

# Ranking Twitter Discussion Groups

James Cook

Abhimanyu Das



Krishnaram  
Kenthapadi



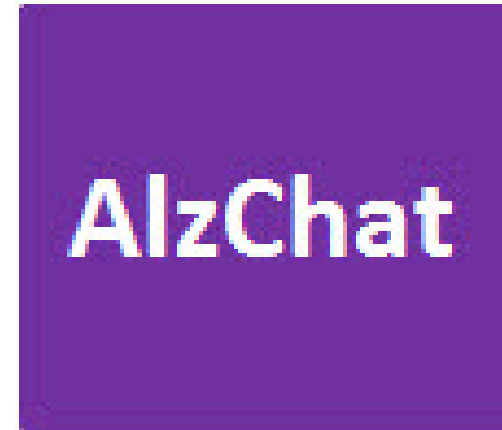
Nina Mishra



COSN 2014

# Outline

- Twitter discussion groups
- Our algorithm
- Theoretical results
- Evaluation



Group Chats on Twitter  
[C, Kenthapadi, Mishra 2013]

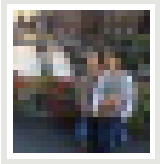
**MOVIE TALK  
ON  
SUNDAY**

**#MTOS**



raghavmodi Raghav

#MTOS hosted by @NitrateDiva in one hour. the topic is suspense <http://t.co/8bvRl6wd>



raghavmodi Raghav

#MTOS hosted by @NitrateDiva in one hour.

the topic is suspense: <http://t.co/8bvRl6wd>



raghavmodi Raghav

#MTOS hosted by @NitrateDiva in one hour. the topic is suspense: <http://t.co/8bvRl6wd>

<http://nitratediva.wordpress.com>

## The Suspense Is Killing Me



1. How do you define suspense in the cinema? As a viewer, do you consider suspense a desirable trait in a film?
- 2A. What is the greatest “suspense film” you’ve ever seen? Why?
- 2B. What’s the best, most suspenseful movie scene or sequence you can think of?

**nitratediva** The Nitrate Diva

**2A. What is the greatest "suspense film" you've ever seen? Why? #MTOS**



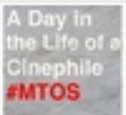
**jimsfilmmodules** James Aston

**2A:**Harakiri (Kobayashi's version)-an ending that ranks amongst 1 of best made all the more memorable by the growing tension throughout **#MTOS**



**kevrockcity** Kevin Koehler

Vertigo because it's perfect. RT @NitrateDiva **2A. What is the greatest "suspense film" you've ever seen? Why? #MTOS**



**movietos** #MTOS

Thank you everyone. **Next week's #MTOS** will have host @Thompson\_film with the topic Film Noir. Do follow him and spread the word. Cheers!

**nitratediva** The Nitrate Diva

**2A. What is the greatest "suspense film" you've ever seen? Why? #MTOS**



**jimsfilmmodules** James Aston

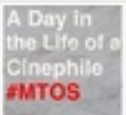
**2A:**Harakiri (Kobayashi's version)-an ending that ranks amongst 1 of best made all the more memorable by the growing tension throughout **#MTOS**



**kevrockcity** Kevin Koehler

Vertigo because it's perfect. RT @NitrateDiva **2A. What is the greatest "suspense film" you've ever seen? Why? #MTOS**

**Next week's #MTOS will have host @Thompson\_film with the topic Film Noir.**



**movietos** #MTOS

Thank you everyone. **Next week's #MTOS** will have host @Thompson\_film with the topic Film Noir. Do follow him and spread the word. Cheers!





**StephenS**

@StephenSType1

It's that time for me. Closing: I'd be a very unhappy person without this hour every week, topic or not. Thanks @SweeterCherise #dsma



**Elaine Cook**

@vivizaraz

Social media enables me to access a support network, both those I know in RL and online contacts. Having child with ASD v isolating #mhchat

Find group discussions about:

movies



1. #MTOS
2. #FilmCurious
3. #DriveInMob

Sort by...

# tweets with “movie”?

Fraction of tweets with “movie”?

# users who tweet “movie”?

# Related Work

- Group Chats on Twitter  
[CKM 2013]

*Algorithms for finding group chats*

This work: *Ranking*

# Related Work

- Group Chats on Twitter

[CKM 2013]

- Search in Online Forums

[Elsas, Carbonell 2009] [Cong et al. 2008]

*Finding forum threads*

This work: *Finding discussion groups.*

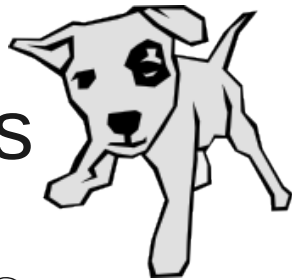
# Related Work

- Group Chats on Twitter  
[CKM 2013]
- Search in Online Forums  
[Elsas, Carbonell 2009] [Cong et al. 2008]
- PageRank [Brin, Page 1998], HITS [Kleinberg 1998]

