NSLAB TSINGHUA UNIVERSITY

Personalized Recommendation of Social Software Items Based on Social Relations

Ido Guy, Naama Zwerdling, David Carmel,
Inbal Ronen, Erel Uziel, Sivan Yogev,
Shila Ofek-Koifman

Present by *Feng Xie*

March 13, 2012

Ido Guy

- Manager, Social Technologies group,
  IBM Haifa Research Lab, Israel
- Co-author of over30 conference and journal Papers on DM, HCI, IR, RS
- Research work mostly on social media, especially on social network mining and analysis and social recommender systems.
- Organizing and program committees of leading conferences including CHI, WWW and CIKM.
- Currently, Workshop co-chair for CHI 2012, Industry co-chair for UMAP 2012, and program co-chair for RecSys 2012
- http://domino.research.ibm.com/comm/research_people.nsf/pages/ido.index.html

RecSys 2012

For Conference

- **Deadline for abstracts (mandatory for long/short papers): April 2, 2012, 11.59 pm (PST)**
- **Paper submission deadline: April 9, 2012, 11.59 pm (PST)**
- Paper acceptance notifications: June 1, 2012
- Camera-ready version: June 22, 2012

Content

- Background & Introduction
- Social Network-based Recommendation
  - Familiarity vs. Similarity vs. Overall
  - Explanation vs. Non-explanation
- Experiments Evaluation & Analysis
- Conclusions & Future Work

Background & Introduction

- Social Software
  - Social Software encompasses a range of software systems that allow users to interact and share data
- Example
  - Blog Systems
  - Wikis
  - Social Bookmarking
  - Social Network Sites

Why Personalized Recommendation?

- More and more social software sites keep popping up
  - Harder for sites to attract users
  - Harder for users to select which sites to keep track with
  - Information overload
- Personalized Recommendation
  - Adapt the content based on the individual user's characteristics
  - Present the most attractive and relevant items to the user

Typical Method of PR

- Content-based
  - Similar items to user
  - Lose diversity
- Collaborative Filtering
  - Similar users share mutual interests
  - With diversity

Social Network-based Recommendation

- Collaborative Filtering
- Assumptions
  - Seek advice from friends-familiarity
  - Similar users share same activity-similarity

Lotus Connections

- Enterprise Social Software Application Suite
- Contains 7 social software applications and a homepage (v2.5)
- We focus on bookmarks, blogs, wikis, files, and communities

- Bookmarks (Dogear)–900K bookmarks, 2M tags, 21K users
- Blogs–130K blog entries, 350K tags, 17K users
- Wikis–3K wikis, 10K tags, 5K editors
- Files–15K publicly shared files, 24K tags, 8K users
- Communities–6K communities, 19.5 tags, 174K members

SONAR

- Social network aggregator across the enterprise
- Harvests and weights relationships from different public sources
  - Familiarity: org chart, 2 SNSs, people tagging, membership in a project wiki, co-authorship of paper and patents
  - Similarity: co-usage of tags, co-bookmarking of web pages, co-commenting
  - Overall: aggregating all relationship types

Relationship Score

- Familiarity
  - 0 stands for no relationship, 1 stands for the strongest relationship. Average per relationship
- Similarity
  - Jaccard's index
  - Dividing the number of items in the intersection by the number of items in the union set.
- Overall
  - Average above two scores

