



What is Twitter, a Social Network or a News Media?

Haewoon Kwak Changhyun Lee Hosung Park Sue Moon
Department of Computer Science, KAIST, Korea

19th International World Wide Web Conference (WWW2010)

Twitter, a microblog service



twitter

Home Profile Find People Settings Help Sign out

What's happening? 140

|

Latest: <http://gigglehd.com/zbxe/3998449> 구글 역시 머리 좋다능. 클라우드를 거쳐서 프린터로 인쇄. 이렇게 되면 크롬 OS 클라이언트가 점점 가벼워질 수 있겠지;; 근데 이러면 프린터는 무조건 외부랑 통신되어야 되나?; 29 minutes ago Tweet

Home

 **red_nude** 지금 봤어요. 작은 경당이네요. 그 자체로 기도인듯..RT @EGGPUDING: #holybread #photokr <http://twitpic.com/1gssrb>
less than 5 seconds ago via twtkr

 **sanghoon_kim** 아아악. 커피 n잔으로 충분히 충전했다 생각했는데 bitmap 인덱스 하나를 계산하지 못하겠어 π 벌써 몇시간째야 π
1 minute ago via web

 **minariboy** 이런 이름의 의원도 있었군. 이란 말을 아까 빼먹었어서; RT @jinalsi: 한나라당 조전혁 의원이 기어이 전교조 명단을 공개했습니다. 법원의 공개금지명령을 어겼는디 본인은 헌법기관이니 법 위에 있다는 게 공개이유 = <http://j.mp/b0BYAU>
2 minutes ago via Osfoora

 **haewoon**
1,656 tweets

278 following 329 followers 33 listed

one·forty
n. a place to find Twitter apps.

Home

@haewoon

Direct Messages 30

Favorites

Retweets

Search Q

Saved Searches ▲

kaist

Lists ▲

kshs

collaborators

New list | View all

Twitter, a microblog service

Home Profile Find People Settings Help Sign out

What's happening? 140

Latest: <http://gigglehd.com/2bxe/3998449> 구글 역시 머리 좋다능. 클라우드를 거쳐서 프린터로 인쇄. 이렇게 되면 크롬 OS 클라이언트가 점점 가벼워질 수 있겠지;; 근데 이러면 프린터는 무조건 외부랑 통신되어야 되나?; 29 minutes ago

Tweet

Home

Home

@haewoon

Direct Messages 30

Favorites

Retweets

Search

Saved Searches

kaist

Lists

kshs

collaborators

New list View all

write a short message

Twitter, a microblog service

twitter

Home Profile Find People Settings Help Sign out

What's happening? 140

Latest: <http://gigglehd.com/zbxe/3998449> 구글 역시 머리 좋다능. 클라우드를 거쳐서 프린터로 인쇄. 이렇게 되면 크롬 OS 클라이언트가 점점 가벼워질 수 있겠지;; 근데 이러면 프린터는 무조건 외부랑 통신되어야 되나?; 29 minutes ago

Tweet

Home

red_nude 최근 봤어요. 작은 경당이네요. 그 자체로 기도인듯..RT @EGGPUDING: #nolybread #photokr <http://twitpic.com/1gss1b> less than 5 seconds ago via twtkr

sanghoon_kim 아아아. 커피 n잔으로 충분히 충전했다 생각했는데 bitmap 인덱스 하나를 계산하지 못하겠어 π 벌써 몇시간째야 π 1 minute ago via web

minariboy 이런 이름의 의원도 있었군. 이란 말을 아까 빼먹었어서; RT @jinalsi: 한나라당 조전혁 의원이 기어이 전교조 명단을 공개했습니다. 법원의 공개금지명령을 어겼는디 본인은 헌법기관이니 법 위에 있다는 게 공개이유 = <http://j.mp/b0BYAU> 2 minutes ago via Osfoora

haewoon 1,656 tweets

278 following 329 followers 33 listed

one·forty *n.* a place to find Twitter apps.

Home

@haewoon

Direct Messages 30

Favorites

Retweets

Saved Searches

kaist

Lists

kshs

collaborators

New list View all

read neighbors' tweets



In most OSN



“We are friends.”



In Twitter



“I follow you.”

Following on Twitter

“Unlike most social networks, following on Twitter is **not mutual**. Someone who thinks you're interesting can follow you, and you don't have to approve, or follow back.”

<http://help.twitter.com/entries/14019-what-is-following>

Following = subscribing tweets

twitter

Home Profile Find People Settings Help Sign out

What's happening? 140

Latest: <http://gigglehd.com/zbxe/3998449> 구글 역시 머리 좋다능. 클라우드를 거쳐서 프린터로 인쇄. 이렇게 되면 크롬 OS 클라이언트가 점점 가벼워질 수 있겠지;; 근데 이러면 프린터는 무조건 외부랑 통신되어야 되나?; 29 minutes ago

Tweet

Home

red_nude 최근 봤어요. 작은 경당이네요. 그 자체로 기도인듯..RT @EGGPUDING: #nolybread #photokr <http://twitpic.com/1gssrh> less than 5 seconds ago via twtkr

sanghoon_kim 아아아. 커피 n잔으로 충분히 충전했다 생각했는데 bitmap 인덱스 하나를 계산하지 못하겠어 π 벌써 몇시간째야 π 1 minute ago via web

minariboy 이런 이름의 의원도 있었군. 이런 말을 아까 빼먹었어서; RT @jinalsi: 한나라당 조전혁 의원이 기어이 전교조 명단을 공개했습니다. 법원의 공개금지명령을 어겼는디 본인은 헌법기관이니 법 위에 있다는 게 공개이유 = <http://j.mp/b0BYAU> 2 minutes ago via Osfoora

haewoon 1,656 tweets

278 following 329 followers 33 listed

one·forty *n.* a place to find Twitter apps.

Home

@haewoon

Direct Messages 30

Favorites

Retweets

Search

Saved Searches

kaist

Lists

kshs

collaborators

New list View all

recent tweets of followings



jkrums

Follow

<http://twitpic.com/135xa> – There's a plane in the Hudson. I'm on the ferry going to pick up the people. Crazy.





©2009 Ian D. Marsden - www.marsdencartoons.com

<http://blog.marsdencartoons.com/2009/06/18/cartoon-iranian-election-demonstrations-and-twitter/marsden-iran-twitter72/>

The goal of this work

We analyze how directed relations of following set Twitter apart from existing OSNs.

Then, we see if Twitter has any characteristics of news media.

me · di · a [*mee-dee-uh*]



1. a pl. of medium

2. the means of communication, as radio and television, newspapers, and magazines, that reach or influence people widely

The goal of this work

We analyze how directed relations of following set Twitter apart from existing OSNs.

Then, we see if Twitter has any characteristics of news media.

Summary of our findings

- 
- 
1. Following is mostly not reciprocated (not so “social”)
 2. Users talk about timely topics
 3. A few users reach large audience directly
 4. Most users can reach large audience by WOM* quickly

*WOM: word-of-mouth

Data collection (09/6/1~9/24)

- 41.7M user profiles (near-complete at that time)
- 1.47B following relations *publicly available
- 4262 trending topics
- 106M tweets mentioning trending topics
 - ▶ Spam tweets removed by CleanTweets

How we crawled

- Twitter's well-defined 3rd party API
- With 20+ 'whitelisted' IPs
 - ▶ Send 20,000 requests per IP / hour



Recent studies

- Ranking methodologies [WSDM'10]
- Predicting movie profits [HYPERTEXT'10]
- Recommending users [CHI'10 microblogging]
- Detecting real time events [WWW'10]
- The 'entire' Twittersphere unexplored

Part I.

1. Following is mostly not reciprocated (not so “social”)
2. Users talk about timely topics
3. A few users reach large audience directly
4. Most users can reach large audience by WOM* quickly

Why do people follow others?

- Reflection of offline social relationships

otherwise,

- Subscription to others' messages

Sociologists' answer

- “Reciprocal interactions pervade every relation of primitive life and in all social systems”



Is following reciprocal?

- Only 22.1% of user pairs follow each other
- Much lower than
 - ▶ 68% on Flickr
 - ▶ 84% on Yahoo! 360
 - ▶ 77% on Cyworld guestbook messages

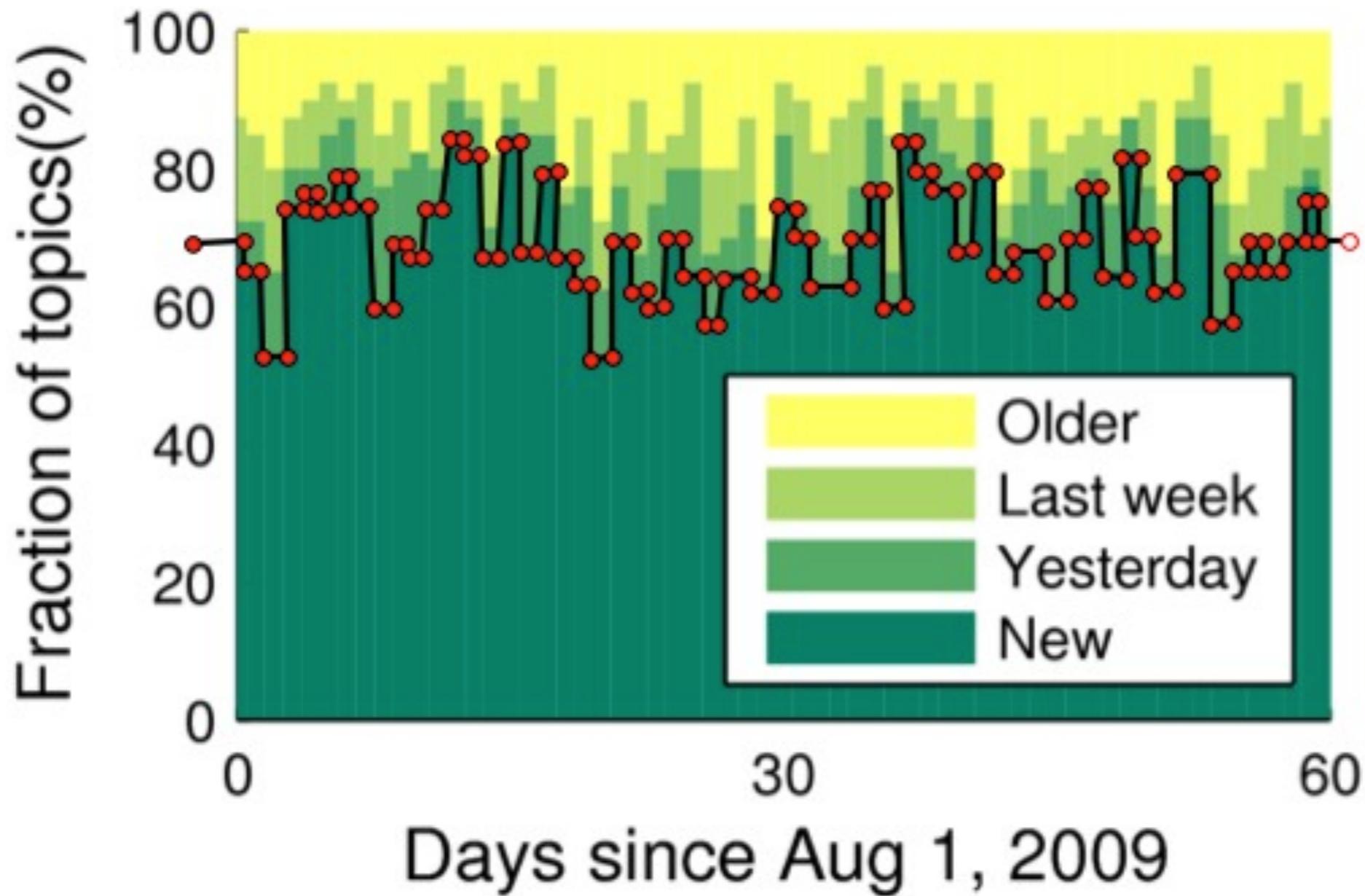
Low reciprocity of following

- Following is not similarly used as friend in OSNs
 - ▶ Not reflection of offline social relationships
- Active subscription of tweets!

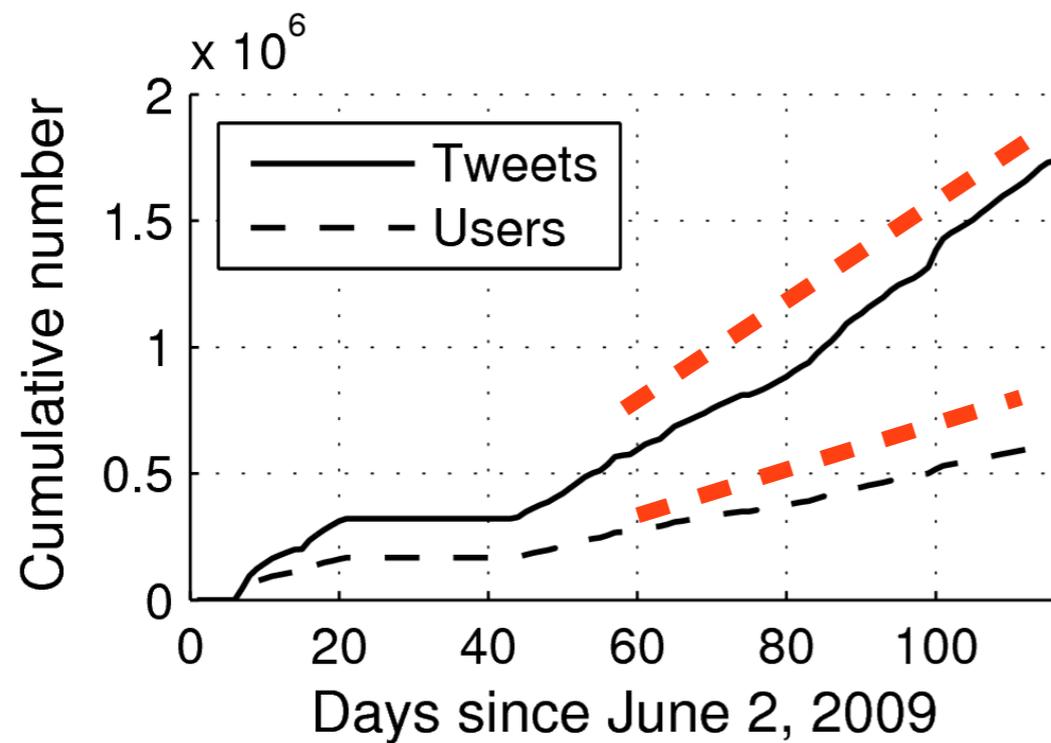
Part II.

1. Following is mostly not reciprocated (not so “social”)
- 2. Users talk about timely topics**
3. A few users reach large audience directly
4. Most users can reach large audience by WOM* quickly

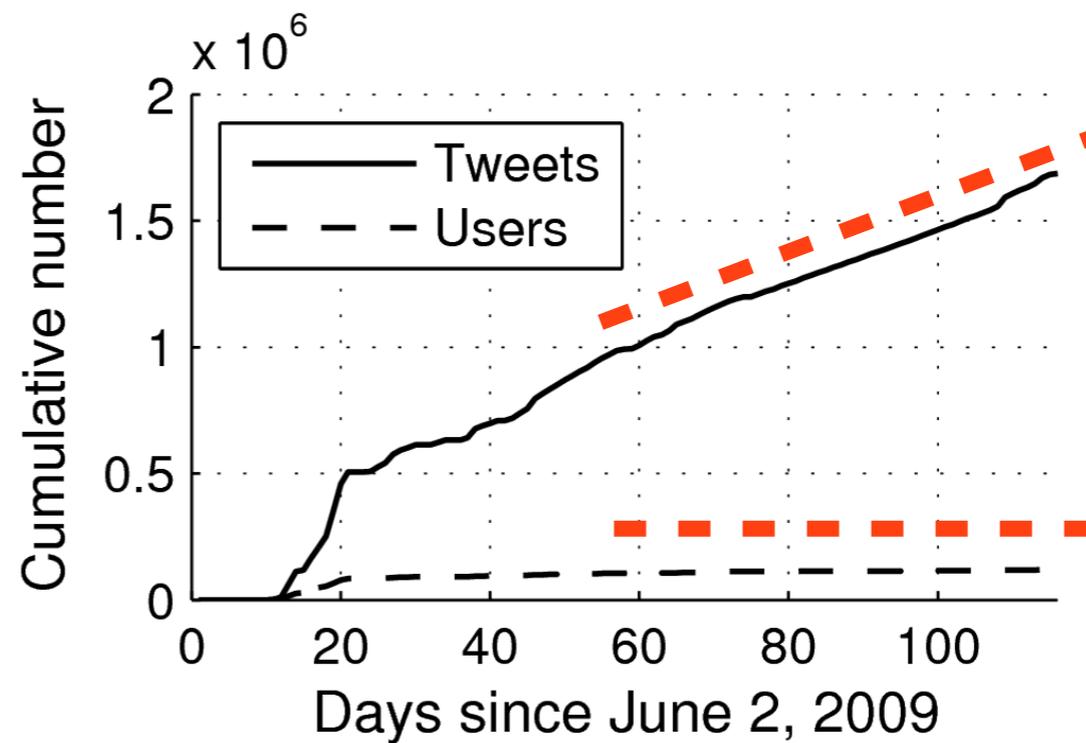
Dynamically changing trends



User participation pattern can be a signature of a topic



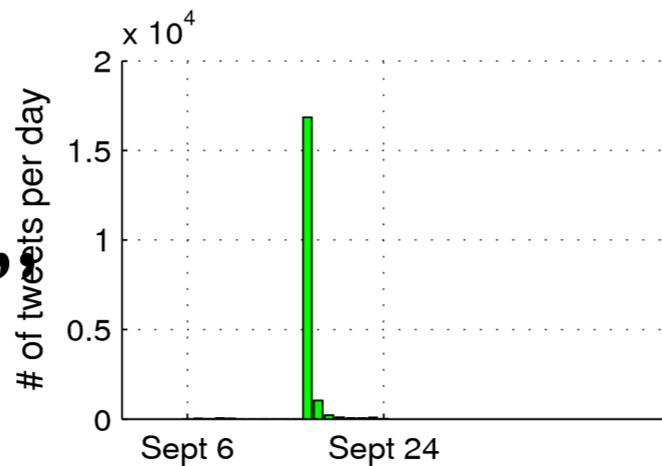
(a) Topic 'apple'



(b) Topic '#iranelection'

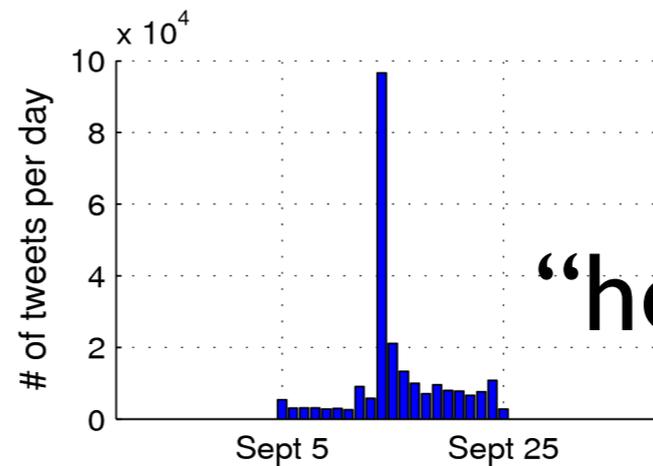
Majority of topics are headline

31.5%
“ephemeral”



