

Taming the long tail

Identify Filtering in Social Media

Avner May, Nitish Korula, Silvio Lattanzi
A. Chaintreau
Columbia University, Google

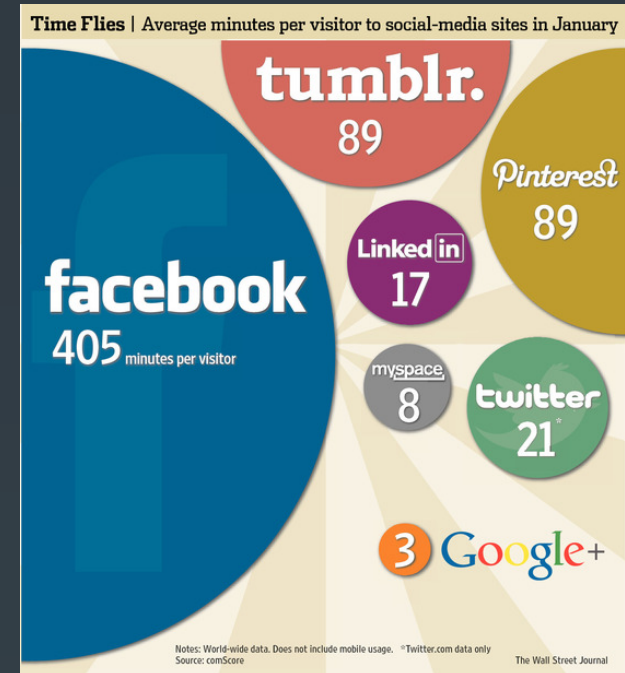
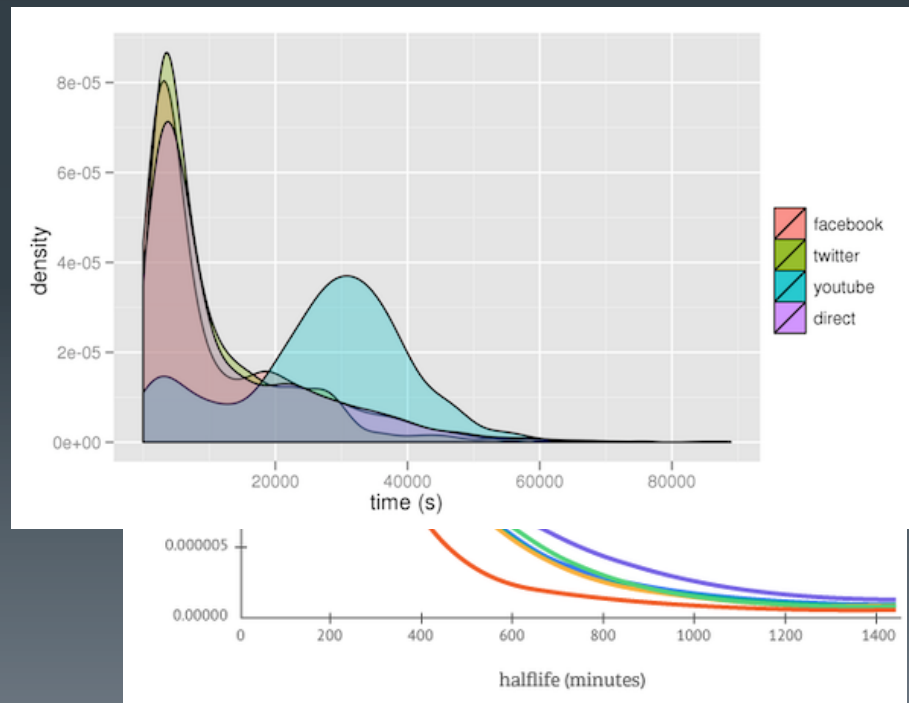


joint work with Avner, Nitish and Silvio

Are social media sustainable?

From the trenches: no!

- Users's dilemma
 - May I be missing something?



