Mutual contextualization in tripartite graphs of folksonomies

Ching-man Au Yeung, Nicholas Gibbins, Nigel Shadbolt
{cmay06r, nmg, nrs}@ecs.soton.ac.uk

Intelligence, Agents and Multimedia Group,
School of Electronics and Computer Science
University of Southampton

Overview

Objectives

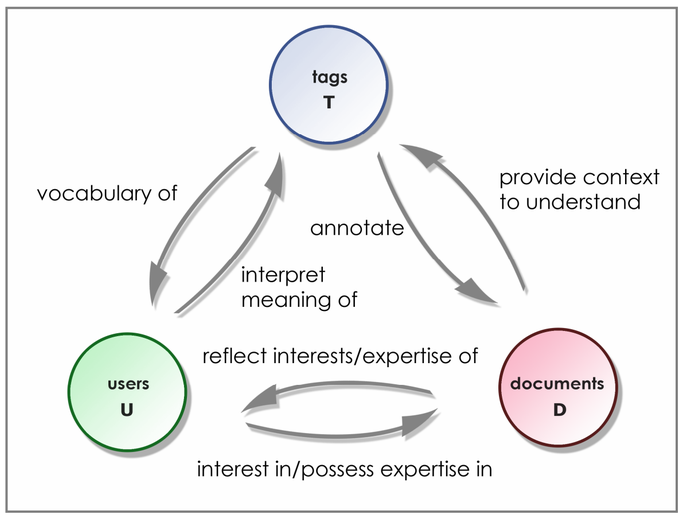
- To understand the dynamics in folksonomies
- To investigate how semantics can be extracted from folksonomies
- To study the potential applications of the information extracted from folksonomies

Methods

- Apply network analysis techniques on graphs of folksonomy
- Apply results on various applications: e.g. tag disambiguation and recommender systems

Mutual Contextualization

- What does a user ID, a tag or a URL represents?
- Folksonomies contain interrelations between these entities.
- The entities provide a context for each other so that their semantics can be understood.
- Hence, we can acquire a better understanding of the entities by exploring the associations between them.



A folksonomy $F = (U, T, D, A)$

A is a set of annotations: $A \subseteq U \times T \times D$

Study Individual Elements in a Folksonomy

- To study individual elements in a folksonomy, we extract bipartite graphs from the folksonomy hypergraph.
- For example, for a particular user, we can extract all the tags used and all the documents tagged by this user; this is similar for a tag or a document.
- We can then fold the bipartite graphs into different one-mode networks for analysis.
- Findings are expected to be useful and beneficial to various applications on the Web.

A user (*u*)

Bipartite graph TD_u

$TD_u = \langle T \cup D, E_{TD} \rangle$

$E_{TD} = \{ \{t,d\} \mid \{u,t,d\} \in A \}$

A tag (*t*)

Bipartite graph UD_t

$UD_t = \langle U \cup D, E_{UD} \rangle$

$E_{UD} = \{ \{u,d\} \mid \{u,t,d\} \in A \}$

A document (*d*)