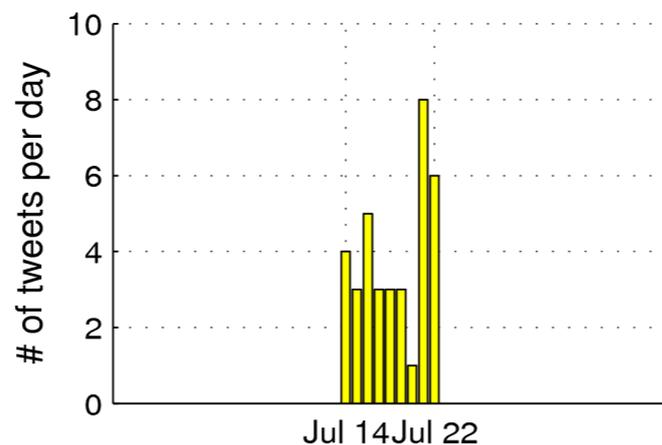
(a) Exogenous subcritical
(topic ‘#backintheday’)

54.3%
“headline news”



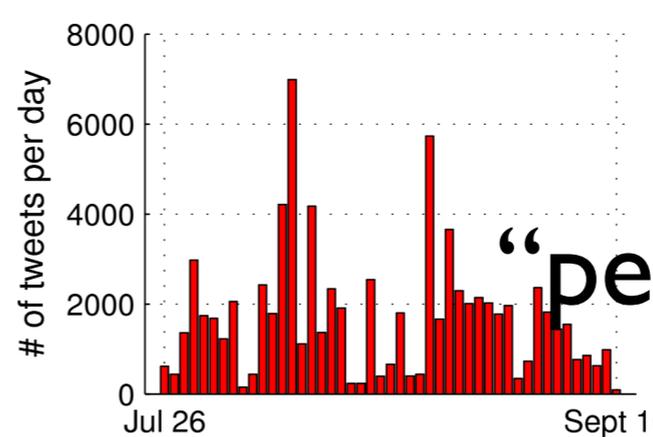
(b) Exogenous critical
(topic ‘beyonce’)

6.9%



(c) Endogenous subcritical
(topic ‘lynn harris’)

7.3%
“persistent news”



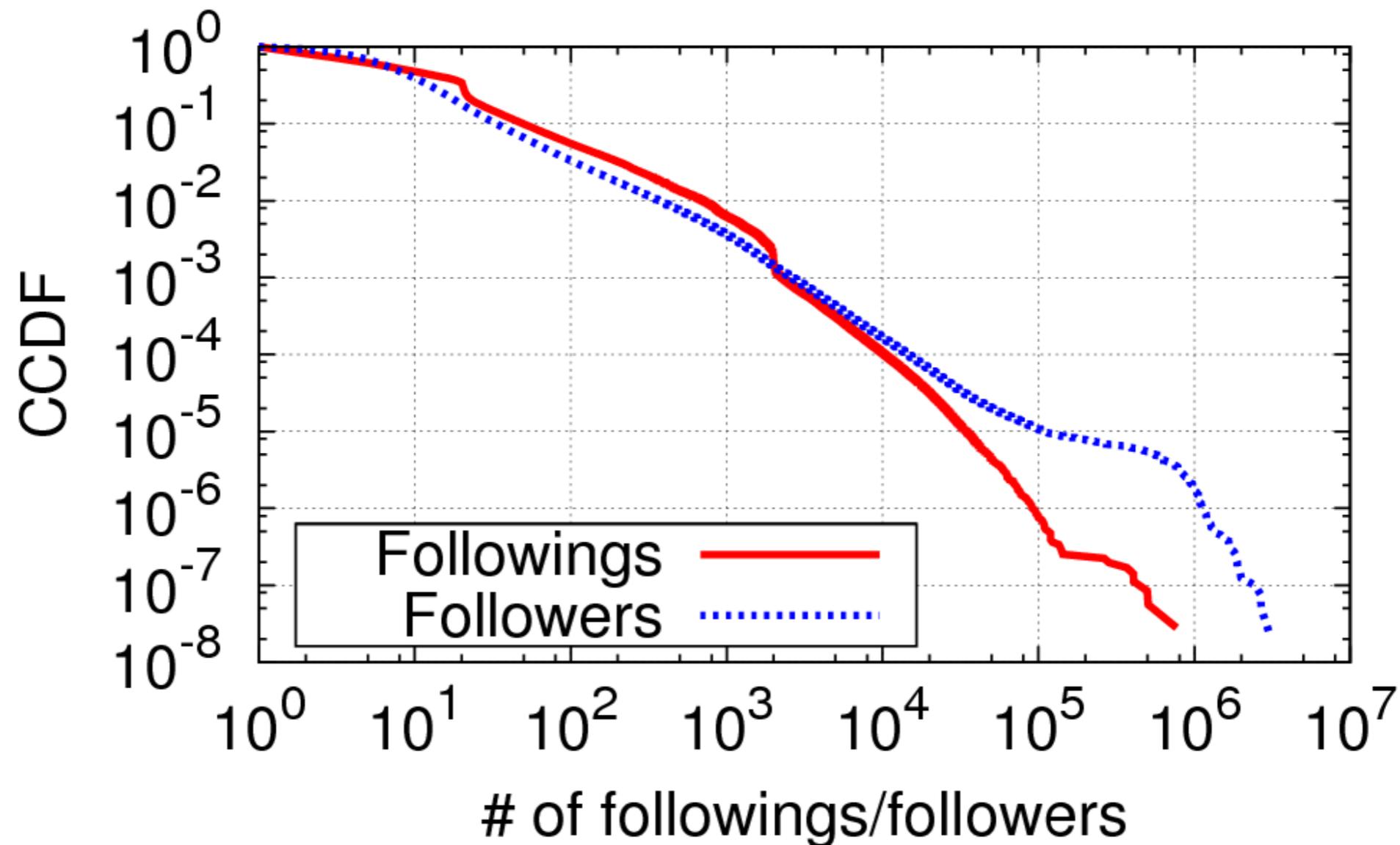
(d) Endogenous critical
(topic ‘#redsox’)



Part III.

1. Following is mostly not reciprocated (not so “social”)
2. Users talk about timely topics
- 3. A few users reach large audience directly**
4. Most users can reach large audience by WOM* quickly

How many followers a user has?

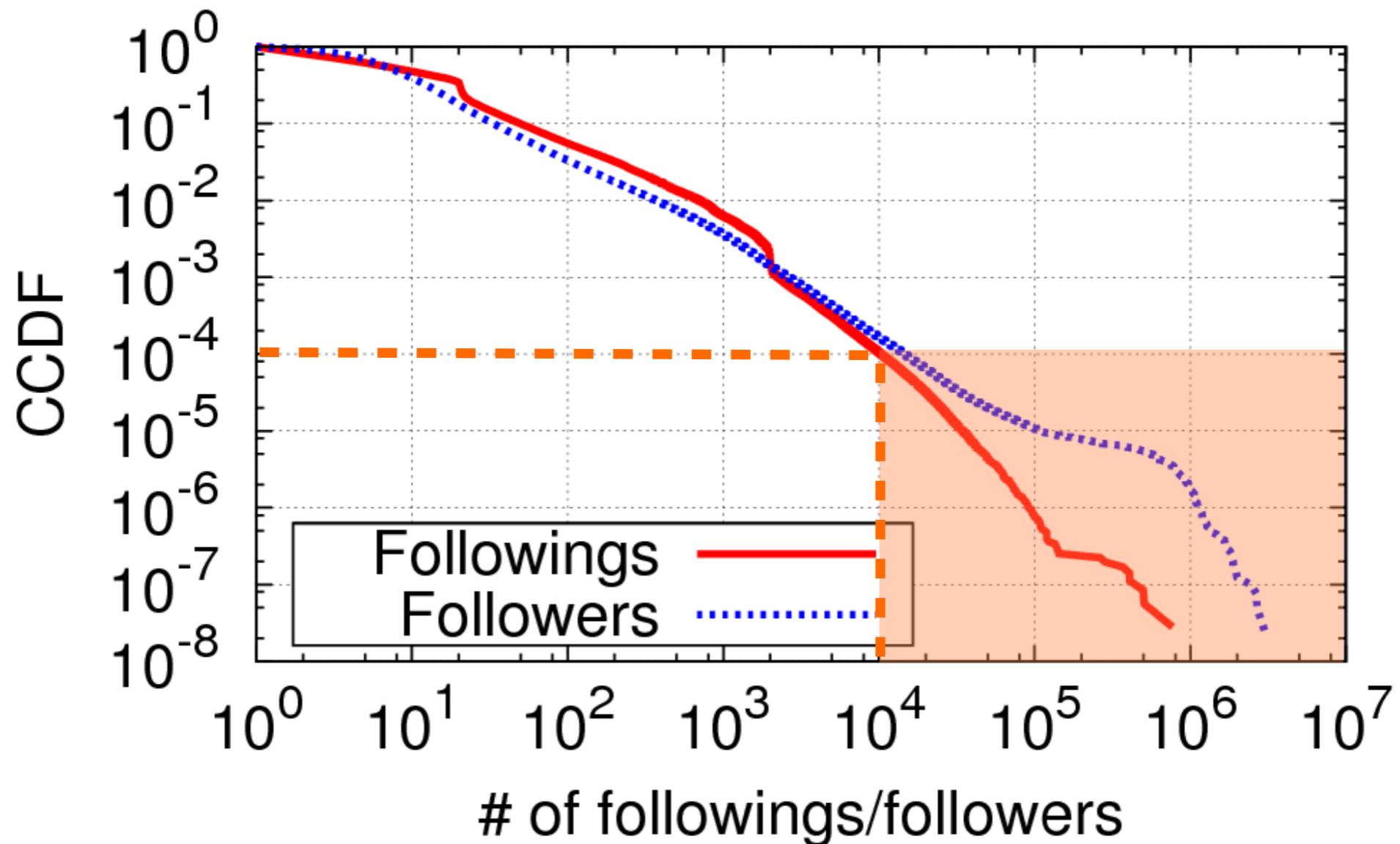




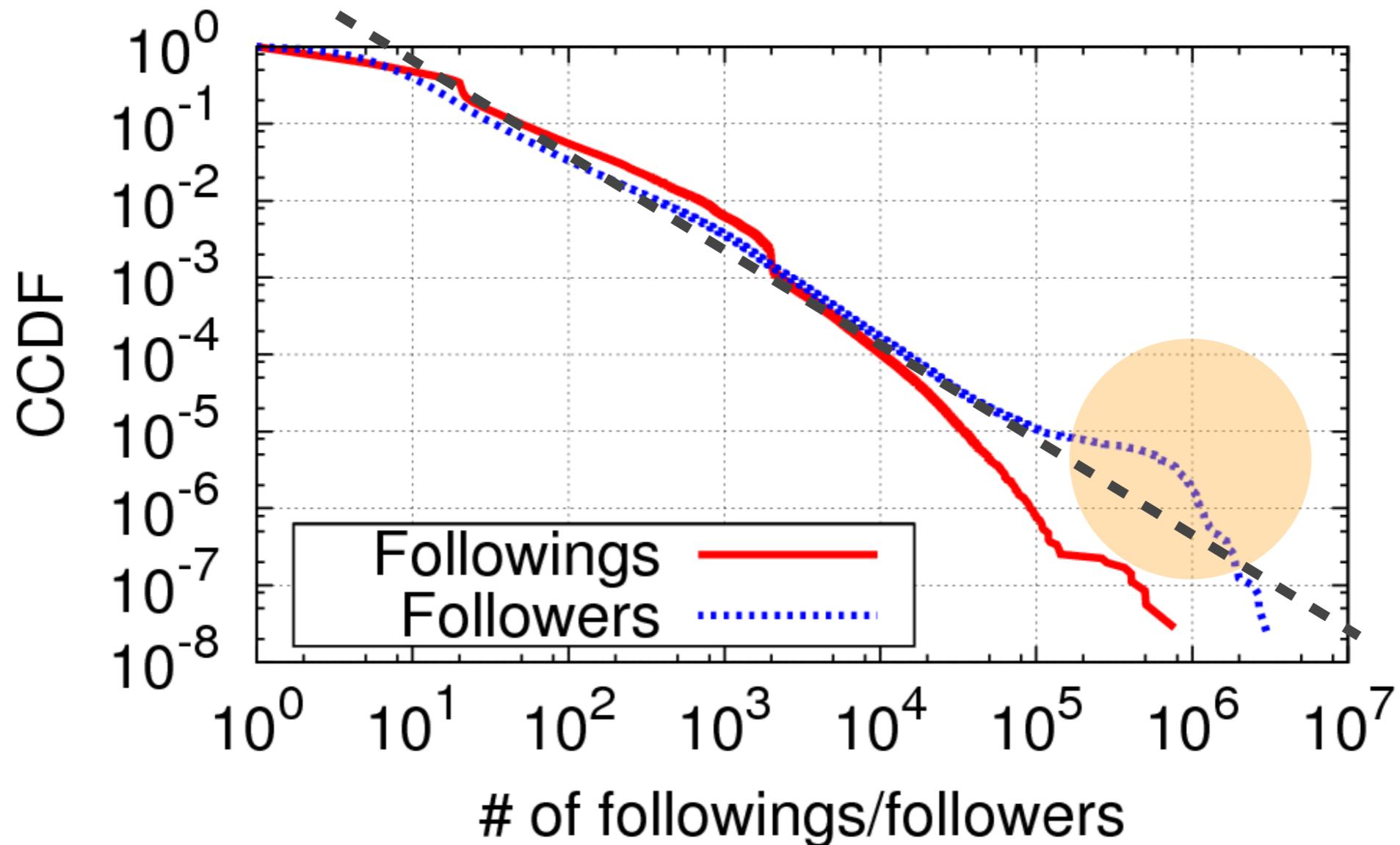
CCDF

- Complementary Cumulative Density Function
- $CCDF(x=k) = \int_k^{\infty} P(x) dx$

Reading the graph



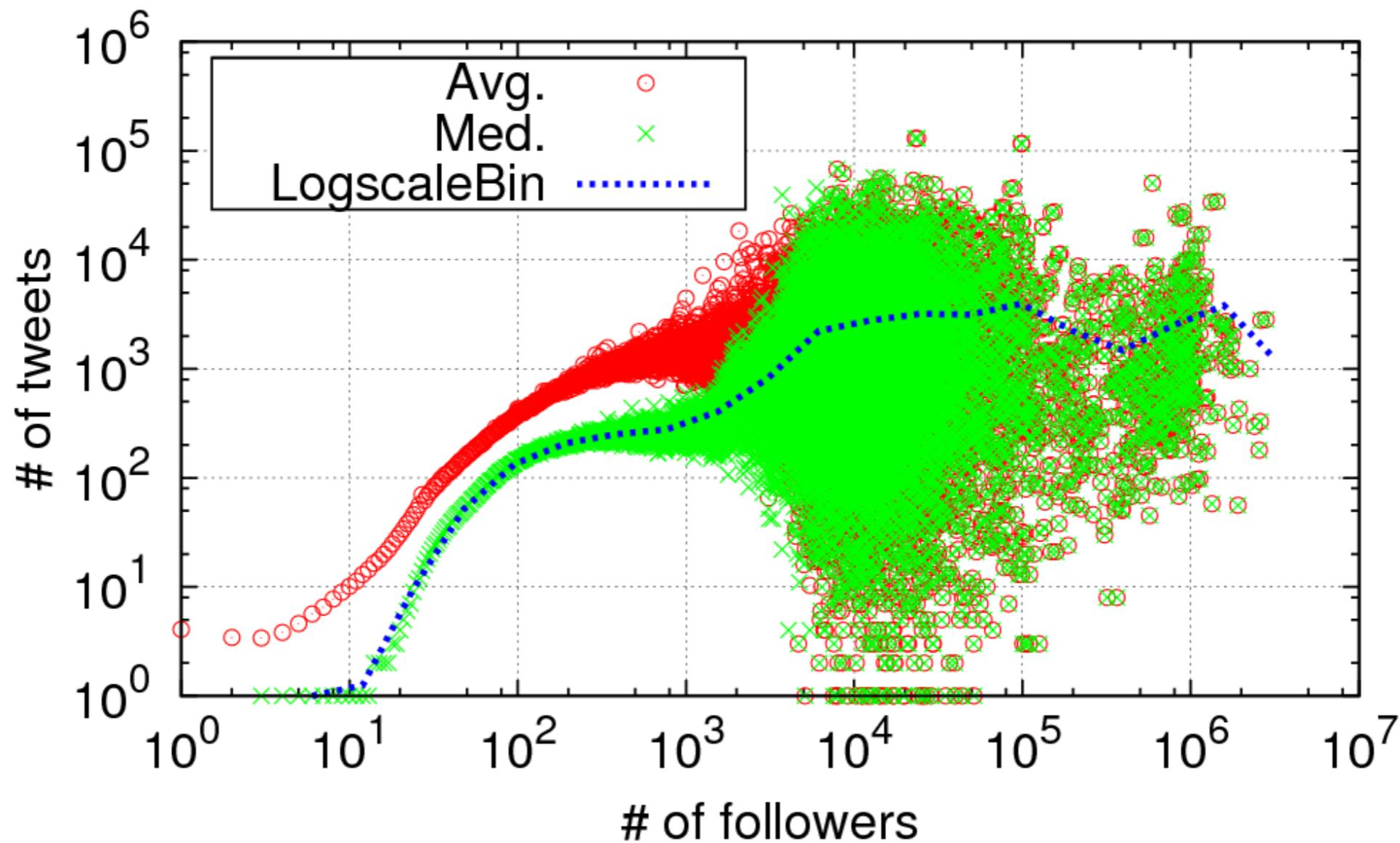
Plenty of super-hubs



More super-hubs than projected by power-law

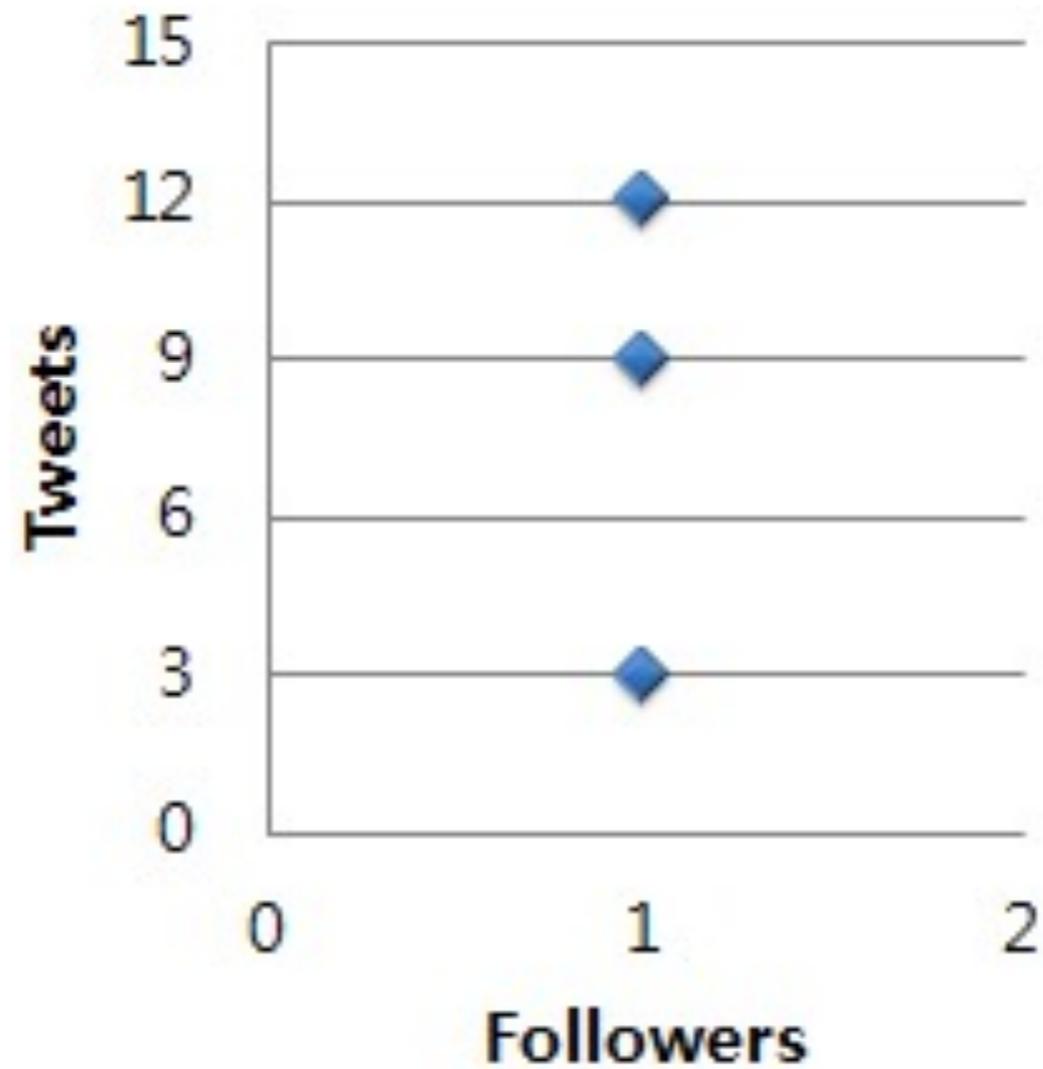
- Where do they get all the followers? Possibly from...
 - ▶ Search by 'name'
 - ▶ Recommendation by Twitter
- They reach millions in one hop

Are those who have many followers active?



