

Contextual Matching-based Blog Overlay Network for Information Sharing on Blogosphere

Jason J. Jung · Chul-Mo Koo

Trackback RSS Feed 가 가
Context 가 가
Network Isolation 가 가
BlogRoll Context Community
BON(Blog Overlay Network) Tag
identification Mental
Matching

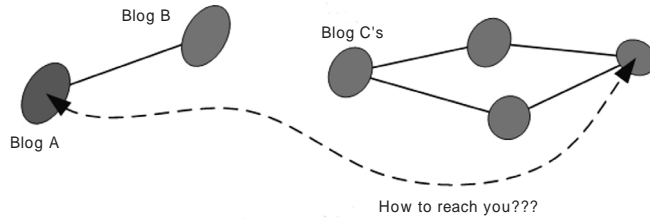
: , , Community of Practice, Context matching

Blog can provide users efficient information sharing services through Tackback and RSS feeds. However, they are not taking into account the context, so that the recommended information is not useful at all. More seriously, not only the increasing number of blogs but also network isolation phenomena on blogosphere make the information sharing among blogs more difficult. In order to solve these problems, in this paper, we want to improve the information sharing services by exploiting linkage information like blogroll. In particular, for employing social network analysis methods, we propose a novel blog platform, called blog overlay network (BON), to identify communities with respect to the discovered contexts. By using this platform, we want to match tags, whcih are represented by blogger's mental model, for supporting collaborations exchanging information contained the corresponding blogs.

Keywords: Blogs, Information Sharing, Community of Practice, Context matching

I. 1 (Publishing)
Web 2.0 가 Blog (,)
 (Blogs Weblogs) [18].
Web 2.0 (Posting Information)

:
:
:



1. Network Isolation

(Repository)

[1]. 가 (Platform)

가 (Social Interaction)

(Knowledge Sharing)

(Collaboration) 가 (Virtual Space)

[2]. 가 [11].

- (Comment) (Trackback)
- (Aggregation)
- RSS(RDF Site Summary)[13] Feeding (Propagation)

(Collective Intelligence)[12] (Knowledge Creation)

Integration) (Knowledge 가 (Pivot)

(Knowledge Transformation) (Dissemination)

(Exploitation) Cohen and Levinthal (Absorptive Capacity) -- "an ability to recognize the value of new information, assimilate it, and apply it to commercial ends"

[14].

(Bulletin Board)

(Connectivity), (Centrality), (Cliques) [15]. (Social Inclusion) (Agreement) [16].

가 1 Blog A B Blogosphere Blogrolling 가 가

Blog A Blog C (Dimension)

Feeds RSS (Web Syndication)[17] (Mailing List)

RSS (Notification) Context (

Tag

Contextual Heterogeneity[19]

Network) (Social
 (Post) (Hyperlink)
 BlogRoll
 가
 (Social Network Analysis)
 (Hub), (Authoritative)
 (Centrality) (Social
 Features) [5],[6].
 (Centrality)
 (BON) Blog Overlay Network
 Overlay
 Contextual Matching
 Context Community
 Community
 II Blog
 Overlay Network Tag Contextual Matching
 Community III Blog
 Overlay Network 가 IV
 V

Blogroll
 $R_K^{Roll} B_K \times B_{K'}$
 $P_K \times T_K$
 ,
 Trackback
 $R^{Comment}$ $R^{Trackback}$
 Comment
 Blogroll Graph
 Barabasi
 (Degree of Centrality)가
 가
 (SPD:
 Shortest Path Distance) 가
 • Closeness:
 가

$$Closeness(B_K) = \frac{N-1}{\sum_{K'=1, K'}^N SPD(B_K, B_{K'})}$$
 N 가
 Closeness 가 가
 가
 가
 • Betweenness:

$$Betweenness(B_K) = \frac{|\{ \langle B_i, B_j \rangle \mid SPD(B_i, B_k) + SPD(B_k, B_j) = SPD(B_i, B_j) \}|}{N(N-1)}$$

II. Blog Overlay Network

(Multi-layered social network)

1.

