

# Cyworld Growth: Sheer Volume vs User Activity

2007.12.7. Friday  
Online Social Networks Workshop  
MSR Cambridge

Hawoong Jeong and Sue Moon  
KAIST

in collaboration with  
Yong-Yeol Ahn, Hyunwoo Chun,  
Seungyeop Han, Yong-Ho Kim, Haewoon Kwak

# Overview

- Cyworld Growth in Sheer Volume
  - Tipping point vs Saturation
- Cyworld Guestbook Log
  - Growth in activity
- Cyber-only relationship
  - Technology-assisted relationship management
- Evolution of social networking services
  - Ongoing/Future work

# CyWorld

- Largest SNS in South Korea
  - Started in September 2001
  - 10 million users in 2004
  - Over 20 million users out of 49 million population
- Front runner of many features
  - Friend (*il-chon*) relationship
  - Guestbook
  - Testimonial (*il-chon-pyung*)
  - Photos - scraps
  - Avatar in cyber home

# My CyWorld "Mini-Homepage"

http://minihp.cyworld.nate.com - 싸이월드>문수복님의 미니홈피 - Microsoft Internet Explorer

TODAY 0 | TOTAL 535



제 블로그 주소는 게시판에 공지되어 있습니다. 이 메일은 sbmoon@cs.kaist.ac.kr로 주시면 됩니다!

EDIT HISTORY

문수복 (1965.9.10(+))  
sbmoon@cs.kaist.ac.kr

일촌 클러

★일촌 파도타기

## 수~ 의 미니홈피

사생활보호설정 | 도메인신청

최근 2주간 새 게시물이 없습니다.  
소식이 뜸한 친구에게 마음의 한마디 남겨주세요~

사진첩 0/18	게시판 0/2
방명록 1/35	

Home

- 프로필
- 사진첩
- 게시판
- 방명록
- 관리

미니룸 Storyroom default



보로 선물 도착! 손님방에 있습니다.

저는 주로 블로그 합니다.

내 미니룸 | 미니미설정 | 미니룸설정 | 답글수 | 메인설정

What friends say 한마디로 표현해보세요 | 1촌평관리

- 세상에서 제일 에너지가 많은 친구!!^^ (어여쁜 친구 진취연)

싸이월드 홈

액티브 | 프렌들리 | 센서티브

심볼을 자랑해 주세요

스크랩 | 즐겨찾기

차용 중인 아이템 | 선물하기 | 소망상자

음악등록하기

랜덤 미니홈피 가기 | 로그아웃

@siesta

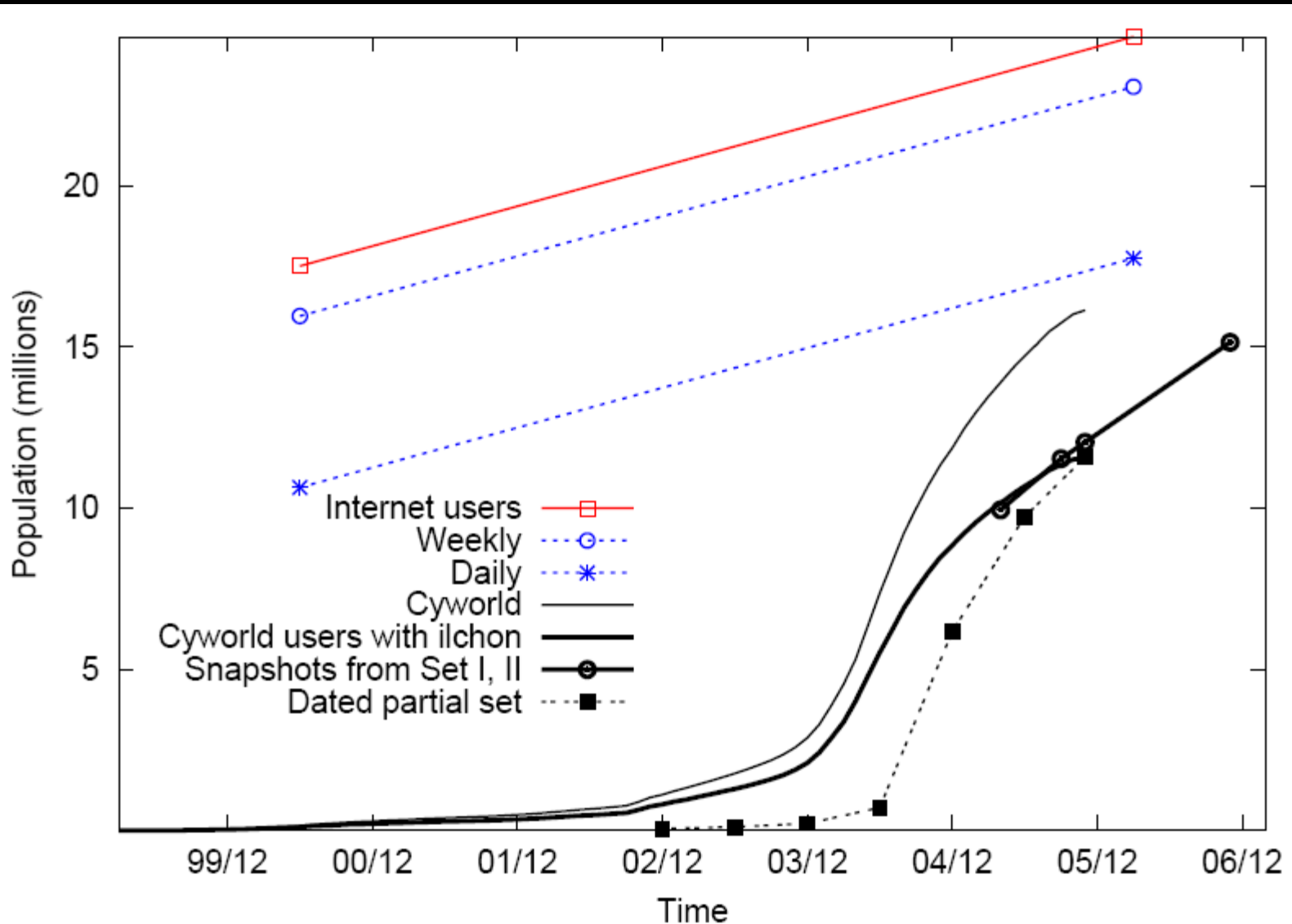
인터넷

**LET'S BEGIN WITH  
CYWORLD GROWTH IN VOLUME**

# Cyworld Data Sets

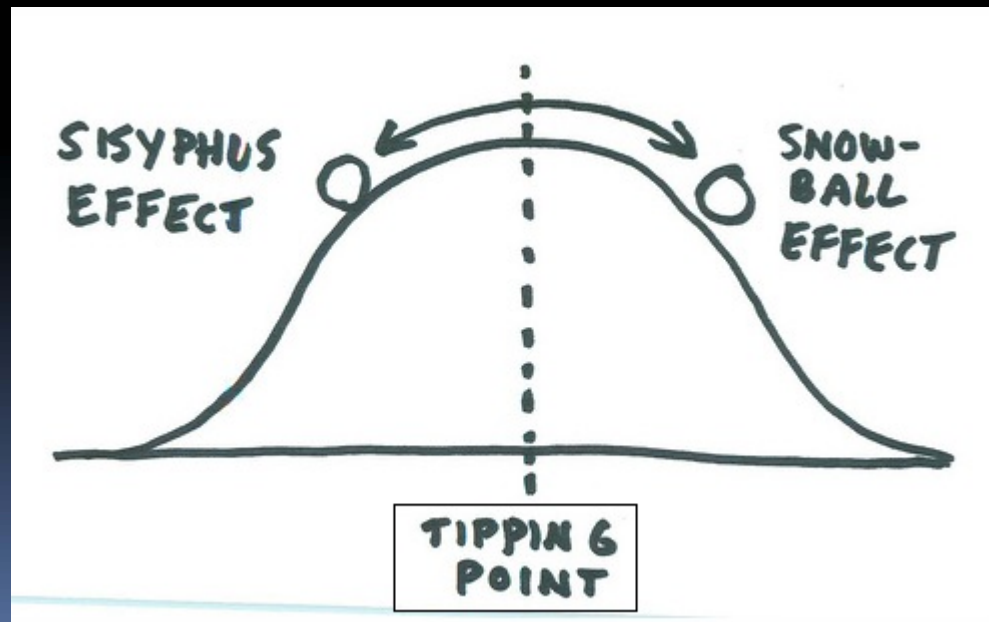
- 4 snapshots of complete ilchon relationships
  - April, September, November of 2005
  - November of 2006
- Guestbook logs
  - June 2003 to October 2005
  - (writer, guestbook owner, timestamp)
  - 8.4 billion messages from 17 million users

# Cyworld Growth in Sheer Volume



# Tipping Point

- Definition
  - Derived from “White Flight” social phenomenon
  - Point of “phase transition” due to “peer pressure”