Sprockets

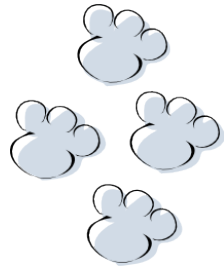


# talkSprockets



@ alice

# sprockz



@ bob

# sprocketChat



@ carol

Sprockets



**Stationary Distribution:**

**Final Ranking:**

$\Pr[\#\text{talkSprockets}] = 0.3$

$\Pr[\#\text{sprockz}] = 0.2$

$\Pr[\#\text{sprocketChat}] = 0.5$

$\#\text{sprocketChat}$

$\#\text{talkSprockets}$

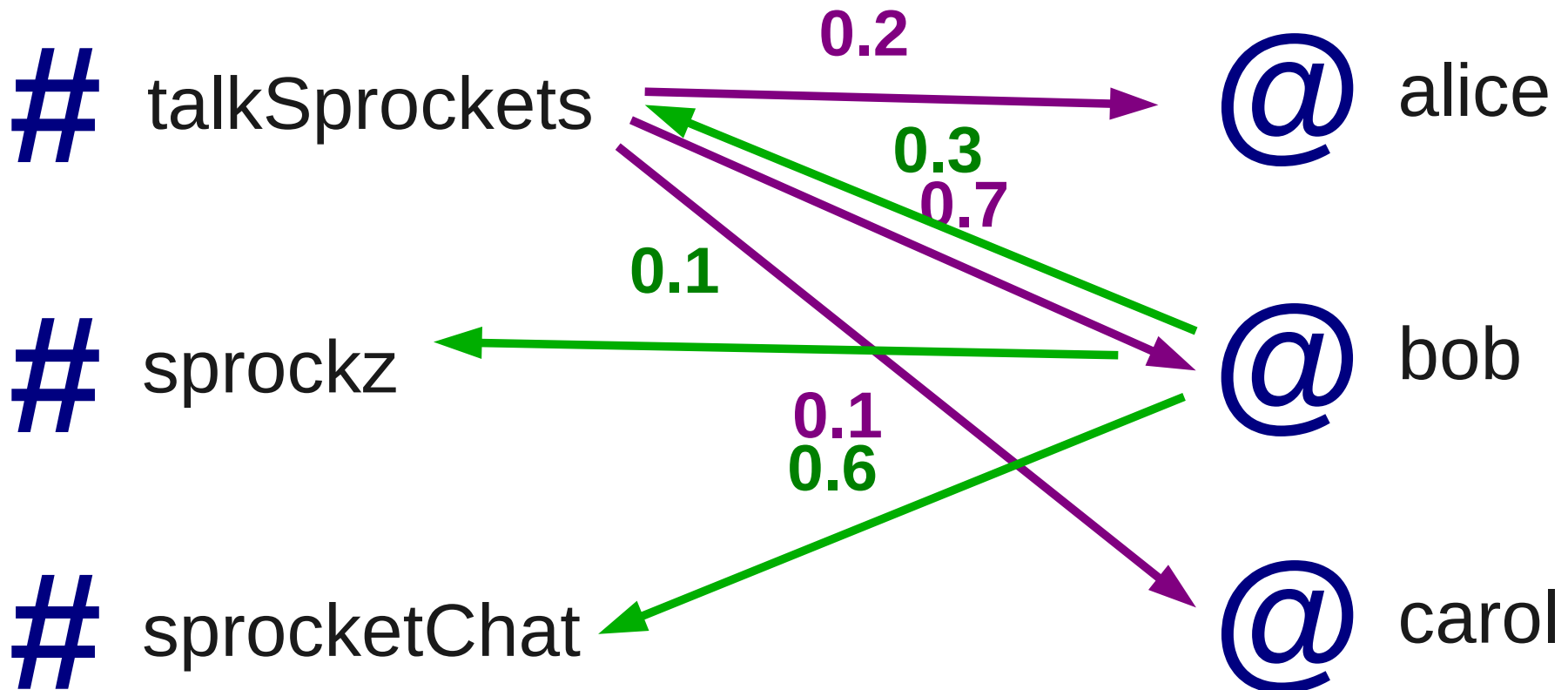
$\#\text{sprockz}$

$$M_{gh} = \begin{bmatrix} \dots \end{bmatrix}$$

$$\sum_u A_{gu} P_{guh}$$

Authority Scores  $A_{gu}$

