

A Tag-Based Blog Recommendation System for Smart-M3 Multi-Blogging

Diana V. Zaiceva, Dmitry G. Korzun

Petrozavodsk State University
Department of Computer Science



10th FRUCT Conference, November 11th, Tampere, Finland

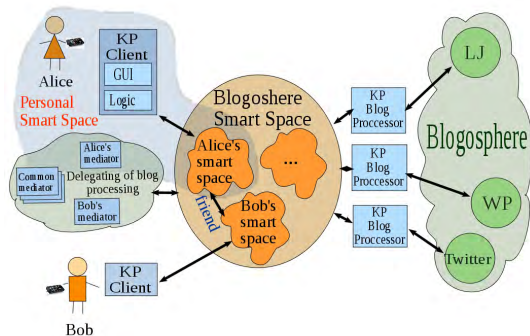
Table of Contents

- 1 Multiblogging with SmartScribo
- 2 Blog Recommendation
- 3 Tags Ranking
- 4 Conclusion



SmartScribo: Smart-M3 Application

- Distributed multi-agent architecture
- Blogosphere is shared in the smart space
 - ▶ partially
 - ▶ dynamically
- Interaction with multiple blogs at many blog services
- Agents are KPs:
 - ▶ clients
 - ▶ processors
 - ▶ **mediators**
- Synchronization by notification (subscription)
- Proactive blog retrieval, e.g., based on **recommendation**
- Individual and collaborative activity



Personal Smart Spaces

■ Person

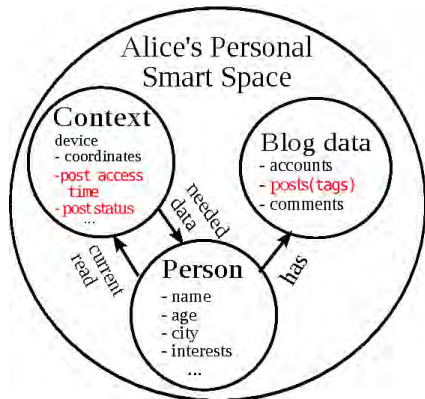
- ▶ personal (long-term) interests

■ Blog data

- ▶ **post** starts discussion
- ▶ discussion evolves with commentaries
- ▶ **tags** describe post semantics

■ Context

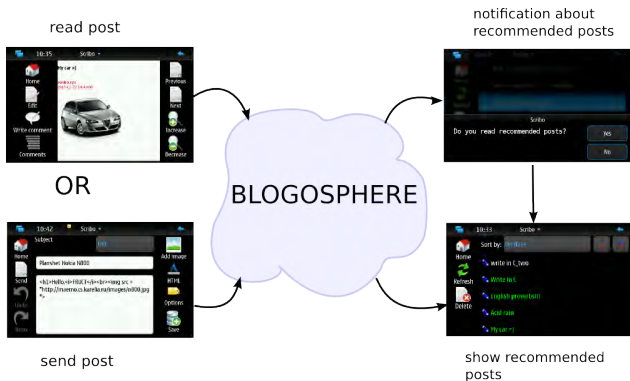
- ▶ coordinates & surroundings
- ▶ ...
- ▶ post read&write: **post status**
- ▶ given post p and user u ,
 T_{pu} is the latest post access time



Objective: Proof-of-the-Concept

Searching personally-interesting discussions in the blogosphere

- Are smart spaces an appropriate paradigm?
- user u context for post p
 - ▶ read or send: **post status**
 - ▶ latest post access time T_{pu}

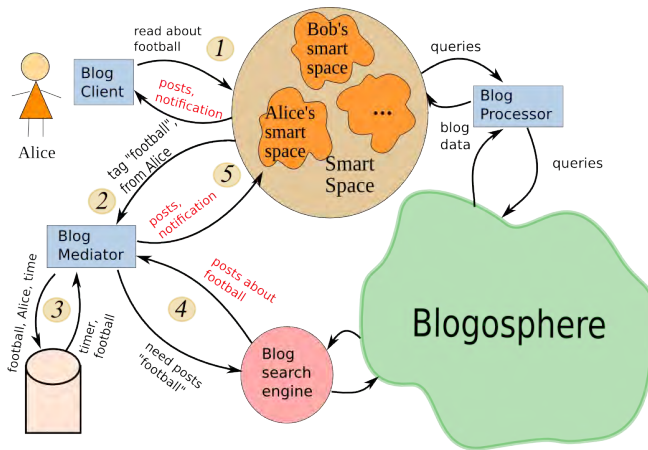


Our testbed: **Smart-M3** and **SmartScribo**



Basic Steps

- 1 User u activity: post access time and other attributes
- 2 Tracking post status
- 3 Tag index $data(t, u)$ and rank R_{tu}
- 4 Blog discovery for top-ranked tags
- 5 Publishing blog recommendation