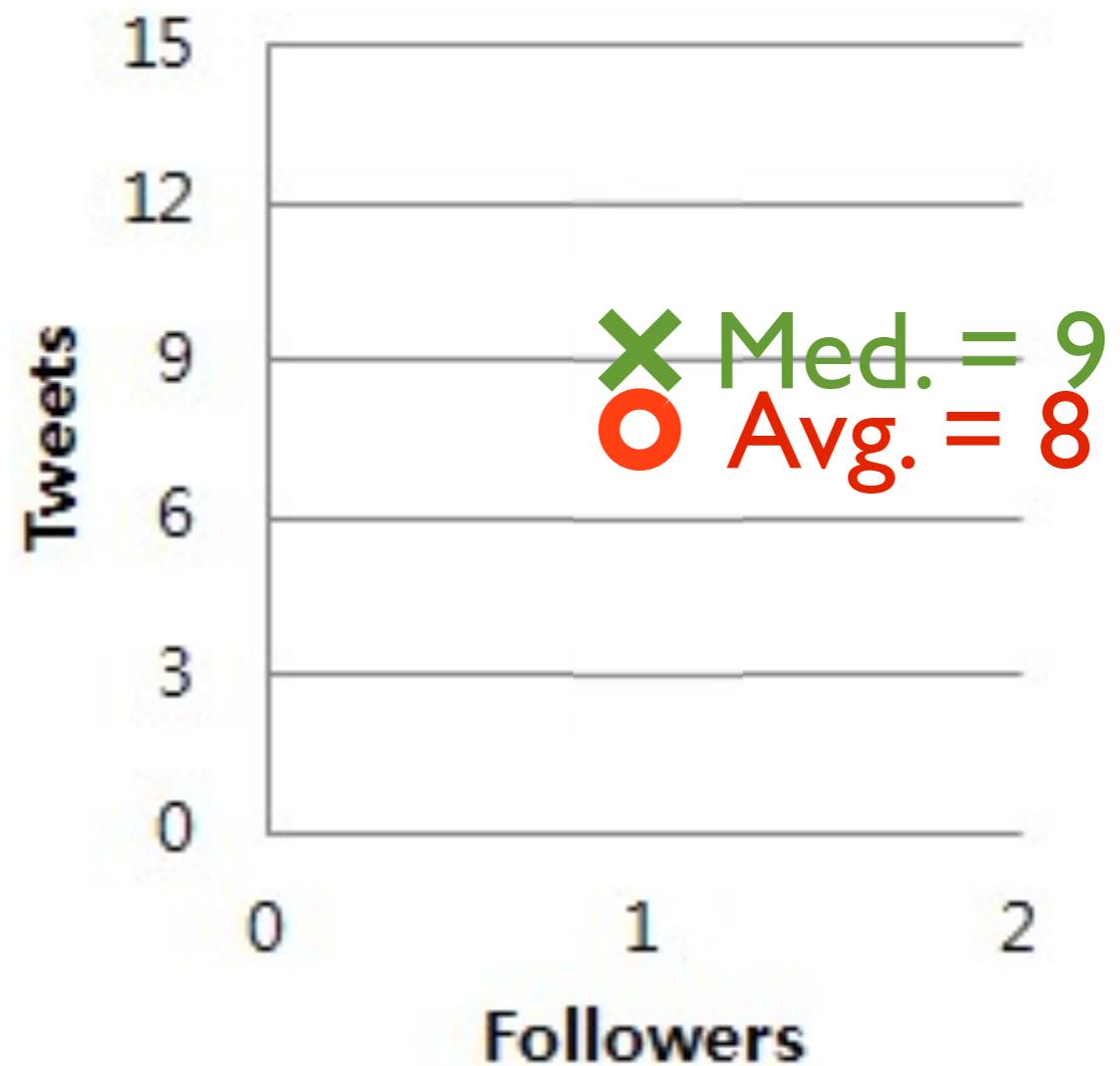
How we plotted

followers	tweets
1	3
1	9
1	12

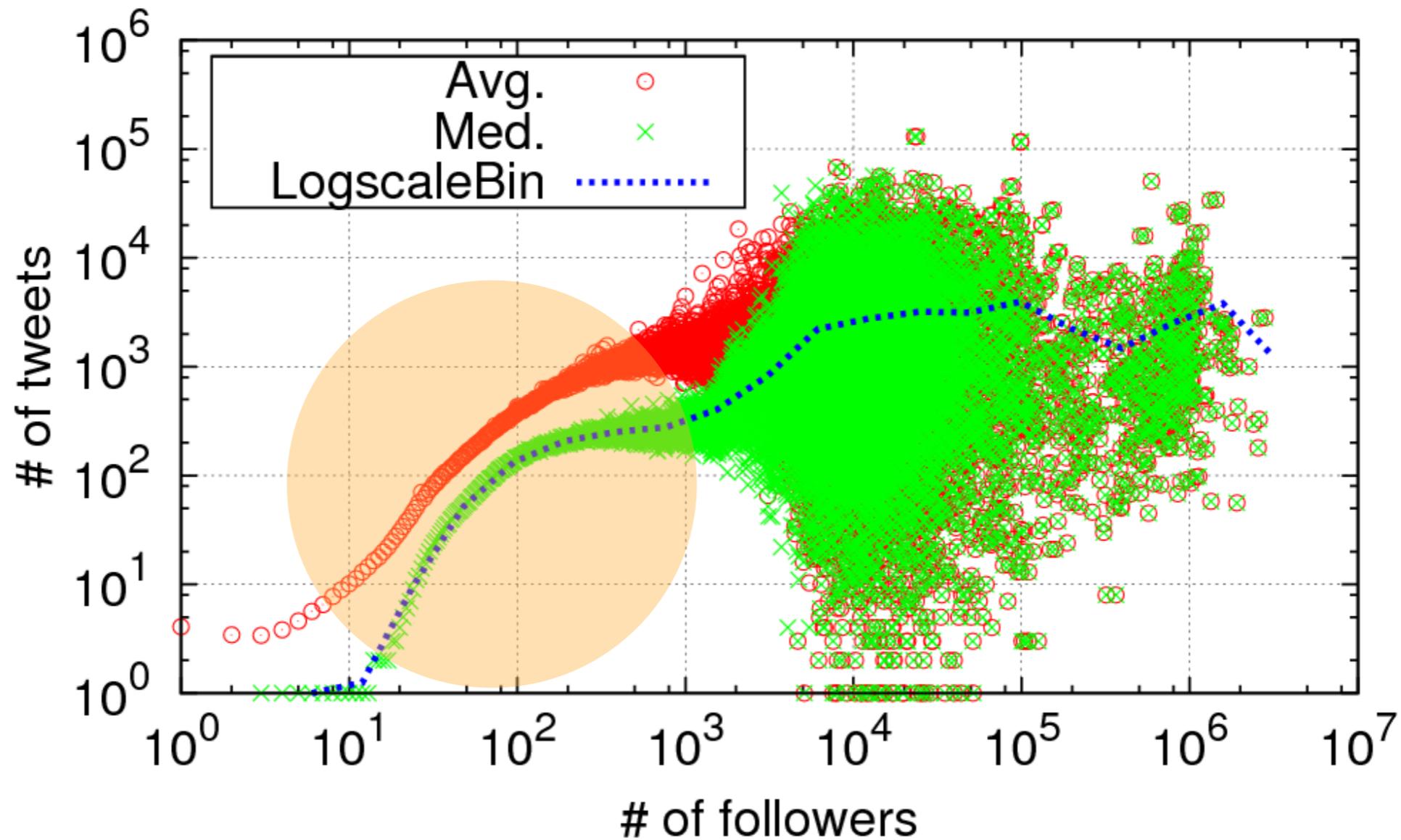


How we plotted

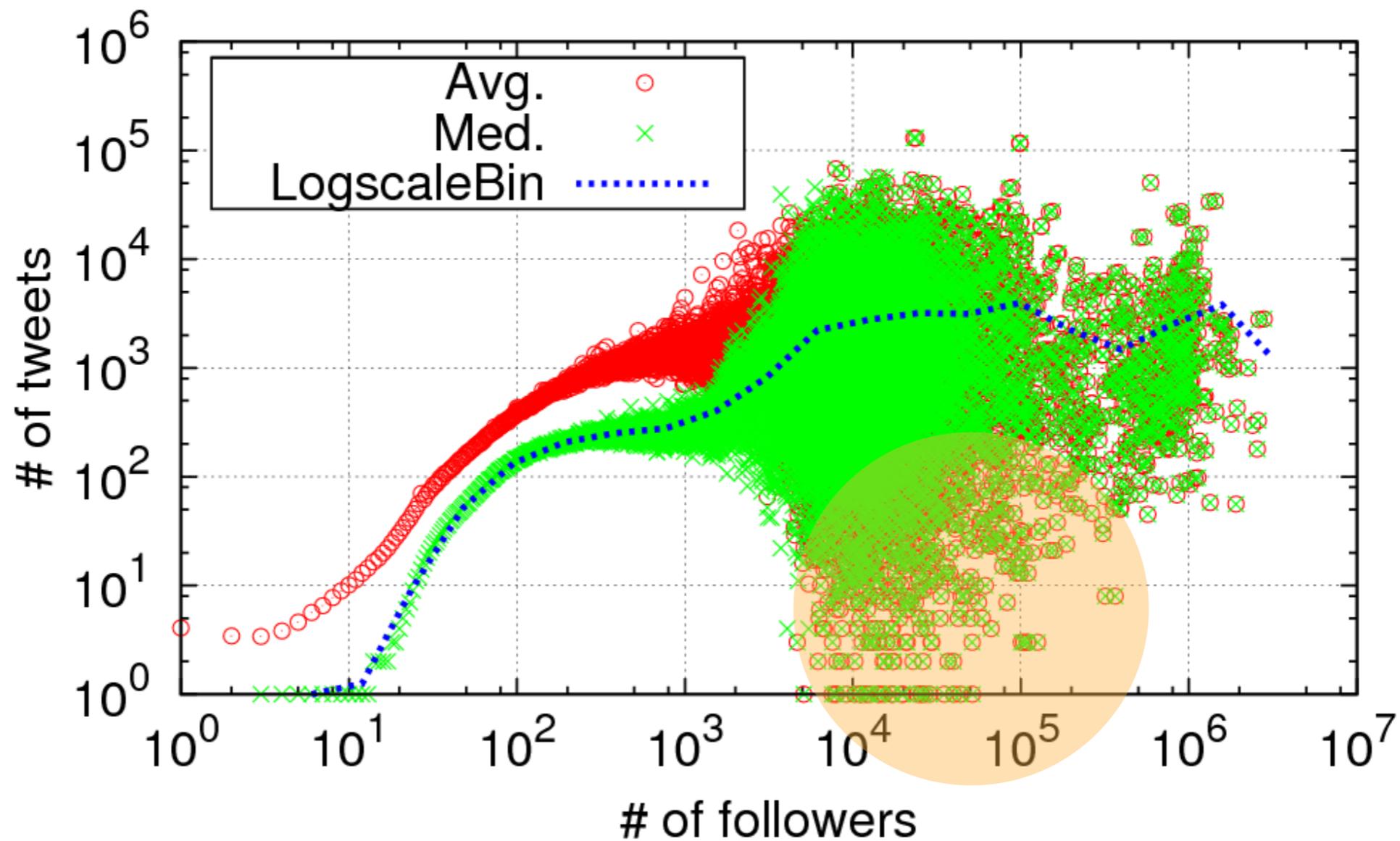
followers	tweets
1	3
1	9
1	12



More followers, more tweets



Many followers without activity



Twitter user rankings by Followers, PageRank and RT

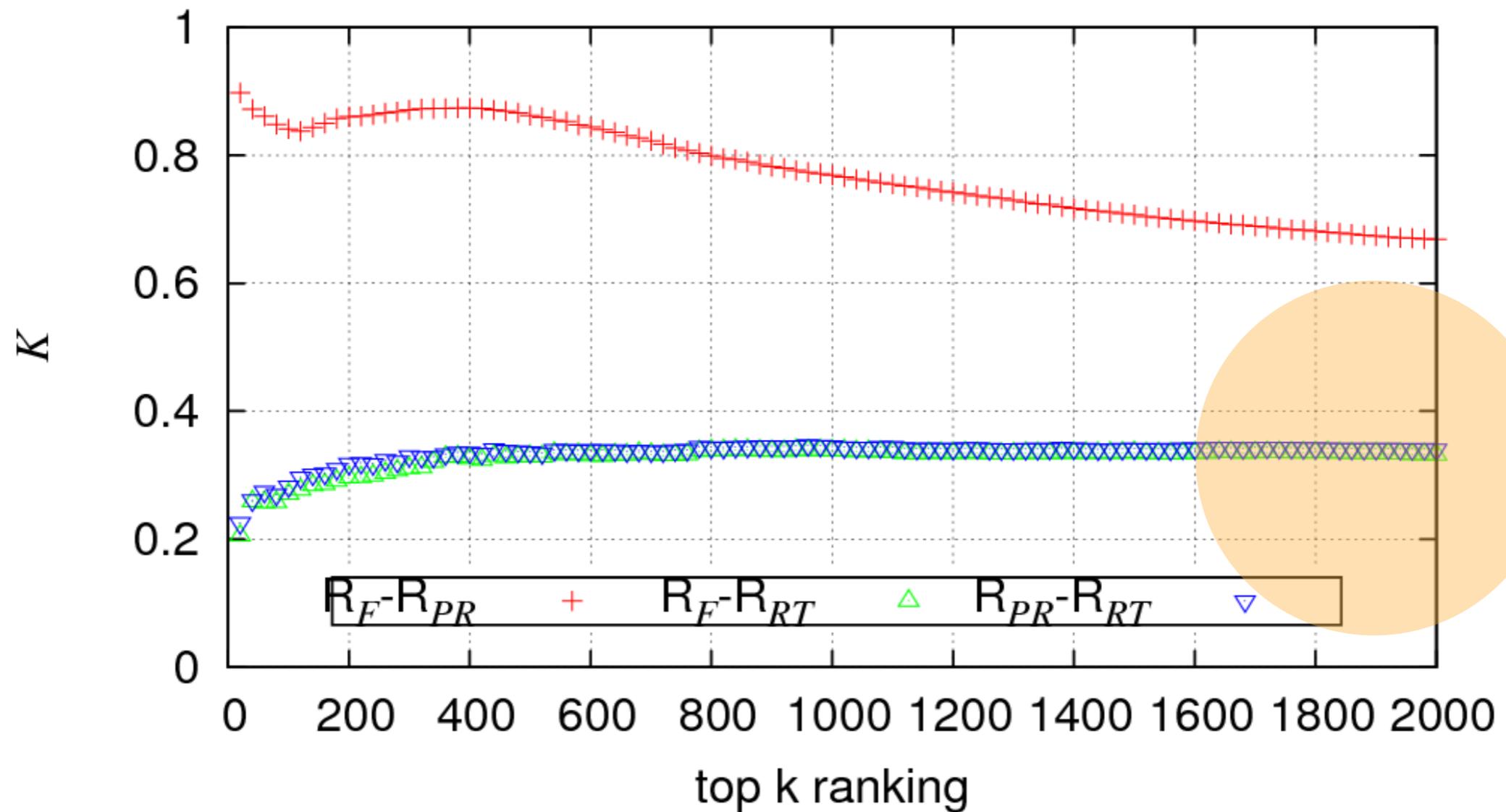


Rank	Followers	PageRank	RT
1	actor	actor	news
2	musician	president	news
3	show host	news	news
4	news	show host	journalist
5	show host	musician	news
6	twitter	show host	news
7	president	sports star	musician
8	show host	musician	news
9	sports star	twitter	comedian
10	model	show host	about iran

Twitter user rankings by Followers, PageRank and RT

Rank	Followers	PageRank	RT
1	actor	actor	news
2	musician	president	news
3	show host	news	news
4	news	show host	journalist
5	show host	musician	news
6	twitter	show host	news
7	president	sports star	musician
8	show host	musician	news
9	sports star	twitter	comedian
10	model	show host	about iran

Great discrepancy among rankings

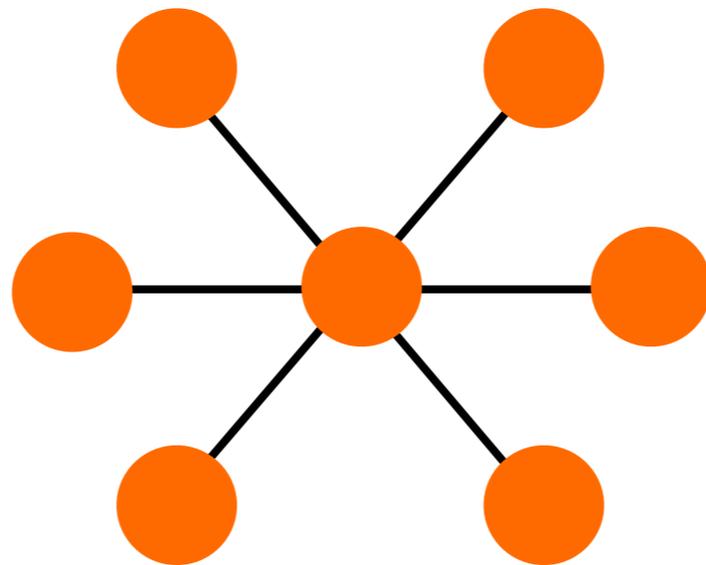
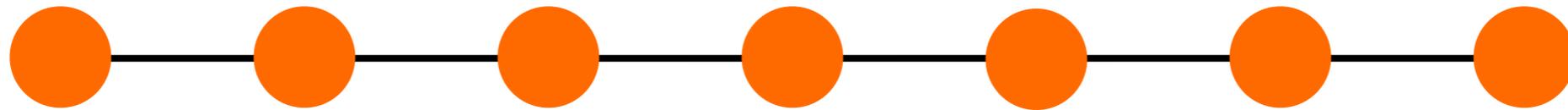


Part IV.