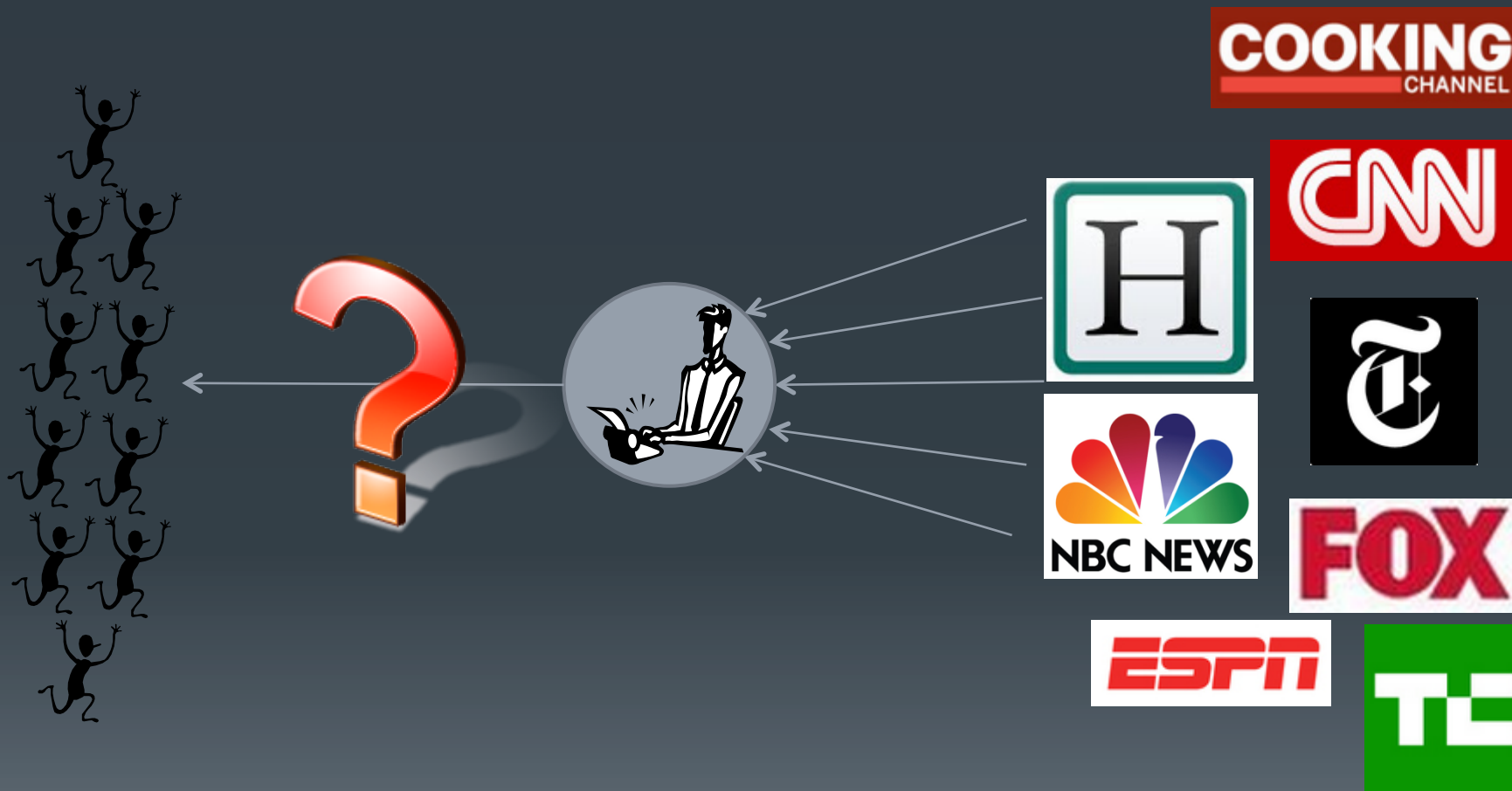
- Content producers
 - May I be missing my audience?

From the faculty lounge: of course!

- **Socializing** is essential for information
 - To find about jobs [Gr74], innovation [CKM57]
“It pays to know / It hurts to be unaware.”
- When looking for good content, most of the time is wasted, but some gems are priceless
 - This process is more efficient **collectively**
 - And curating is at least informally **rewarded**
- In this talk, we focus on **news** dissemination

What is the role of intermediaries?