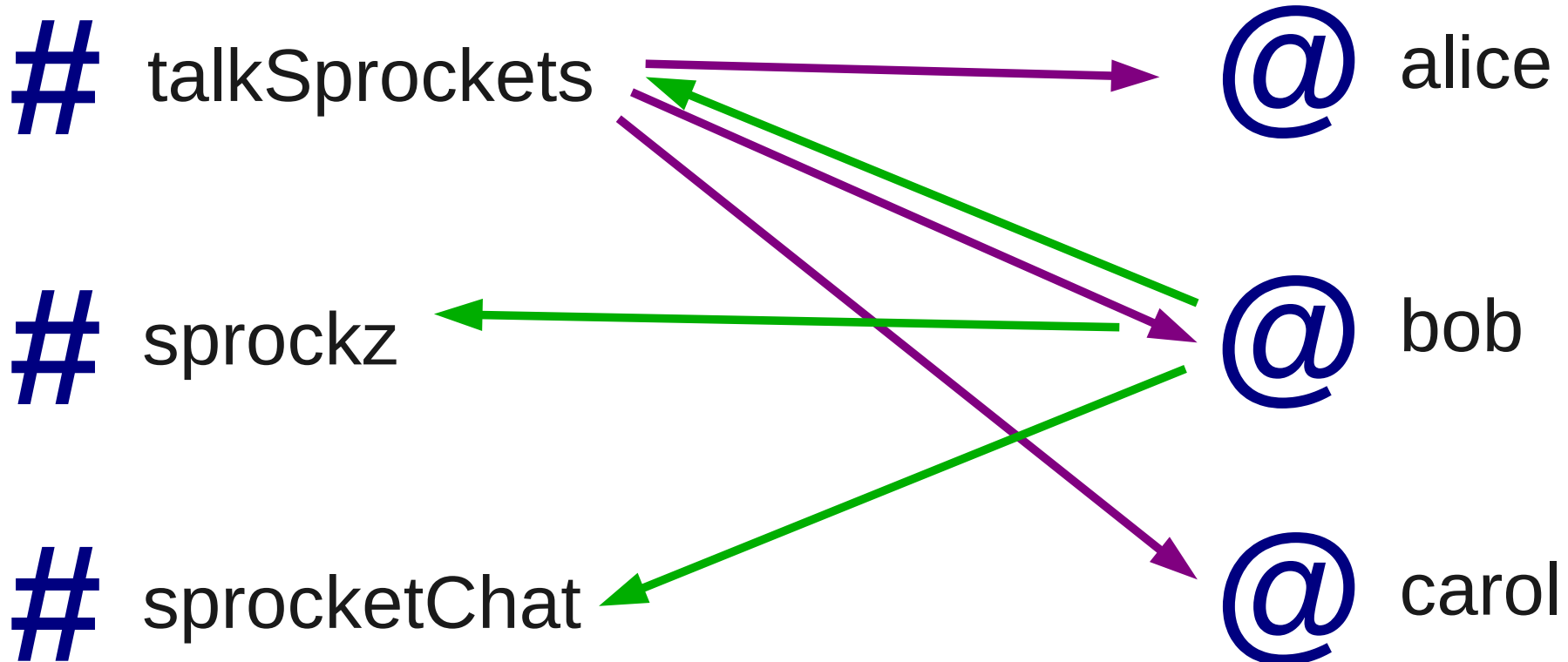
Preference Scores  $P_{guh}$





$$M_{gh} = \lambda \boxed{D_h} + (1 - \lambda) \sum_u A_{gu} P_{guh}$$

Teleport Distribution  $D_h$  Authority Scores  $A_{gu}$   
 Preference Scores  $P_{guh}$



# Group Preference Model

$$M_{gh} = \lambda D_h + (1 - \lambda) \sum_u A_{gu} P_{guh}$$

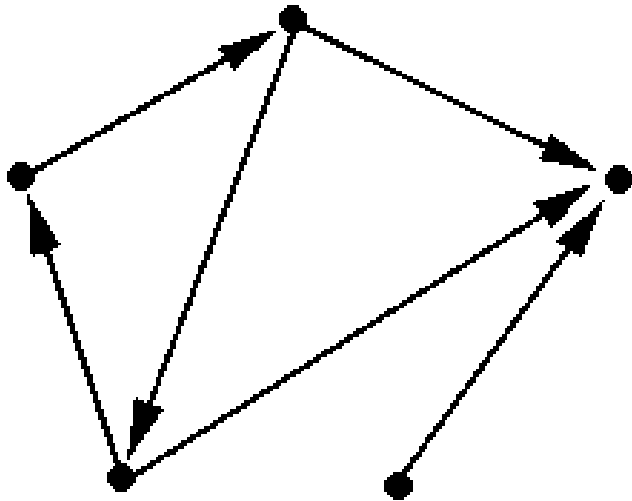
Find stationary distribution  $\pi$

Rank  $g > h$  if  $\pi_g > \pi_h$

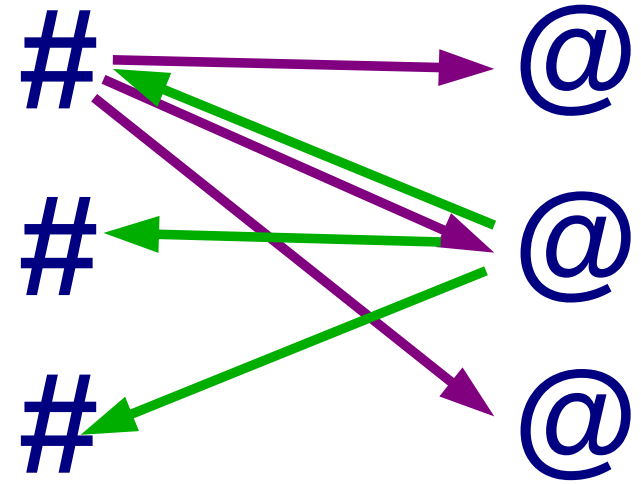
**DISCLAIMER:**

Use only for ranking.  
Not a model of reality.