Item Recommendation Algorithm

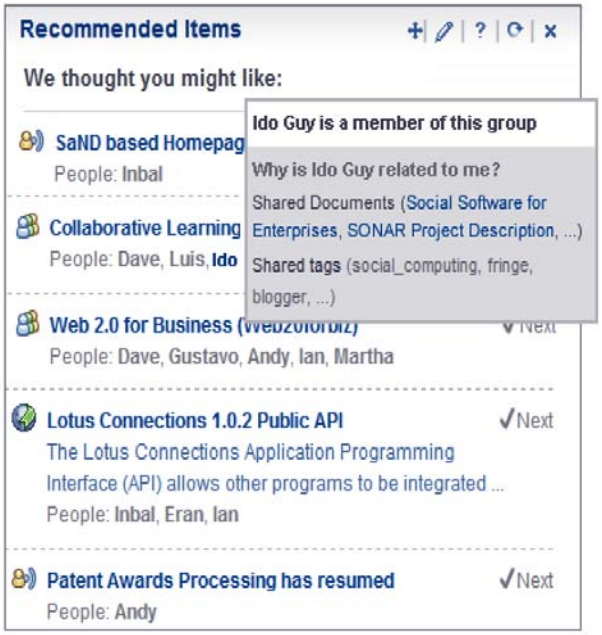
- Retrieve top 30 related people (either familiarity, similarity, or overall)
- Retrieve related items to these people based on a unified index – authorship (0.6), membership (0.4), commenting (0.3), tagging (0.3)

$$RS(u,i) = e^{-\alpha t(i)} \cdot \sum_{v \in N^T(u)} S^T[u,v] \sum_{r \in R(v,i)} W(r)$$

- Ultimately, recommender score will increase due to:
  - More people and/or tags within the user's profile are related to the item
  - Stronger relationships of these people and/or tags to the user
  - Stronger relationships of these people and/or tags to the item
  - Freshness of the item (update/creation date)

Item Recommendation Widget

- Deployed on Lotus Connections Homepage
- Suggests 5 mixed items – bookmarked web pages, blogs, or communities (icon represents the type)
- Title is a link to the original document
- Short description (optional)
- User can remove and get the next item



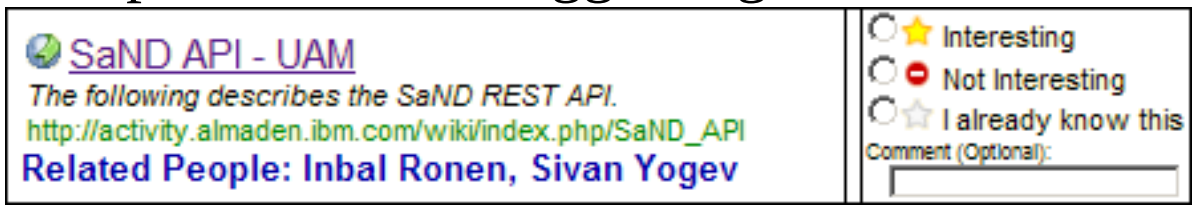
- Explanations:
  - Names of related persons (the implicit recommenders for this item)
  - When hovering over a name – see the relationship of this person to the user and to the item

Main Research Questions

- Familiarity vs. Similarity (vs. Overall)
- The value of Explanations
- Bookmarks vs. Blogs vs. Communities