Understanding these intermediaries



2013: two interesting works

Twitter “precision”
40.5% average
— Encouraging!



On the Precision of Social and Information Networks

Reza Bosagh Zadeh
Stanford University
rezab@stanford.edu

Ashish Goel
Stanford University
ashishg@stanford.edu

Kamesh Munagala
Duke University
kamesh@cs.duke.edu

Aneesh Sharma
Twitter, Inc
aneesh@twitter.com

ABSTRACT

The diffusion of information on online social and information networks has been a popular topic of study in recent years, but attention has typically focused on speed of dissemination and recall (i.e. the fraction of users getting a piece of information). In this paper, we study the complementary

aspects of information diffusion: the number of users who are interested in a piece of information every day. There are many hypotheses as to the source of their popularity, and one popular hypothesis relates to the effectiveness of these networks as information dissemination mechanisms [10, 28]. In particular, a fundamental question about effectiveness is one of personalization: given the large number of users, one would expect them to be interested in a diverse set of content, and the network

Self-organizing Flows in Social Networks

Nidhi Hegde^{1,*}, Laurent Massoulié², and Laurent Viennot^{3,**}

¹ Technicolor, France

nidhi.hegde@technicolor.com

² Microsoft Research - INRIA Joint Centre, France

laurent.massoulie@inria.fr

³ INRIA - Paris Diderot University, France

laurent.viennot@inria.fr

Abstract. Social networks offer users new means of accessing information, essentially relying on “social filtering”, i.e. propagation and filtering of information by social contacts. The sheer amount of data flowing in these networks, combined with the limited budget of attention of each user, makes it difficult to ensure that

Homogeneous or
structured
interests leads to
efficient networks

Can we find evidence of filtering?

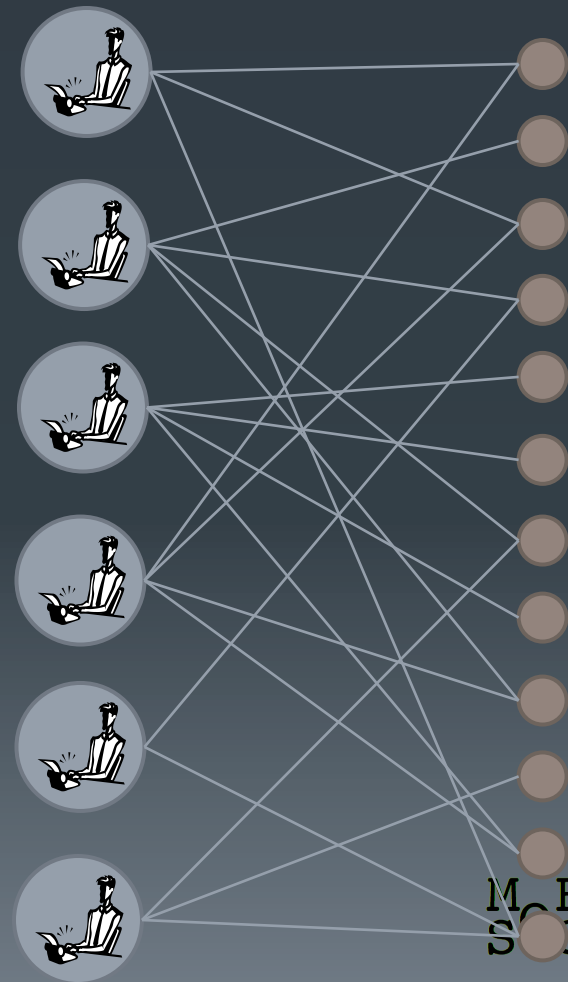
Looking for filtering

Data Sets	Source	Users	URLs
NY Times Links	Twitter	330k	33k
Bin Laden Death	Twitter	700k	545k
Occupy Wall Street	Twitter	354k	316k
Steve Jobs Death	Twitter	719k	251k
iPhone 5 Launch	Twitter	81k	37k
iPhone 5 Launch	Facebook	330k	193k
All Spinn3r blogs	Spinn3r	68k	441k
Obama	Spinn3r	13k	85k
Facebook	Spinn3r	12k	70k
Euro	Spinn3r	10k	53k
Mubarak	Spinn3r	7k	43k