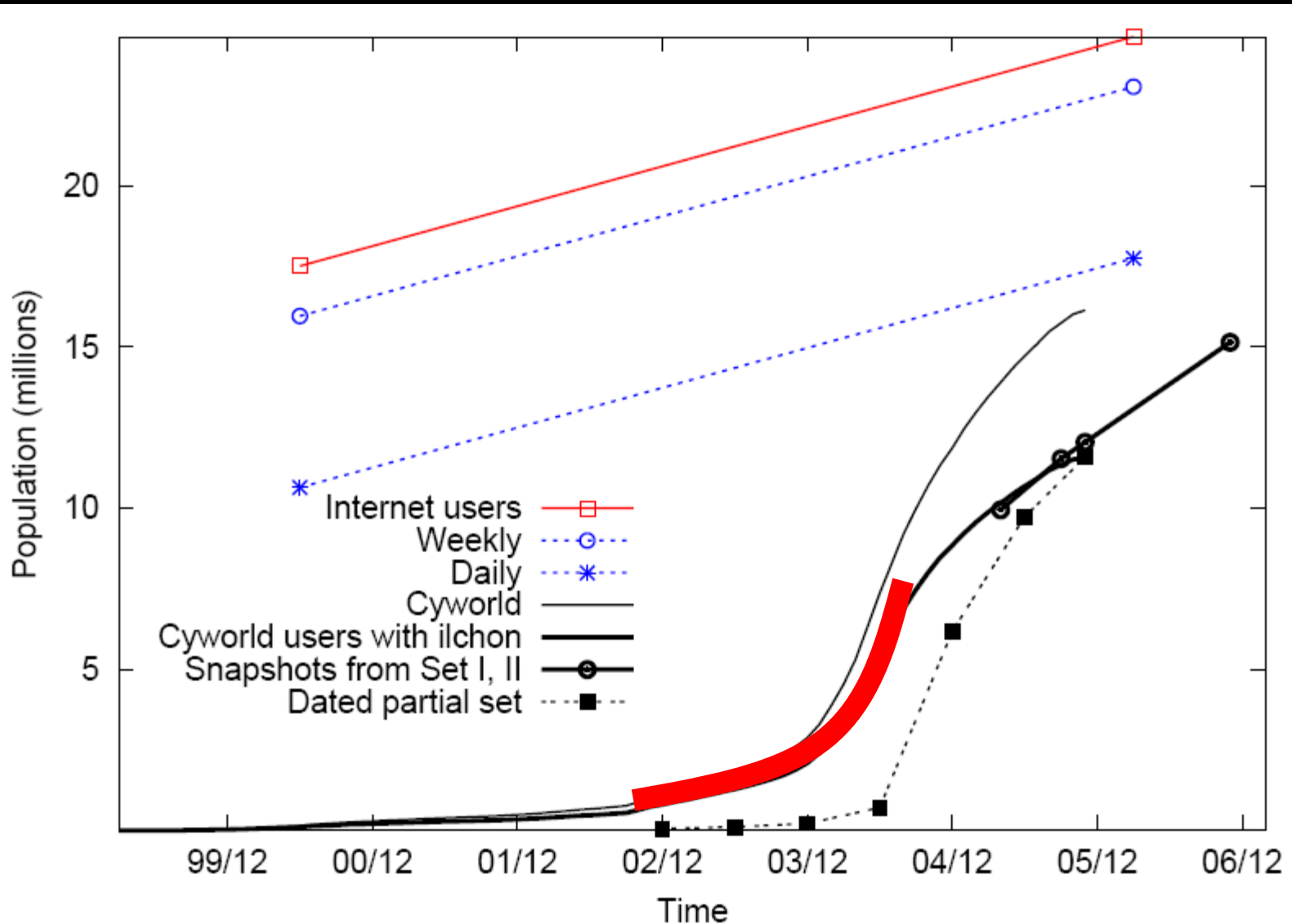




# Existing Tipping Point Models

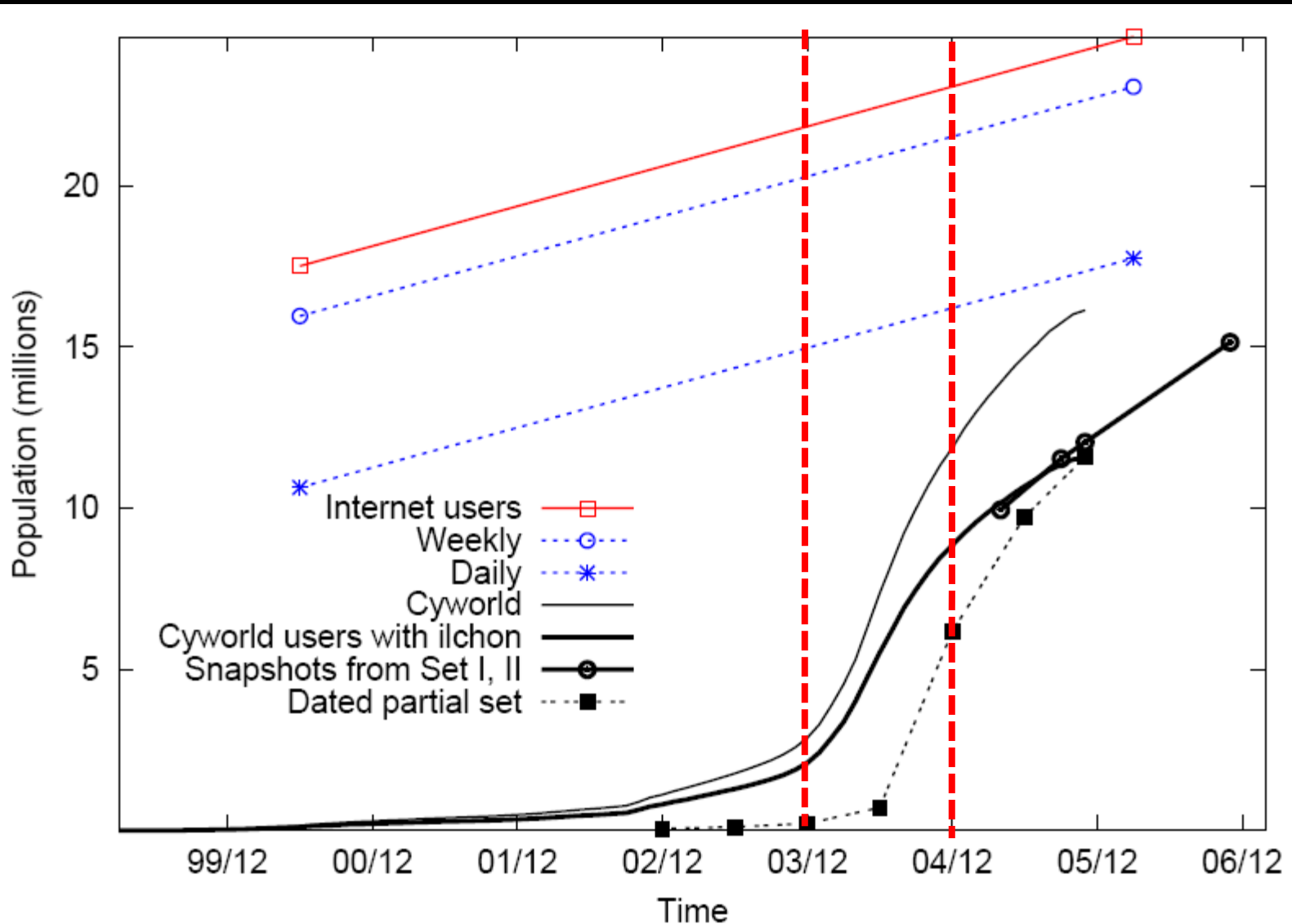
- Input parameters
  - Peer pressure
- Insufficient for our data
  - Lack of resistant factors
    - Delay in adoption
    - Individual variation in adoption
  - Still only explains our data partially