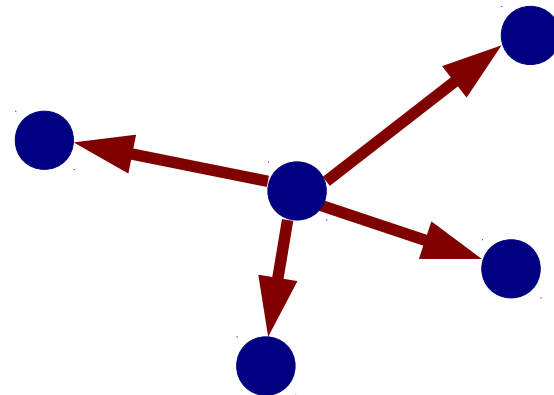
# Group Preference Model



# Hubs and Authorities



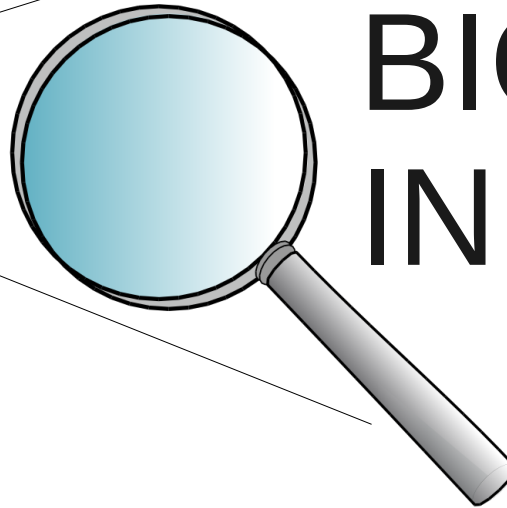
# Random Surfer Model (PageRank)



# Stability

# Stability

small change in input



## BIG CHANGE IN RANKING?

- PageRank and HITS are unstable.



- Our algorithm is also unstable.

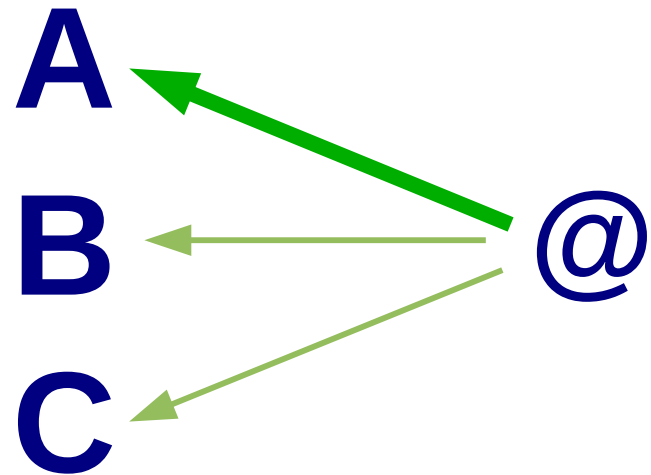


# Stability

## Theorem

If we increase one user's preference for group A (at the expense of other groups) then A's rank will not go down.

[Chien, Dwork, Kumar,  
Simon, Sivakumar 2003]