Intermediaries

URLs Posted



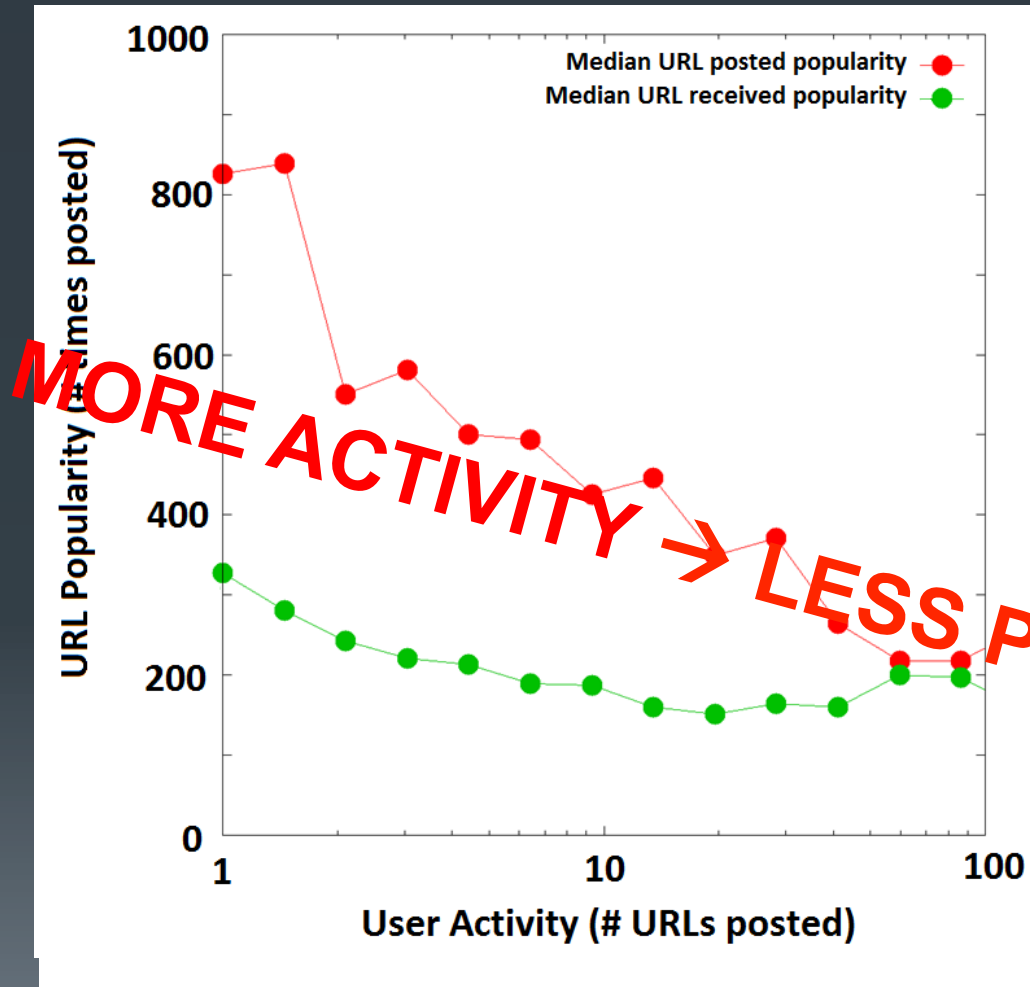
M B L E
S C I A L
B

Evidence of information filtering

“Filtering law”

Not an artefact of

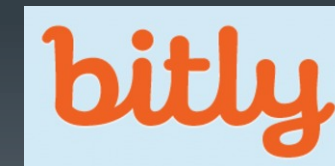
- replacement
- exposure



Many open questions

- Can we find more evidence of precision?
 - Using click (Twitter data grant, more partners)
 - Does selectivity correlate with success?
- Current models somewhat at odds
 - Discrete topics + continuous popularity range
 - Are there more general models
- Can crowd-curation be improved?
 - In principle (no friction etc.), already efficient.
 - With incentive? With new mechanism?