# Cyworld Growth in Sheer Volume

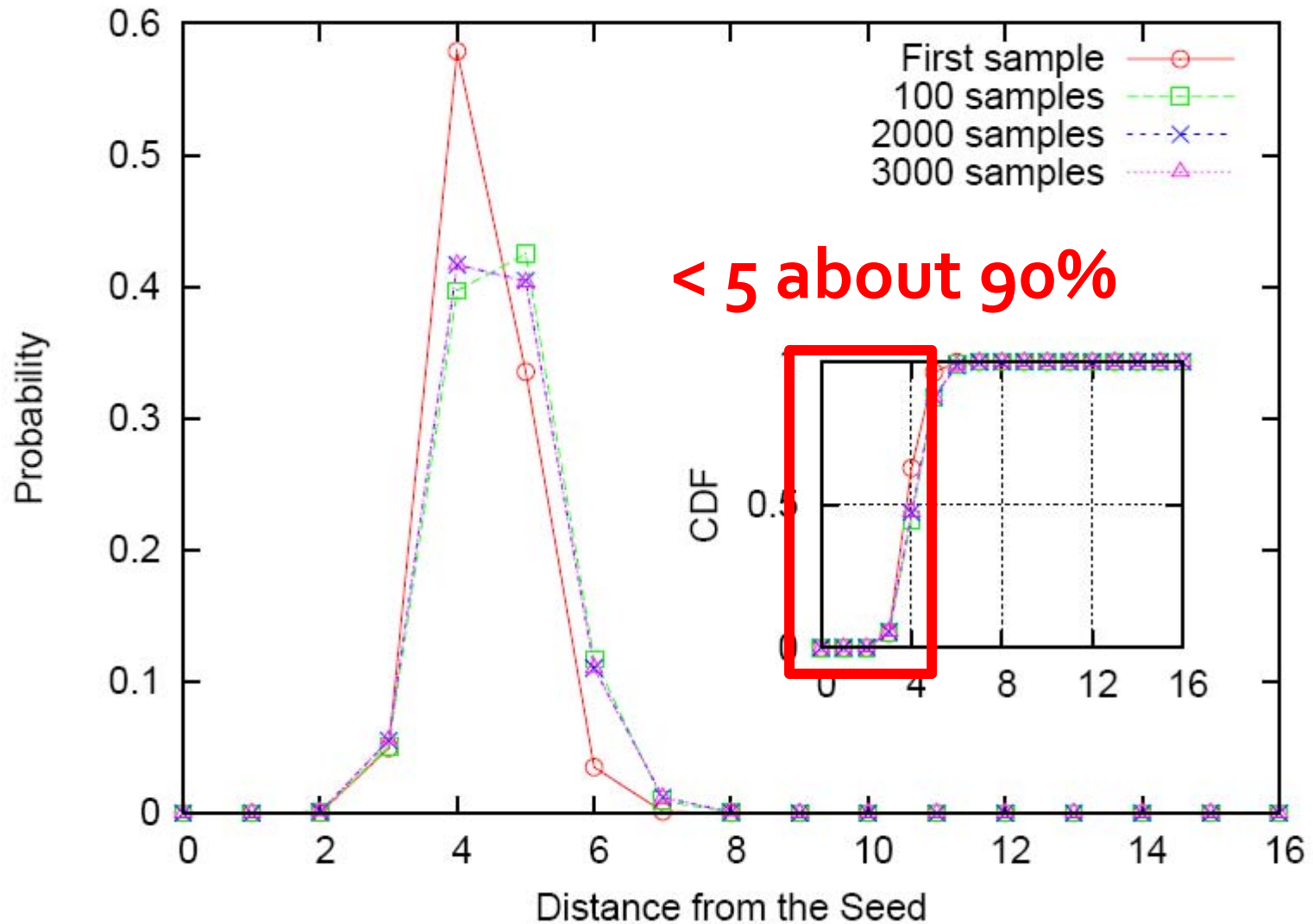


**HOW LONG CAN CYWORLD  
MAINTAIN EXPONENTIAL GROWTH?**

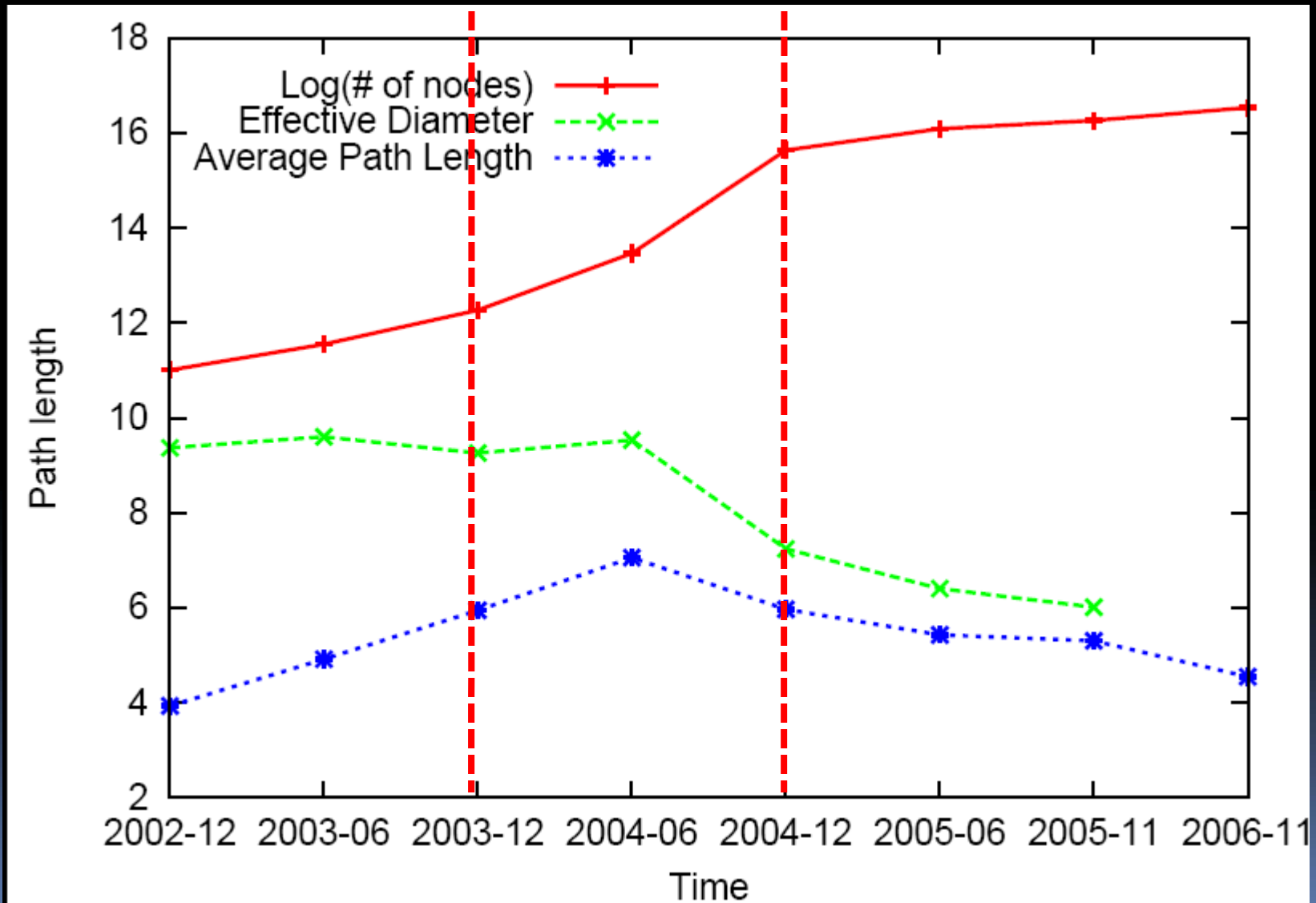
# Cyworld Growth in Sheer Volume



# Average Path Length

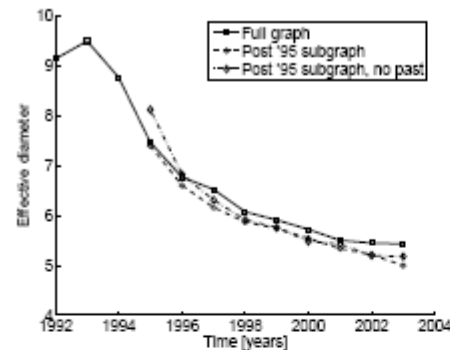


# Evolution of Avg Path Length

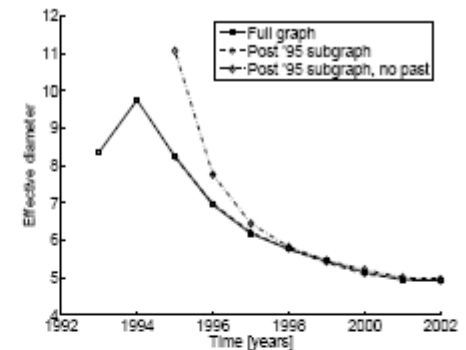


# Shrinking Diameter [KDD2005]

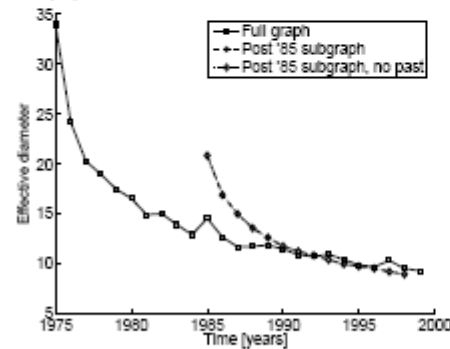
- Constant or slowly growing diameter assumption  
=> Not true



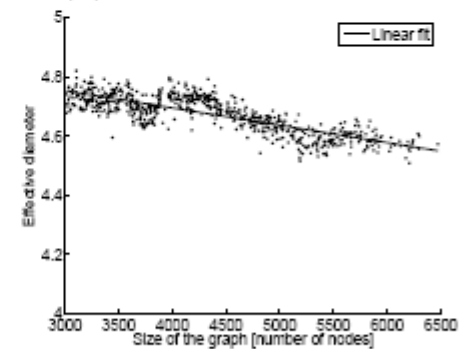
(a) arXiv citation graph



(b) Affiliation network



(c) Patents



(d) AS

Figure 3: The effective diameter over time.