1. Following is mostly not reciprocated (not so “social”)
2. Users talk about timely topics
3. A few users reach large audience directly
4. **Most users can reach large audience by WOM* quickly**

***WOM: word-of-mouth**

Which is more efficient for WOM?

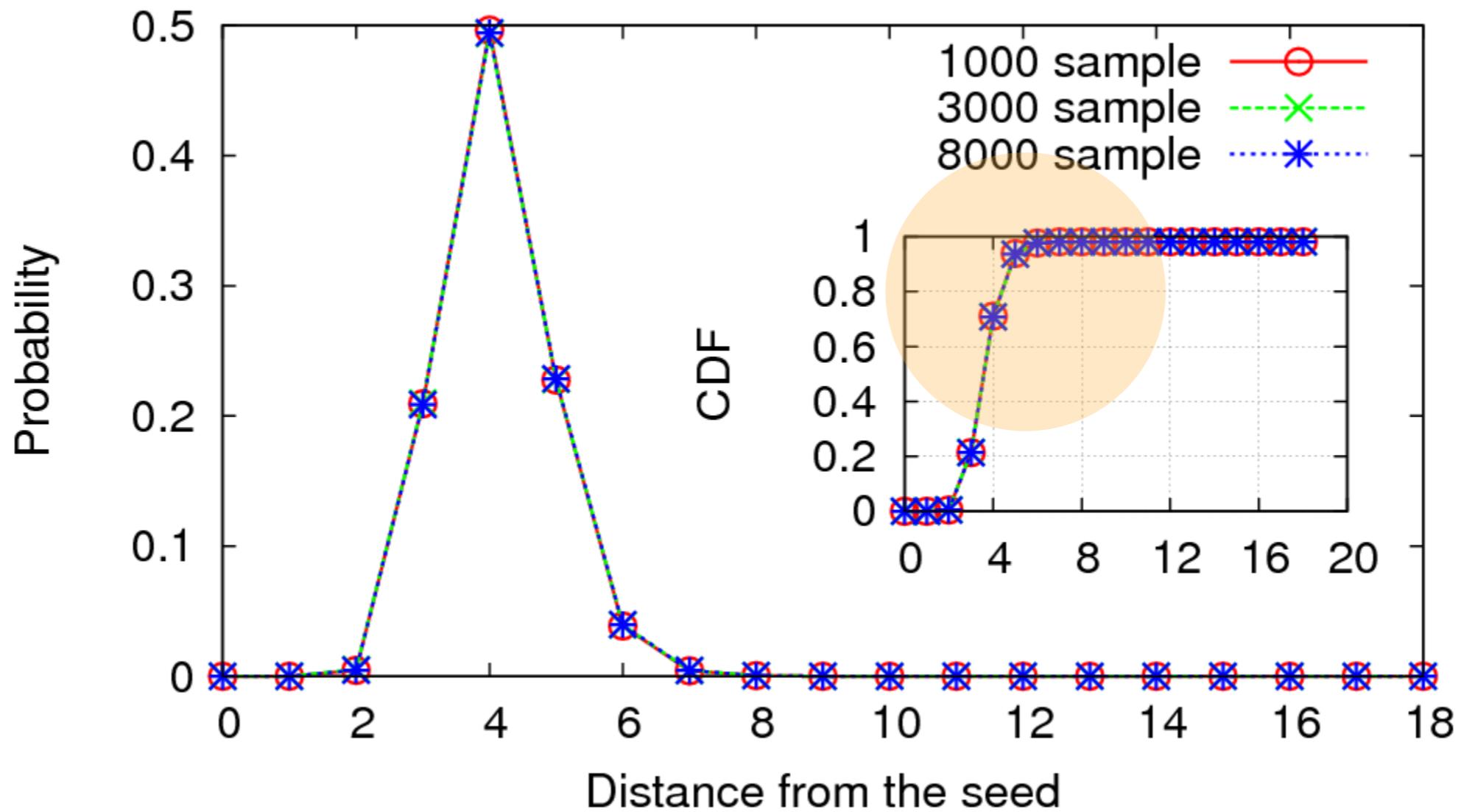




In Twitter

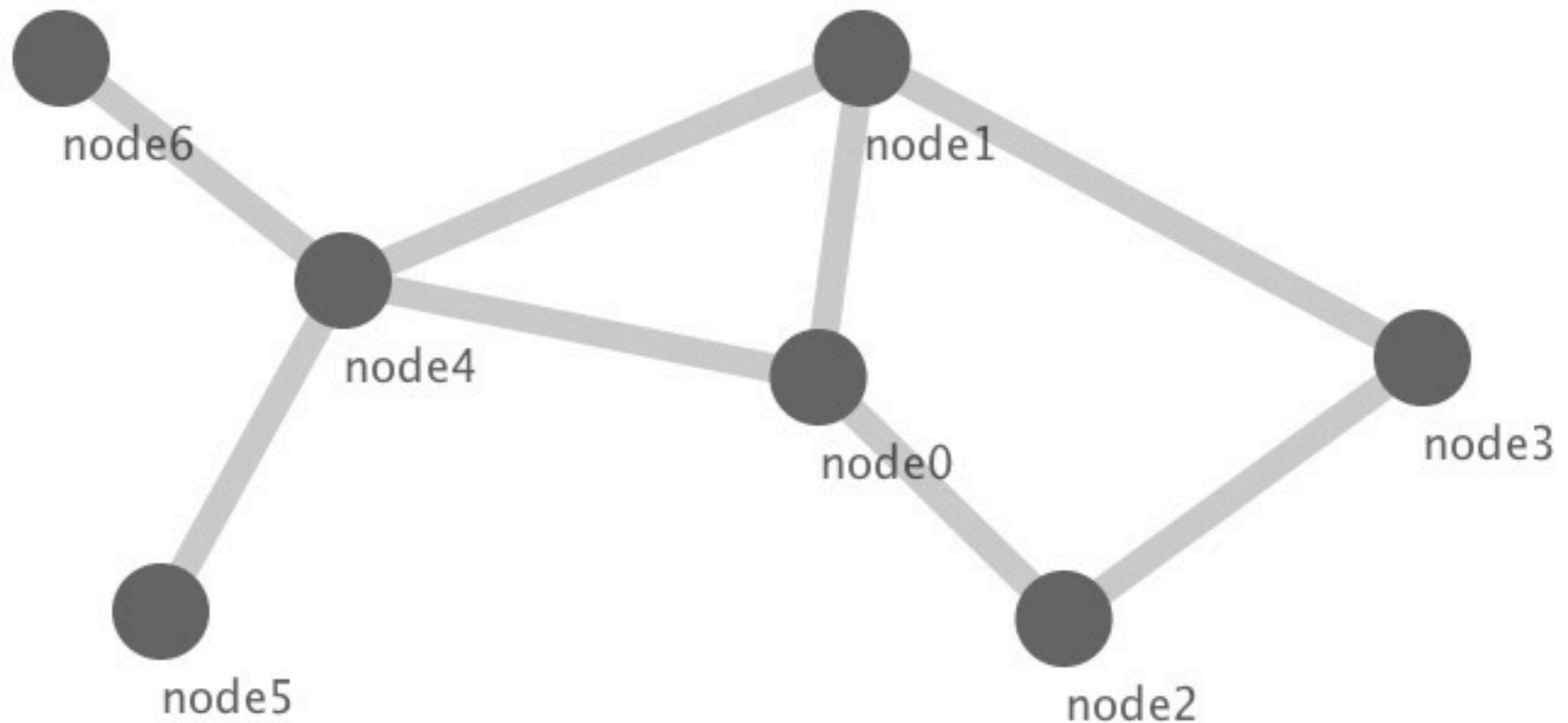


Average path length: 4.1



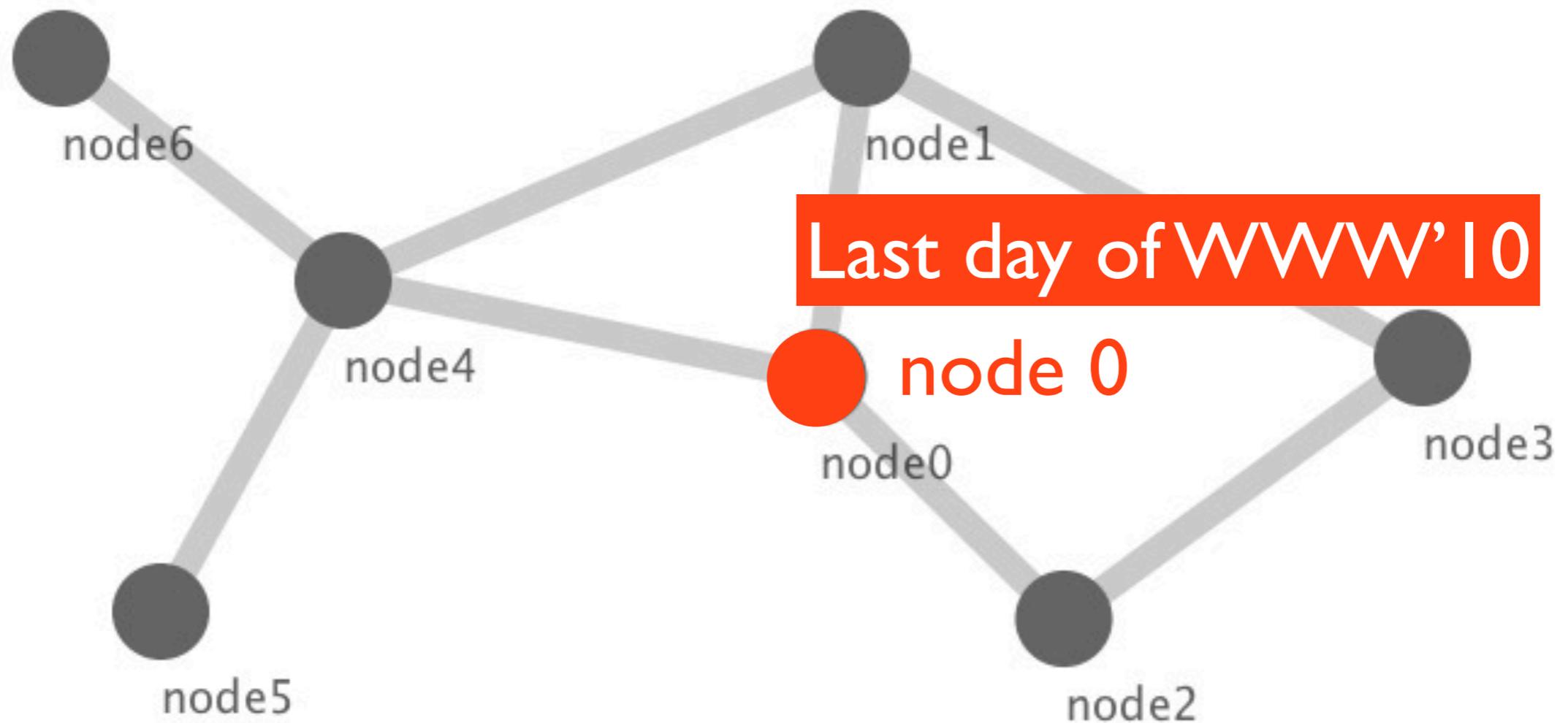
Retweet (RT)

- Relay tweets from a following to followers



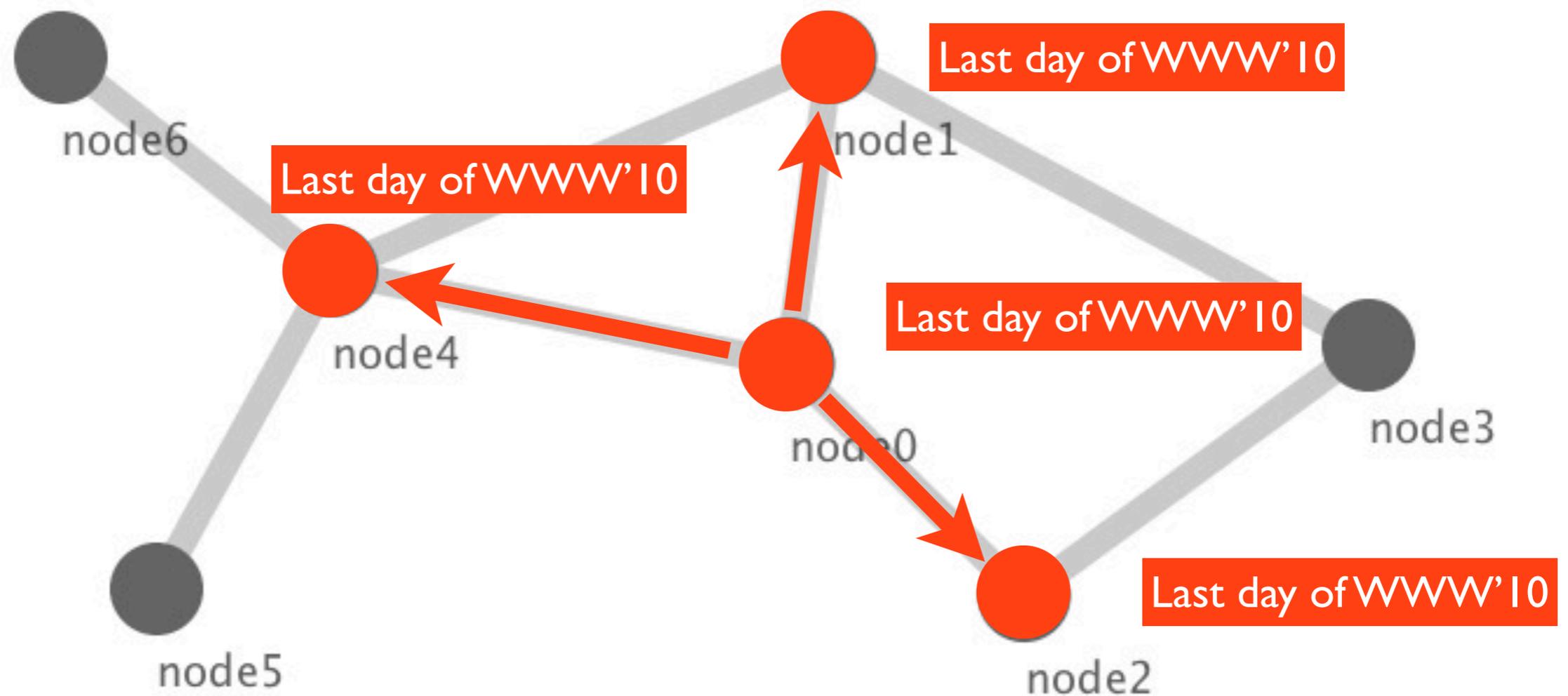
Retweet (RT)

- Relay tweets from a following to followers



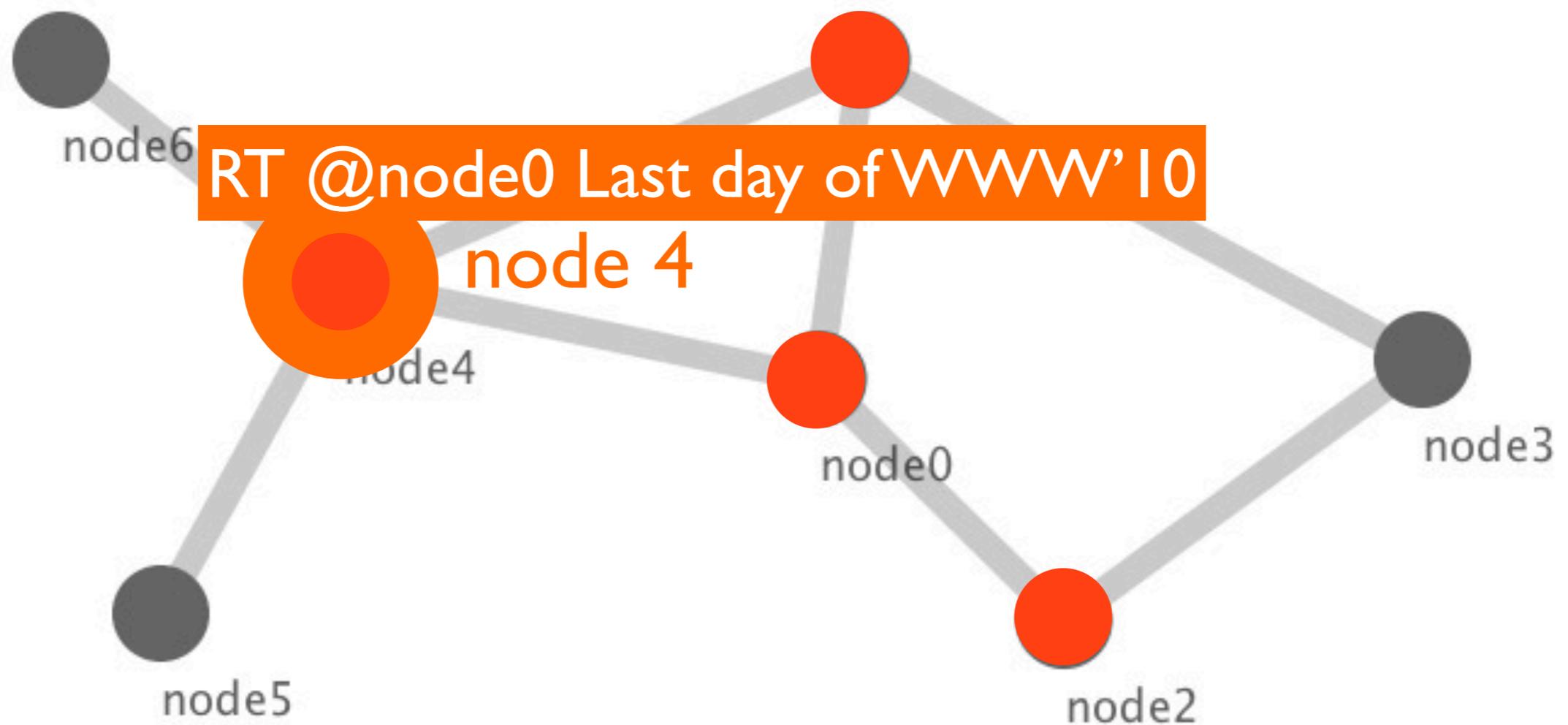
Retweet (RT)