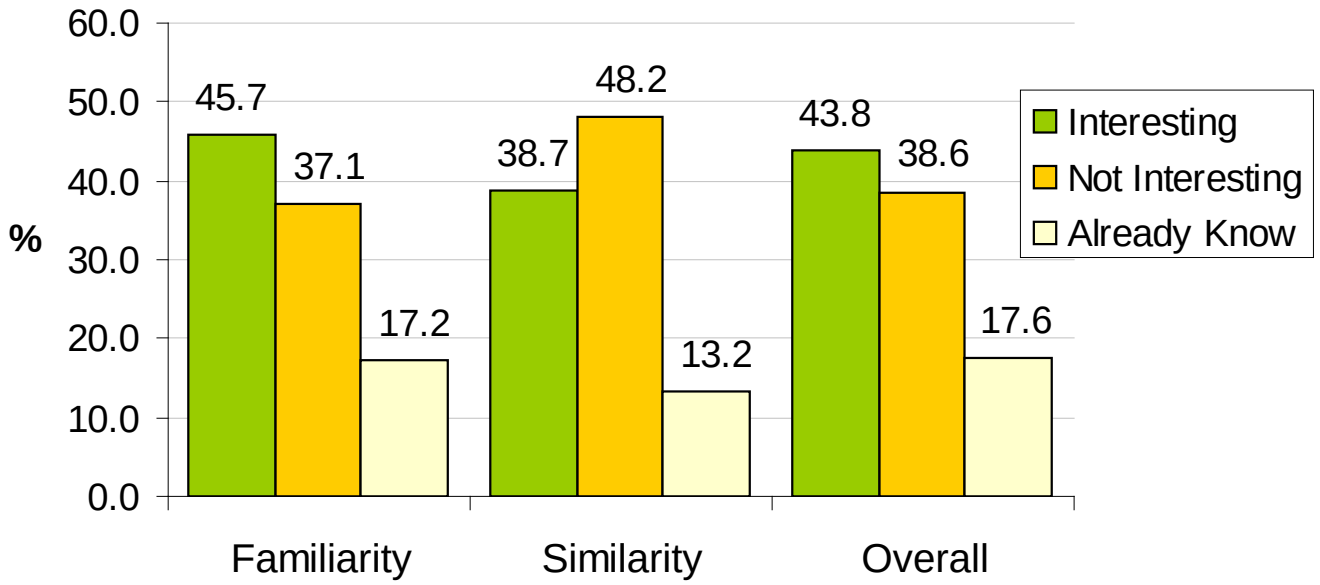
Main Evaluation-User Survey

- User Survey sent to 757 LC users who have rich networks
  - 290 responded (28 countries across all IBM divisions)
- Two phases – each suggesting six items to rate



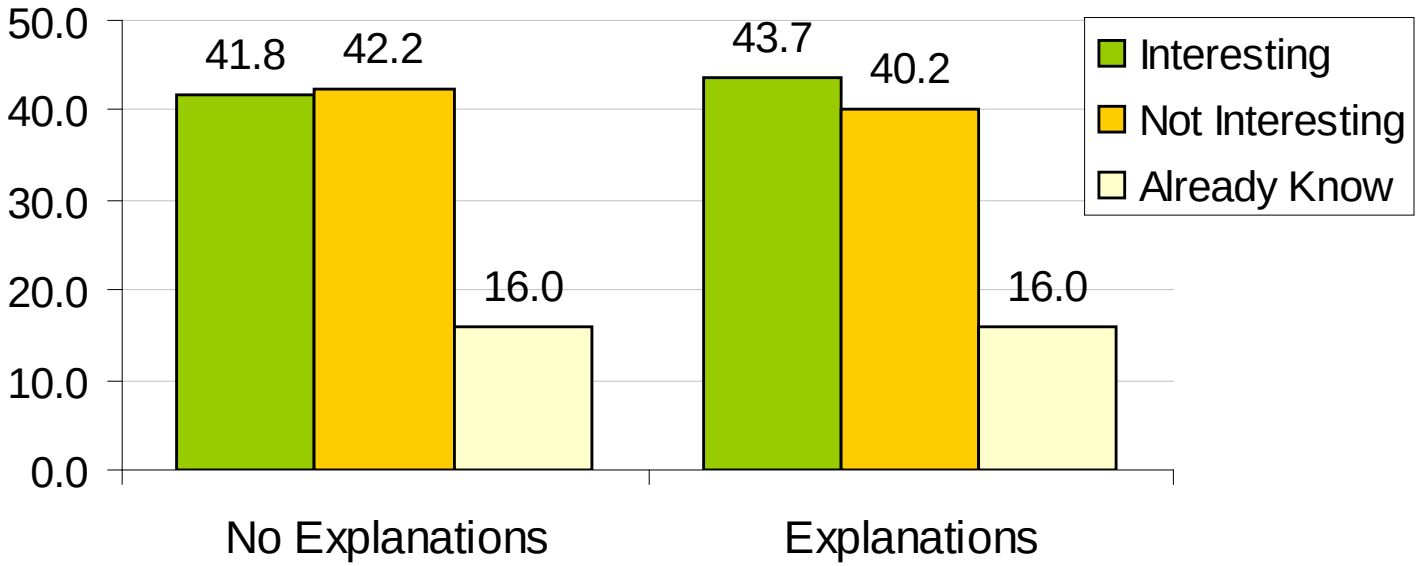
  - 2 of each item type (randomly ordered)
  - One phase without explanations, the other with
  - Not explicitly told whether to click an item to evaluate it
- Three groups – familiarity, similarity, overall