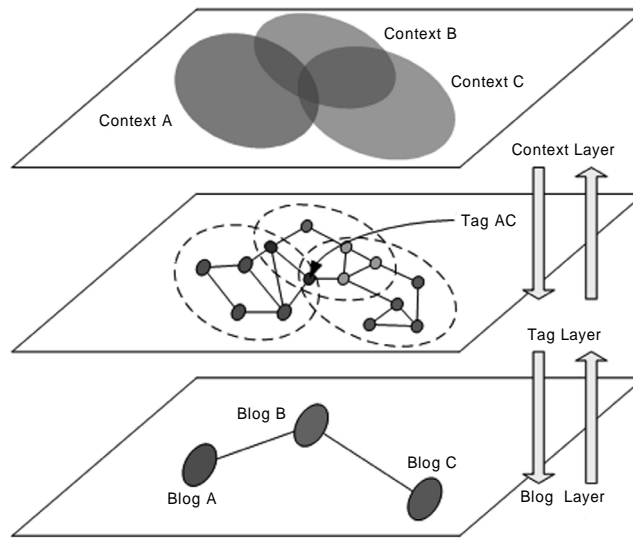
B_K

(Pivot)

$$B_K = \{P_K, T_K, R_K^{Roll}, R_K^{Comment}, R_K^{Trackback}, R_K^{Tag}\}$$

• Hub and Authoritative: Hub
 Authority 가
 Authority
 Hub Kleinberg[6]
 Iteration Blogosphere

$$P_K = \{p_1, p_2, \dots, p_N\} \quad T_K = \{t_1, t_2, \dots, t_M\}$$



2. Blog Overlay Network

Hub Authority

Mental Model

$$Hub_{t+1}(B_K) = \sum_{B_{K'}} R_K^{Roll} Auth_t(B_{K'})$$

• Context Layer: Matching (Similarity)

$$Huth_{t+1}(B_K) = \sum_{B_{K'}} R_K^{Roll} Hub_t(B_{K'})$$

가

Authority 가
Betweenness 가

Hub Authority가

Social feature
Blog layer
Network Isolation
Tag Layer Mapping
Tag (Potential)

2.

Blog Overlay Network

Blog A Blog C 가
Blog B Mediation 2
Tag Layer “Tag AC” Context
Mapping 가

• Blog Layer:

R^{Roll}

Blogroll
Blog Layer
Blog B

3. Contextual matching

가

Blog A C

Tag layer
Tag Matching

• Tag Layer:

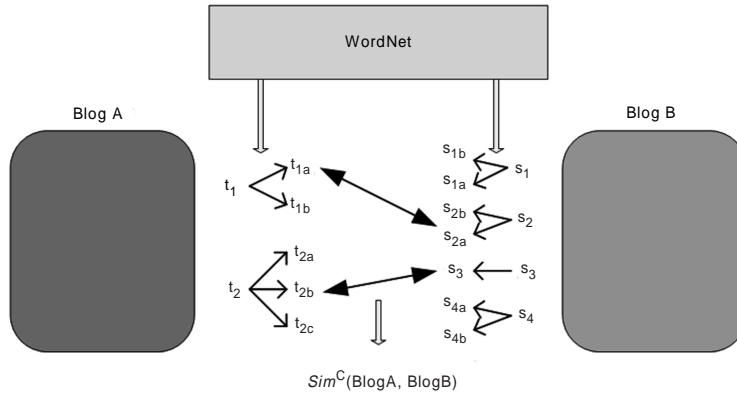
R^{Tag}

Labeling

가
Blog Contextualuzation 가
Context

Context

가



3. Ontology Tag Matching

(network substance)'

tion

WordNet Contextualiza-
Context Matching
(Similarity) Sim^C
 B_i, B_j Context Matching
Semantic Enrichment Tag
Matching
가

가

3.1. Contextualization

$$Sim^C(B_i, B_j) = \max \sum_{T_i, t'} T_i Sim^T(t, t')$$

Semiotic context Tag
Lexical Form Heterogeneity ("apple" "apples")
Contextual Mismatching
"Apple" Tag 가
"Computer"
"Fruit"

$$\max \sum_{T_i, t'} T_i \left(1 - \frac{Dist(t, t')}{\max(|t|, |t'|)} \right)$$

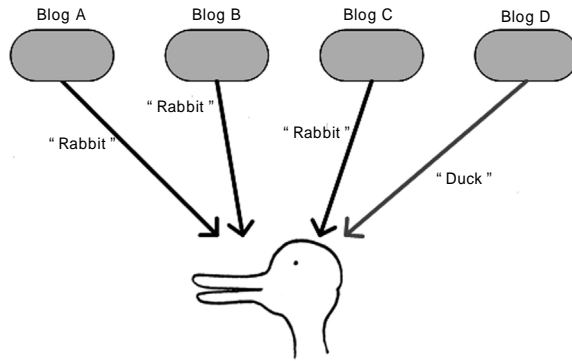
$Dist$ Edit
Distance (Substring Distance)
Similarity 2 Context Layer
Community Identification
Weighting