- Relay tweets from a following to followers



Retweet (RT)

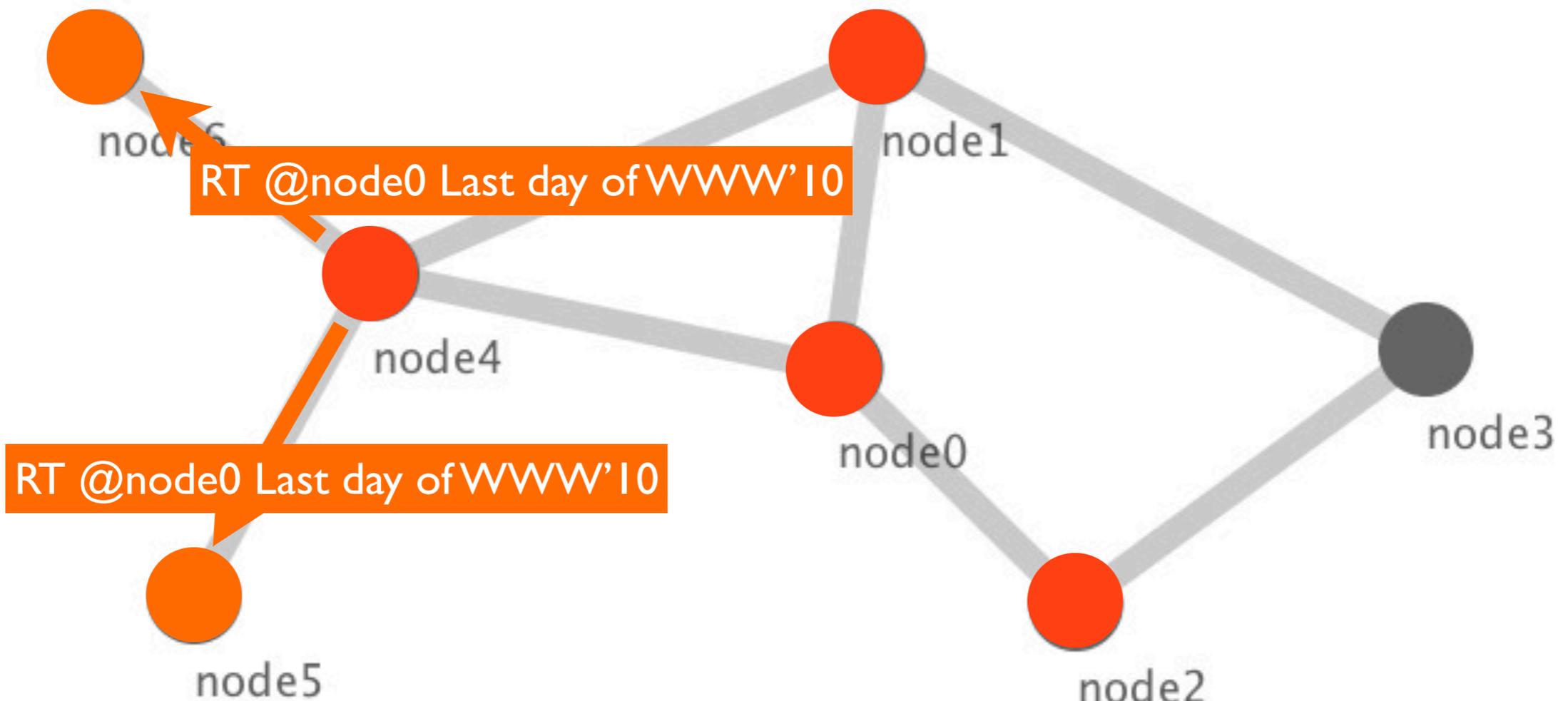
- Relay tweets from a following to followers



Retweet (RT)

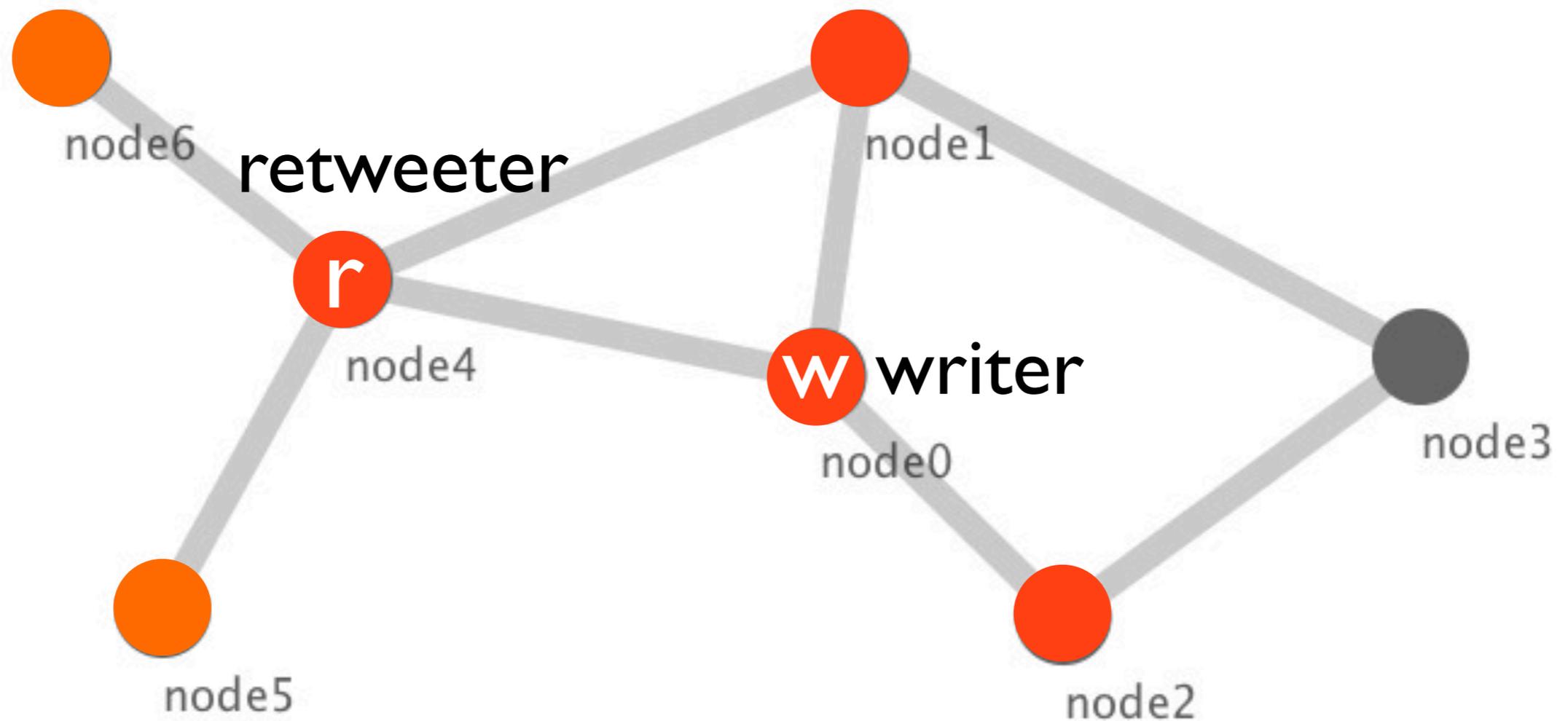
- Relay tweets from a following to followers

RT @node0 Last day of WWW'10



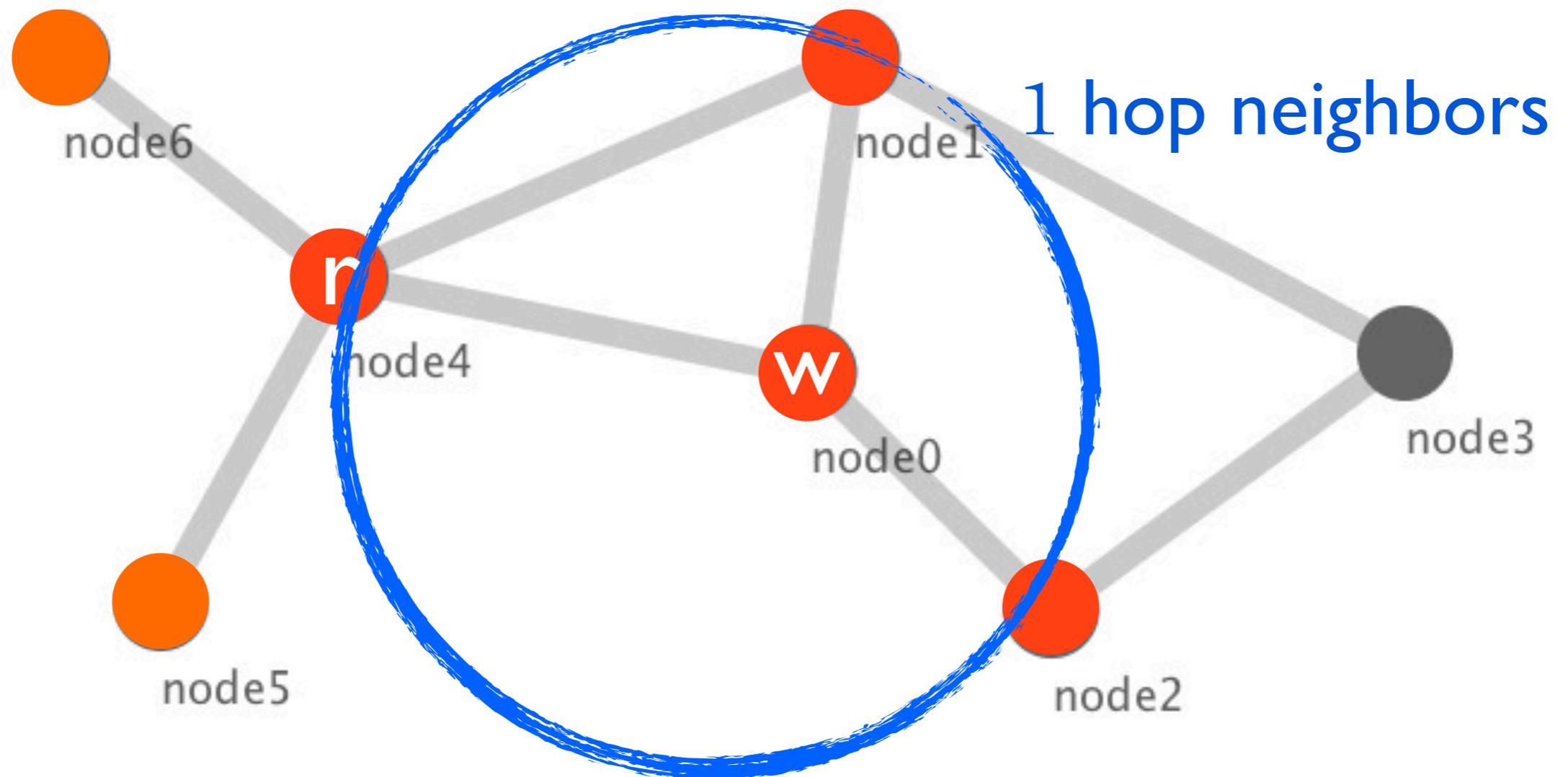
Retweet (RT)

- Relay tweets from a following to followers



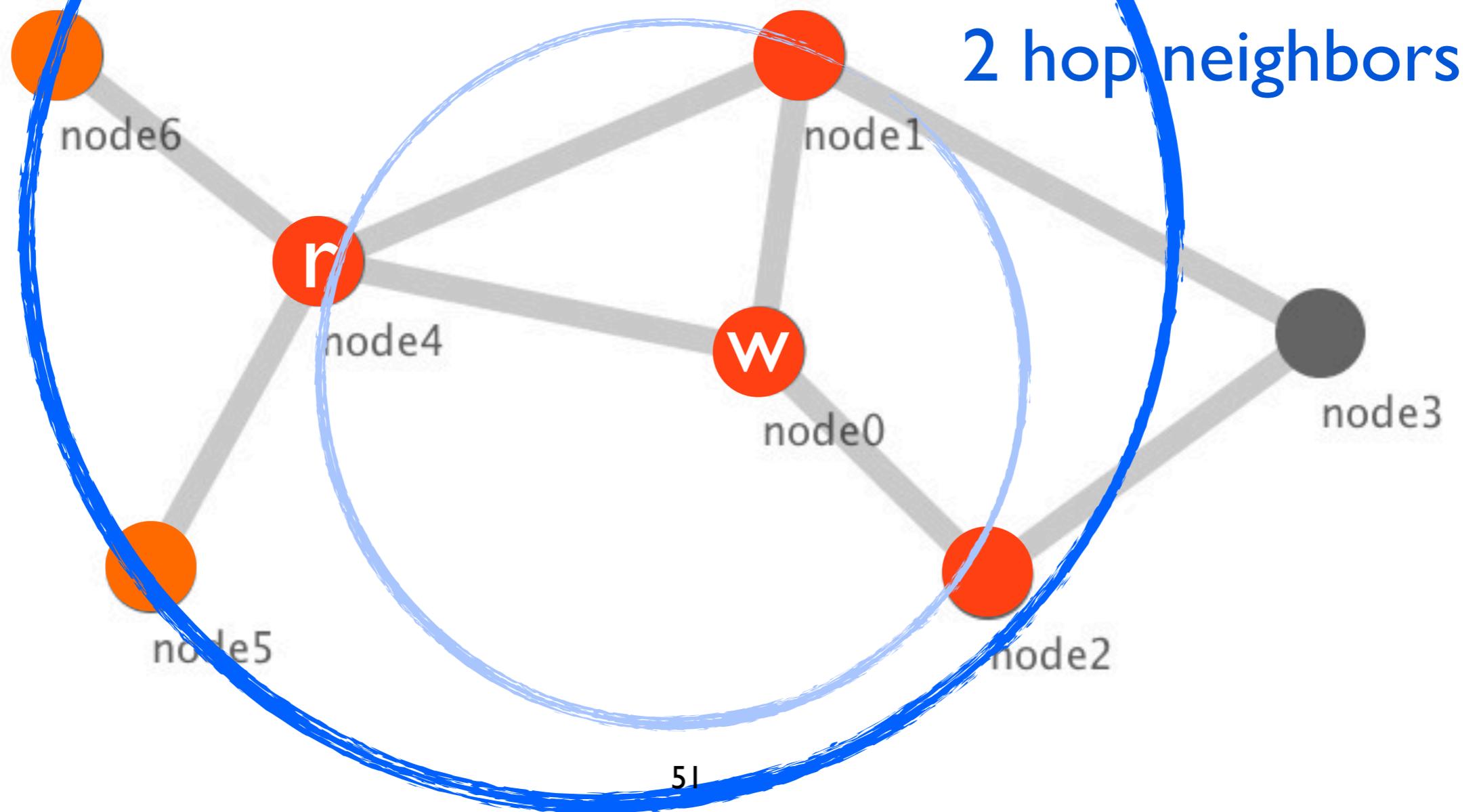
Retweet (RT)

- Not only 1 hop neighbors



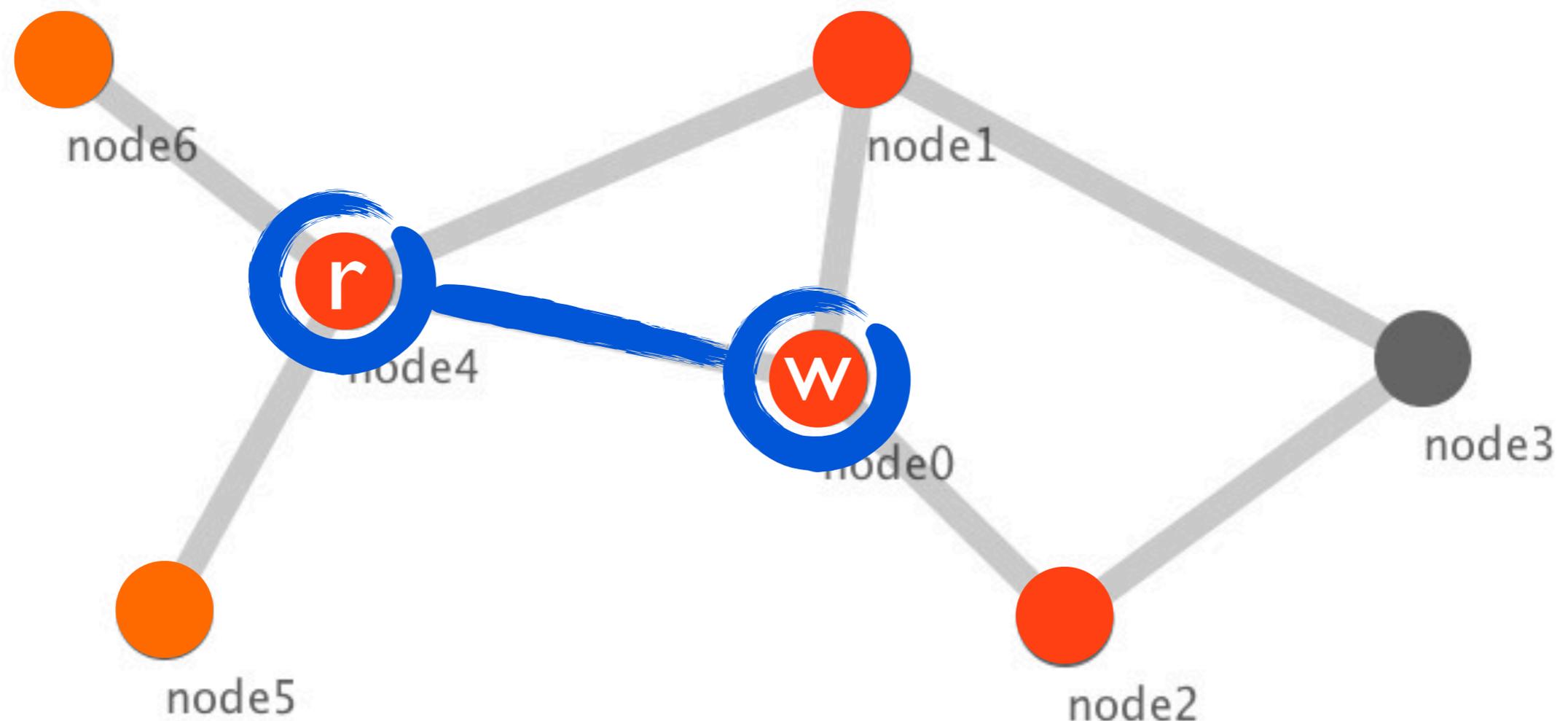
Retweet (RT)