Familiarity vs. Similarity



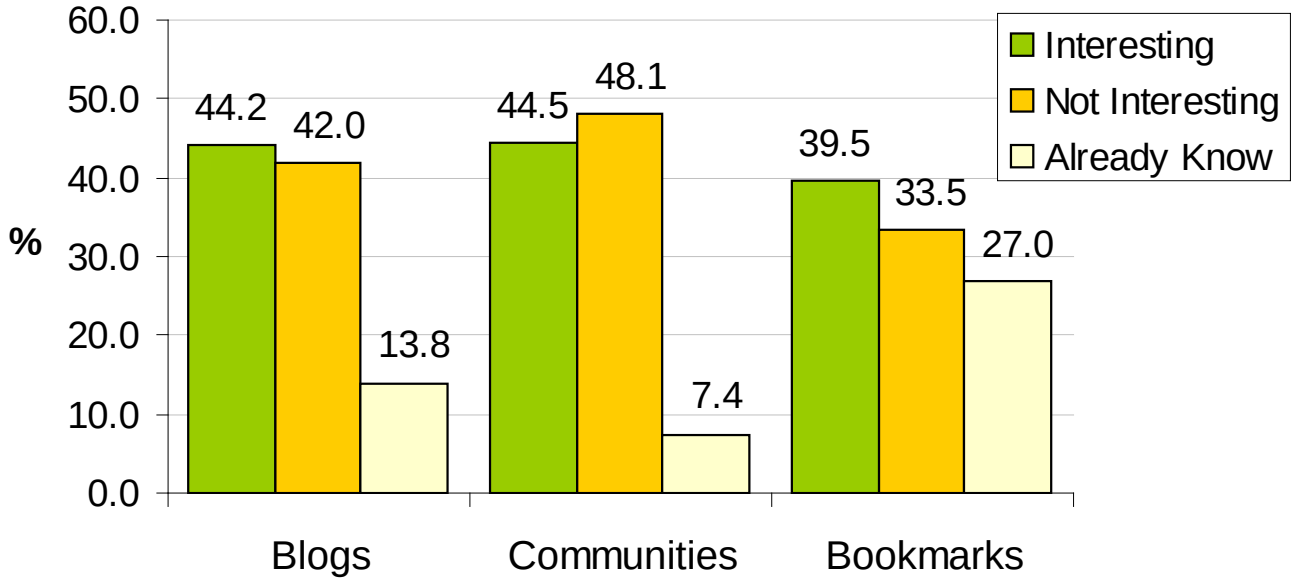
- Familiarity results are significantly higher than similarity
- Familiarity results are slightly higher than overall

Explanations



- Items with explanations are rated higher than items w/o explanations
- Differences are not significant, but consistent over the three groups
- *"The explanations let me understand the reason for an item showing up even if I did not find it interesting"*

Item Type Comparison



- Bookmarks have the highest interest ratio
- Communities have the highest absolute interest percentage
- "Already Know" difference is most notable – bookmarks are more expected, while communities are more surprising

Conclusions

- Both familiarity and similarity relationships can produce interesting recommendations, however familiarity is more effective
- There is no “whole greater than the sum of its part” effect -overall network does not improve over familiarity
- Explanations have instant effect in increasing interest rate in recommendations
- Bookmarks, blogs, and communities all yield interesting recommendations
  - Bookmarks are most effective yet expected, communities are more surprising

Future Work

- Combining social networks with content-based filtering
- Feedback mechanism
  - Both at the item level and at the explanation (person) level
- Experiment outside the enterprise
  - Friends rather than colleagues
- Examine benefit of using more complex algorithms over aggregated data

Any Questions?

Thanks