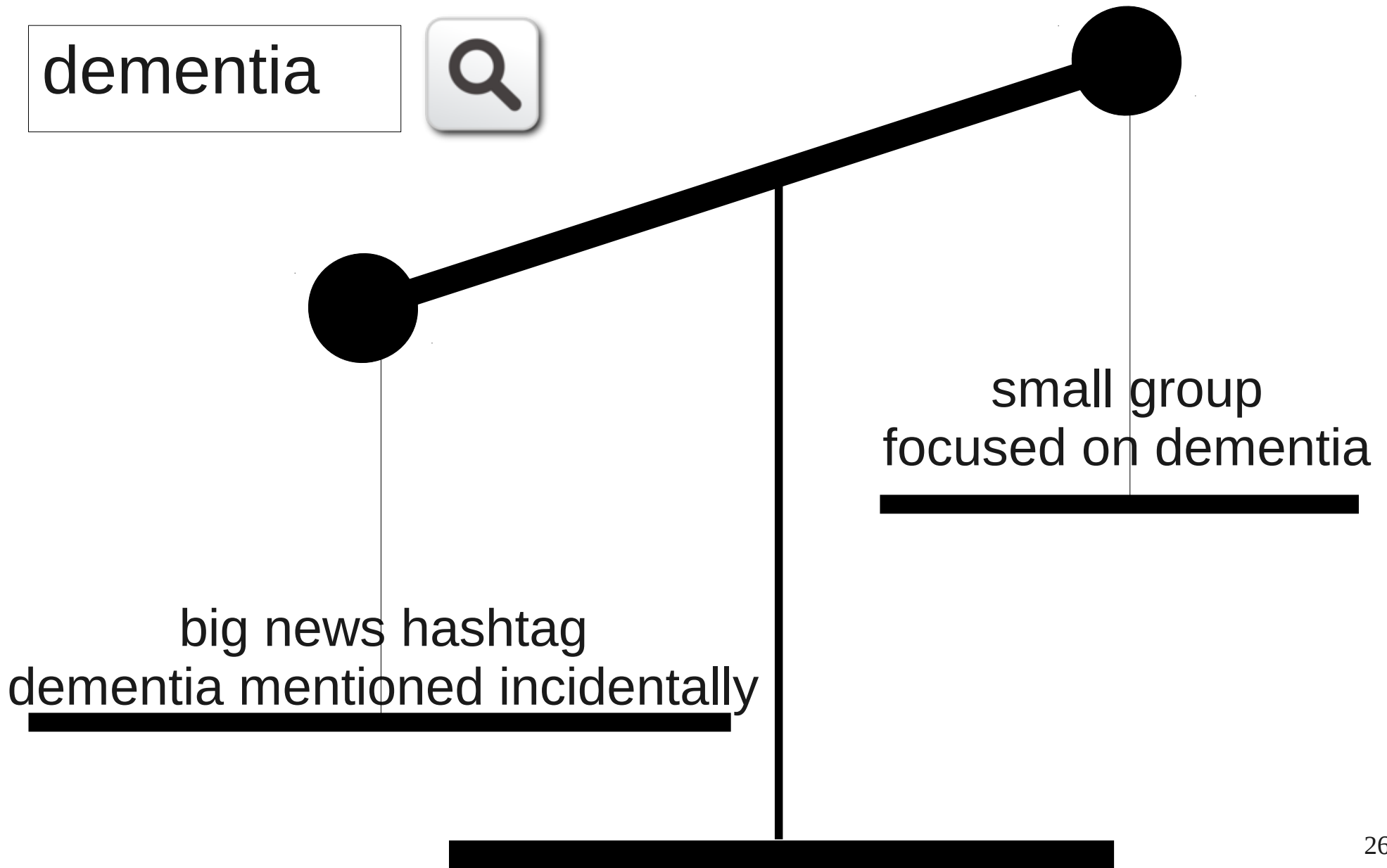
Rank by # times query occurs?