Thank you!

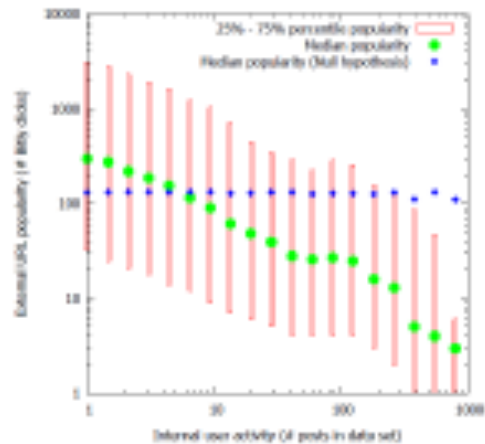


Back-Up Slides

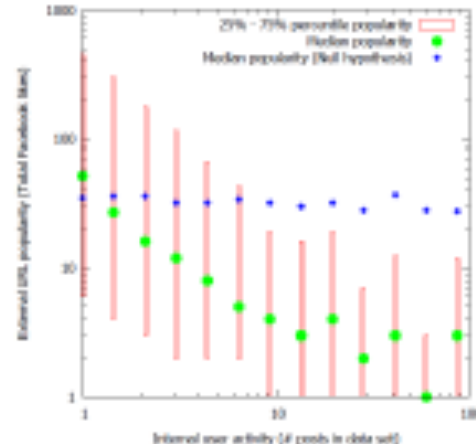
Theoretical Results

Audience Strategy	Pure Strategy Equilibrium?	Price of Anarchy
Greedy	No	--
Satisficing	Yes	2
Satisficing w/ blogger ability	Yes	2

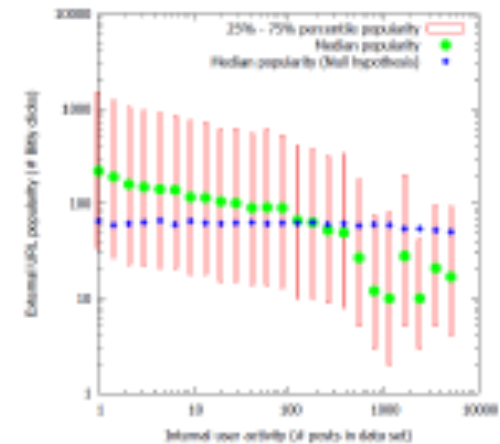
Filtering Law Consistent Across Data Sets



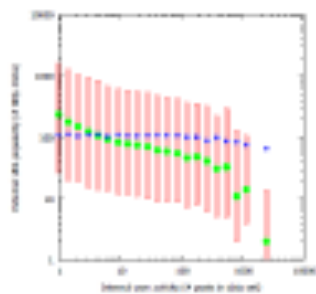
(a) TW Bin Laden



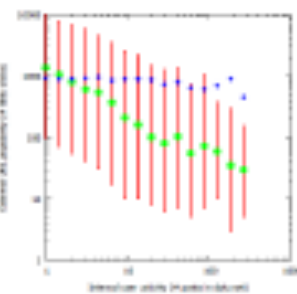
(b) FB iPhone5



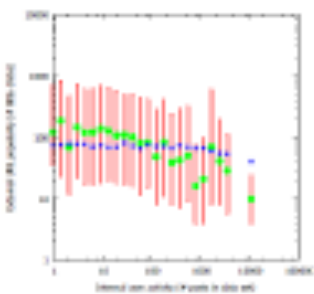
(c) Blog All



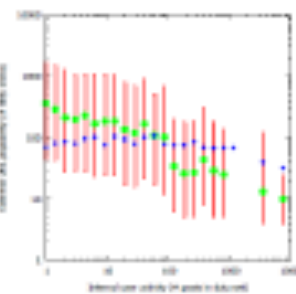
(TW Occupy)



