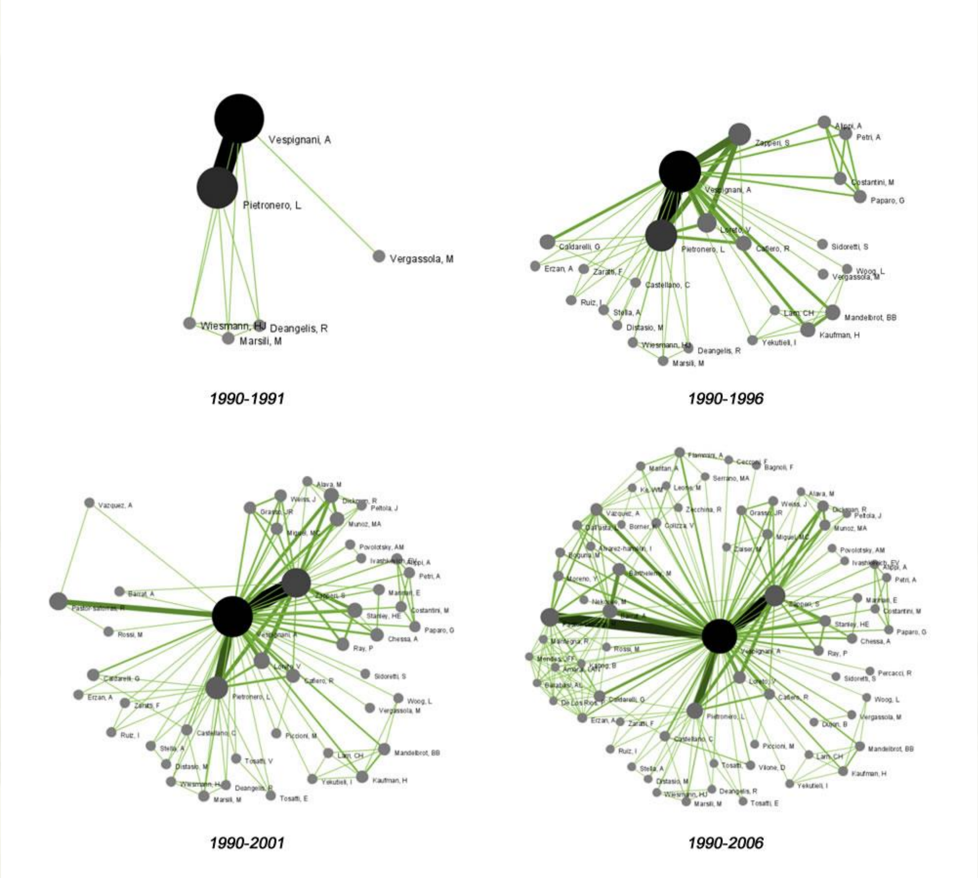
Snowballing Effects in Preferential Attachment: The Impact of The Initial Links

Huanyang Zheng and Jie Wu

Computer and Information Sciences
Temple University

1. Introduction

- Evolving networks
  - Citation networks
  - Internet
  - Social networks
  - P2P networks
  - Road networks
  - Amazon networks