- More goes further

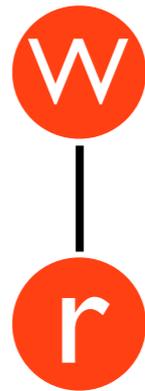


We construct RT tree

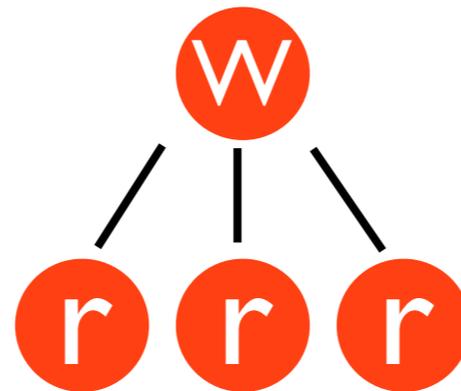
- A tree with writer and retweeter(s)



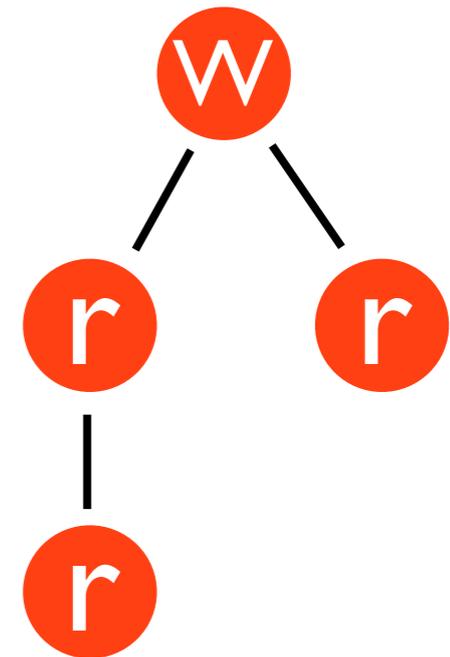
Height of RT trees



1

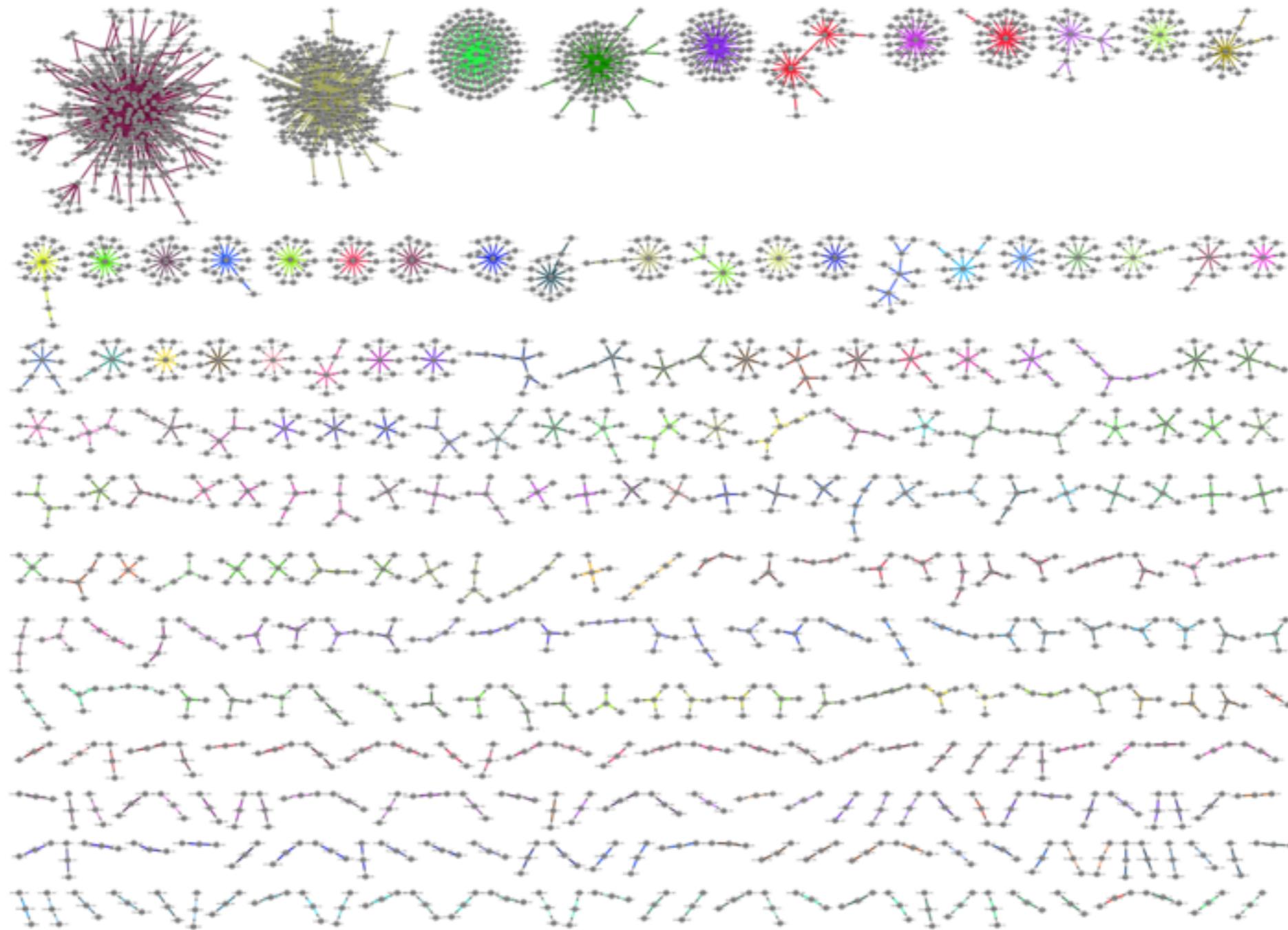


1

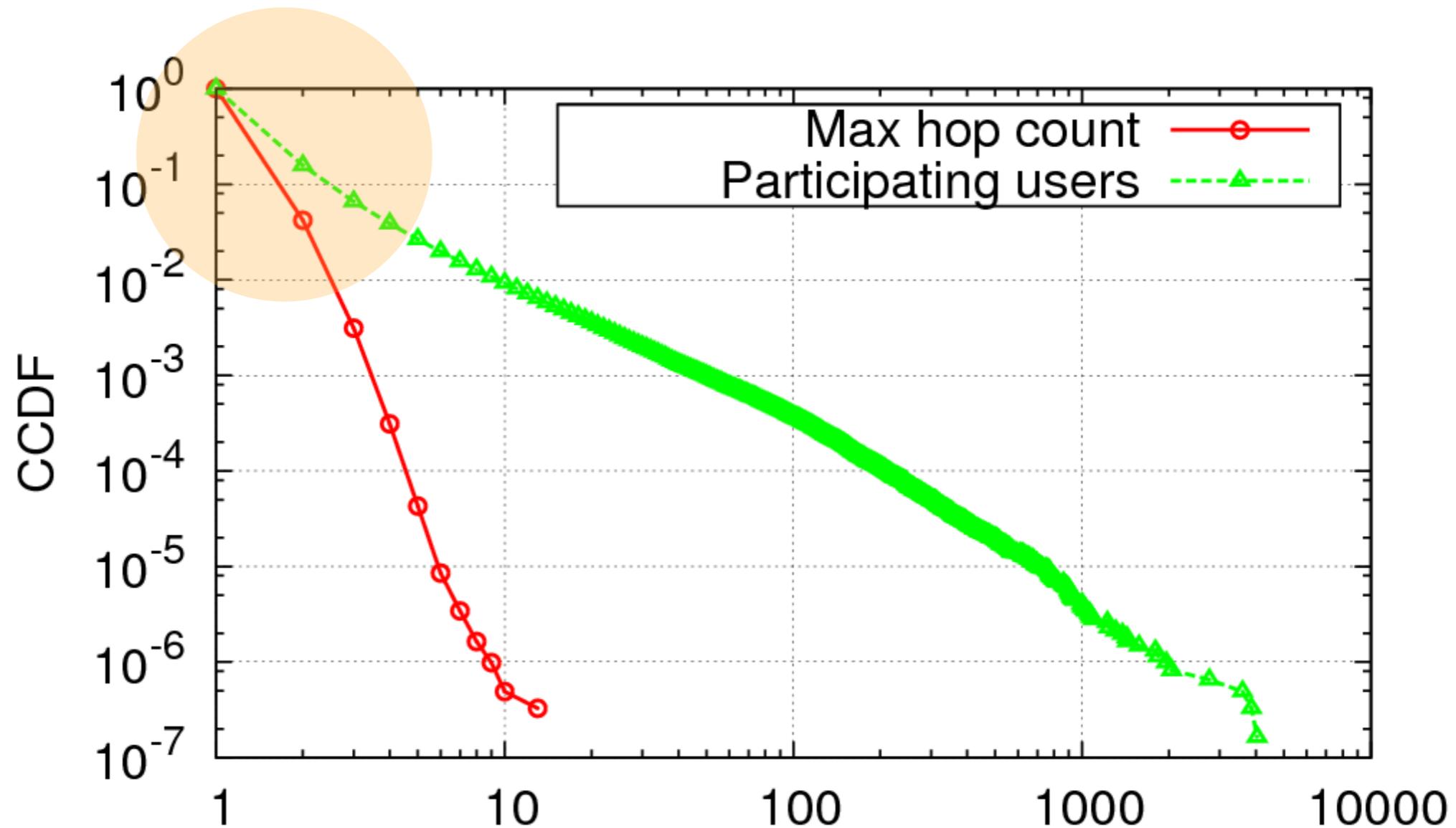


2

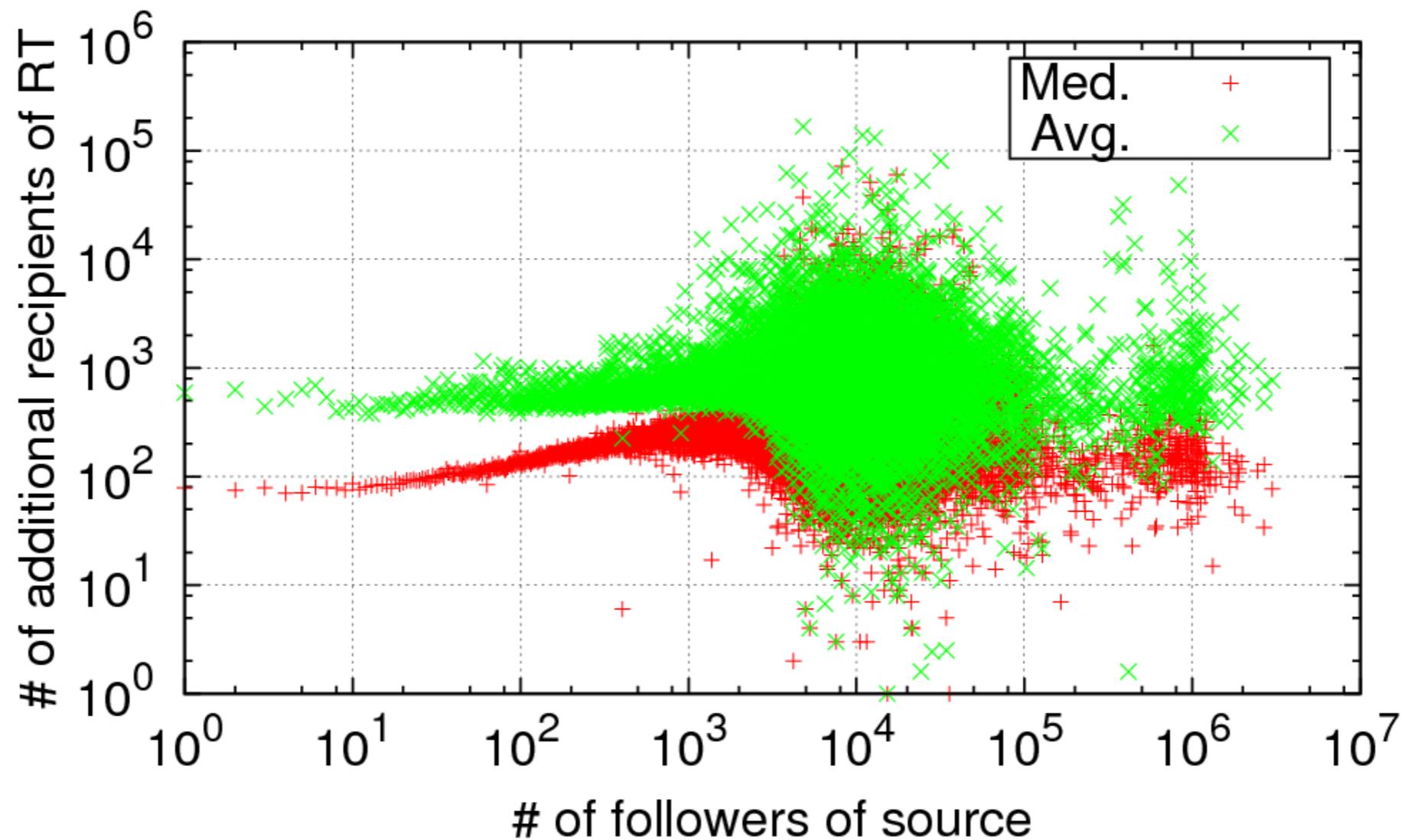
Empirical RT trees



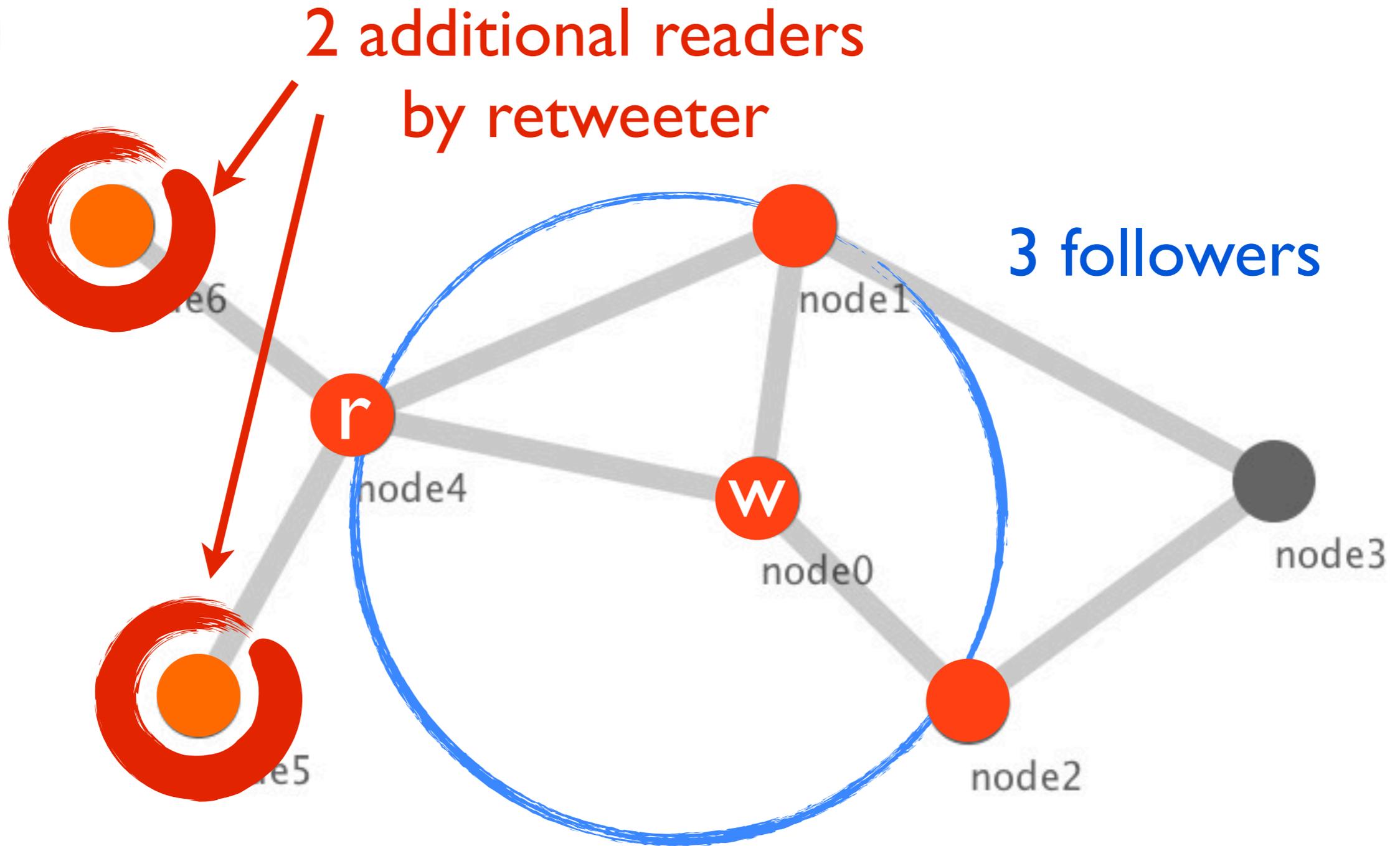
96% of RT trees = Height 1



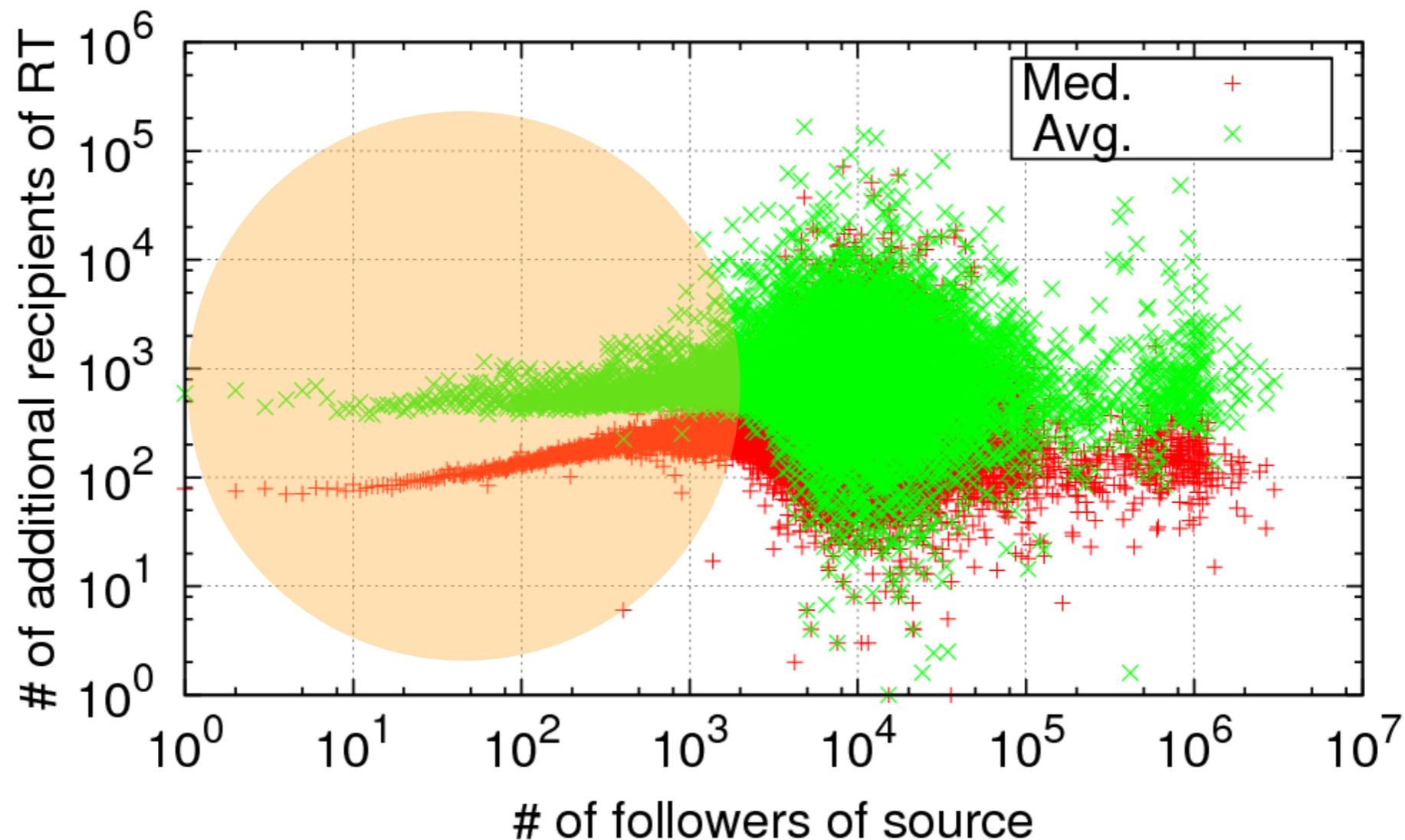
Boosting audience by RT



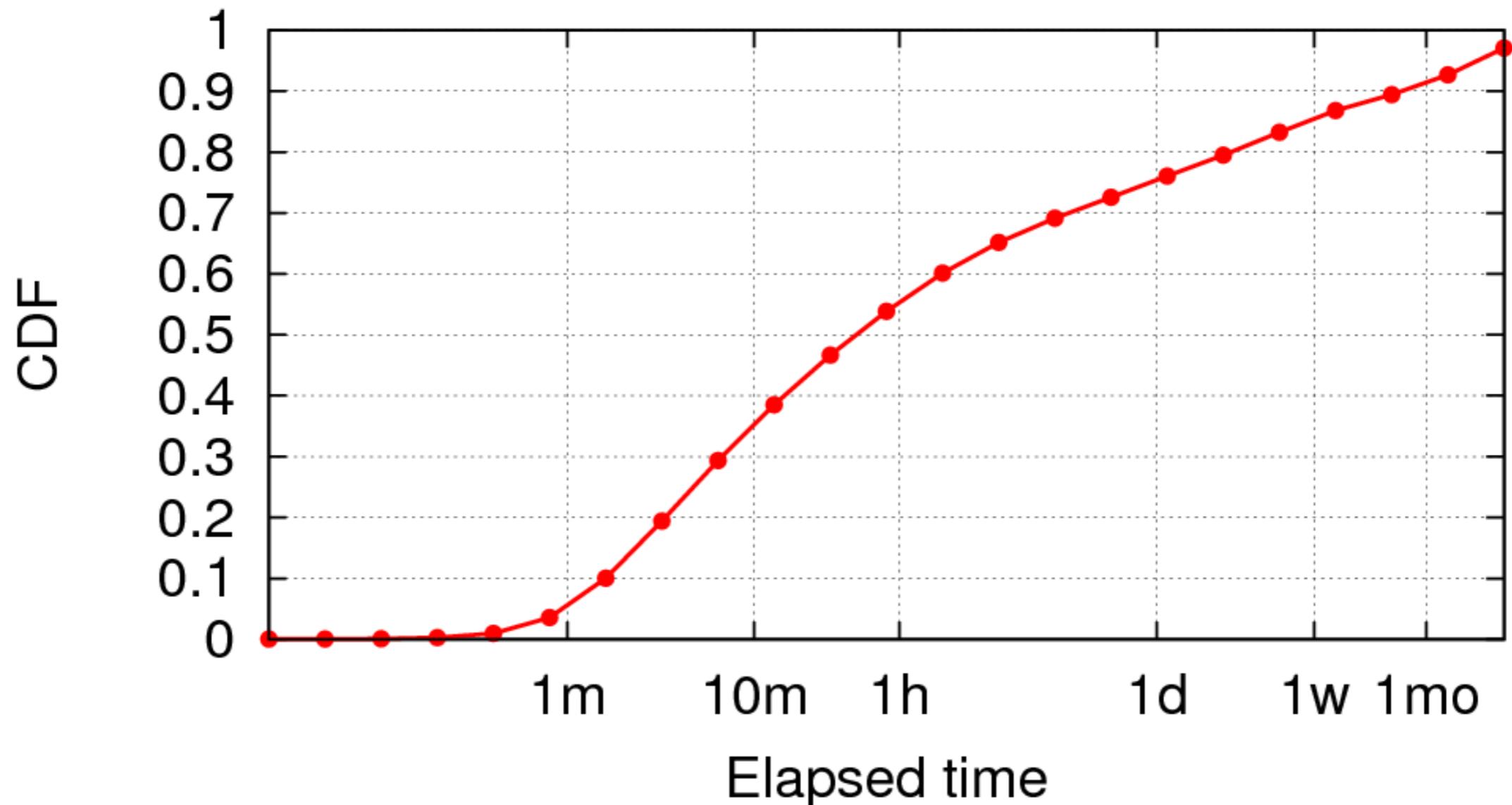
Additional readers



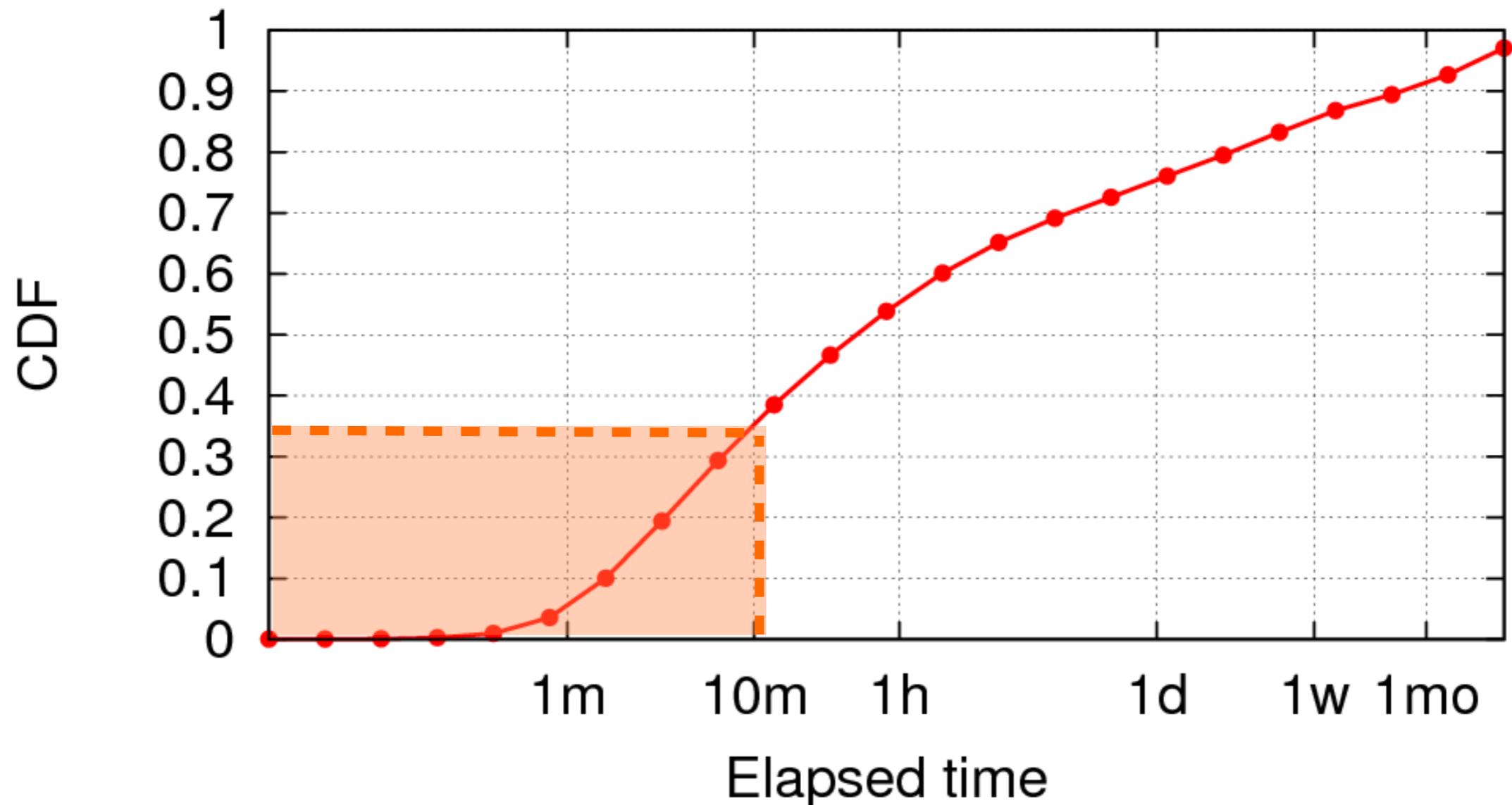
A retweet brings a few hundred additional readers



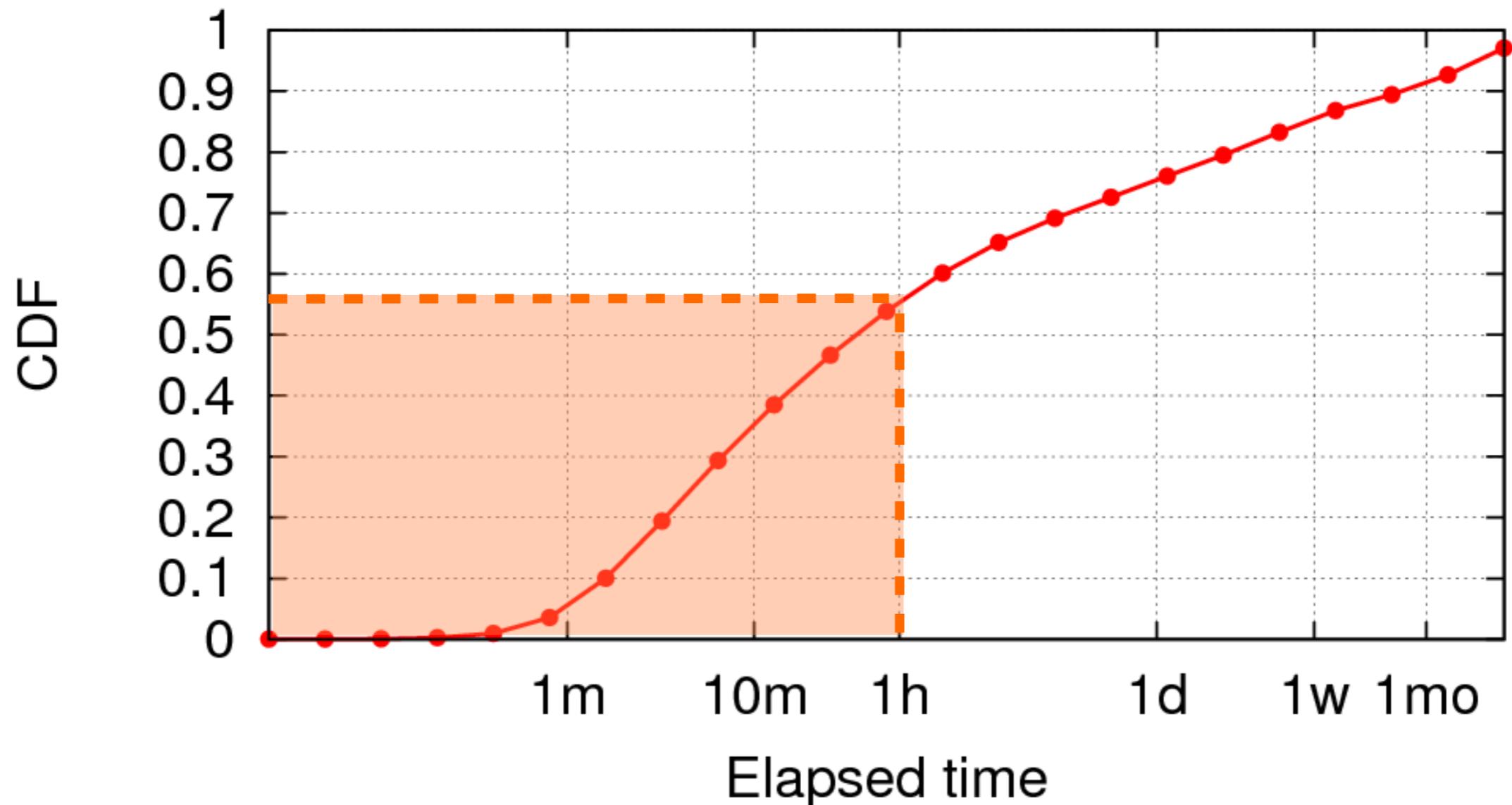
Time lag between hops in RT tree



Fast relaying tweets by RT: 35% of RT < 10 min.



Fast relaying tweets by RT: 55% of RT < 1hr.



Summary

1. We study the entire Twittersphere
