32nd Annual ACM SIGIR '09
Boston, USA, Jul 19-23 2009

Telling Experts from Spammers

Expertise Ranking in Folksonomies

Michael G. Noll
Christoph Meinel

(Albert) Ching-Man Au Yeung
Nicholas Gibbins
Nigel Shadbolt

Hasso Plattner Institute
Uni Southampton

Introduction

HPI Hasso Plattner Institut

UNIVERSITY OF Southampton
School of Electronics and Computer Science

Folksonomies and **Collaborative Tagging**

- Large and still increasing popularity in the WWW today

Web pages

photos music books videos

Delicious.com – social bookmarking service by Yahoo! with 5+ million users

- Idea: Freely annotating resources with keywords aka “tags”
- Result: bottom-up “categorization” by end users, aka “folksonomy”
- Used for organizing resources, sharing, self-promotion, …
- Additional effect: **new means of resource discovery**

Motivation

Two related goals for our work on expertise in folksonomies:

1. **Identifying and promoting experts for a given topic**
   Weighting user input, giving (better) recommendations, identify trendsetters for marketing/advertising/product promotion, etc.

   Topic *:= conjunction or disjunction of one or more* ***tags***

2. **Demoting spammers**
   Reduce impact of spam and junk input thereby improving system quality, performance, operation

Models

Model of expert users

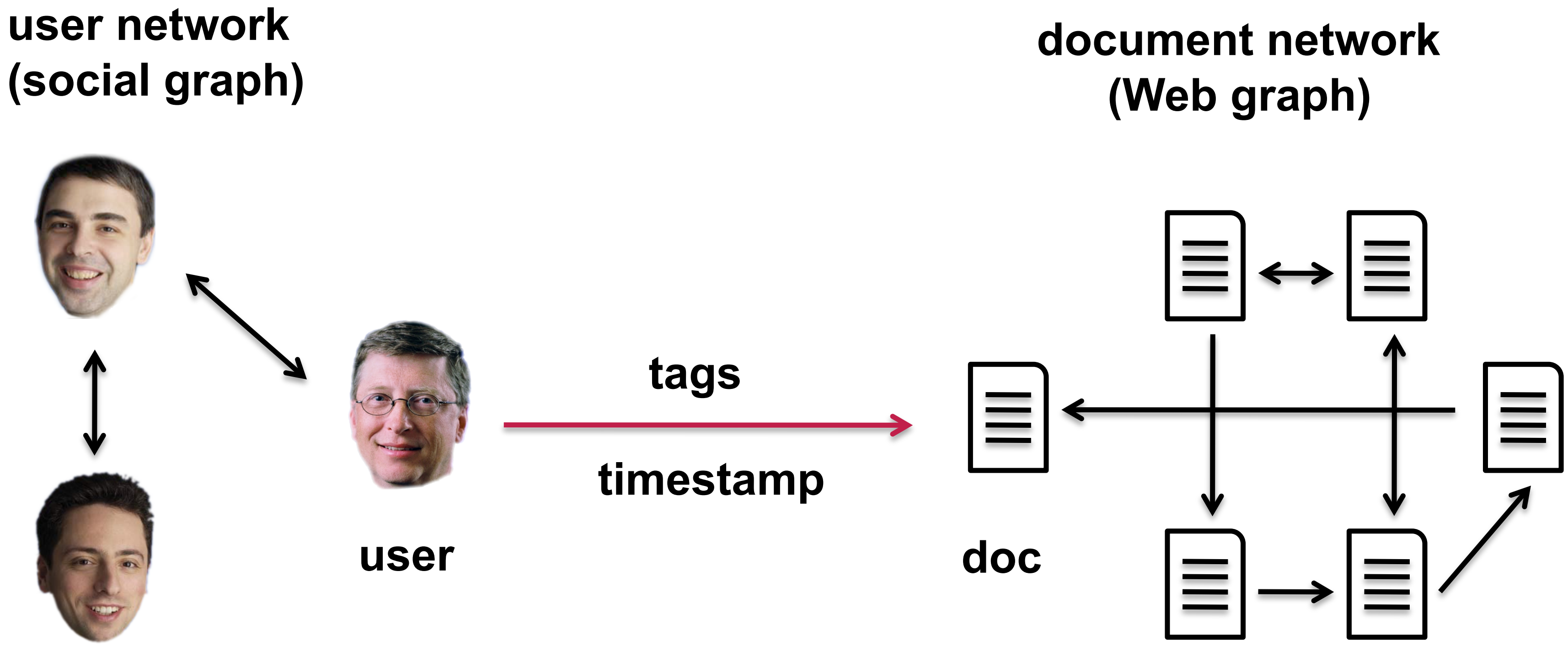
What makes an expert an expert?

Postulation of two assumptions of expertise for resource discovery, grounded on literature from **computer science** (that's you) and **psychology**

1. **Mutual reinforcement of user expertise and document quality**
   Expert users tend to have many high quality documents,
   and high quality documents are tagged by users of high expertise.