dementia



# Rank by # times query occurs?

dementia

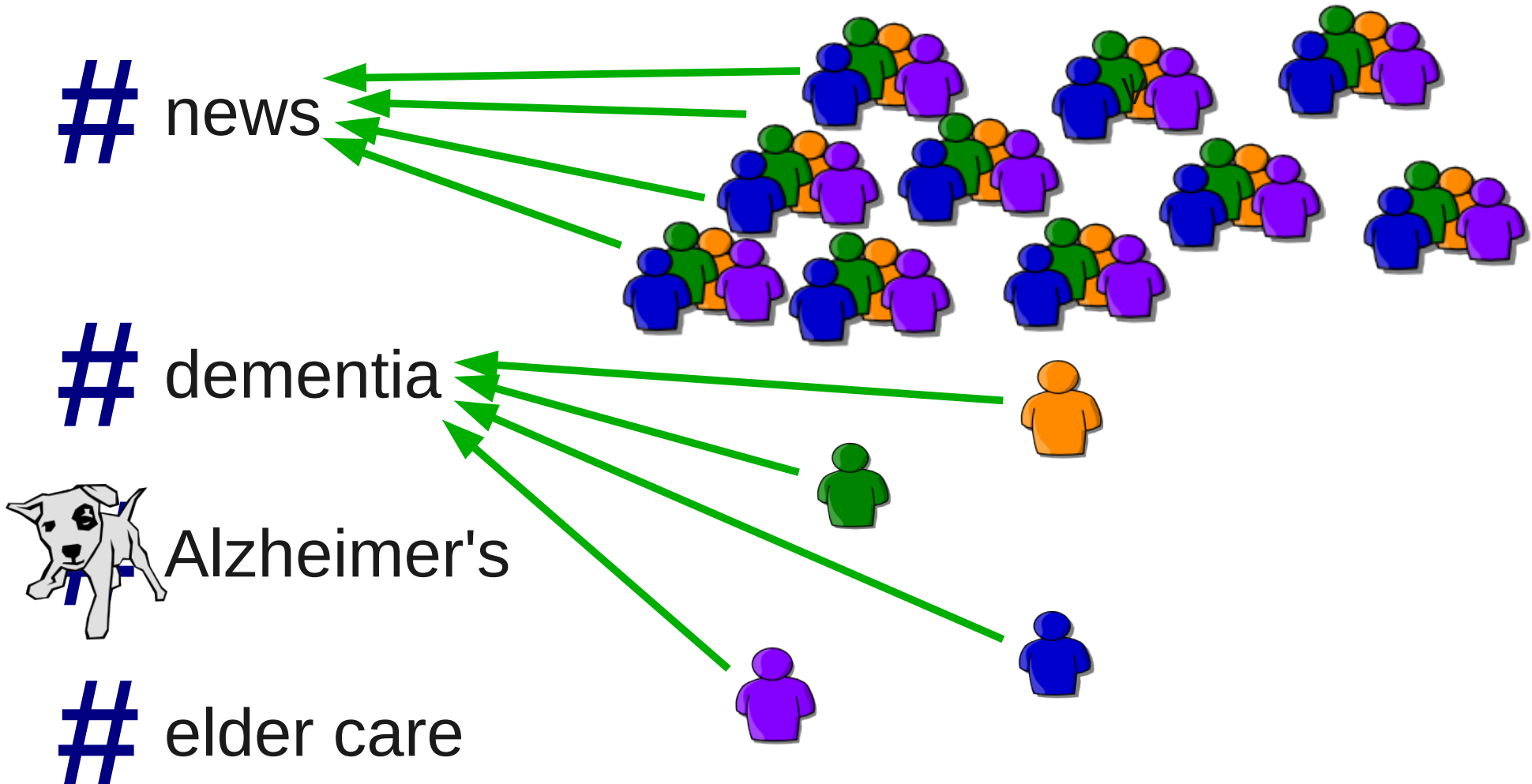




# Example

**Theorem:** Dementia chat ranked at top.\*

\*(Assuming the teleport distribution is uniform.)



# Evaluation

Baseline algorithms:

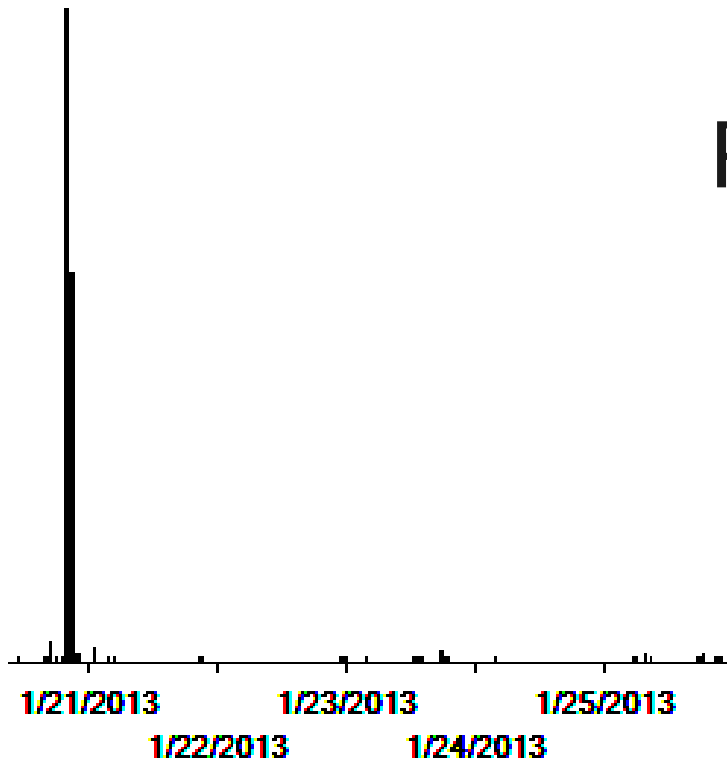
- # tweets with query
- Fraction of tweets with query
- # users who tweet with query

