



# *Pharos*

## *Social Map-Based Recommendation For Content-Centric Social Websites*

W. Zheng, M. Zhou, S. Zhao, Q. Yuan, X. Zhang, C. Chi  
IBM Research – China

Presenter: Quan Yuan (Leo)

## Two Challenges in Recommender System

- Content-centric social websites (e.g., forums, wikis, and blogs) have flourished with the exponential growth of UGC. But people feel difficult to navigate these sites and locate desired information
- Researchers developing recommender systems to solve this problem
  - E.g. Blog/News/Webpage recommender
- However, current recommenders must address two challenges:
  - difficult to make effective recommendations for new users (the cold start problem) due to a lack of user information
  - difficult to explain recommendation rationales to end users to make the recommendation more trustworthy

## Pharos - Social map-based recommender system

- Social map: a dynamically generated *marauder's map* of a content-centric social website, it summarizes users' content-related social behavior (e.g., reading, writing, and commenting) over time as a set of "*latent communities*"
- A *latent community* characterizes the **implicit** connections among a set of users and the content they generated.
- Then based on Social Map, we recommend
  - "hot" communities
  - "hot" content/people in each community
- Pharos on line (Internal access only):
  - <http://pharos.dyn.webahead.ibm.com/>



[Jul/1 ~ Jul/15](#)

[Jul/16 ~ Jul/31](#)

[Aug/1 ~ Aug/15](#)

Entries

Bloggers

Feature

networking  
wireless apple  
access vpn  
ocdc jun wifi  
**iphone**  
works connect  
gen leap inotes  
certificate

url link post  
blogcentral  
aprs dogear  
bcv4 wiki **social**  
lcw hams **blog**  
connection  
widget lotus  
custom

update  
change plugins **consulting**  
download linux **software**  
windows **specialist**  
xp **install** **social**  
**sametime**  
features notes  
lotus version  
ubuntu c4eb  
creates

juni manager  
pe assets **specialist**  
itdelivery **project**  
issl business  
chief soa  
sap iram um  
integrator  
sehr service

security  
web camp **analyst**  
por boot **insights**  
oracle **principal**  
cloud **leader**  
una os uma  
computing  
office como  
servidor banda

performance  
drive backup **storage**  
system virtual **services**  
tsm medium **data**  
**storage**  
dmx-4 data  
manager disk  
vmware  
emc support

mac custom  
cattail facebook  
pbc2009 profile **software**  
products note **lotus**  
mashup widget **social**  
connection  
plugins lotus  
api integrator  
iphone quickr  
sametime demo



BENITEZ, LUIS A.

Consulting Social Softwa...

- [And Now We Have An Official Wi...](#)
- [Making Customizations Easier i...](#)
- [My Second 24 Hours With Lotus ...](#)
- [Lotus Connections Goes Mobile](#)



Cogan, Steve

Certified IT Specialist ...

- [Widget wednesday - Notes8.01 e...](#)
- [Widget-Wednesday - catalogue a...](#)
- [Widget-Wednesday & "Twit-Notes...](#)
- [my favourite widget for notes8...](#)



Muecke, Gerald

Advisory IT Specialist -...

- [Cattail or Files?](#)



Feinstein, Jason W. (JAS...

Software Engineer - Tivo...

- [bluto 0.8.1b - Now With 100% M...](#)