# Cyworld Saturation

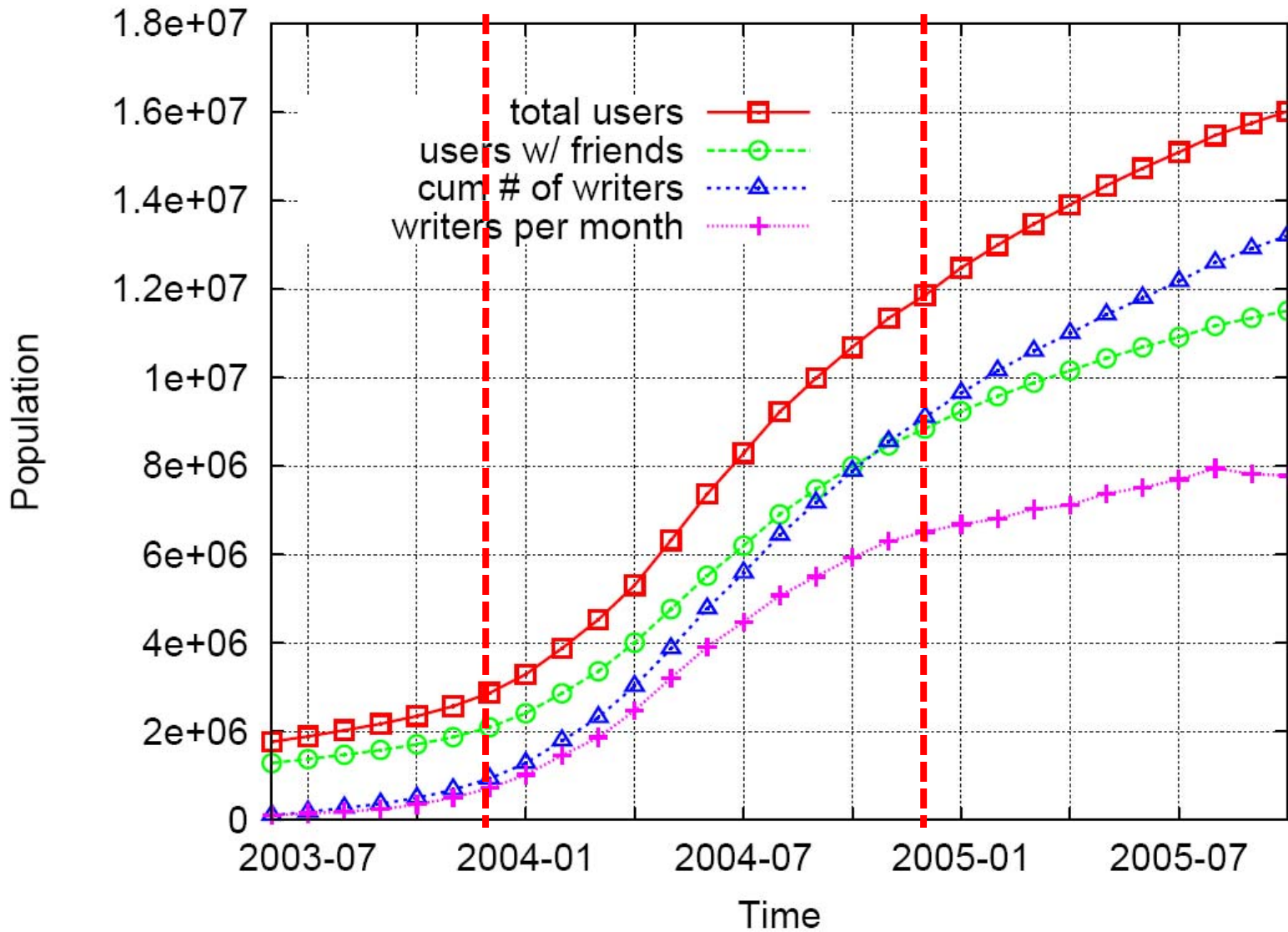
- Volume-wise
  - Exponential growth from 2003 to 2004
  - Continues to grow in 2005 and on
- Topologically
  - Network has started to “densify”
  - No network growth model fits SNS
    - Delay in ilchon building
    - Bounded population



**SO FAR WE HAVE SEEN ONLY  
GROWTH IN SHEER VOLUME**

**NOW IN-DEPTH LOOK AT  
CYWORLD USER ACTIVITY**

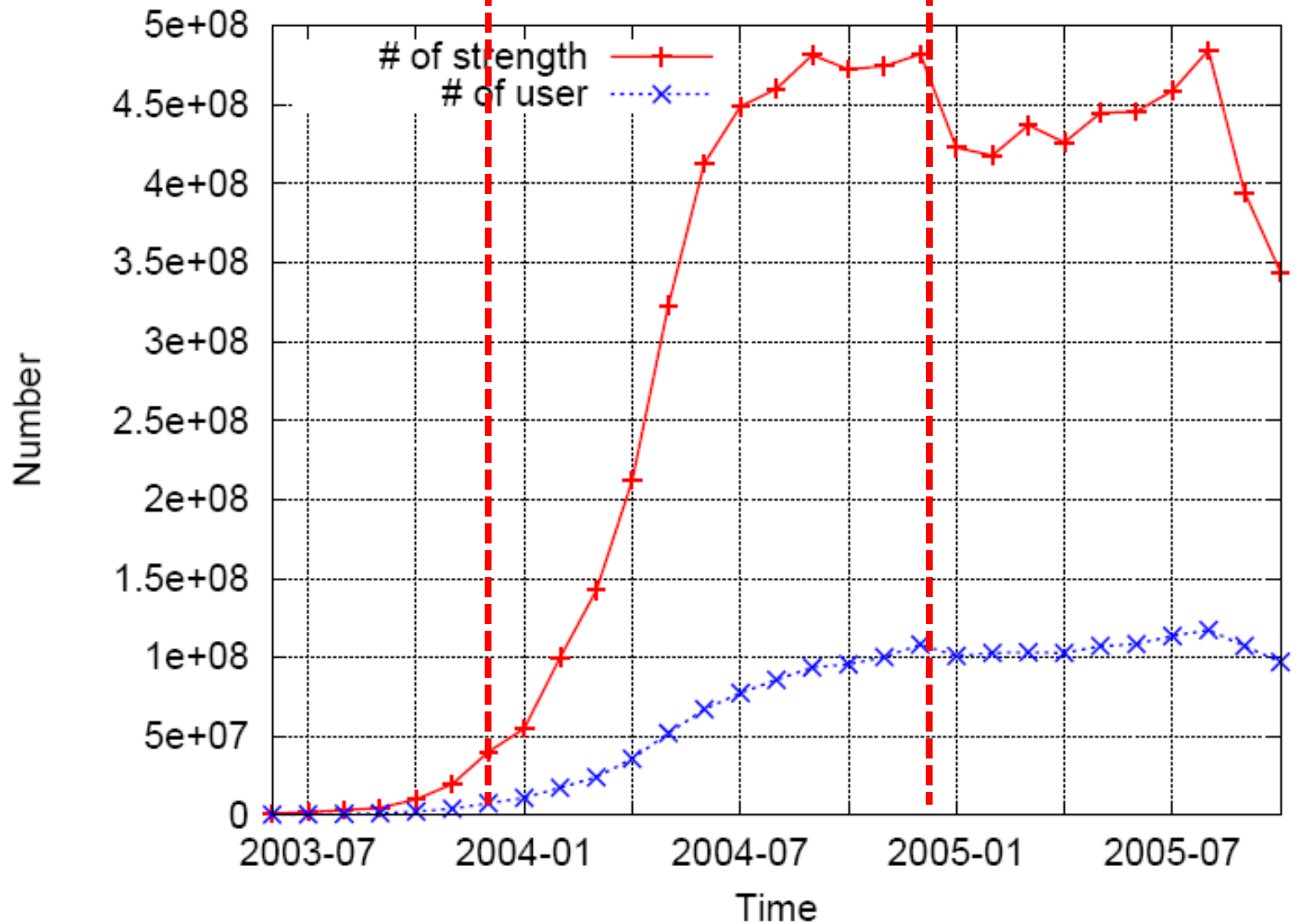
# Cyworld Guestbook Activity



# Terms for Activity

- Node Degree
  - # of guestbooks a user visited and wrote comments
- Node Strength
  - # of total comments written