(Ontology) Semantic
Contextualization

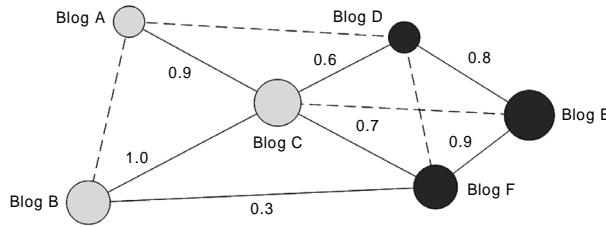
3.2. Consensus Contextualization

B_K Tag (T_K)
WordNet[8] Lexical
3 WordNet
Tag
Semantic
Enrichment "Computer"
Tag가 "Computing Machine," "Computing
Device," "Data Processor," "Electronic Computer,"
"Information Processing System," "Calculator"

Contextualization
Consensus context
Tag Tag
Reliable Trustworthy
4
Tagging
(Blog A, B, C) "Rabbit" Tag
Blog D "Duck" Tag



4. Consensus Tag Contextualization



5.

Factor(CF) Assign Tag Tagging Consensus Similarity Similarity Context Community Newman[9] Modularity function

$$CF(R_K^{Tag}) = \left\{ \langle p_i \times t_j, cf \rangle \mid cf = \frac{|\{B_H | p_i \times t_j, R_H^{Tag}\}|}{|\{B_G | p_i \times t_x, R_G^{Tag}\}|} \right\}$$

, Consensus Factor가 Tag Context

$$Q(CL) = \sum_{i=1}^K \frac{\sum_{B_a, B_b \in Comm_i} Sim^C(B_a, B_b)}{|Comm_i|}$$

4. (Dynamic Community Identification)

K Community Q Community formation 가

Context Layer Community Context layer propagation Posting Notification Context Layer 5 Tag Matching information 가 Context

Similarity 가 가 Community1={A, B, C} Community2={D, E, F} Community identification 가 Aggregation Tag가 Community (Centrality)가

1. Tag Matching (Relevant deliveries/Total deliveries)

Bloggers	Simple Blogrolling	Tag matching			
		Simple String Matching		WordNet	
1	0.27	0.31	114.02%	0.40	146.86%
2	0.26	0.28	107.28%	0.35	134.10%
3	0.28	0.33	116.96%	0.37	129.68%
4	0.27	0.29	110.11%	0.39	145.32%
5	0.32	0.29	90.88%	0.39	121.70%
6	0.35	0.31	89.88%	0.39	113.29%
7	0.29	0.34	113.95%	0.39	130.95%
8	0.35	0.32	90.91%	0.40	113.92%
9	0.23	0.34	150.00%	0.40	176.75%
10	0.32	0.36	112.73%	0.41	128.57%
11	0.36	0.32	89.30%	0.41	116.06%
12	0.31	0.34	109.48%	0.39	126.80%
13	0.34	0.30	87.72%	0.44	129.24%

2. Community (K = Number of communities)

K	Tag matching			
	Simple String Matching		WordNet	
2	0.34	107.13%	0.40	101.42%
3	0.35	110.28%	0.35	114.10%
4	0.41	129.18%	0.37	88.75%

III. Context

Blog Overlay Network
 Tag Matching Community
 ()
 BlogGrid[10] 1)
 Tagging , 206 Post 194
 Tag
 Tag Matching 1 3가
 Tag Matching 가

Tag Context Matching Tag 가
 WordNet
 Tag Matching 가 32% 가
 Simple string matching 6.4%
 가
 Community
 2 13 K=2,
 3, 4 , Community 가
 K=4 , Community

IV.

Context matching Overlay network
 Context Context

1) BlogGrid et al [10] Jung

2.0 가 Open API
 Mash-up Application
 SEMKEY[7] Conceptualization
 WorkNet Wikipedia
 Jung and Euzenat[4]
 Peer-to-Peer
 (Personal Ontology) Alignment
 Semantic Social network
 Multi-layered architecture
 Ontology Context
 heterogeneity
 Jung[3] Hyperlin-
 ked Focused Crawling
 Web Information System
 Context
 Tag

V.