# Evaluation

- Queries
- Ground Truth
- Dataset of group discussions

# Evaluation: Dataset

One year of tweets



Require at least 10 meetings

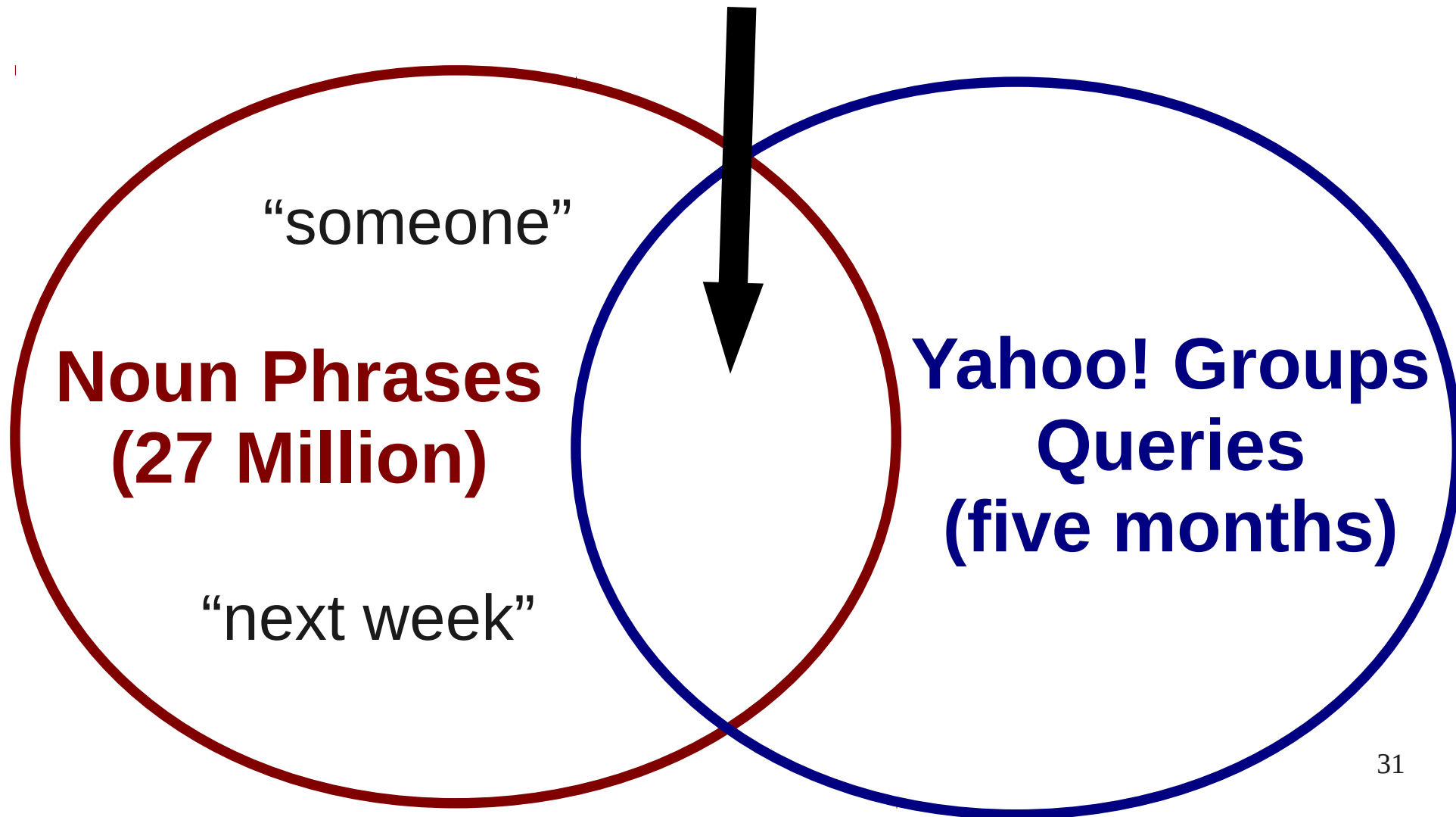


27K group discussions

One week of #MTOS

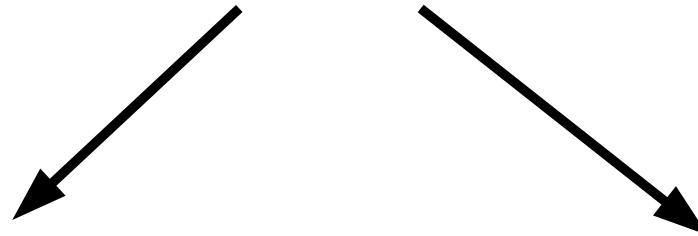
# Evaluation: Queries

2000 Test Queries



# Evaluation: Ground Truth

“Experts” — Query appears in profile text



~~2000~~ 600 Queries

Poor Quality

# Evaluation: Ground Truth

“Experts”