2. Low reciprocity distinguishes Twitter from OSNs
3. Twitter has characteristics of news media:
 - ▶ Tweets mentioning timely topics
 - ▶ Plenty of hubs reaching a large public directly
 - ▶ Fast and wide spread of word-of-mouth



Resources

- <http://an.kaist.ac.kr/traces/WWW2010.html>

Supplementary info.

About Twitter

“Twitter is a real-time information network powered by people all around the world that lets you share and discover what’s happening now.”

Twitter asks “what’s happening” and makes the answer spread across the globe to millions, immediately.



About Twitter

“Twitter is a real-time information network powered by people all around the world that lets you share and discover what’s happening now.”

Twitter asks “what’s happening” and makes the answer spread across the globe to millions, immediately.

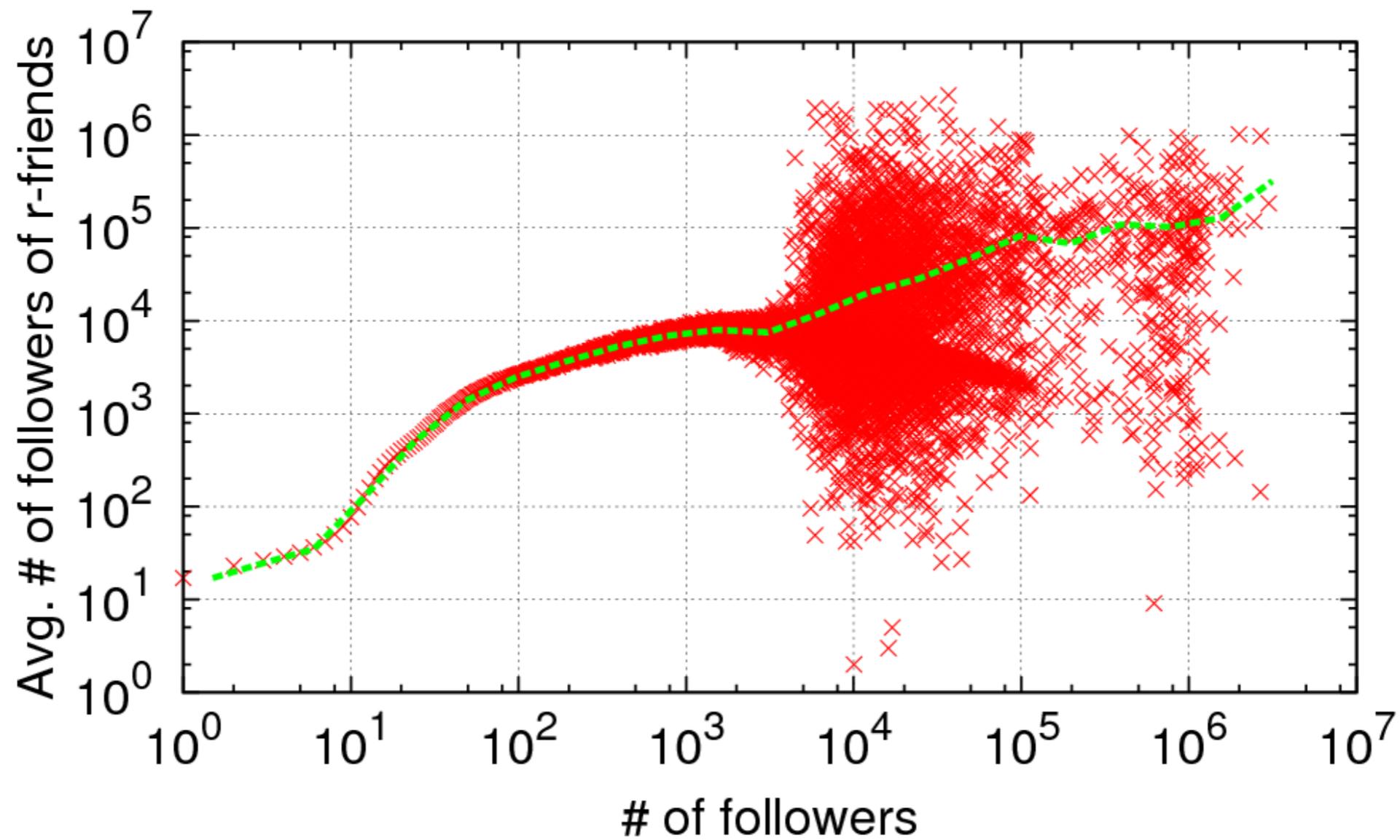


makes the answer spread across the globe to millions, immediately.