Multi-layered Blog Overlay
 Network
 • Tag Context Matching
 • Tag Matching Blog Similarity
 • Context (Mediation)
 Jung[3] Heavy
 Expensive Ontology Alignment
 가 Web 2.0
 Scalability
 Community (Social
 capital) 가 가
 (Semantic
 blog) Personal Ontology
 Tag
 Term Semantic Relationship
 User Friendly
 e-Learning Blog
 BON

- [1] B.A. Nardi, D.J. Schiano, M. Gumbrecht, and L. Swartz, "Why we blog," *Communications of ACM*, Vol. 47, No. 12, 2004, pp. 41-46.
- [2] W. Treese, "Open systems for collaboration," *netWorker*, Vol. 8, No. 1, 2004, pp.13-16.
- [3] J.J. Jung, "Ontological Framework Based on Contextual Mediation for Collaborative Information Retrieval," *Information Retrieval*, Vol. 10, No. 1, 2007, pp. 85-109.
- [4] J.J. Jung and J. Euzenat, "Towards Semantic Social Networks," In E. Franconi, M. Kifer, and W. May (Eds.): *Proceedings of the 4th European Semantic Web Conference*, Vol. 4519 of *Lecture Notes in Computer Science*, 2007, pp. 267-280.
- [5] S. Wasserman and K. Faust, *Social network analysis: methods and applications*. Cambridge University Press, Cambridge (UK), 1994.
- [6] J.M. Kleinberg, "Authoritative sources in a hyperlinked environment," *Journal of the ACM*, Vol. 46, No. 5, 1999, pp. 604-632.
- [7] A. Marchetti, M. Tesconi, and F. Ronzano, "SemKey: A Semantic Collaborative Tagging System," *Proceedings of the World Wide Web Conference (WWW2007)*, 2007.
- [8] WordNet, <http://wordnet.princeton.edu/>
- [9] M.E.J. Newman, "Fast algorithm for detecting community structure in networks," *Physical Review E*, Vol. 69, 2004, pp. 66-133.
- [10] J.J. Jung, I. Ha, G.-S. Jo, "BlogGrid: towards an efficient information pushing service on blogspace," *Proceedings of the 4th International Conference on Grid and Cooperative Computing (GCC 2005)*, Vol. 3795 of *Lecture Notes in Computer Science*, 2005, pp.178-183.
- [11] I. Nonaka and N. Konno, "The Concept of "Ba": Building Foundation for Knowledge Creation," *California Management Review*, Vol. 40, No. 3, 1998, pp. 40-54.
- [12] Blog of Collective Intelligence, <http://www.community-intelligence.com/blogs/public/>
- [13] <http://web.resource.org/rss/1.0/>
- [14] W.M. Cohen and D.A. Levinthal, "Absorptive Capacity: A New Perspective on Learning and Innovation," *Administrative Science Quarterly*, Vol. 35, 1990, pp. 128-152.
- [15] M.S. Rajasekaran and P. Zaphiris, "*Social Network Analysis of a Participatory Designed Online Foreign Language Course*," In J. Jacko, C. Stephanidis (Ed.), *Human-Computer Interaction, Theory and Practice*, Lawrence Erlbaum, 2003, pp. 218-222.

- [16] B. Madsen, "The vision of social inclusion - a risky process for the professionals and the excluded," 2003, Accessed online on August 2 2004 at:
http://www.cvastork.dk/filer/key_note.doc
- [17] http://en.wikipedia.org/wiki/Web_syndication
- [18] <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
- [19] P. Bohannon, E. Elnahrawy, W. Fan and M. Flaster, "Putting context into schema matching," Proceedings of the 32nd international conference on Very large data bases (VLDB 2006), 2006. pp. 307-318.



(Jason J. Jung)

1999. 2:
 2002. 2:
 2005. 8:
 2006. 9: INRIA Rhone-Alpes,
 2006. 11 ~ 2007. 8:
 2007. 8~ :
 : Semantic Web, Ontology,
 Description Logic, Web 2.0
 E-mail: j2jung@intelligent.pe.kr,
j2jung@gmail.com
 Tel: + 82-32-875-5863
 Fax: + 82-32-874-1435



(Chul-Mo Koo)

1997. 2:
 1999. 2:
 2003. 8:
 2004.3 ~ 2005. 2: Associate Researcher,
 MIS Research Center,
 University of Minnesota
 2004.8 ~ 2006. 5: MIS Faculty, Lewis College
 of Business, Marshall University
 2007.1 ~ : BK21

: Knowledge Management, Electronic
 Commerce Strategy, Mobile Commerce,
 Social Communication Media,
 Supply Chain Management,
 IT Outsourcing & SLA

Email: helmetgu@inha.ac.kr
 Tel: + 82-32-872-7445
 Fax: + 82-32-872-7446