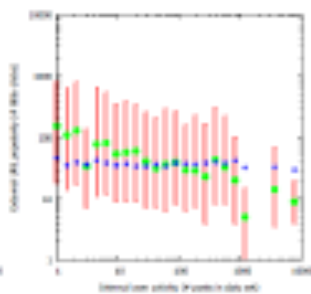
(TW Steve Jobs)



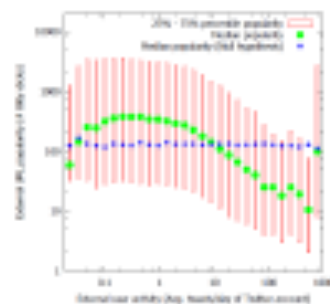
(Blog Obama)



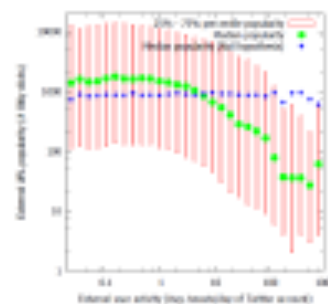
(Blog Mubarak)



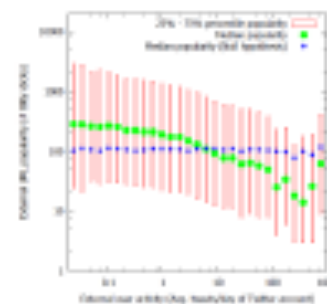
(Blog Euro)