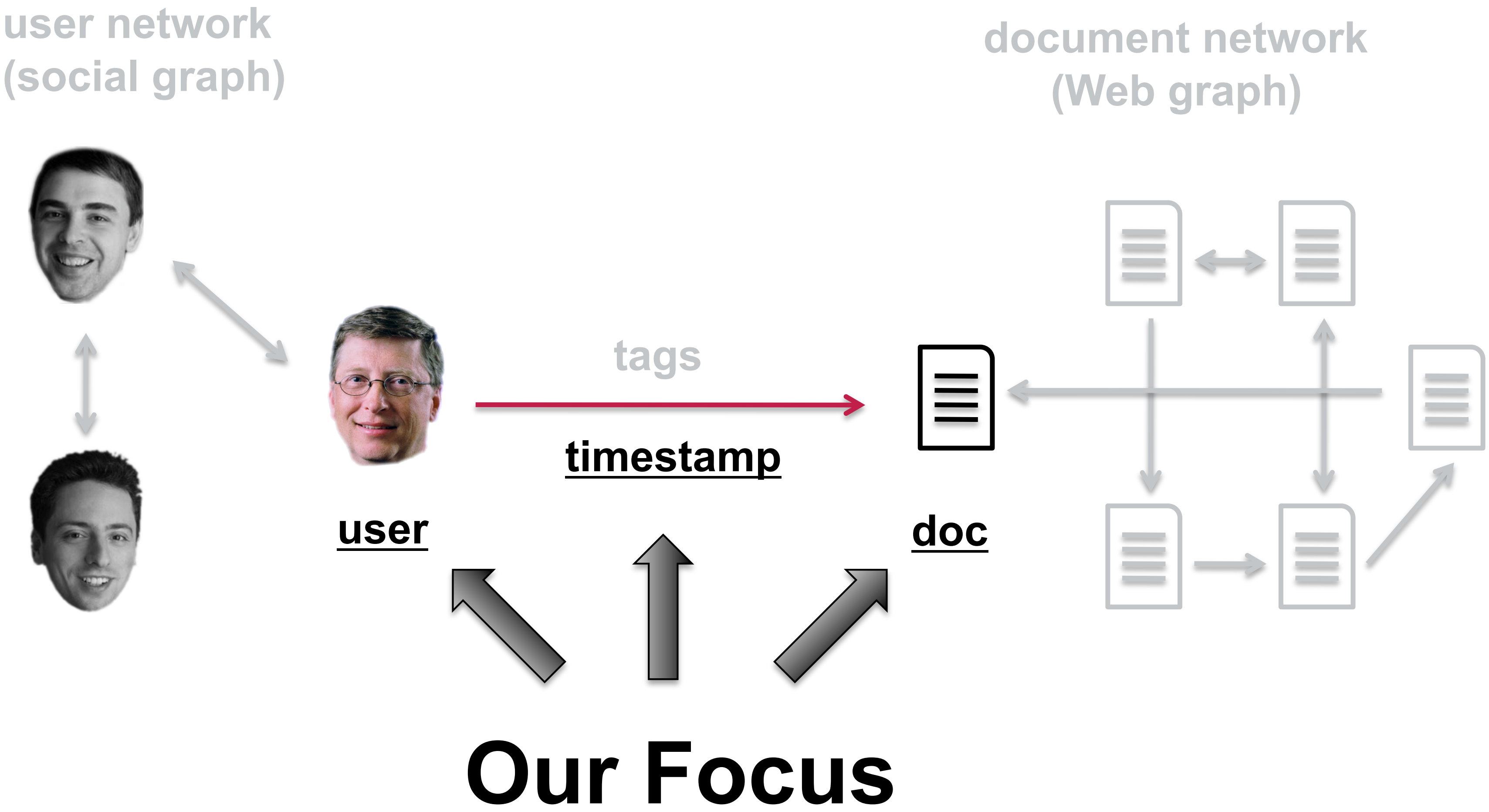
2. **Discoverers vs. followers**
   Expert users are discoverers – they tend to be the first to bookmark and tag high quality documents, thereby bringing them to the attention of the user community. Think: researchers in academia.

