

Biomedical Text Understanding and User Modeling

Hyoil Han

Assistant Professor, College of Information Science and Technology,
Drexel University

Abstract:

Biomedical researchers who search for documents via PubMed, a service of the National Library of Medicine (NLM) and the National Institutes of Health (NIH), use indexing tools that rely heavily on published abstracts and self-reported keywords. Currently, PubMed contains over 16 million citations. The number of citations in PubMed is expected to grow at a rate of over 500,000 articles annually. Coupled with the sheer volume of articles added every year, PubMed has become a relatively hostile online research environment. The answer to a simple query in PubMed may contain over 10,000 responses. Although users may find documents related to their research, the number of articles retrieved requires a tremendous amount of review, which may prevent them from discovering significant scientific information. The first part of my talk is about Biomedical Text Understanding. To improve such a hostile online research environment, we are working to build an intelligent system to summarize biomedical articles. The main idea for the system is to develop natural language understanding algorithms based on concepts rather than terms, concept chaining and frequencies in the articles. The second part of my talk is about User Modeling. Semantically enhanced user modeling is used to build a personalization system based on Web content visited by each user. The main idea is to use implicit content-based user modeling based on sense disambiguation to improve personalization.

Bio Sketch:

Hyoil Han is an Assistant Professor at Drexel University. She obtained an MSc Degree from Korea Advanced Institute of Science and Technology (KAIST).

She then worked for the Samsung Electronics and Korea Telecom before obtaining a PhD in Computer Science and Engineering from the University of Texas at Arlington in 2002. Her research areas lie in the merging of techniques from the fields of databases (DB) and artificial intelligence (AI) and applying the new combined techniques to biomedical informatics and the Semantic Web with an emphasis on text mining, and data integration and management. She has published 30 papers in refereed literature.