# Evolution of Node Strength/Degree



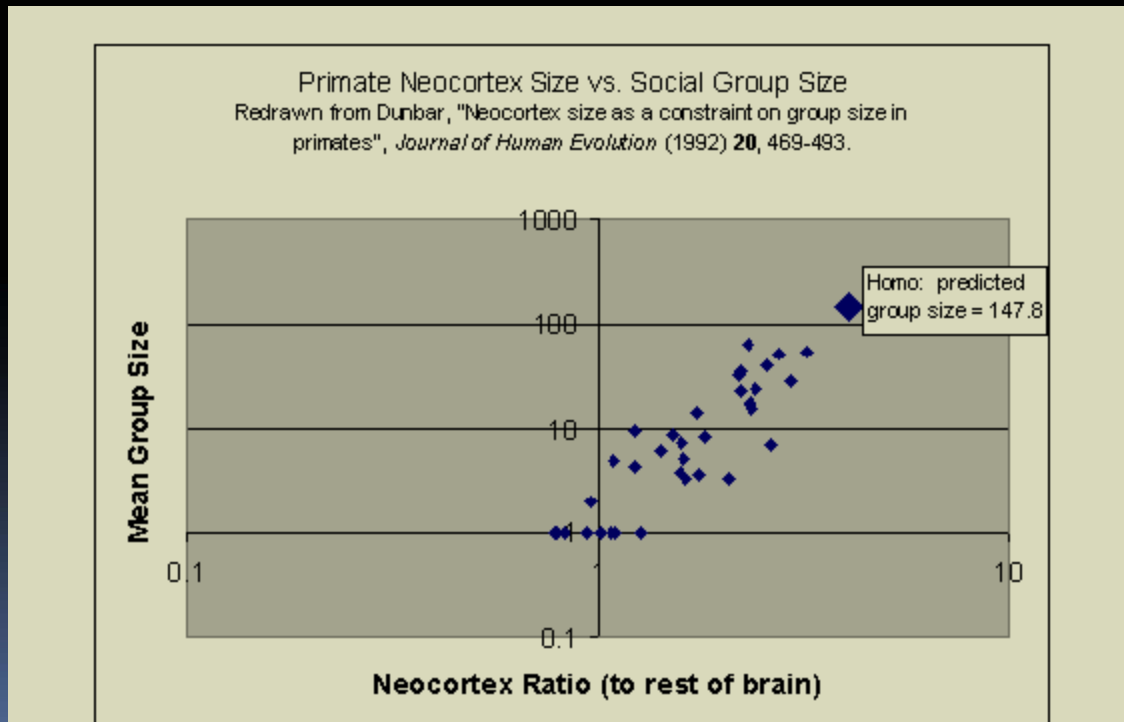
# Cyworld Activity

- Activity data
  - True representation of “livelihood” of Cyworld
- Saturation in activity more prevailing

# OFF-LINE VS CYBER-ONLY RELATIONSHIPS

# Dunbar's Number = 150

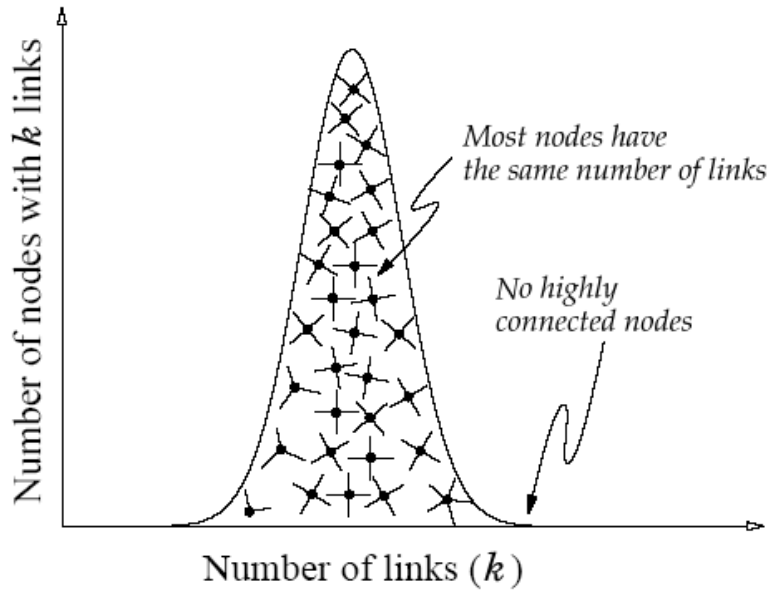
- Theoretical Max #
  - of individuals with whom a set of people can maintain a social relationship



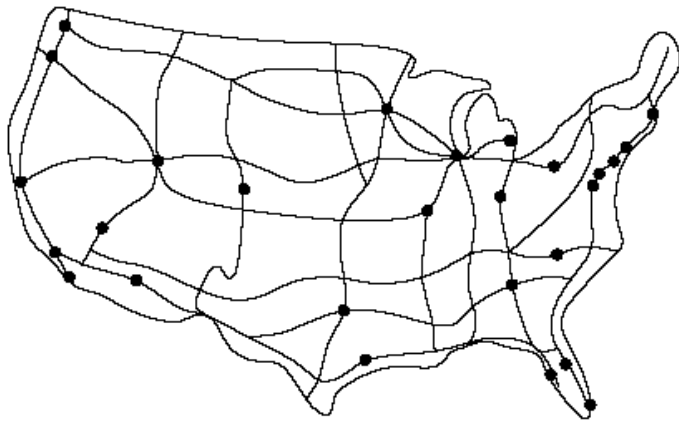
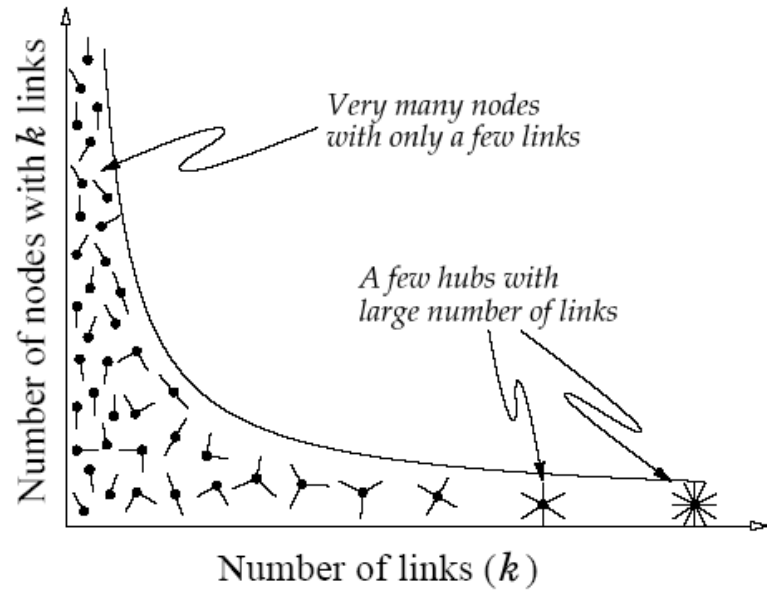