Model of expert users



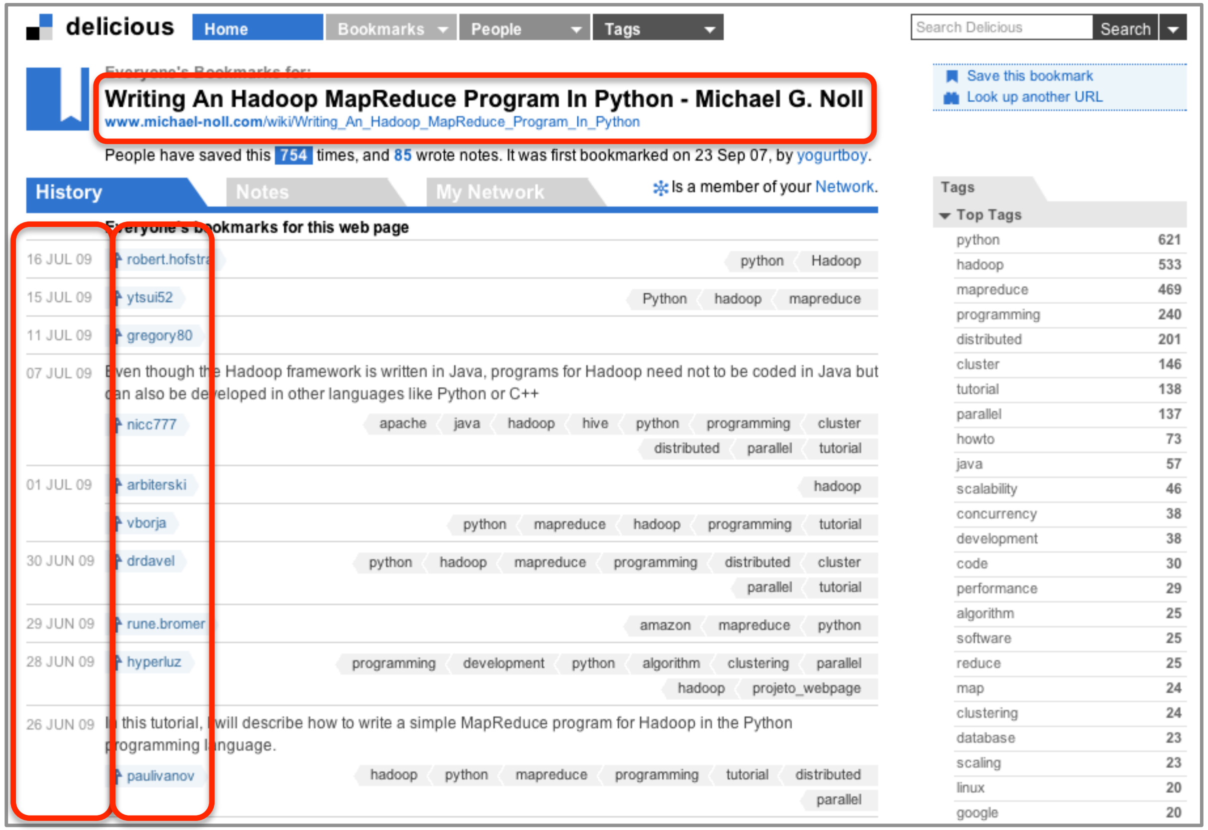
Context of social bookmarking / collaborative tagging