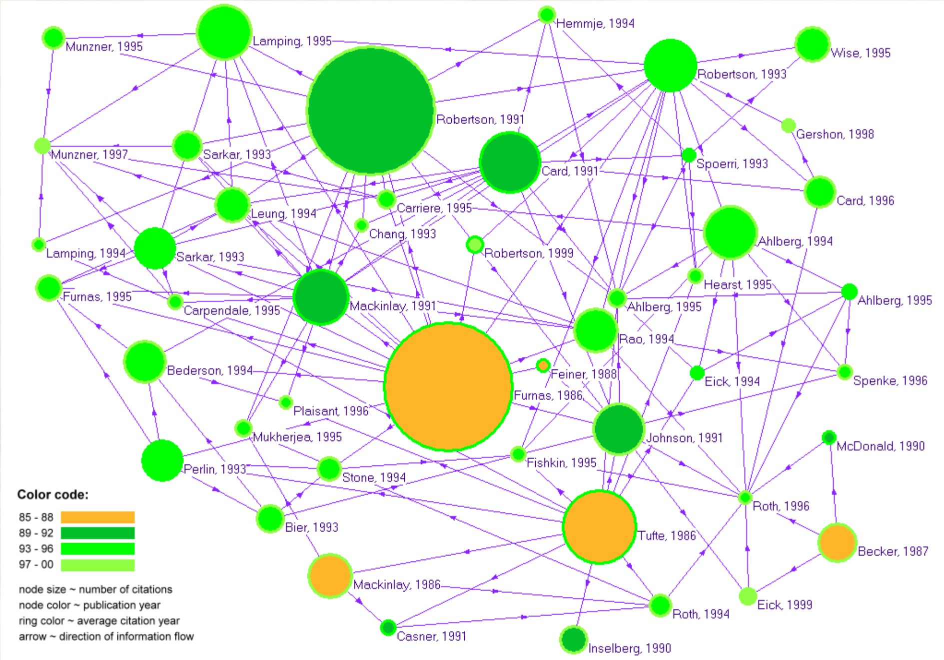


- Node degrees
  - Snowballing effects
  - Rich gets richer

- The impact of initial node degree in snowballs?

2. Preferential Attachment

- Nodes come into the network one by one
  - For example, papers come into the citation network one by one



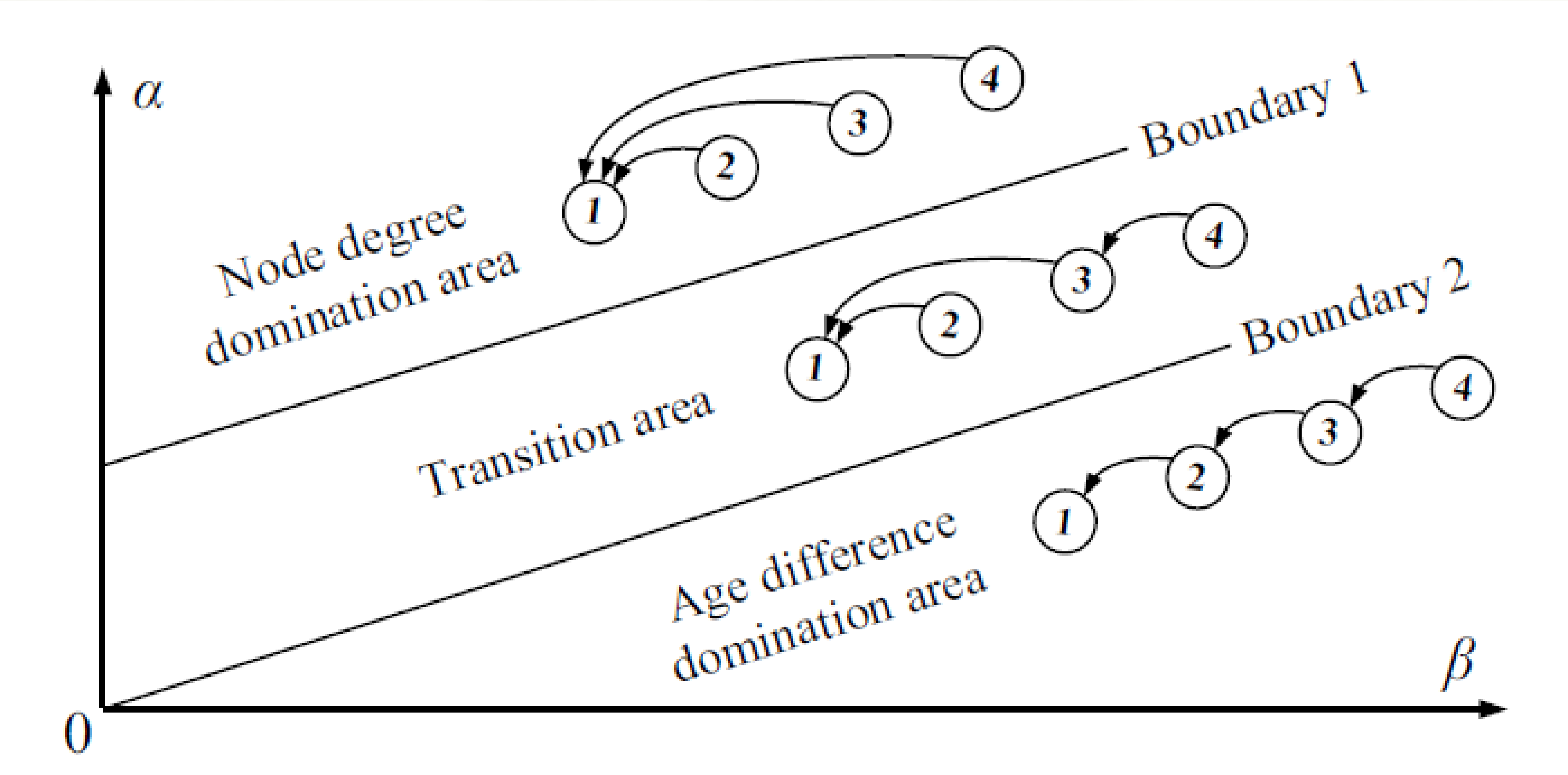
- Newly income node attach to previously existing nodes
- Attach probability depends on degree
  - For example:
  - We prefer to cite papers with more citations