# Two kinds of networks!

## Bell Curve



## Power Law Distribution

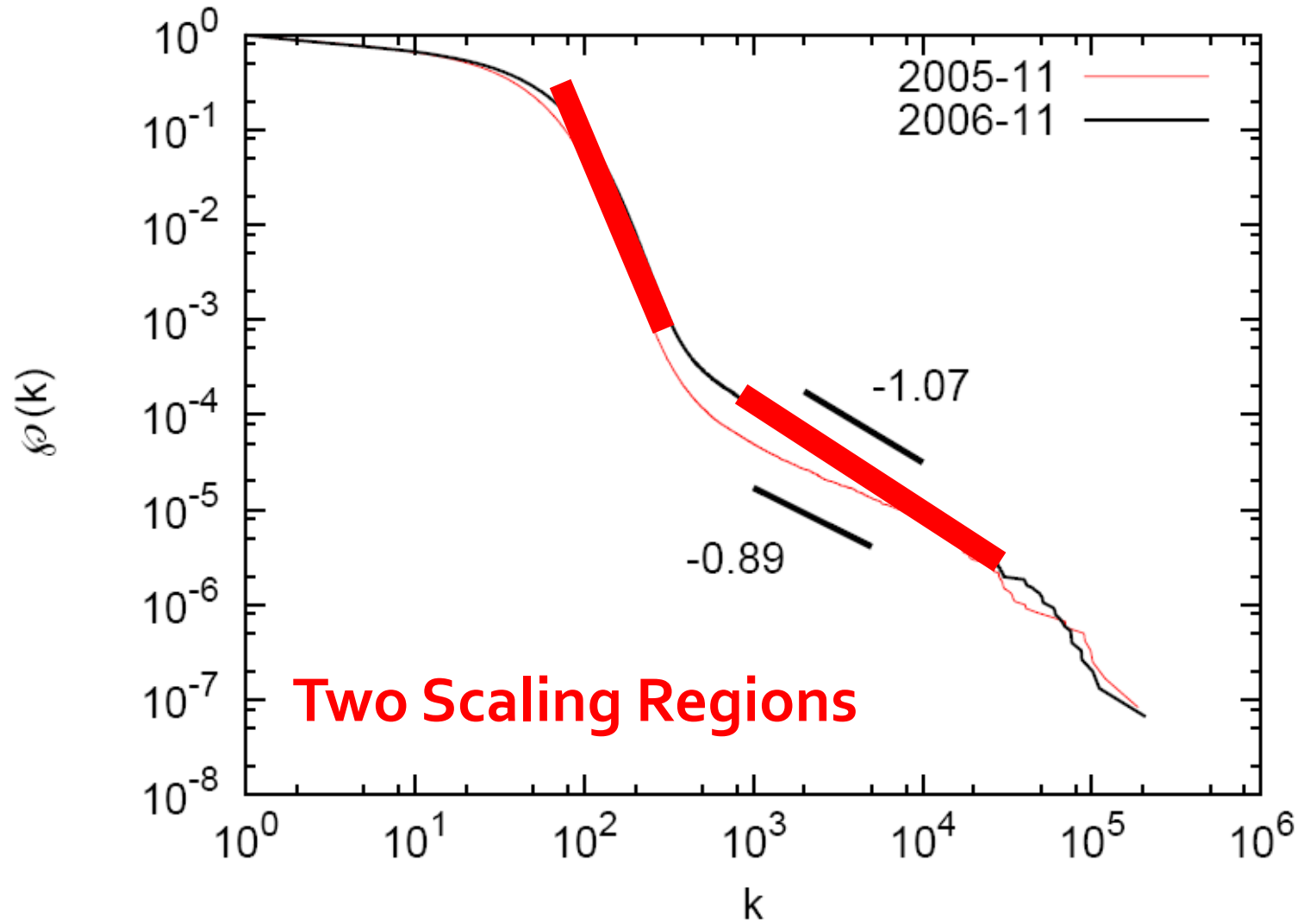


Small-World Network



Scale-Free Network

# Ilchon Degree Distribution



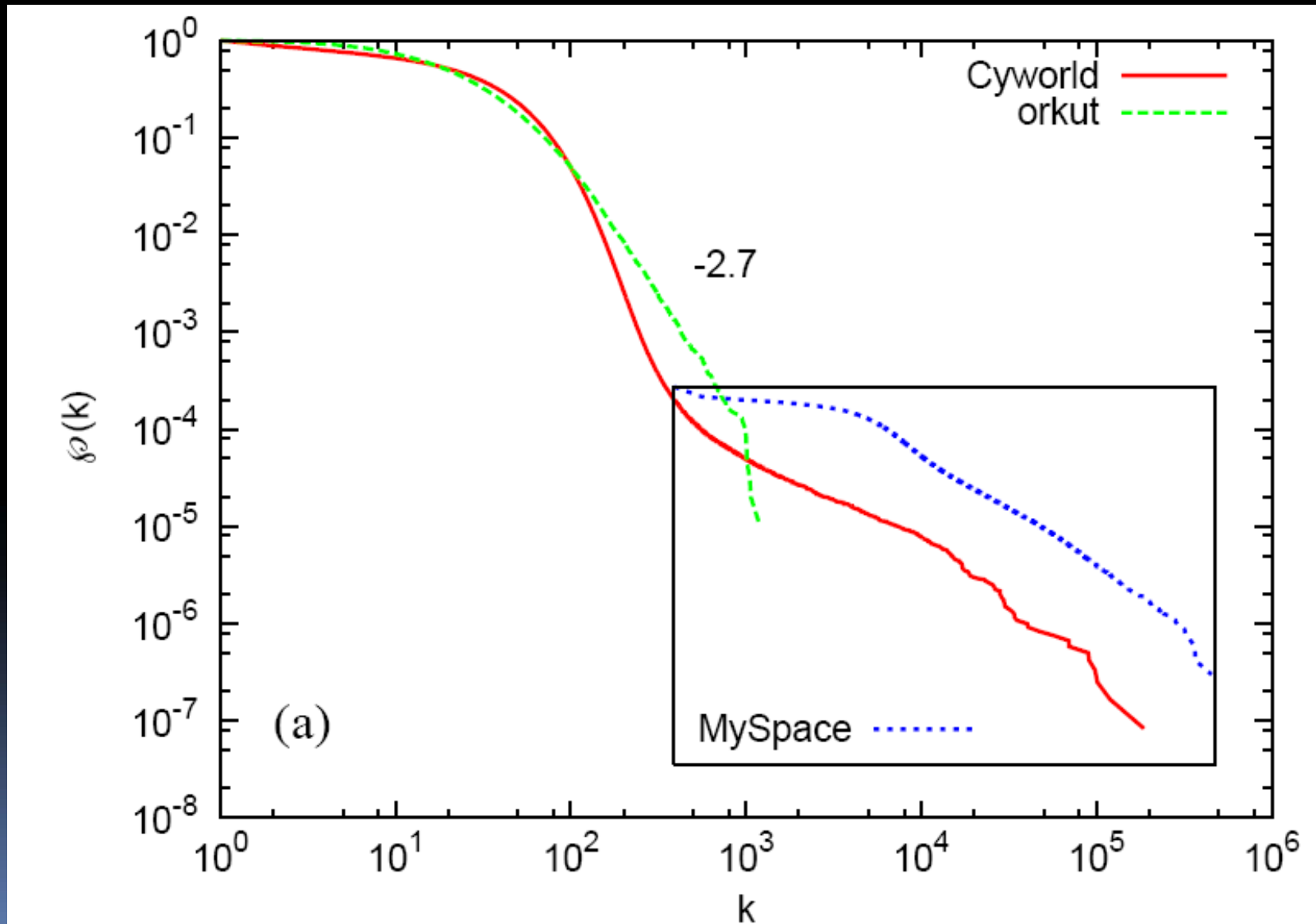
# MySpace Data Set

- Largest in the world
  - Began in Jul 2003
  - Has 200 million by Sep 2007
- Snowball sampled
  - During Sep/Oct 2006
  - Random seed to 100,000 users
  - About 23% of users had friend list hidden

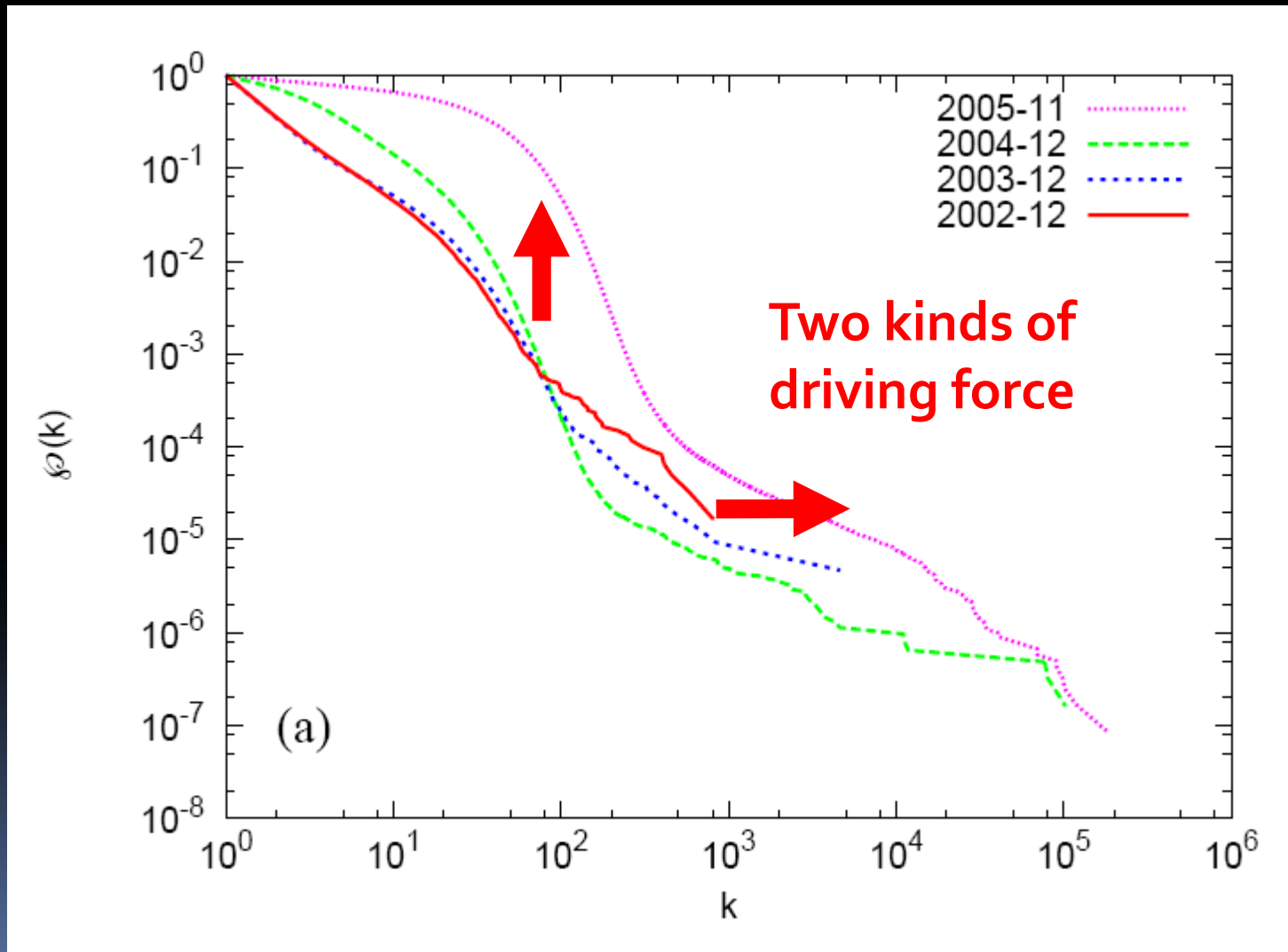
# Orkut Data Set

- Google SNS
  - Began in Sep 2002
  - Became official Google service in Jan 2004
  - Began as invitation-only; open now
  - Has 67 million users in August 2007
  