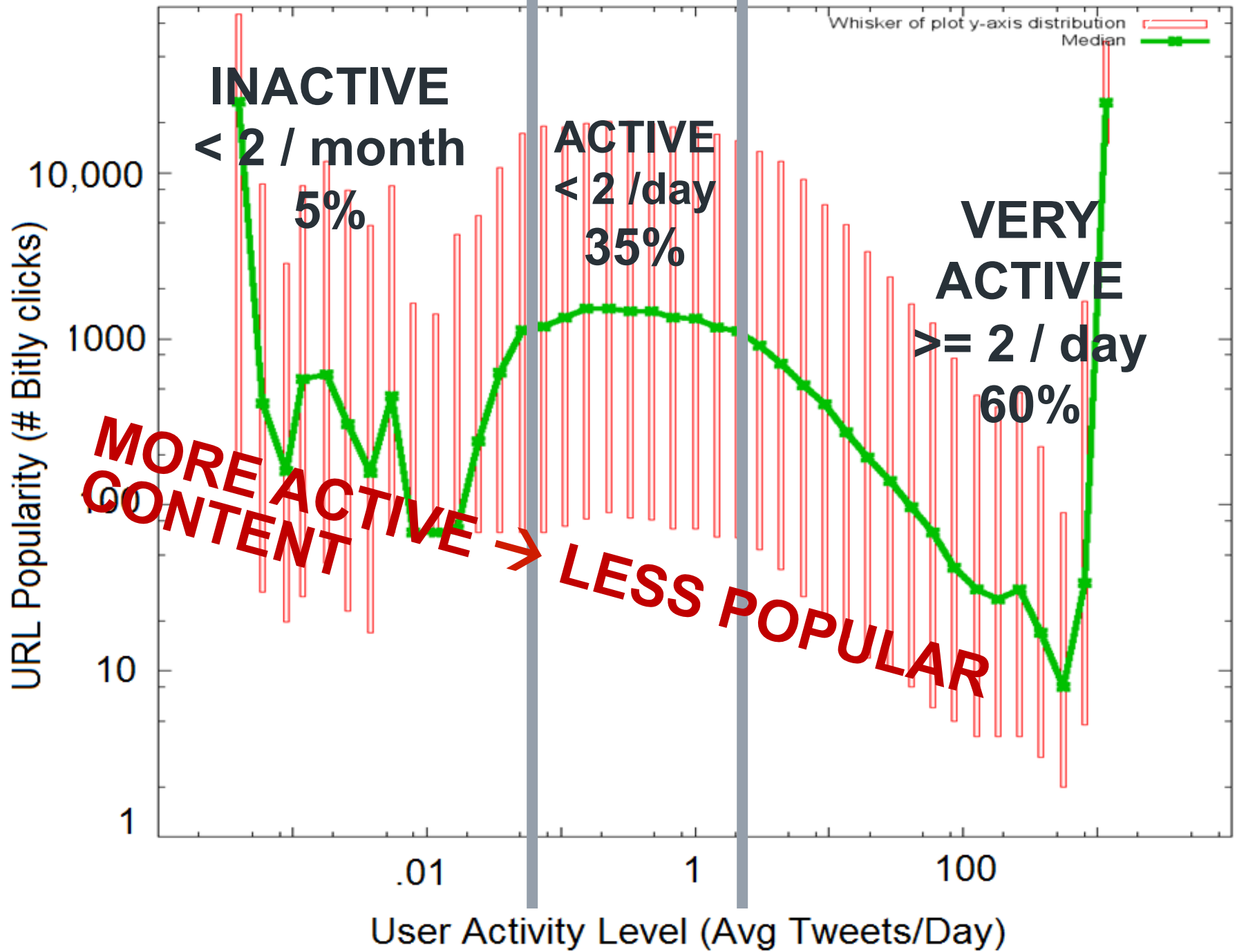
(a) TW Bin Laden



(b) TW Steve Jobs



(c) TW Occupy WS

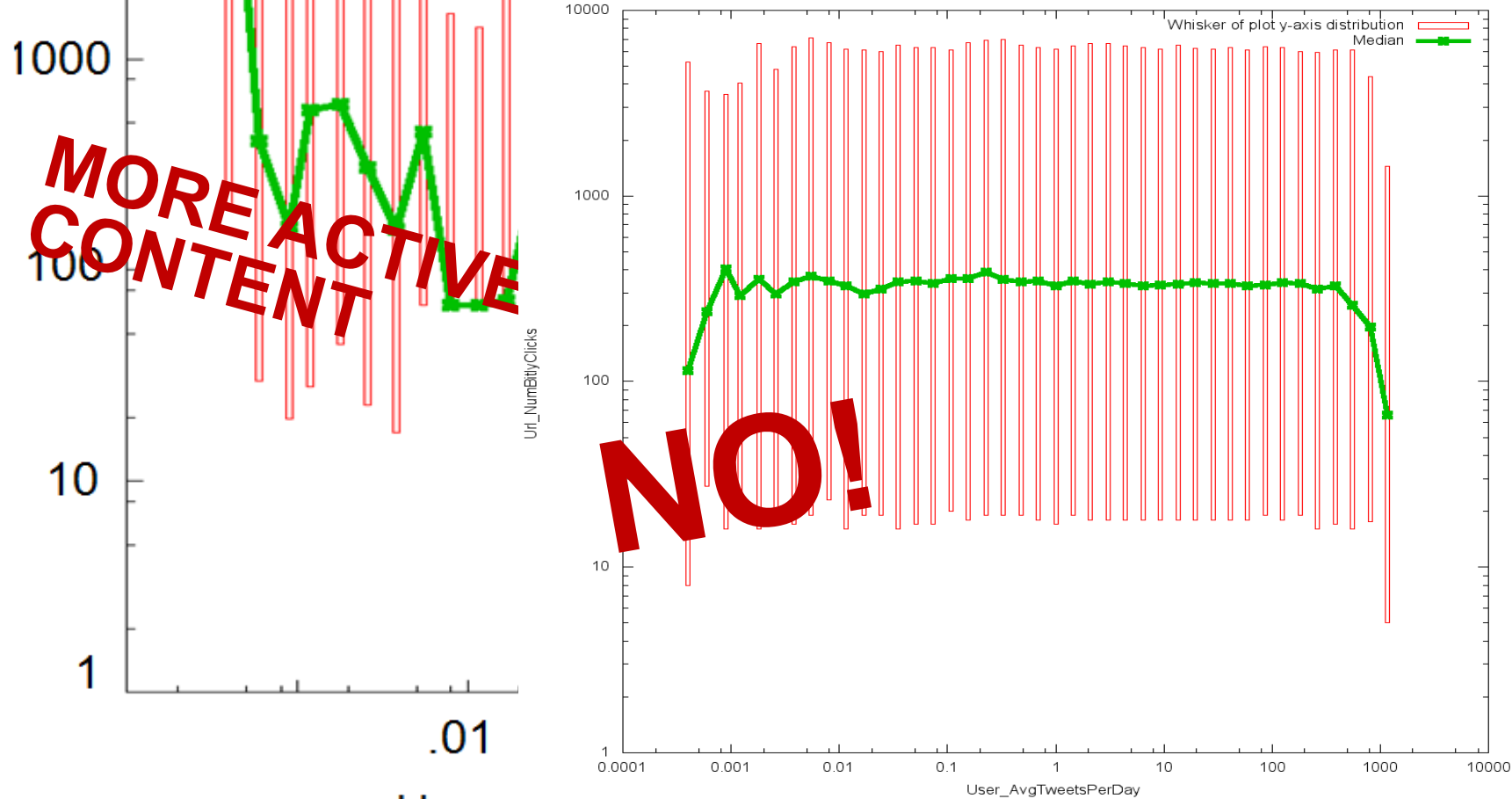
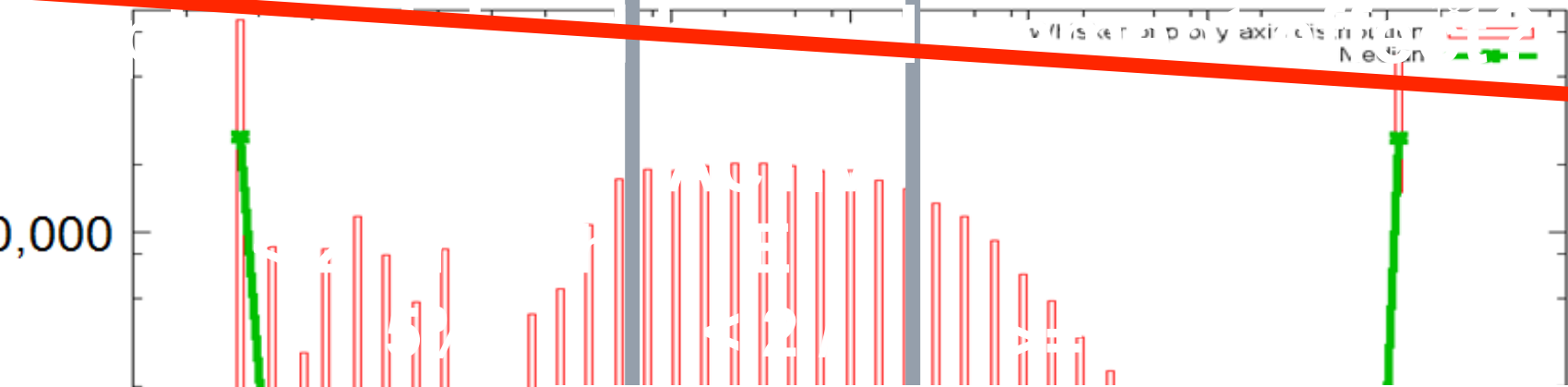


URL Popularity (# Bitly clicks)

MORE ACTIVE CONTENT

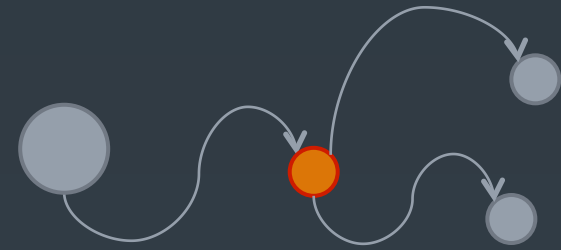
NO!

User Activity Level (Avg 1 tweets/Day)



In Summary...

- Previous work:
Intermediaries play key role in information dissemination.
- We provided theoretical and empirical justification for intermediaries as information filters.
- Come see my poster!
 - Results not shown: Role of filtering on success of intermediary



MORE ACTIVE → LESS
POPULAR CONTENT

M
S
O
C
I
A
L
M
E
D
I
A
B