3. Age-sensitive Preferential Attachment

- Age-sensitive attachment
  - For example, we prefer to cite new papers than old papers, if they have the same citations
- Attachment probability
  - Depends on both node degree and age difference
  - Denoted by $d^{\alpha} \cdot \Delta t^{-\beta}$
- Tradeoff on attachment probability
  - Older nodes have larger degrees by time
  - Larger age difference brings a smaller attach probability
  - Node degree dominates? Age difference dominates?

3. Age-sensitive Preferential Attachment

- Percolation phenomenon



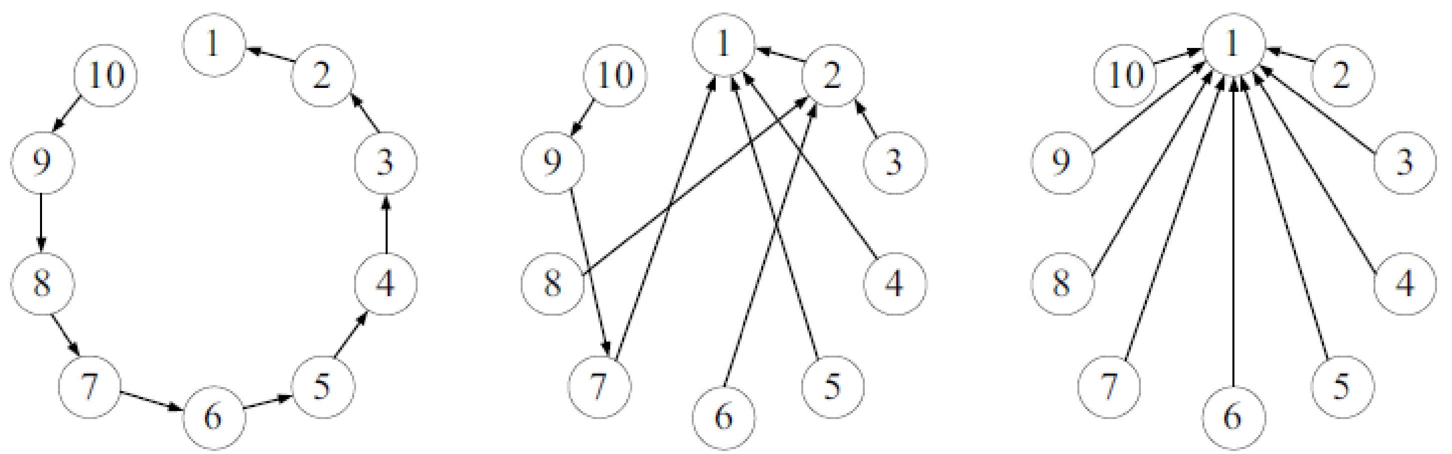
- Dominated factor determines the network structure

3. Age-sensitive Preferential Attachment

- Degree evolving equation:

$$d(s,t+1) = d(s,t) + m \times \frac{d(s,t)^{\alpha}(t-s)^{-\beta}}{\sum_{s=1}^{t} d(s,t)^{\alpha}(t-s)^{-\beta}}$$

- Two boundaries are α = β and α = β + 1.5



(a) $\alpha < \beta$ (b) $\beta \leq \alpha \leq \beta + 1.5$ (c) $\alpha > \beta + 1.5$

4. Node Degree Snowballing

- When $\alpha < \beta$ or $\alpha > \beta + 1.5$
  - Initial node degree is not important
  - Domination of node degree / age difference

- Study the degree snowballing pattern
  - r_i denotes ratio of the additional initial degree to the normal initial degree
  - r_g denotes ratio of increased eventual node degree brought by the additional initial degree
  - Example: how many additional citations can be eventually brought by the initial self-citations?
  - r_g monotonically increases with respect to r_i