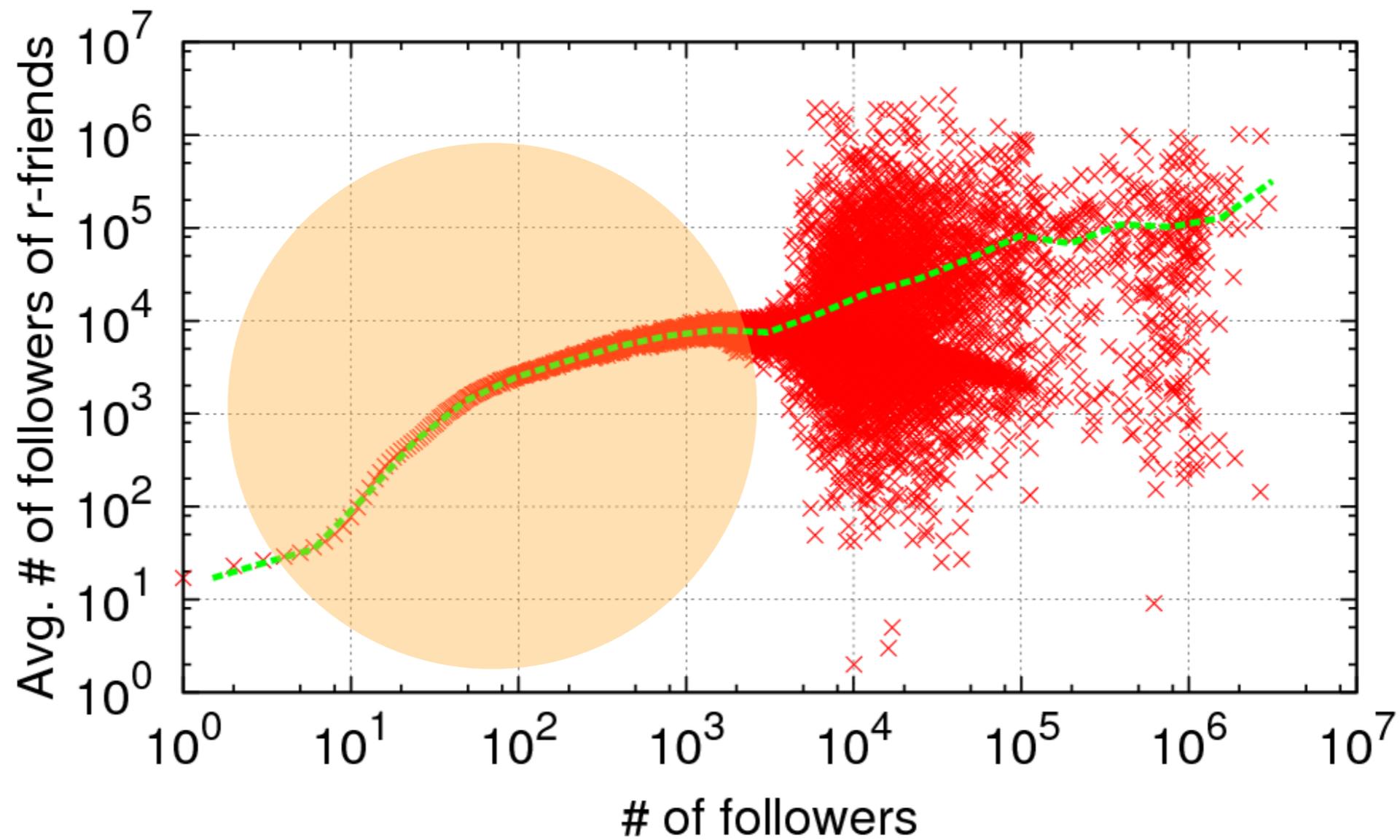
A few numbers

- 105M registered accounts
- 55M tweets a day
- 180M unique visitors a month
- 19B searches a month

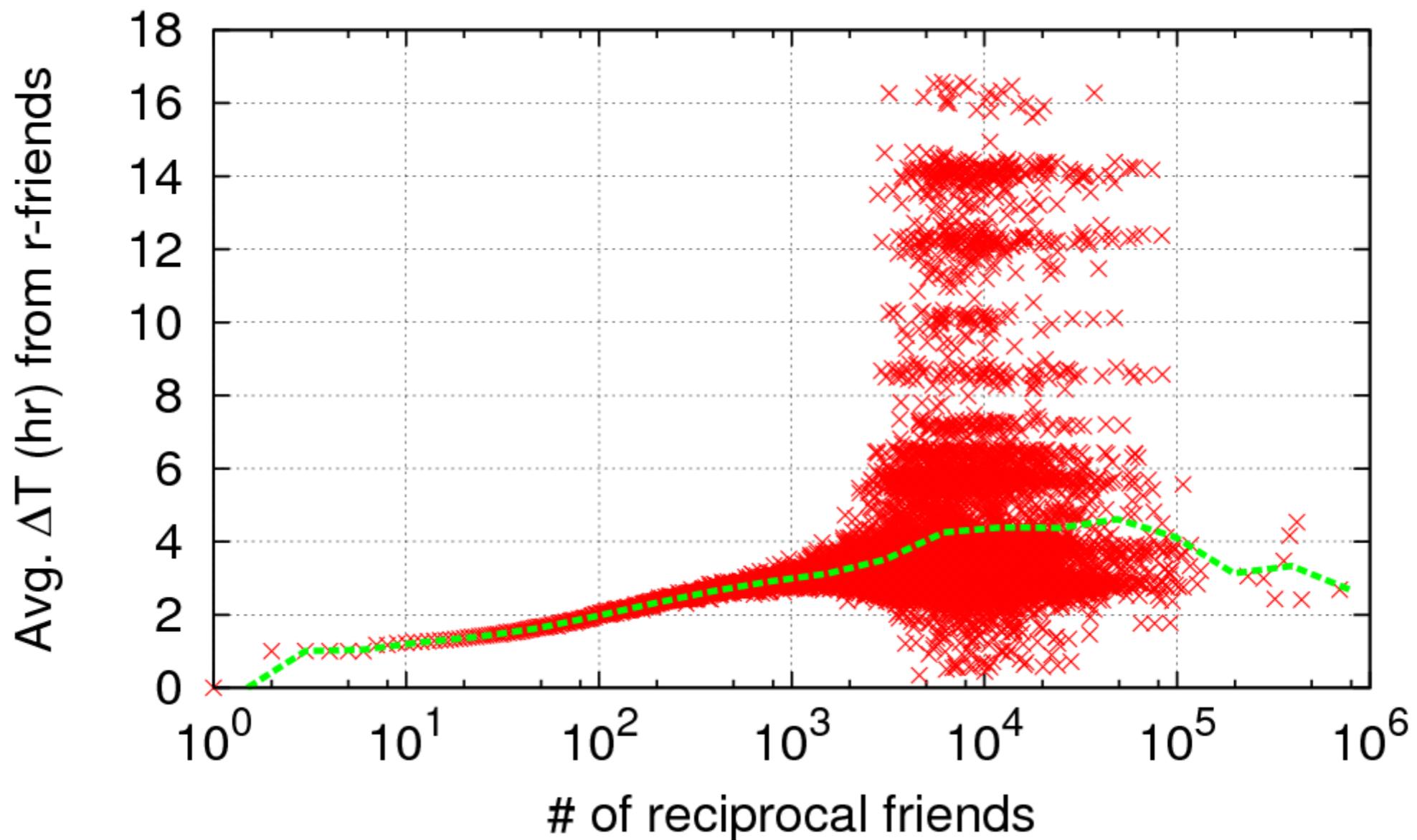
Homophily in terms of followers



Assortative mixing

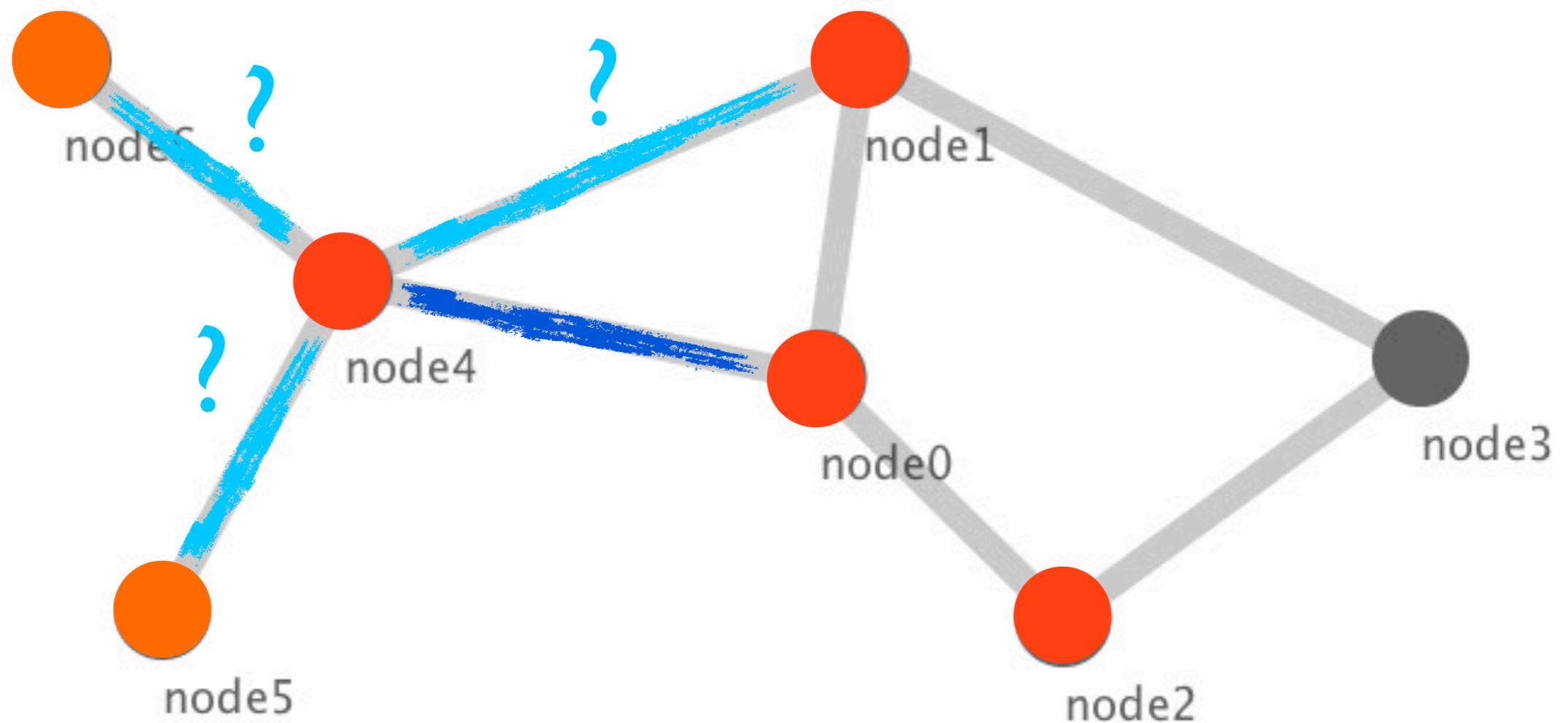


Homophily in terms of location

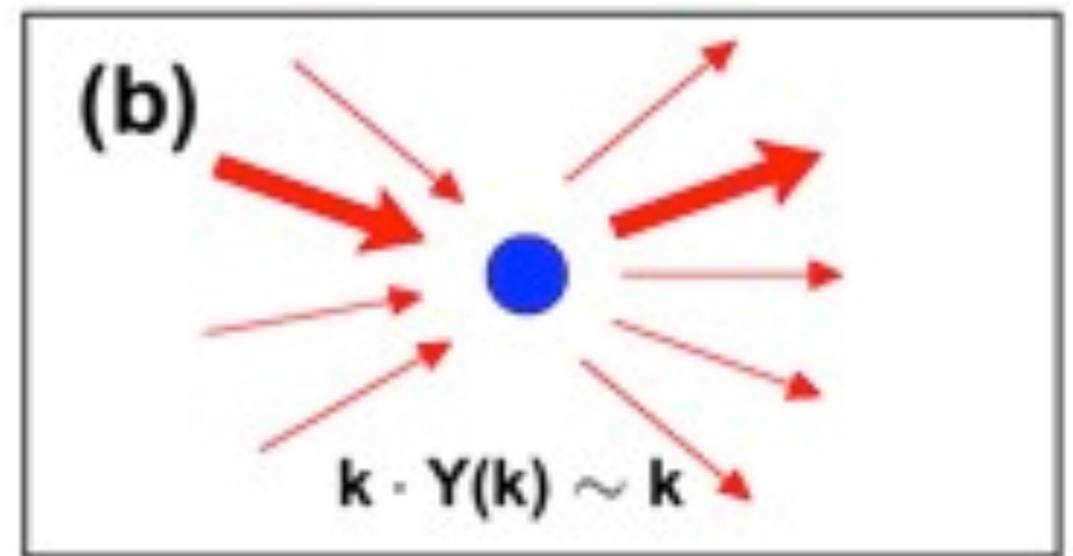
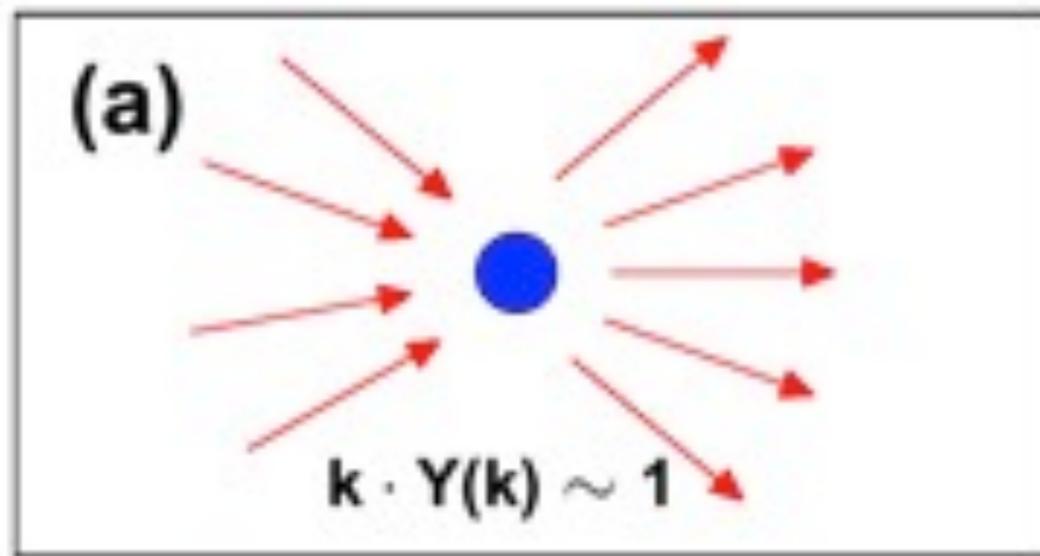


Favoritism in RTs?

- A few informative users?

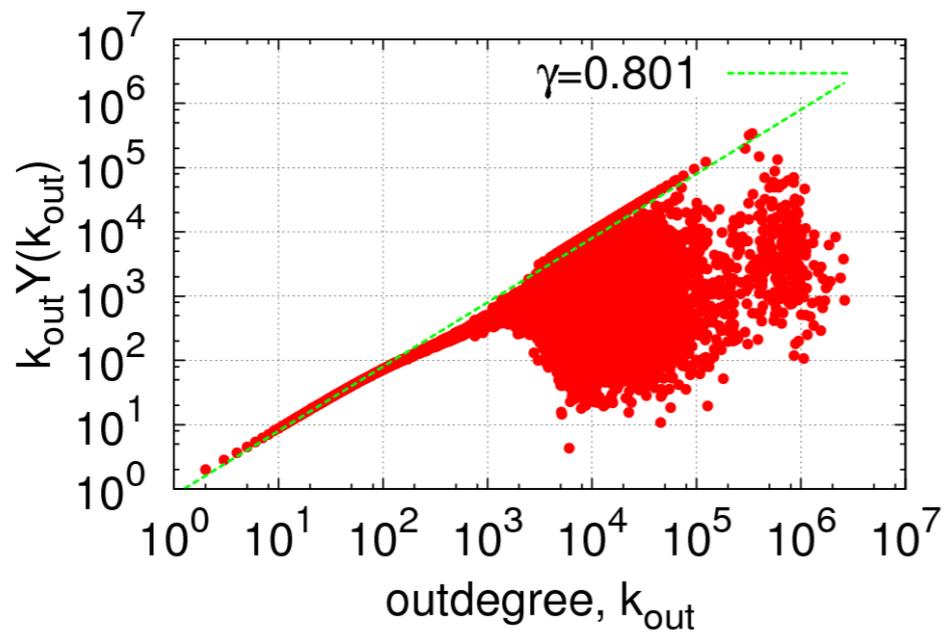


Disparity in weighted network

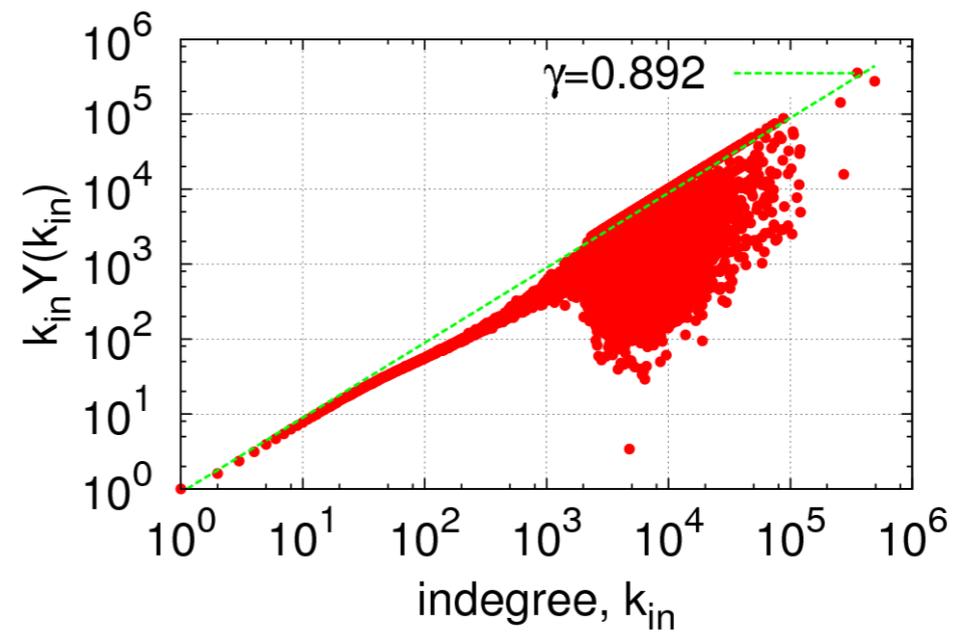


$$Y(k, i) = \sum_{j=1}^k \left\{ \frac{|r_{ij}|}{\sum_{l=1}^k |r_{il}|} \right\}^2$$

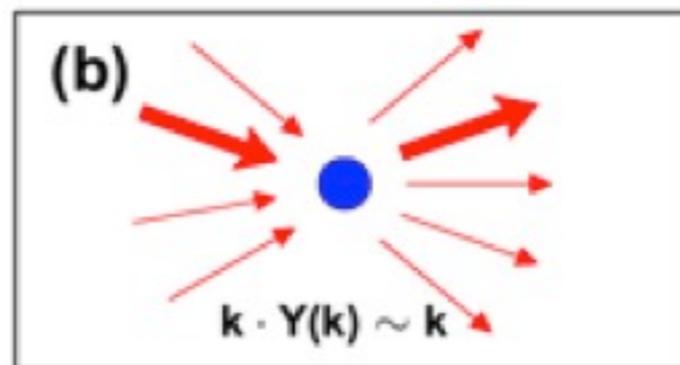
Favoritism in RTs



(a) $k_{out} Y(k_{out}) \sim k_{out}$



(b) $k_{in} Y(k_{in}) \sim k_{in}$



Fast WOM by retweet