Blog Search Engines

Examples

- Technorati
- Google Blog Search
- LjSeek.com
- BlogScope
- BlogPulse
- Regator.com

Search queries

`http://www.blogpulse.com/search?boolean=false&operator=and&keywords=cat+mouse+milk&...`

`http://api.regator.com/posts/search?searchTerm=cat&...`

Blog mediator provides input for search queries

A simple case is query for blogs containing given tags:

$$p \in \text{blogosphere s.t. } t \in p \text{ for some } t \in \mathcal{T}$$



Some Known Classes of Recommendation Methods

- Content-based algorithm
 - ▶ Decision trees and rule induction
 - ▶ Nearest neighbour methods
 - ▶ Relevance Feedback and Rocchios Algorithm
 - ▶ Linear Classifiers
 - ▶ Probabilistic Methods and Naive Bayes
- Collaborative filtering
 - ▶ user based collaborative filtering
 - ▶ item based collaborative filtering
- Demographic recommendation
- Utility-based recommendation
- Knowledge-based recommendation



Simple Model

User u , post p , tag t

$n_p \geq 1$ is the sum number of tags in p

- n_p is a post attribute kept in the blogosphere smart space

Blog mediator tracks status of p and updates tags index

- δ_t is the time elapsed from the latest update (discrete: 0, 1, 2, ...)

R_{tu} shows importance of blogs with t to u

$$R_{tu} = \begin{cases} 1/n_p, & \text{if } \delta_t = 0 \text{ (i.e., } t \text{ is new)} \\ (1 - \alpha)R_{tu}^{\text{old}} + \frac{\alpha}{\delta_t n_p}, & \text{if } \delta_t \geq 1 \end{cases} \quad (1)$$

- $1/n_p$ is topic-focus level of p
- $1/\delta_t$ is freshness of t (tag relevance)
- $0 < \alpha < 1$ is a tradeoff parameter



Tag Relevance

Simple popularity parameter of t personalized for u

$$\delta_t = F\left(\{T_{qv}\}_{t \in q, v \sim u}\right)$$

Factor $1/\delta_t$ shows access activity for posts with t

The higher δ_t the less interest to u in such posts

- Individual activity of u

$$\delta_t = \delta_t(u) = T_0 - \max_{t \in q, q \neq p} T_{qu}$$

- Collective activity: group C influences u 's decisions

$$\delta_t = \delta_t(C) = T_0 - \max \left\{ \max_{t \in q, q \neq p} T_{qu}, \max_{v \in C \setminus \{u\}} \max_{t \in q} T_{qv} \right\}$$

where T_0 is the current time at blog mediator



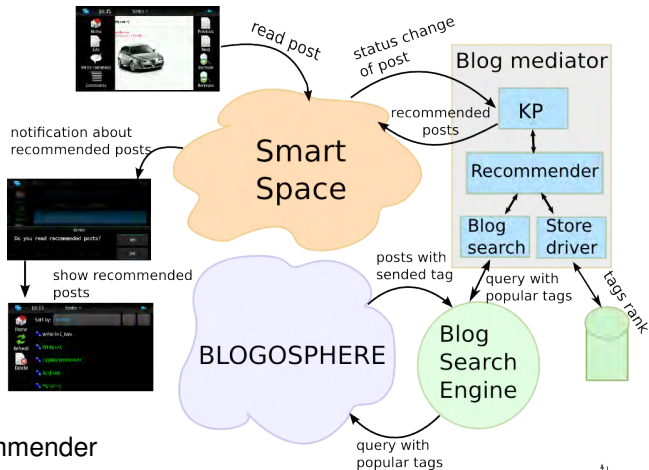
Architectural Design

Blog mediator

- subscribe to post status
- tags indexing and ranking
- query to blog search engine

Blog client

- subscribe to recommender
- proactively receive recommendations
- accept and enjoy



Current Status and Results

- Ongoing research in proof-of-the-concept style
- Usecase scenario and architectural design
 - ▶ Smart-M3 features for smart applications
 - ▶ SmartScribo multi-agent architecture
 - ▶ SmartScribo synchronization mechanism
 - ▶ SmartScribo GUI design rules
- In-progress implementation
 - ▶ Simple model
 - ▶ Use of available services
 - ▶ <http://oss.fruct.org/wiki/SmartScribo>
 - ▶ <http://gitorious.org/smart-scribo/smart-scribo>



Our Plans

- 1 Evaluation experiments
 - 2 From simple model to usable solutions
 - 3 Benefits from known recommendation methods
 - 4 Adaptation to function variety of existing blog search engines
-
- SmartScribo project wiki (in Russian)
<http://oss.fruct.org/wiki/SmartScribo>
 - Open source
<http://gitorious.org/smart-scribo/smart-scribo>
 - Mailing list
smart-scribo@cs.karelia.ru

Q&A