Model of expert users



Bookmarking history of a Web page on Delicious.com

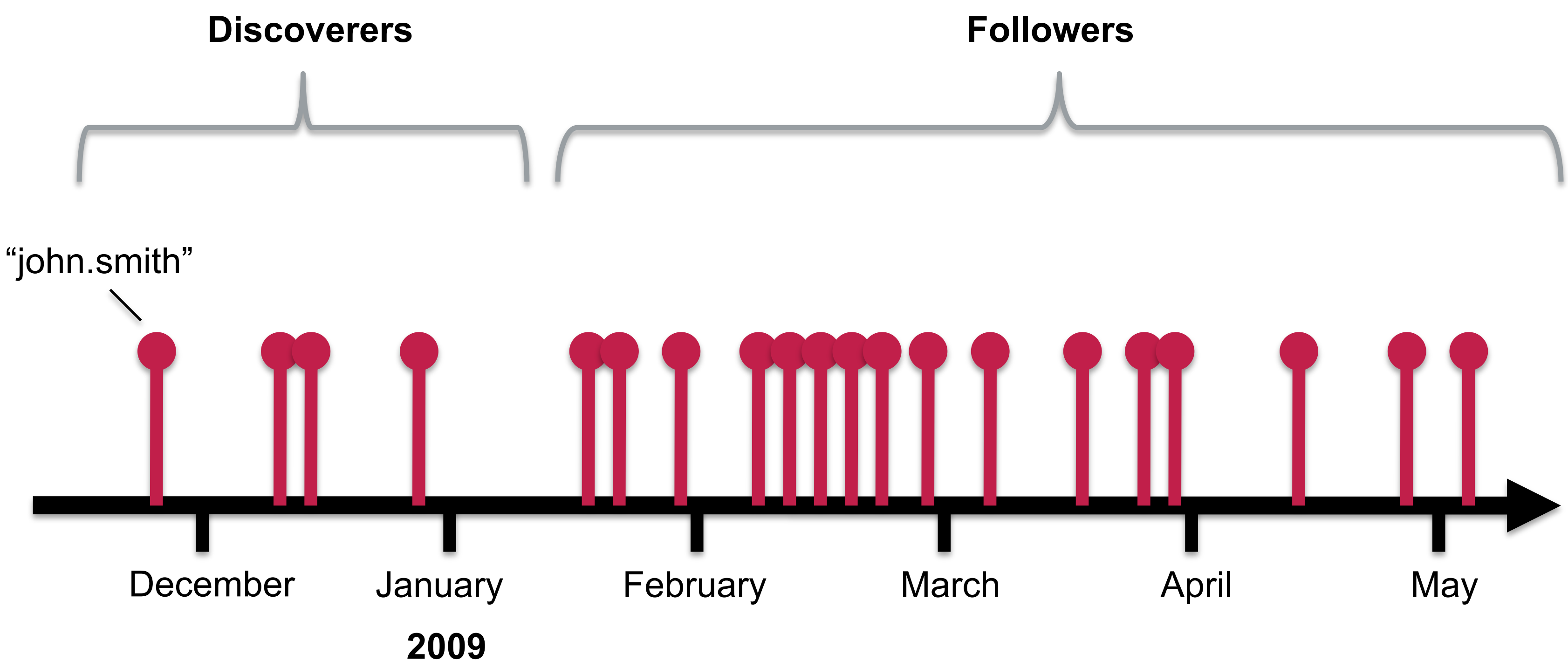
Web page



Timeline

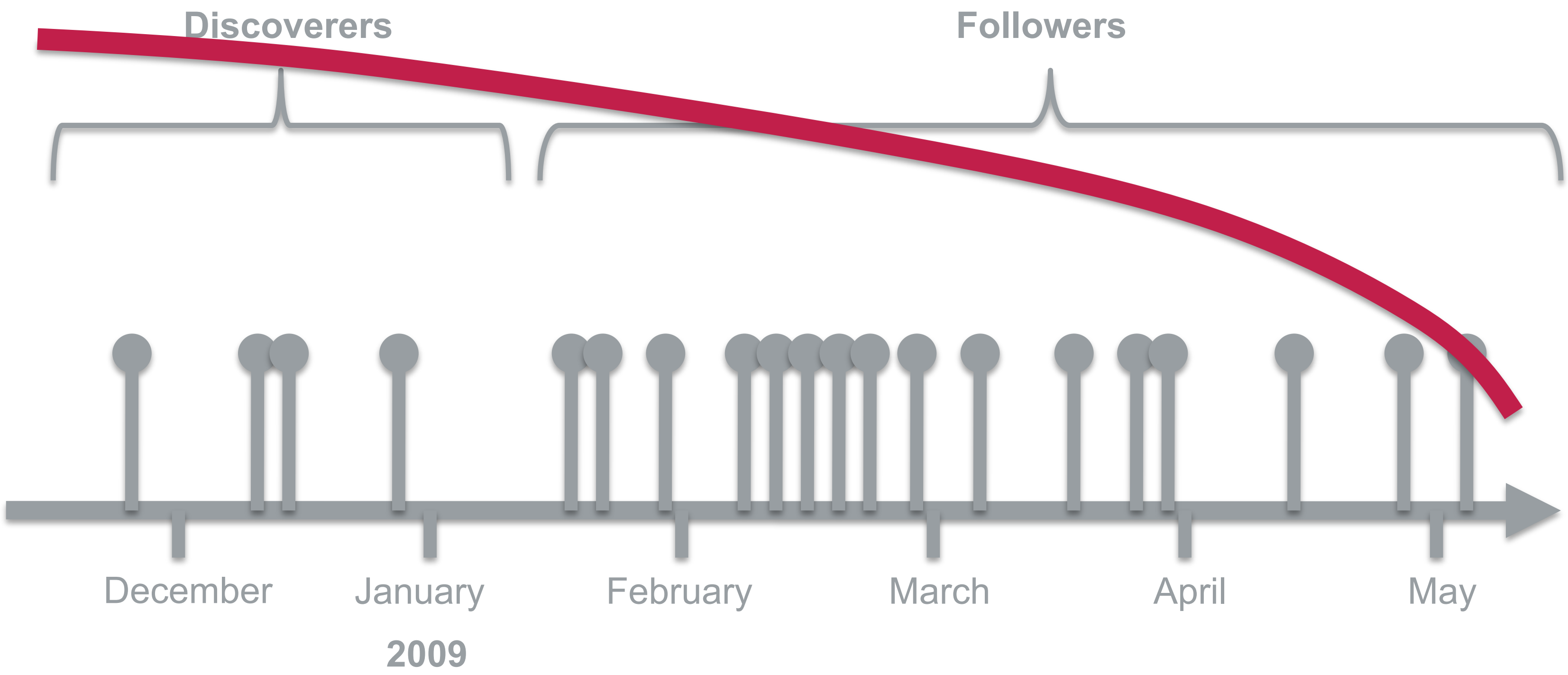
Users

Bookmarking history of a Web page



Model of expert users

Credit score function *C(t)* → **earlier discovery = more credit**



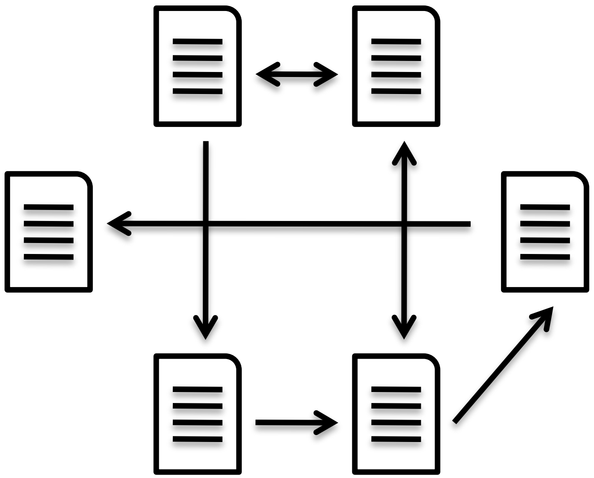
SPEAR Algorithm

Proposed algorithm: SPEAR

SPEAR – SPamming-resistant Expertise Analysis and Ranking

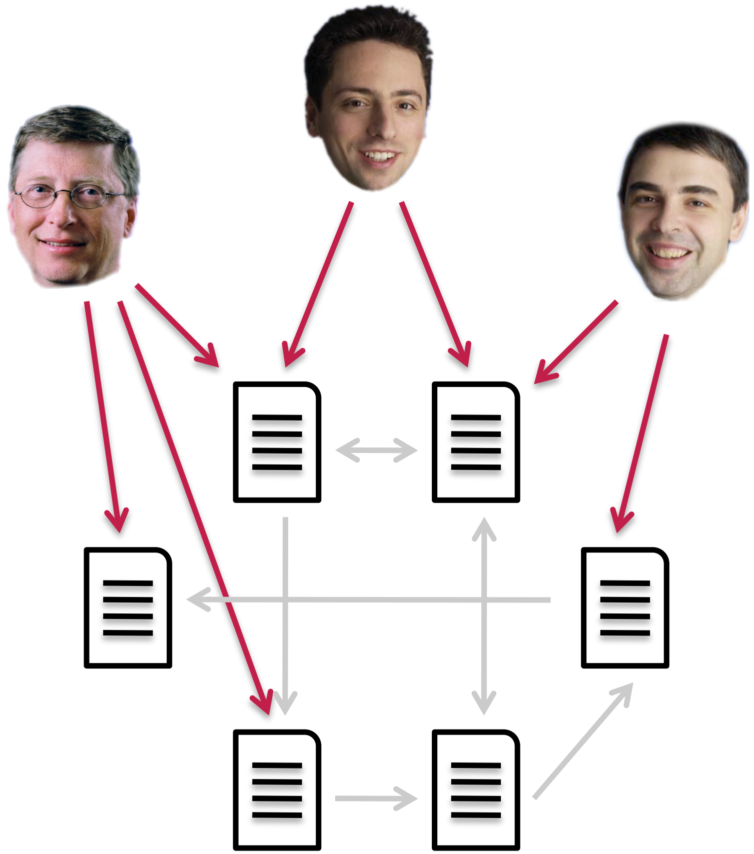
- Based on the **HITS** (Hypertext Induced Topic Search) algorithm
  - *Hubs*: pages that points to good pages
  - *Authorities*: pages that are pointed to by good pages
- *Expertise* and *Quality* (SPEAR) similar to *Hub* and *Authority* (HITS)
  - **Users** are **hubs** – we find useful pages through them
  - **Pages** are **authorities** – provide relevant information
- Difference: only users can point (link) to pages but not vice versa

Proposed algorithm: SPEAR



page ↔ page

HITS / WWW



user → page

SPEAR / Folksonomy

Proposed algorithm: SPEAR