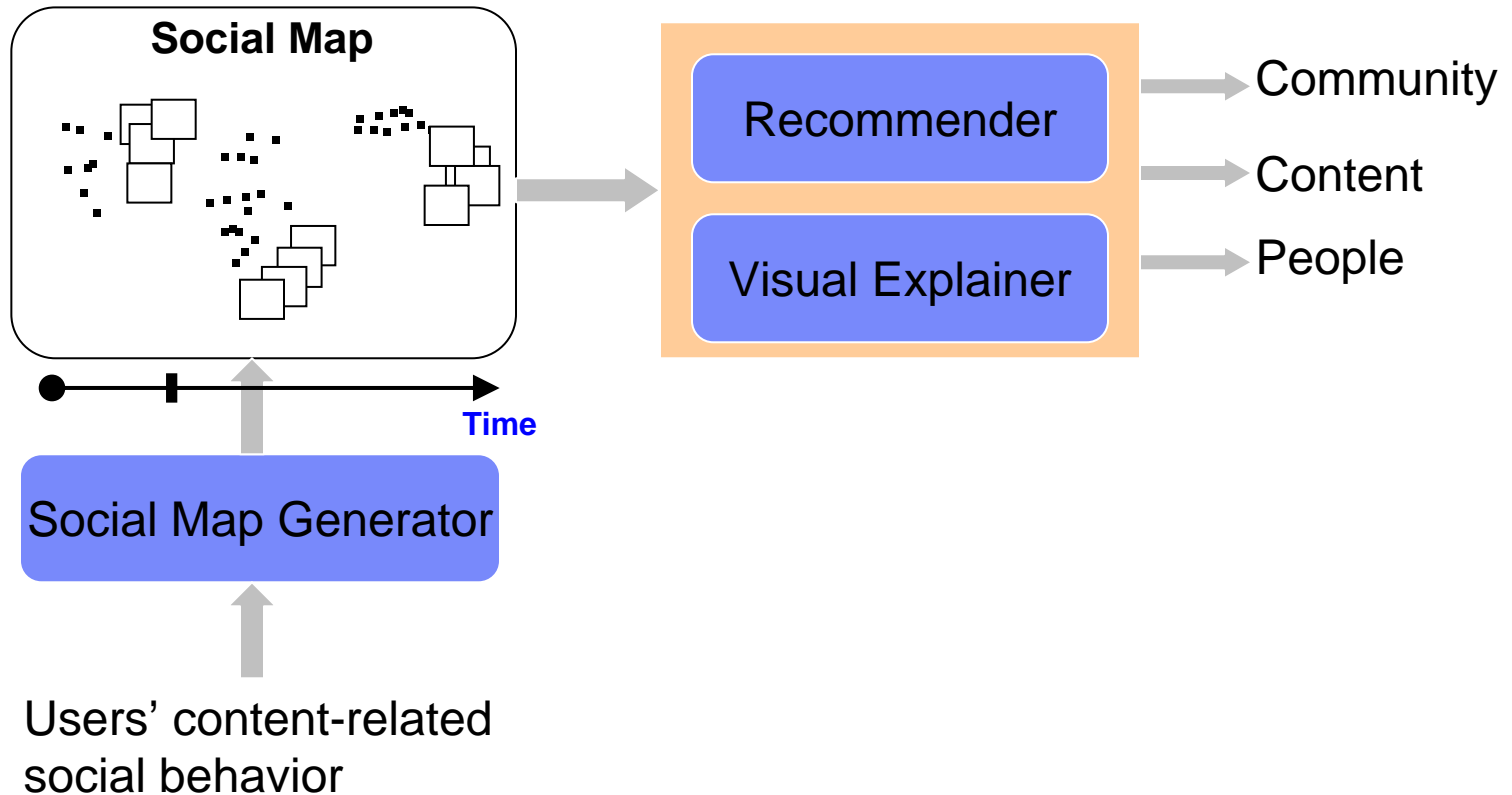
Breaux, Douglas N. (Doug...