4. Node Degree Snowballing

- Degree evolving equation:

$$\frac{\partial d(s,t)}{\partial t} = m \times \frac{d(s,t)^{\alpha}(t-s)^{-\beta}}{\int_1^t d(s,t)^{\alpha}(t-s)^{-\beta}\mathrm{d}s}$$

- Set $\xi = s/t$

$$\frac{d(\xi)^{1-\alpha} - d(1)^{1-\alpha}}{1-\alpha} = \frac{m\int_1^{\xi}\frac{-1}{\xi(1-\xi)^{\beta}}\mathrm{d}\xi}{\int_0^1 d(\xi)^{\alpha}(1-\xi)^{-\beta}\mathrm{d}\xi}$$

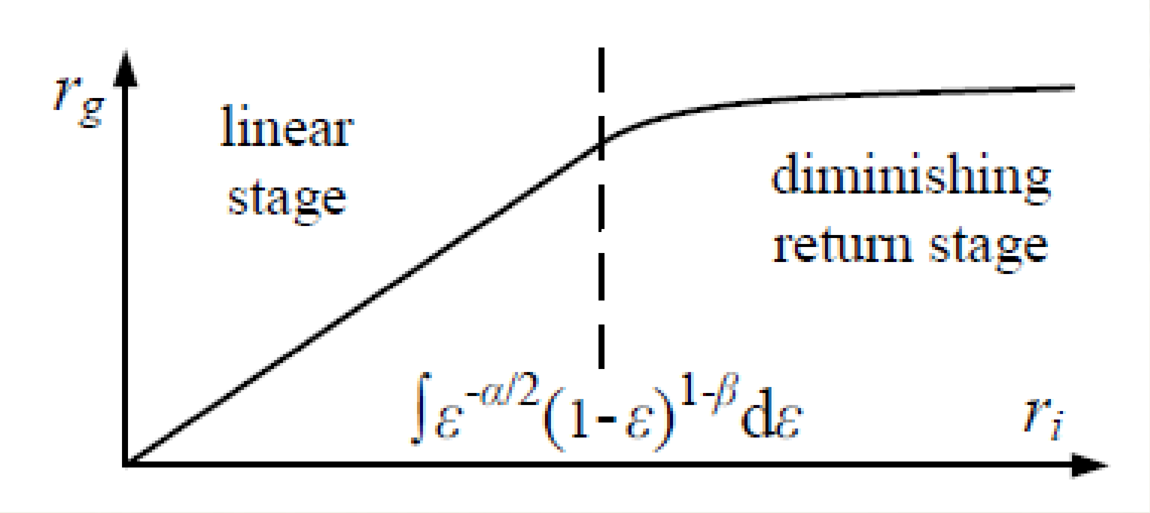
$$\begin{aligned} d'(\xi) &= \left[(m+m')^{1-\alpha} + \frac{m\int_1^{\xi}\frac{-1}{\xi(1-\xi)^{\beta}}\mathrm{d}\xi}{\int_0^1 d(\xi)^{\alpha}(1-\xi)^{-\beta}\mathrm{d}\xi}\right]^{\frac{1}{1-\alpha}} \\ &\approx m \times [(1+\frac{m'}{m})^{1-\alpha} + (1-\alpha)C(\alpha,\beta,\xi)]^{\frac{1}{1-\alpha}} \end{aligned}$$

4. Node Degree Snowballing

* Result for the degree snowballing pattern:

