10] M. Jackson, “A survey of network formation models: Stability and efficiency”, 2005.

Illustrative Example





$\rho = 0.5$ $\quad u_i\left(\boldsymbol{x}, \boldsymbol{g}\right) = v(16) - 2k$ $\quad u_i\left(\boldsymbol{x}, \boldsymbol{g}\right) = v(64) - 2k$

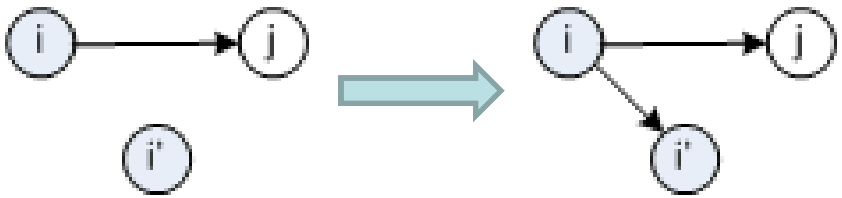
An agent gets higher benefit by consuming content of a wider variety!

Equilibrium with Symmetric Profiles

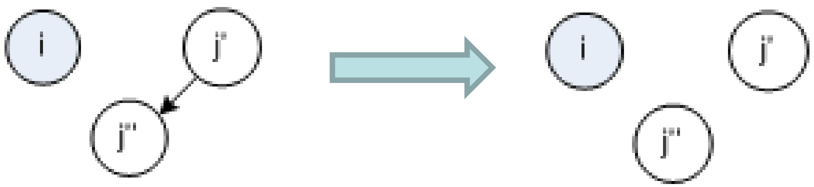
- **Theorem 1.** If a symmetric profile s^* is an equilibrium, then $|N_i(\overline{\boldsymbol{g}}^*)| = |N_j(\overline{\boldsymbol{g}}^*)|, \forall i, j \in N$, i.e. all agents have the same degree.

- *Idea of proof*
  - An agent who accesses more content has less incentive to self-produce
  - All agents have the same production level -> access the same amount of information -> have the same degree

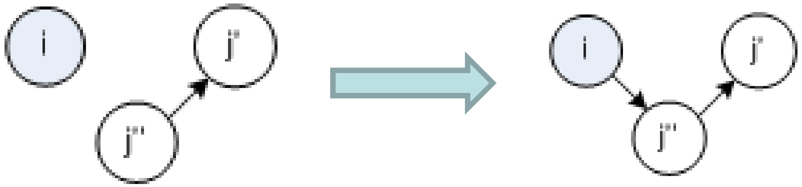
Proof of Lemma 2(ii)



- Agent i is connected with all high producers
- Agent i is not connected with all low producers



- Agent j' does not form links to other low producers



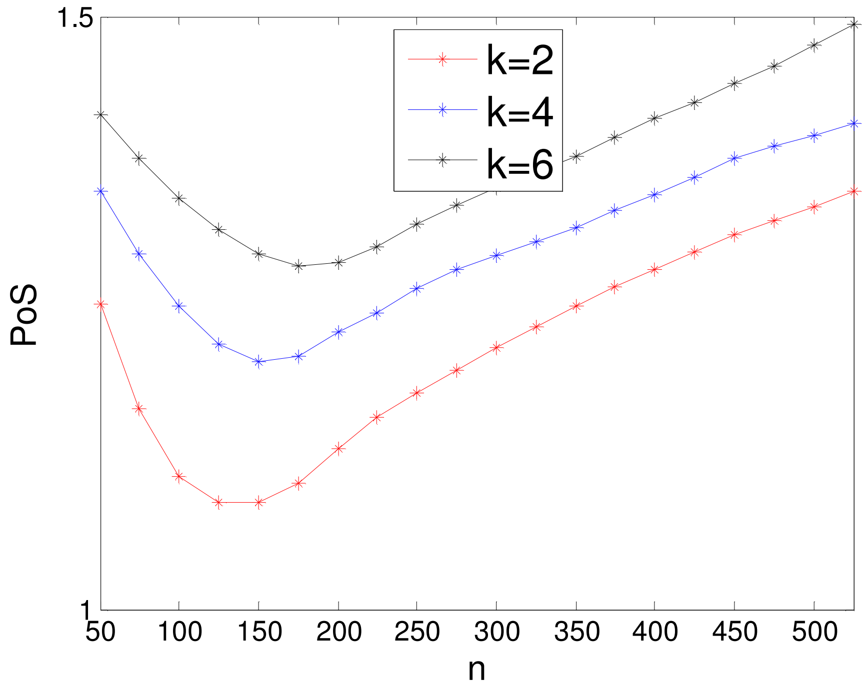
- Each neighbor of agent j' is also the neighbor of agent i

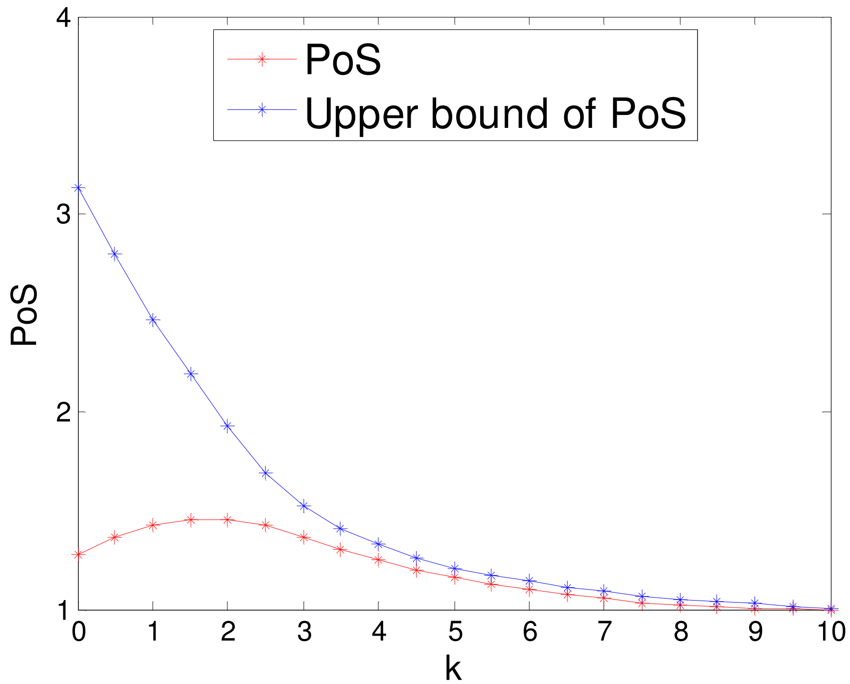
Agent i has a higher effective consumption than agent j'
-> agent i has a smaller production level than agent j'
-> contradiction!

Proof of Theorem 2

- Each agent has a positive production level at equilibrium, which is lower-bounded from 0
- Each link brings negative externality to content production
  - The more neighbor an agent has, the less incentive he has to self-produce content
- The maximum number of neighbors a hub agent wants to maintain is upper-bounded
- More hub agents emerge as the network size grows

Illustration of PoS





- The PoS *monotonically increases* with n when it is sufficiently large.
- The efficiency loss due to the decentralized strategic interactions grows with the network size.

- The upper bound is not tight when *k* is small.
- When *k* increases, the upper bound becomes tight.