- Snowball sampled
  - During Jun to Sep 2006
  - 100,000 users

# Cyworld, MySpace, and Orkut



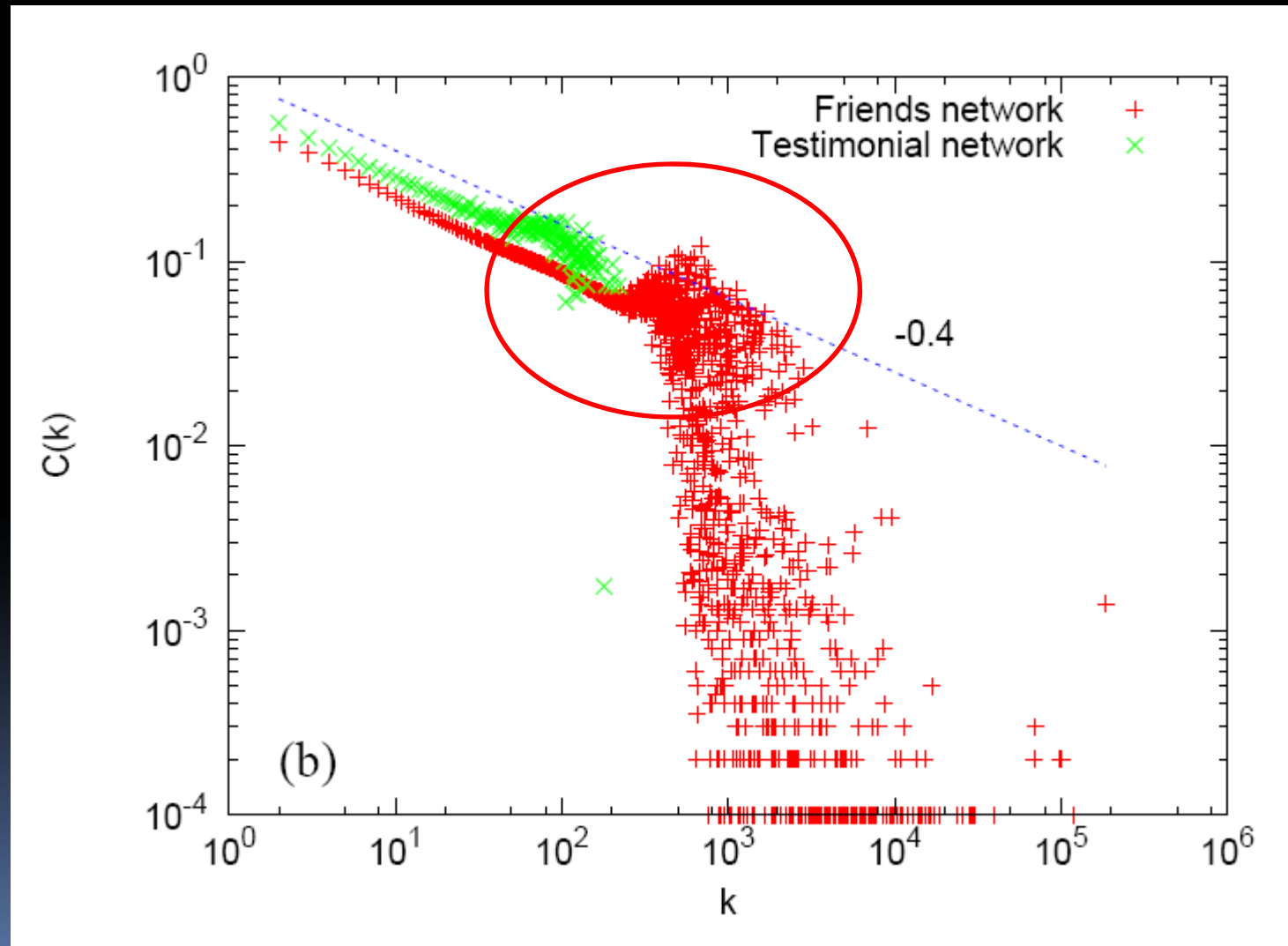
# Evolution of Degree Distributions



# Clustering Coefficient $C(k)$

- # of existing links / # of all possible links between a link's adjacent neighbors
- Close to 1, close to a mesh

# Clustering Coefficient Distribution

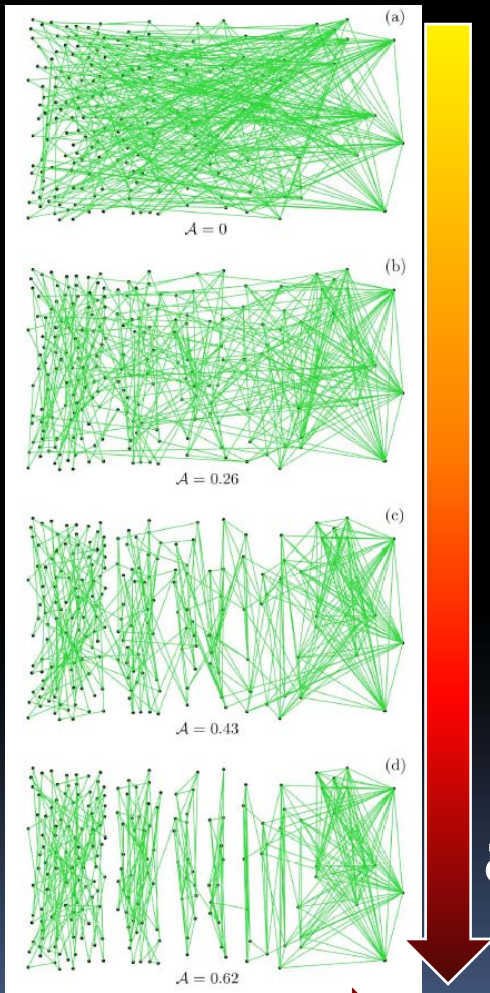




# Degree correlation $k_{nn}$

- Degree  $k$  ~ mean degree of adjacent neighbors of nodes with degree  $k$
- Assortativity: characteristic of  $k_{nn}$  distribution

# Assortative Mixing



“Social”

“non-social”

assortative

degree

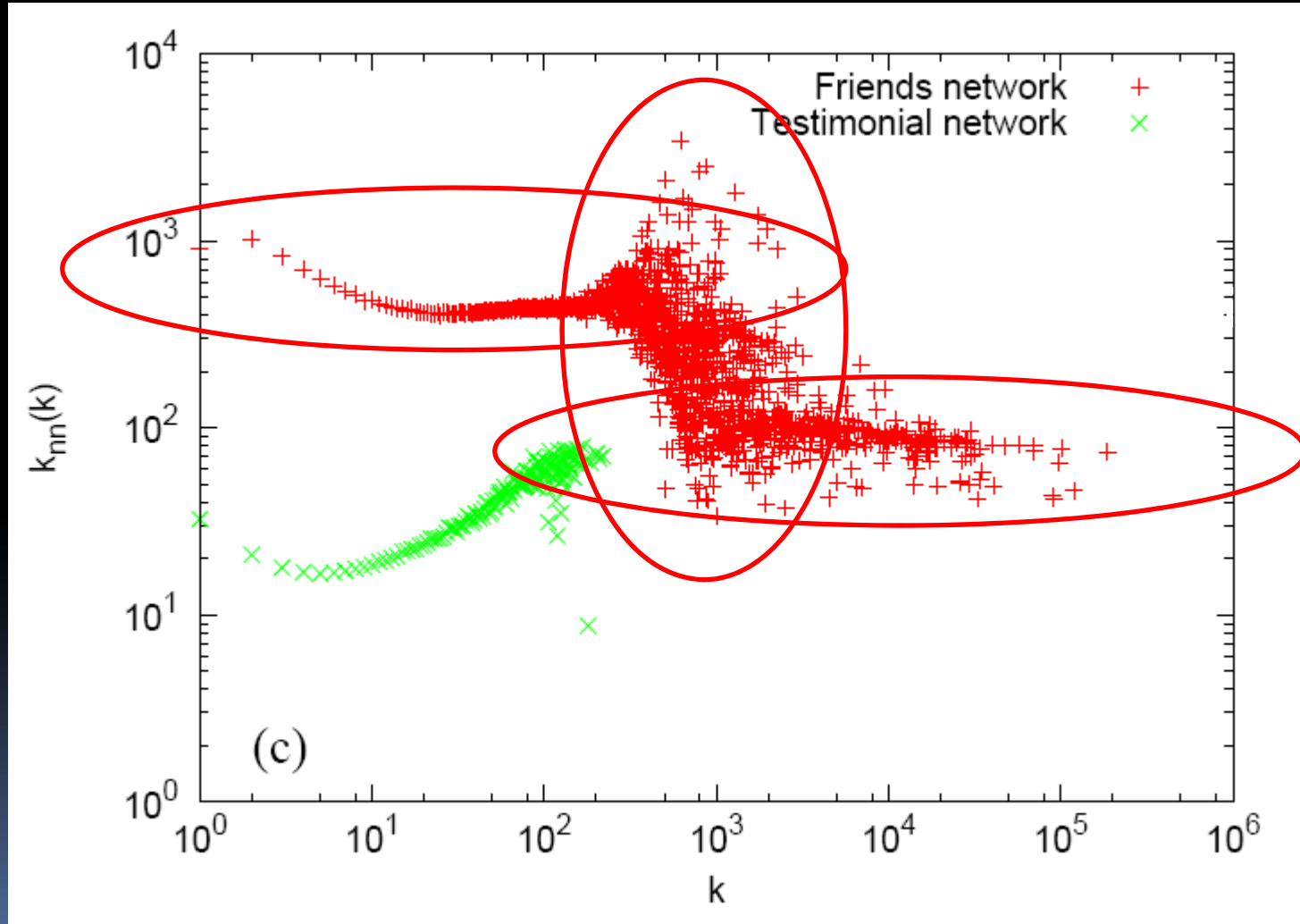
Network	$n$	$r$
Physics coauthorship (a)	52 909	0.363
Biology coauthorship (a)	1 520 251	0.127
Mathematics coauthorship (b)	253 339	0.120
Film actor collaborations (c)	449 913	0.208
Company directors (d)	7 673	0.276
Internet (e)	10 697	-0.189
World-Wide Web (f)	269 504	-0.065
Protein interactions (g)	2 115	-0.156
Neural network (h)	307	-0.163
Marine food web (i)	134	-0.247
Freshwater food web (j)	92	-0.276
Random graph (u)		0
Callaway <i>et al.</i> (v)		$\delta/(1 + 2\delta)$
Barabási and Albert (w)		0