Software Engineer

- [A customer's view of cloud sto...](#)



# Pharos Architecture Overview

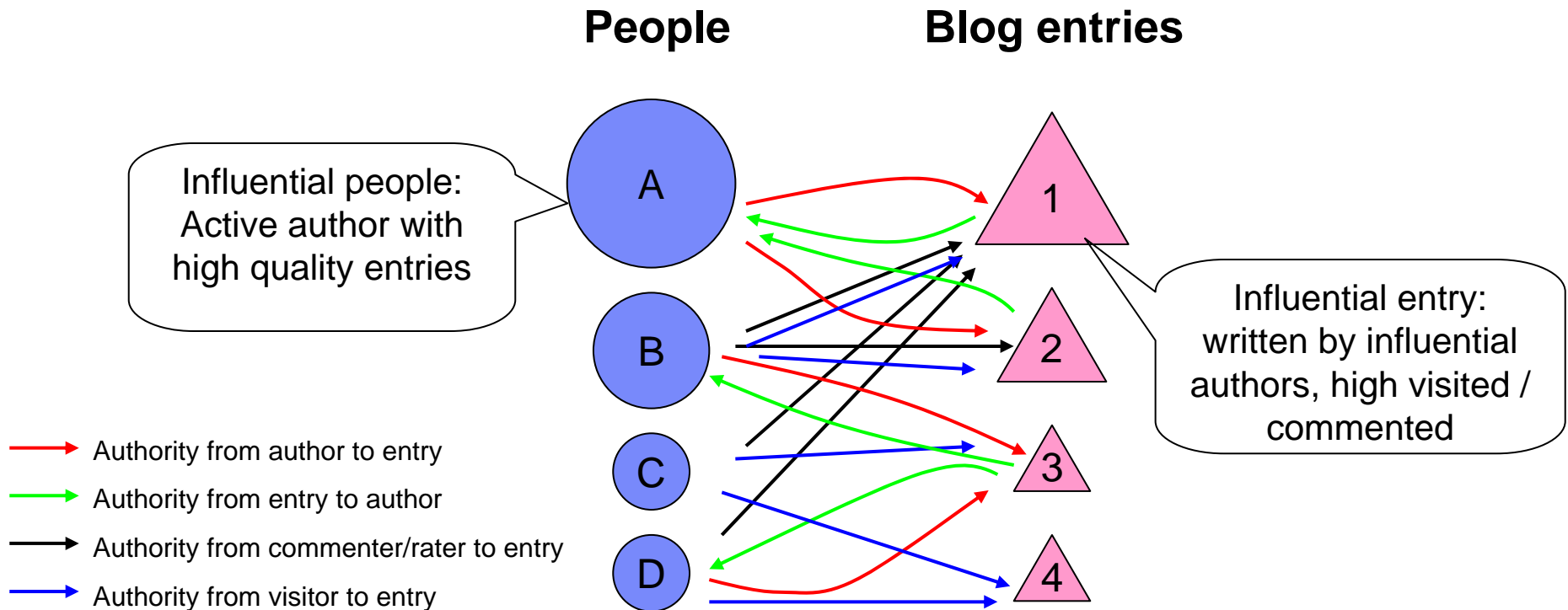


# Social Map Generation

- Community Extraction
  - 3 Approaches
    - User-content by co-clustering (extremely sparse data, co-clustering not work well)
    - Group people firstly, then find associated content (most behaviors from anonymous user, while difficult in distinguish user)
    - Group content firstly, then find associated people (user more likely to be interested in content)
  - LDA for summarizing content into topics, then topics form content clusters.
- Community Ranking
  - Authority-based community ranking
- Community Labeling
  - LDA-extracted keywords for content summary
  - Authority-based TF/IDF for people summary

# Authority-based People/Content Recommendation

- 1. Authority Modeling



$$PR(p_i) = \frac{1-d}{N} + d \sum_{p_j \in M(p_i)} \frac{PR(p_j)}{L(p_j)}$$

- 2. Recommendation

- Recommend authority people
- Recommend authority items

## Visual Explainer

- A bubble chart layout (used by ManyEyes2) to pack top-N communities tightly on the social map
  - bubble's size is determined by community's 'hotness'
- Inside each community, Wordle3 layout used to pack labels tightly



# Summary

- Address the cold start problem in part by using a social map to summarize a social website, which in turn helps new users
- Use the social map to help explain recommendation results by means of content, people, and their relationships, helping improve recommendation trustworthiness.

# Q & A