Hasso Plattner Institut

University of Southampton School of Electronics and Computer Science

Input Number of users ***M***
Number of pages ***N***
Set of taggings $\boldsymbol{R_{tag}}$ = { (user, page, tag, timestamp) | tag = ***tag*** }
Credit score function ***C()***
Number of iterations ***k***

Output: Ranked list ***L*** of users by expertise in topic ***tag***

Algorithm:

Set ***E*** to be the vector (1, 1, …, 1) $\in Q^M$
Set ***Q*** to be the vector (1, 1, …, 1) $\in Q^N$

E: expertise of users
Q: quality of pages

A ← Generate_Adjacency_Matrix($\boldsymbol{R_{tag}}$, ***C***)

A: user → page incl. credits

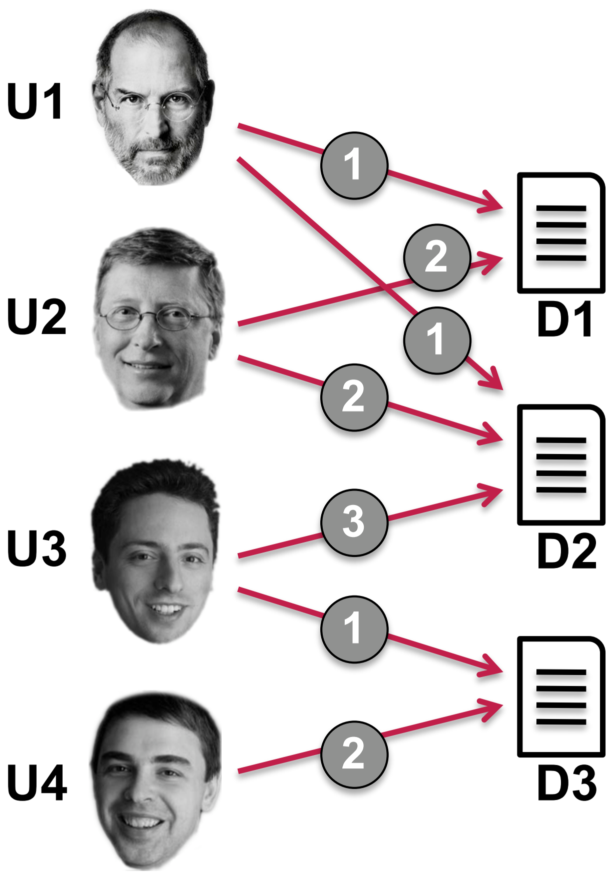
for **i** = 1 to ***k*** do
    E ← ***Q*** x $\boldsymbol{A^T}$
    Q ← ***E*** x ***A***
    Normalize ***E***
    Normalize ***Q***
endfor

mutual reinforcement until convergence

L ← Sort users by their expertise score in ***E***
return ***L***

Proposed algorithm: SPEAR

HPI Hasso Plattner Institut
UNIVERSITY OF Southampton
School of Electronics and Computer Science