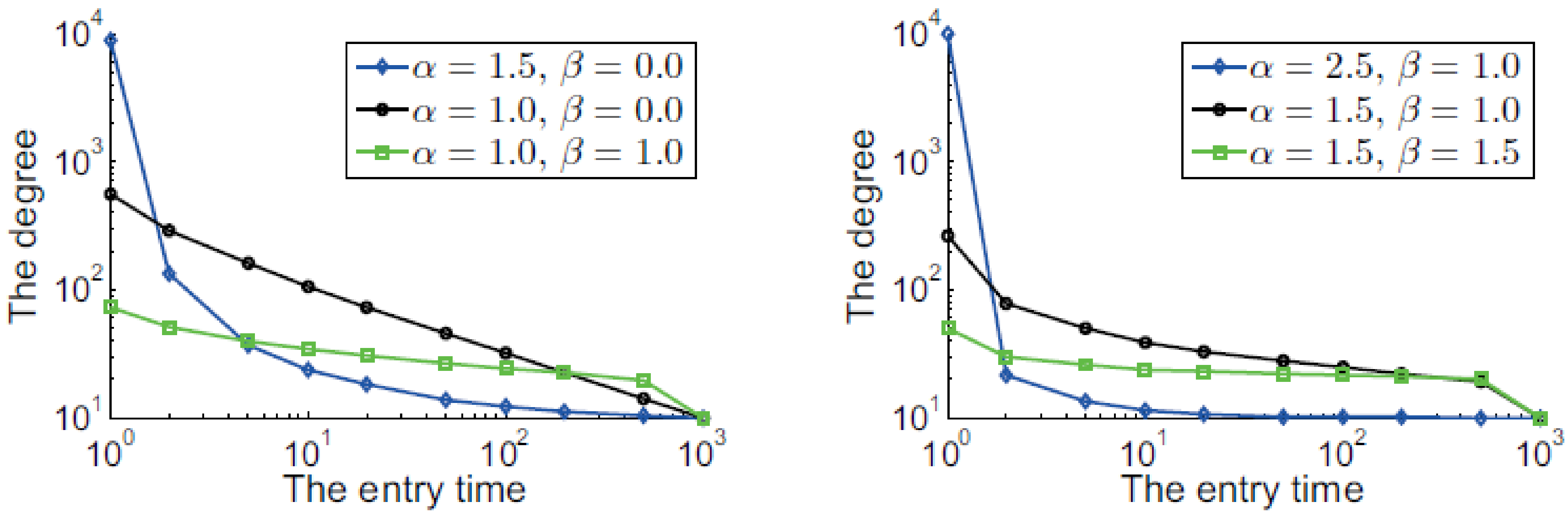
$$C(\alpha, \beta, \xi) = \frac{\int_1^{\xi} \frac{-1}{\xi(1-\xi)^{\beta}} d\xi}{\int_0^1 \xi^{-\alpha/2}(1-\xi)^{-\beta} d\xi}$$

$$r_g = \left[\frac{(1+r_i)^{1-\alpha}+(1-\alpha)C(\alpha,\beta,\xi)}{1+(1-\alpha)C(\alpha,\beta,\xi)}\right]^{\frac{1}{1-\alpha}} - 1$$



5. Experiments

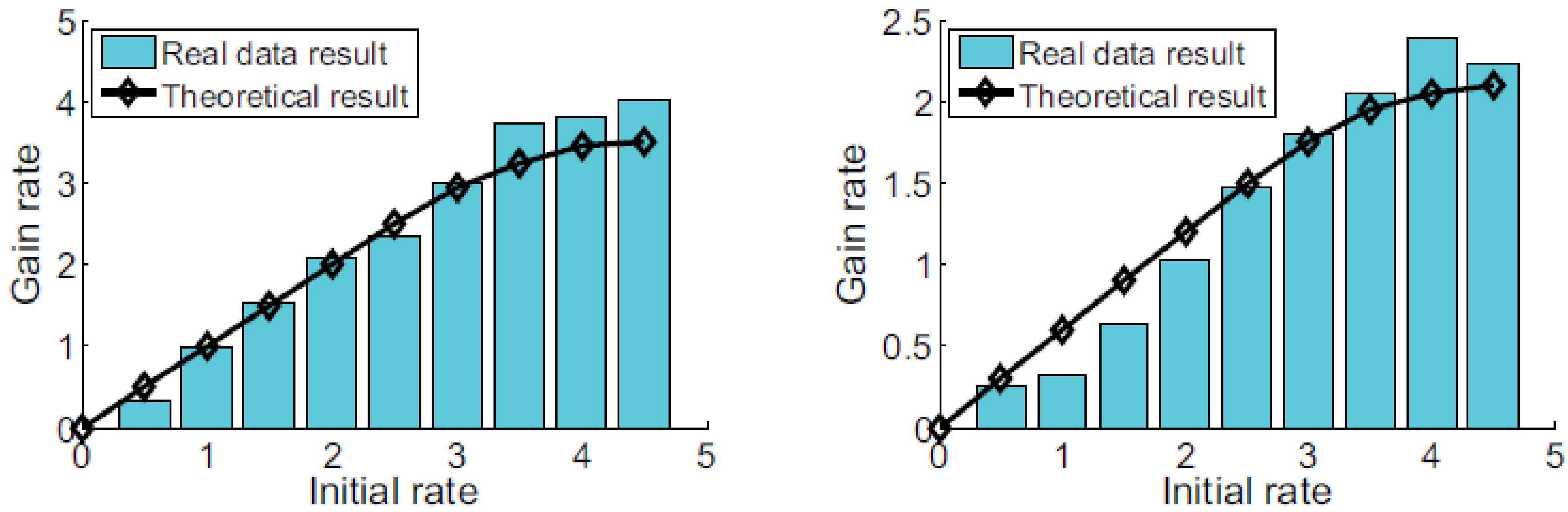
* Simulations on the percolation
  * Boundaries of α = β or α = β + 1.5



(a) Scenario 1. (b) Scenario 2.