M. E. J. Newman, Phys. Rev. Lett. 89, 208701 (2002)

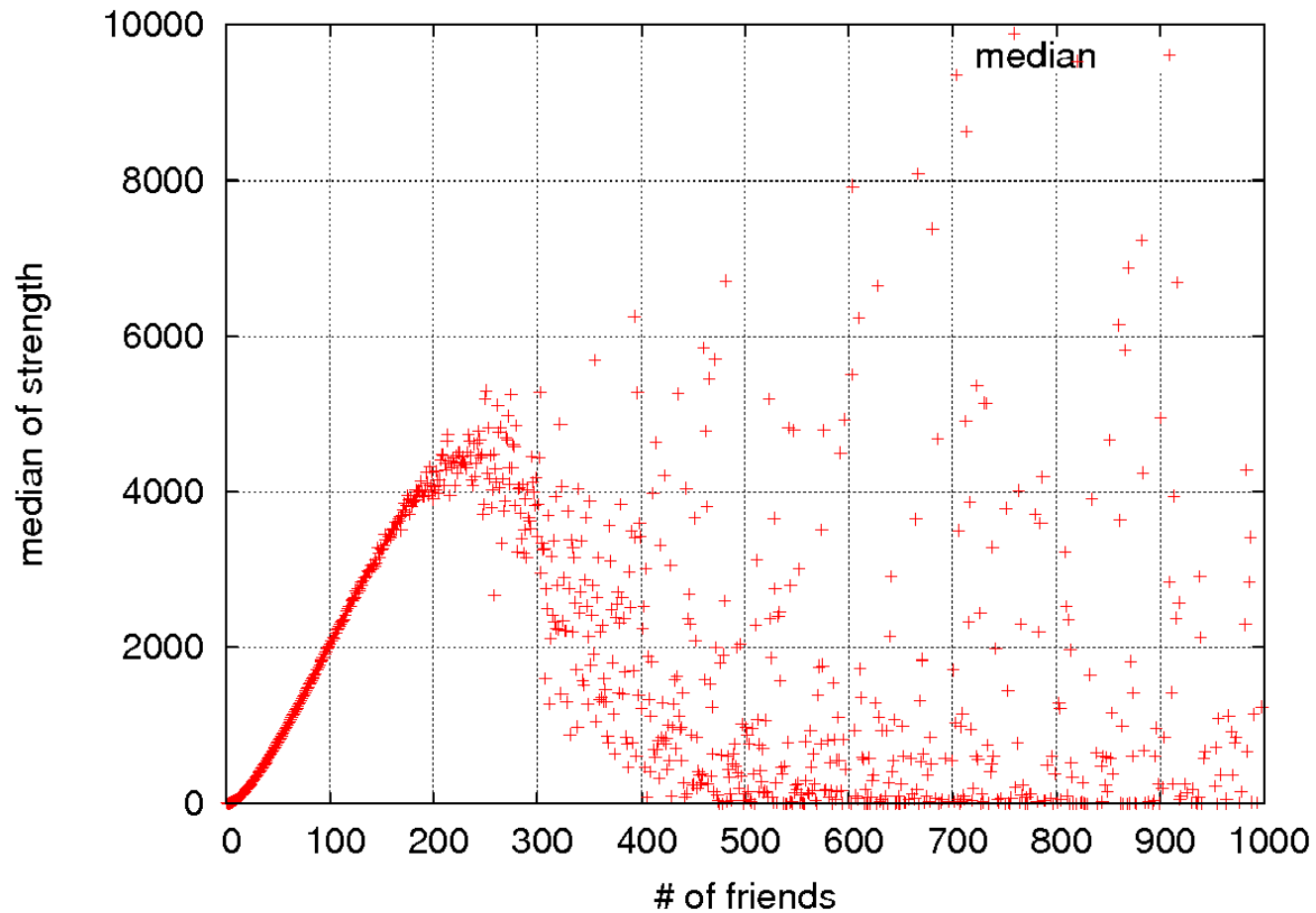
+

-

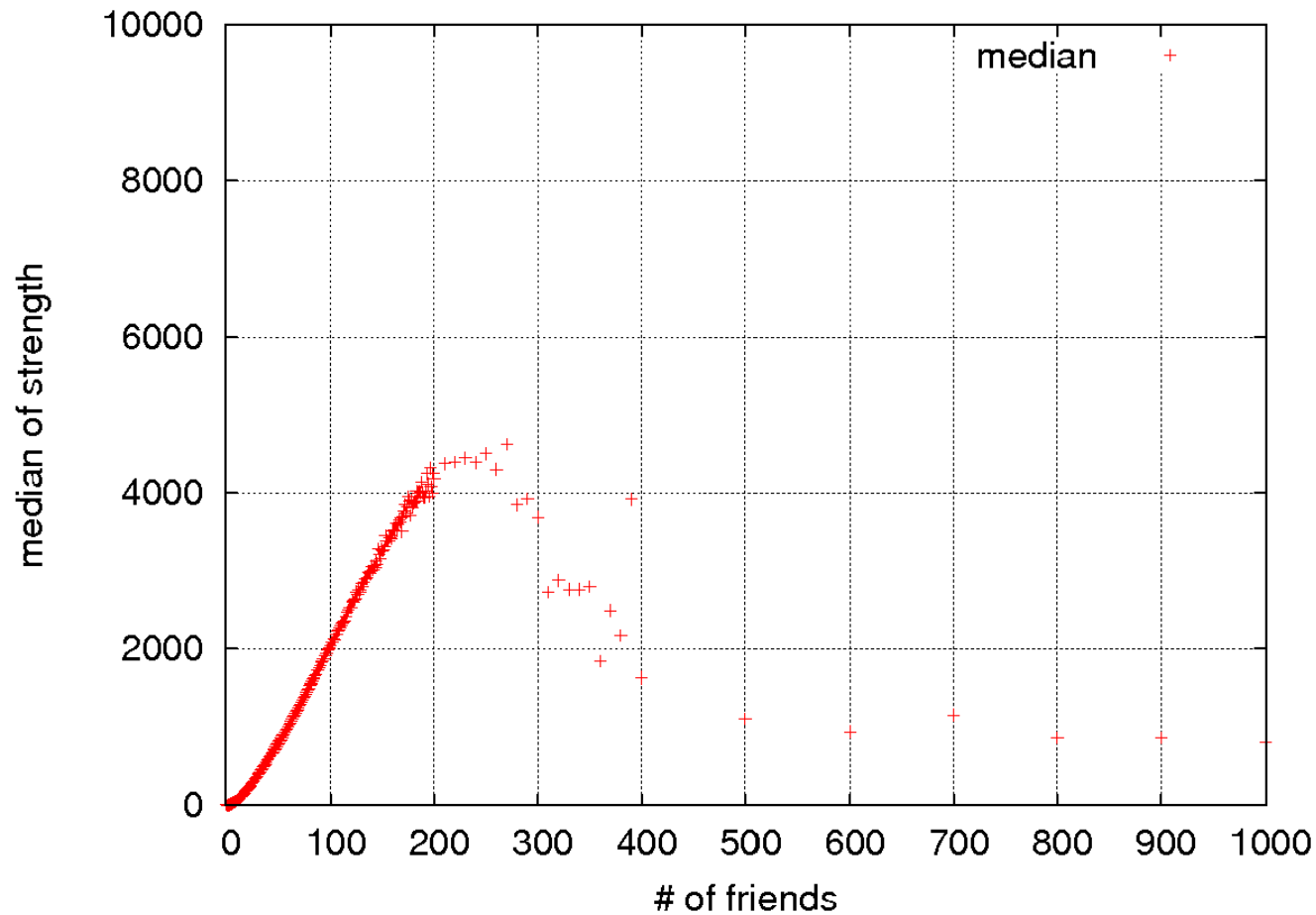
# Degree Correlation



# Median Strength vs Ilchon Degree



200-400 bin=10, >400 bin=100



# Off-line vs Cyber-only

- Why does Cyworld has both user behaviors?
- Significance of 200
  - Push Dunbar's Number to 200 w/ technology
  - "Cyworld Number
- Yet to be valudated
  - Email networks?

# Ongoing/Future Work

- Growth model
  - Existing preferential attachment-based models do not fit
  - Forest fire model? Extensions?
- Cyworld number = 200
  - How much does technology enhance our life?
- Explicit vs implicit communities
  - Clubs, towns vs cliquish behavior

# Thoughts on Evolution of SNS

- Language barriers
- Platforms
  - Applications as “nodes” or “entity of peer pressure”