Folksonomy (simplified)

Adjacency matrix, **credits applied**

	D1	D2	D3
U1	1.4	1.7	0.0
U2	1.0	1.4	0.0
U3	0.0	1.0	1.4
U4	0.0	0.0	1.0

		Rank	Score
Steve	U1	1	0.422
Bill	U2	2	0.328
Sergey	U3	3	0.212
Larry	U4	4	0.038

Ranked list of users by expertise

Evaluation

Experimental Setup

- Problem: lack of a proper ground truth for expertise
- “Who is the best researcher in this room?” ☺
- Workaround: Inserting **simulated** users into **real-world** data from Delicious.com and check where they end up after ranking
- **Real-world data set from Delicious.com** comprising 50 tags with
  - 515,000 real users (and real spammers)
  - 71,300 real Web pages
  - 2,190,000 real social bookmarks

Experimental Setup

- **Probabilistic simulation,** simulated users generated with four parameters
  - **P1:** Number of user’s bookmarks – active or inactive user?
  - **P2:** Newness – fraction of Web pages not already in data set
  - **P3:** Time preference – discoverer or follower?
  - **P4:** Document preference – high quality or low quality?

Experimental Setup

- Simulation of 6 different user types
  Profiles (parameter values) based on recent studies + characteristics of our real-world data sets
- Experts
  - **Geek** – lots of high quality documents, discoverer **(Distinguished Researcher)**
  - **Veteran** – high quality documents, discoverer **(Professor)**
  - **Newcomer** – high quality documents, follower **(PhD student)**
- Spammers
  - **Flooder** – lots of random documents, follower **(found in Delicious)**
  - **Promoter** – some documents (most are his own), discoverer **(found in Delicious)**
  - **Trojan** – some documents, follower **(next-gen spammer)**

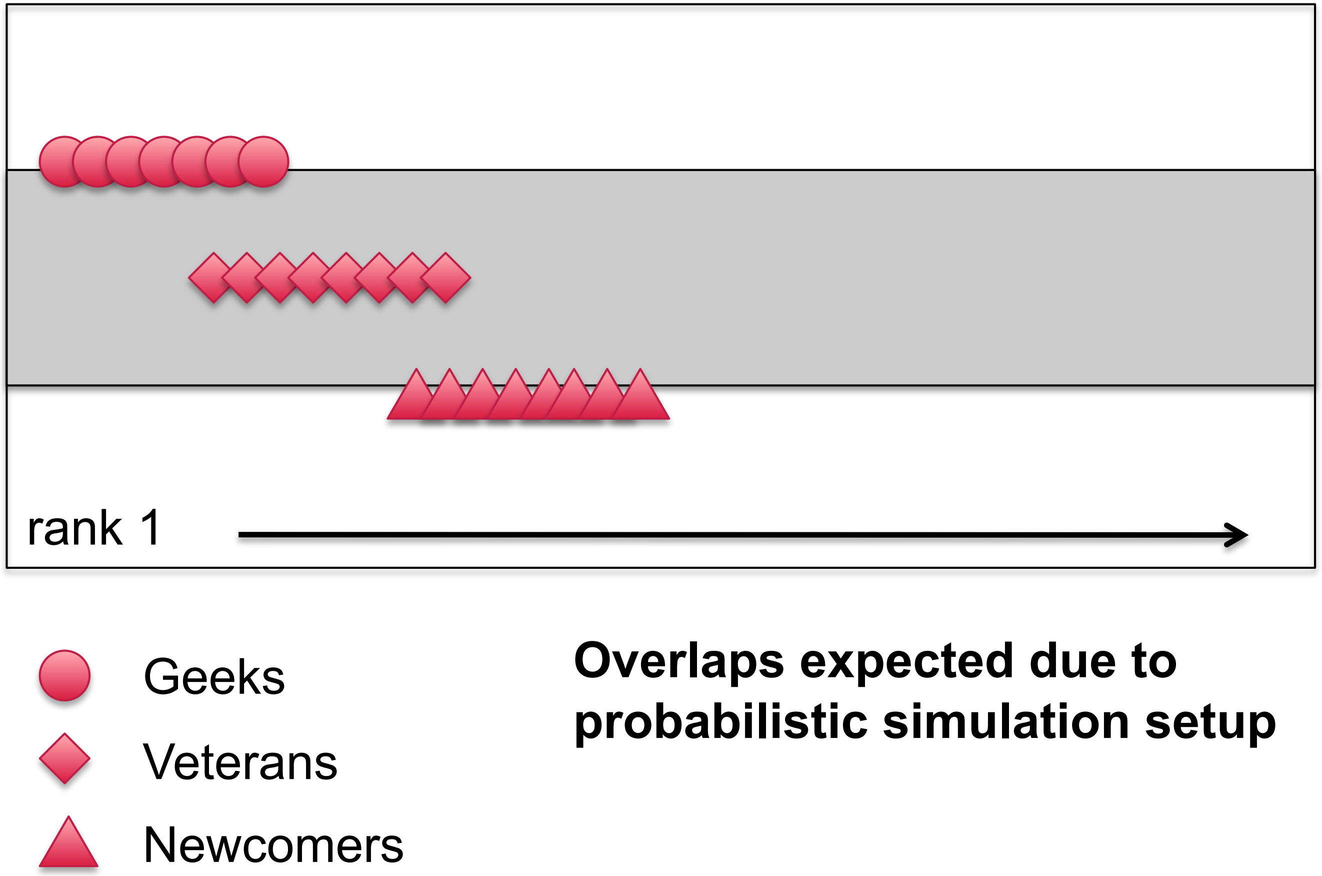
Performance baselines

- **FREQ(UENCY)**
  “Most popular” approach – simple frequency count, looks only at quantity. Seems to be the dominant algorithm in use in practice.
- **HITS**
  Algorithm on which SPEAR is based. Uses mutual reinforcement but *does not* analyze temporal dimension of user activity.