Bipartite graph UT_d

$UT_d = \langle U \cup T, E_{UT} \rangle$

$E_{UT} = \{ \{u,t\} \mid \{u,t,d\} \in A \}$

Folded into one-mode networks by adjacency matrix multiplication

- tag network
- document network

- user network
- document network

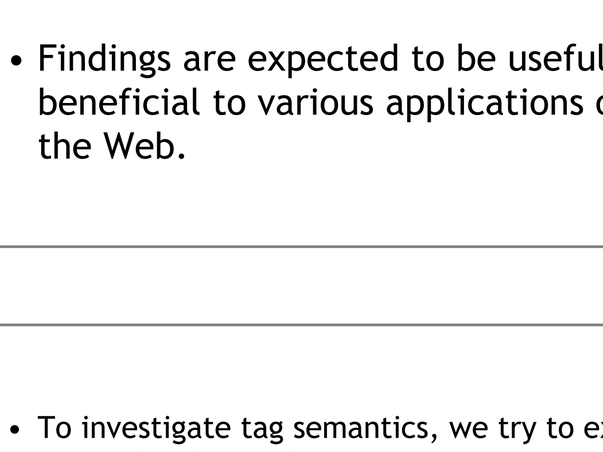
- tag network
- user network

Potential Applications

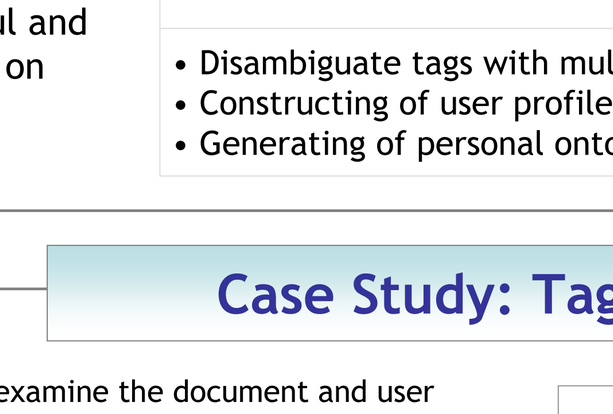
- Disambiguate tags with multiple meanings
- Constructing of user profiles
- Generating of personal ontologies
- Improving recommender systems
- Deriving social networks of shared interests
- Generating annotations for documents

Case Study: Tag Disambiguation

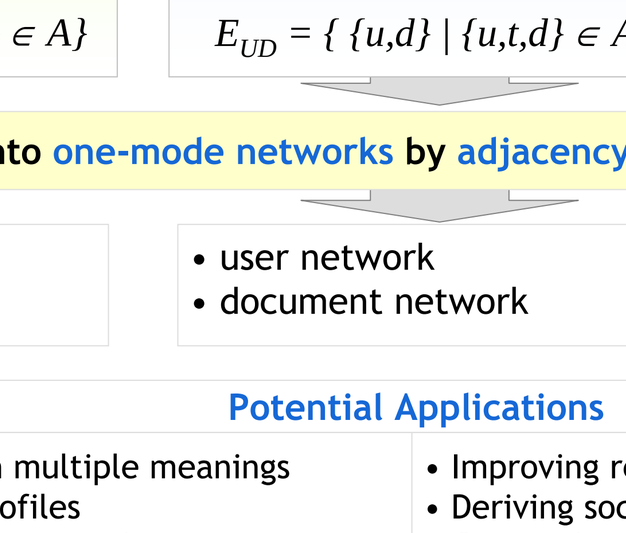
- To investigate tag semantics, we try to examine the document and user networks of a single tag.
- Focusing on the tag ***sf***, we extract the document network and user network.
- We study the possibility of automatic tag disambiguation by examining the topology of the networks.
- Results show that majority of the users use the tag to refer to one meaning only, resulting in different clusters of documents.



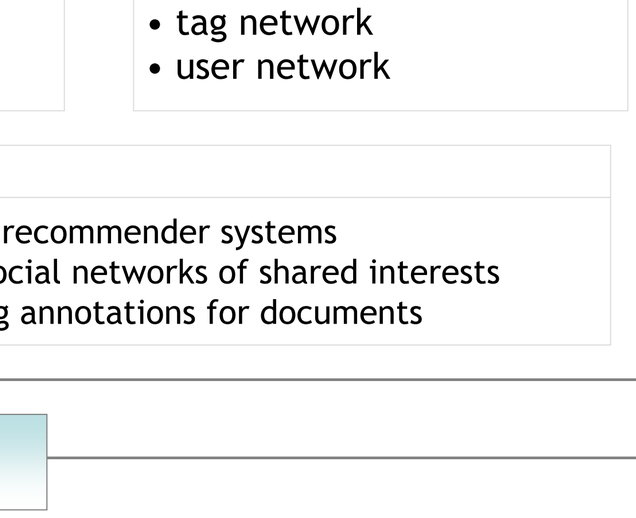
Network of documents being tagged by the tag sf



Network of users who have used the tag sf



Network of documents, edges with weights < 2 removed



Network of most frequently used tags

A paper on this work will be presented in the **International Workshop of Emergent Semantics and Ontology Evolution (ESOE 2007)** on 12 Nov. You can also visit http://www.ecs.soton.ac.uk/~cmay06r/ for more information.

Doctoral Consortium, International Semantic Web Conference 2007, Busan, South Korea, 11-15 Nov 2007