1. #

2. #

3. #

Algorithm 1

1. #

2. #

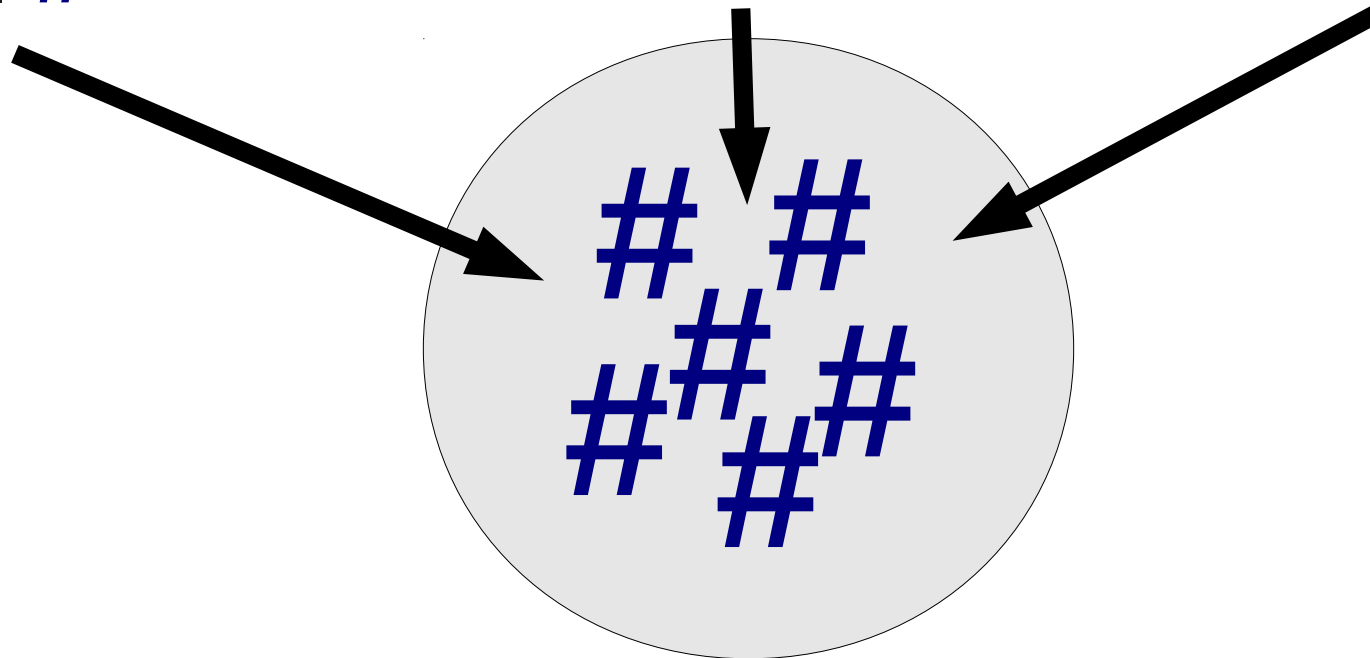
3. #

Algorithm 2

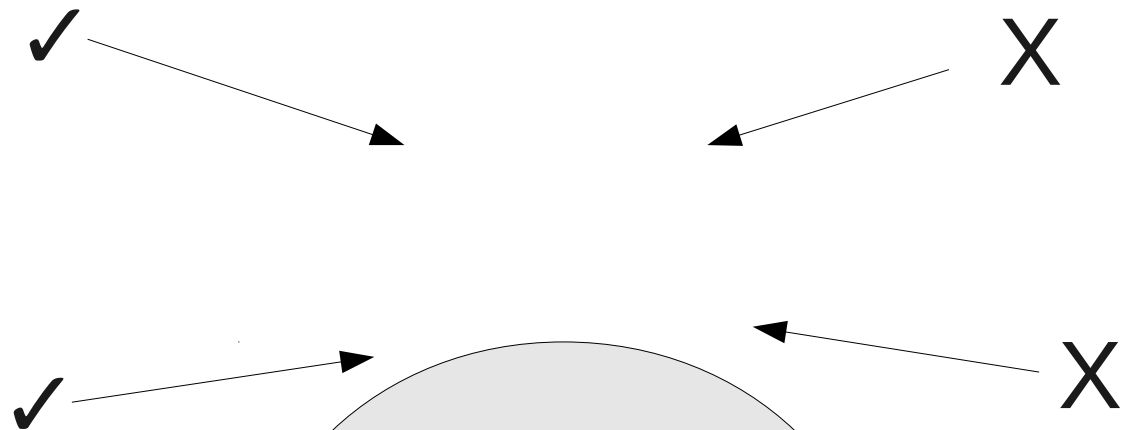
1. #

2. #

3. #



# Evaluation: Ground Truth



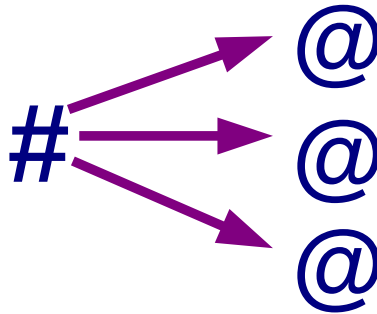
Evaluate by hand: ~~600~~ 50 queries



# Results

	Precision@5	Recall@5
<b>Group Preference Model</b>	<b>0.40</b>	<b>0.49</b>
# distinct users	0.24	0.28
# tweets	0.31	0.36
Fraction of tweets with query	0.27	0.38
( <i>“Experts”</i> )	(0.53)	(0.71)

# Computing Authority Scores



Method	Precision@5	Recall@5
<b># tweets with query</b>	<b>0.40</b>	<b>0.49</b>
# @-mentions with query	0.38	0.47
# followers	0.38	0.47
uniform	0.40	0.48

# Summary

We designed the Group Preference Model, and found good theoretical and experimental results.

# Future Directions

- Which groups are easy to join?
- Different types of query
- Personalized ranking
- Groups are always changing
- Put it online!

Thanks!