Experimental Results

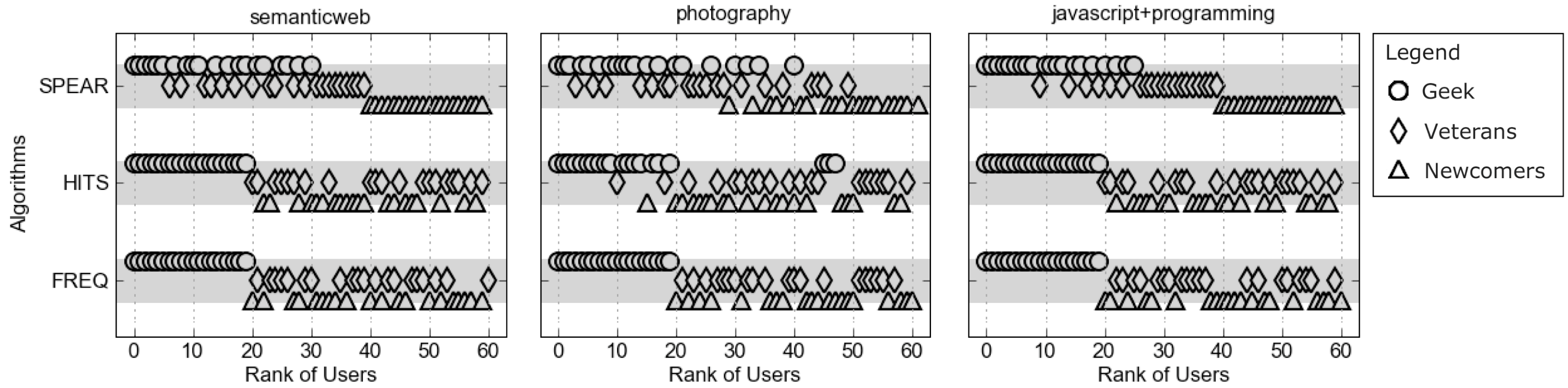
Experts

Experts: “Ideal” result



Overlaps expected due to probabilistic simulation setup

Experimental Results – Promoting **Experts**



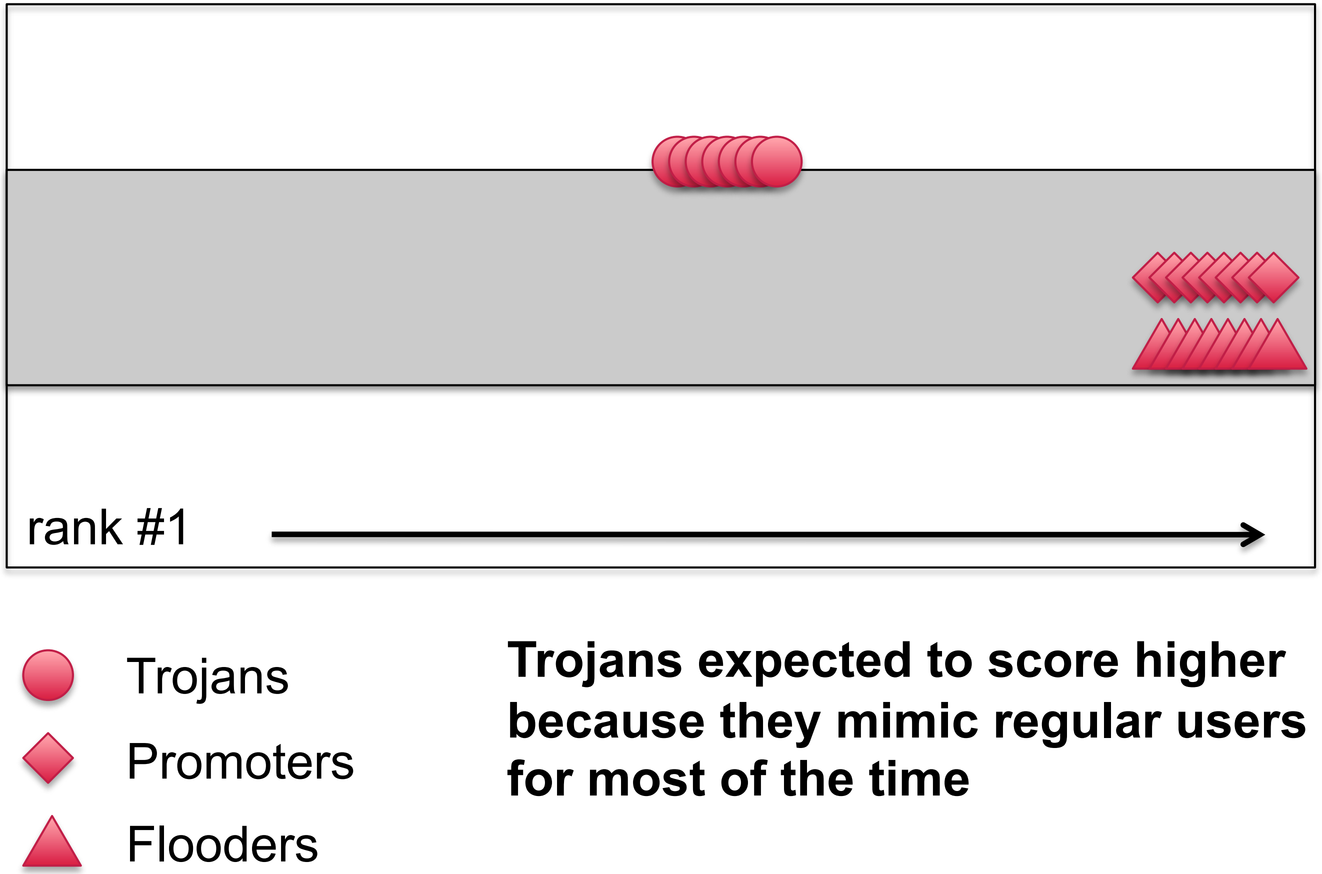
- SPEAR differentiated all expert types better than its competitors
- SPEAR kept expected order of “geeks > veterans > newcomers”
- SPEAR was less dependent on user activity (quality before quantity)