5. Experiments

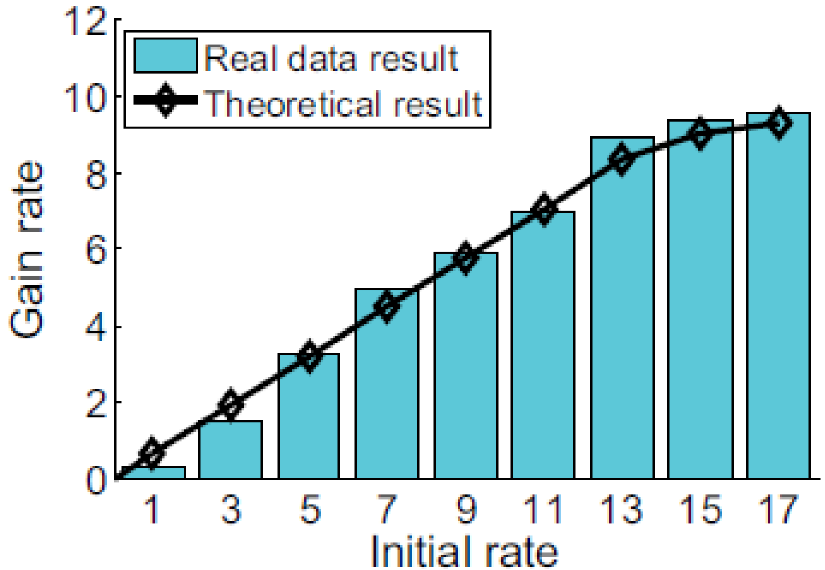
- Real data-driven experiments
  - Arxiv high energy physics phenomenology citation network
  - Include papers published from January 1993 to April 2003
  - Include 34,546 papers with 421,578 citations



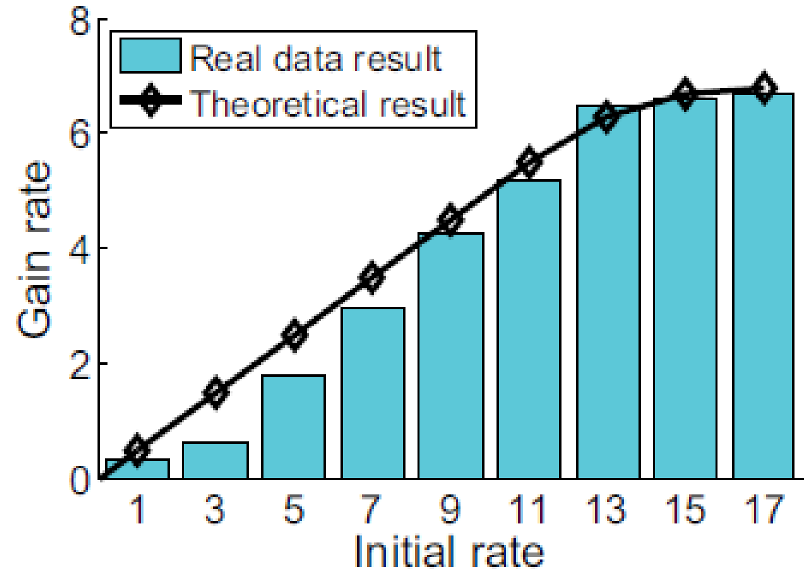
(a) Papers published in 1995.   (b) Papers published in 1998.

5. Experiments

- Real data-driven experiments
  - Flickr photo sharing network
  - Include users from November 2006 to May 2007
  - Include 167,527 users and 526,874 following relationships



(a) Users entered on Monday.



(b) Users entered on Wednesday.

6. Conclusion

- Node degree snowballing effects
  - Evolving networks
- Age-sensitive preferential attachment model
  - Node degree and age difference
- Impact of the initial links
  - Linear stage and diminishing return stage

End

Q & A