Qualitative analysis: manual examination of Top 10 experts for three tags "photography", "semanticweb", "javascript ∩ programming"

- No spammers found (…phew…)
- These users seemed to be more involved or "serious" about their Delicious usage, e.g. provided optional profile information such as real name, links to their Flickr photos or microblog on Twitter
- Their number of bookmarks: from 100's to 10,000's
- "semanticweb": Semantic Web researcher among the experts
- "javascript ∩ programming": Top 2 experts were professional software developers

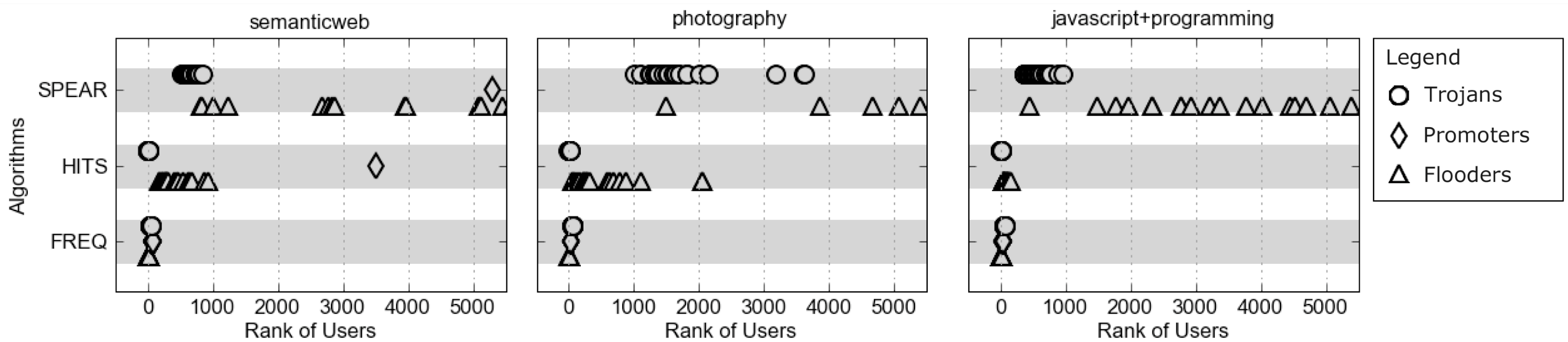
Spammers

Spammers: “Ideal” result



Trojans expected to score higher because they mimic regular users for most of the time

Experimental Results – Demoting **Spammers**



- SPEAR demoted all spammer types significantly more than its competitors
- Only SPEAR demoted all trojans from the TOP 100 ranks
- FREQ completely failed to demote any spammers

Qualitative analysis: manual examination of Top 50 users for the heavily spammed tag “mortgage” (without inserting simulated users)

- Ranked users by their number of bookmarks = FREQ strategy
- 30 out of 50 were (real) spammers, either flooders or promoters
- Compared to FREQ, both SPEAR and HITS were able to remove these spammers from the Top 50
- SPEAR demoted spammers significantly more than HITS

Summary

Summary

Conclusions

- SPEAR demoted all spammer types while still ranking experts on top
- SPEAR was much less vulnerable to spammers due to its reduced dependence on the activeness of users: “quality >> quantity”

Future Work

- Quality score of Web pages deserve more investigation
- Transfer to new problem domains, e.g. blogosphere or music
- Follow-up with user & item recommendation, trend detection

Michael G. Noll
michael.noll@hpi.uni-potsdam.de
Hasso Plattner Institute, LIASIT

Albert Au Yeung
cmay06r@ecs.soton.ac.uk
University of Southampton