Fall 2008 Presentation Schedule

CS591Han Seminar: Advanced Topics on Data Mining (Fall 2008 Presentation Schedule)

Note: This semester's presentation will be focused on the theme of Information Network Analysis. Papers can be selected from this year or previous several year's conference proceedings or journals. We welcome students who would like to present tutorials and/or writing topic survey articles related to this theme.

Two students per unit (20 minutes presentation and 5 minutes discussion for each research paper, i.e., two papers will be covered per class unit.

Week 1 (8/28/08): Class organization and report of KDD'08 conference (Jiawei Han)

Week 2 (9/04/08): Two presentations

1. Yizhou Sun: "An introduction to BibNetMiner and RankClus"

Yizhou Sun, Tianyi Wu, Hong Cheng, Jiawei Han, Xiaoxin Yin, Peixiang Zhao and Min-Soo Kim, "BibNetMiner: Mining Bibliographic Information Networks”, (demo paper), Proc. 2008 ACM SIGMOD Int. Conf. on Management of Data (SIGMOD'08), Vancouver, BC, Canada, June 2008

Paper link

Yizhou Sun et al, "RanClus", submitted

Presentation link

2. Peixiang Zhao:

Antonellis, Ioannis; Garcia-Molina, Hector; Chang, Chi-Chao, "Simrank++: Query Rewriting through Link Analysis of the Click Graph", (VLDB'08),

Abstract: We focus on the problem of query rewriting for sponsored search. We base rewrites on a historical click graph that records the ads that have been clicked on in response to past user queries. Given a query q, we first consider Simrank as a way to identify queries similar to q, i.e., queries whose ads a user may be interested in. We argue that Simrank fails to properly identify query similarities in our application, and we present two enhanced versions of Simrank: one that exploits weights on click graph edges and another that exploits "evidence." We experimentally evaluate our new schemes against Simrank, using actual click graphs and queries from Yahoo!, and using a variety of metrics. Our results show that the enhanced methods can yield more and better query rewrites.

Paper link, Presentation link

Week 3 (9/11/08): (No class due to instructor's meetings at DC and Baltimore)

Week 4 (9/16/08): Two presentations---please note the time is changed to Tuesday, 9/16!

1. Yun Hee Lee

Elisa Bertino, Beng Chin Ooi, Yanjiang Yang, Robert H. Deng. Privacy and Ownership Preserving of Outsourced Medical Data (ICDE'05)

Presentation Slide

2. Tianyi Wu

Kaushik Chakrabarti, Surajit Chaudhuri, Venkatesh Ganti, Dong Xin. An Efficient Filter for Approximate Membership Checking (SIGMOD'08)

Presentation Link (rename the downloaded file name to *.pptx to make it work properly)

Week 5 (9/25/08): Two presentations
1. Cindy Xide Lin
Cindy Xide Lin, Bolin Ding, Jiawei Han, Feida Zhu, Bo Zhao. *Text Cube: Computing IR Measures for Multidimensional Text Database Analysis* (ICDM'08)

Presentation Link

2. Cuiping Li
Nan Wang, Srinivasan Parthasarathy, Kian-Lee Tan, Anthony K. H. Tung. *CSV: Visualizing and Mining Cohesive Subgraphs* (SIGMOD'08)

Presentation Link

**Week 6 (10/2/08): Two presentations**

1. Rui Li
Rui Li, Shenghua Bao, Ben Fei, Zhong Su, Yong Yu. "Towards Effective Browsing of Large Scale Social Annotations", in Proc. of WWW2007 pp. 943-952, May 8-12, 2007, Banff, Alberta, Canada(pdf, slides).

2. Bolin Ding
Xiao, X., Tao, Y. *Output Perturbation with Query Relaxation*. Proceedings of the 34th International Conference on Very Large Data Bases (VLDB), pages 857-868, 2008.

Presentation Slides

**Week 7 (10/9/08): Two presentations**

1. Seb Seith
Fei Wu, Raphael Hoffmann, Daniel S. Weld *Extracting data from wikipedia*. SIGKDD 08, pages 731-739. Presentation Link

2. Zhijun Yin
Yuanyuan Tian, Richard A. Hankins, Jignesh M. Patel *Efficient Aggregation for Graph Summarization* SIGMOD 08, pages 567-578 Presentation Link

**Week 8 (10/16/08): Two presentations**

1. Feida Zhu
Saket Navlakha, Rajeev Rastogi, and Nisheeth Shrivastava, "*Graph Summarization with Bounded Error*" SIGMOD'08 paper SIGMOD 08, pages 419 Presentation Link

2. Xin Jin

**Week 9 (10/23/08): Two presentations**

1. Dustin Botner

2. Chen Chen
Chen Chen, Cindy Xide Lin, Xifeng Yan, and Jiawei Han, *On Effective Presentation of Graph Patterns: A Structural Representative Approach*, Proc. 2008 ACM Conf. on Information and Knowledge Management (CIKM'08), Napa Valley, CA, Oct. 2008. Presentation Link
Week 10 (10/30/08): No presentation due to a travel commitment

Week 11 (11/6/08): Two presentations

1. Jing Gao
   Brian Gallagher, Hanghang Tong, Tina Eliassi-Rad, and Christos Faloutsos, Using Ghost Edges for Classification in Sparsely Labeled Networks, Proc. of KDD 2008. PPT Link (uiuc account and AD password to access), Video

2. Lu An Tang
   Graham Cormode, Divesh Srivastava (AT&T Labs, USA), Ting Yu, Qing Zhang (North Carolina State University, USA). Anonymizing Bipartite Graph Data using Safe Groupings. In VLDB 2008.

Week 12 (11/13/08): Two presentations

1. Yintao Yu
   Gueorgi Kossinets, Jon Kleinberg and Duncan Watts. The structure of information pathways in a social communication network. KDD'08 Presentation Link (use UIUC account and Active Directory password to access) Video

2. Zhenhui Li
   Xiaolei Li, Zhenhui Li, Jiawei Han, Jae-Gil Lee. Temporal Outlier Detection in Vehicle Traffic Data. In ICDE 2009. Slides.

Week 14 (11/20/08): Two presentations

1. Deng Cai
   Deng Cai, Qiaozhu Mei, Jiawei Han, and ChengXiang Zhai, Modeling Hidden Topics on Document Manifold, Proc. 2008 ACM Conf. on Information and Knowledge Management (CIKM'08), Napa Valley, CA, Oct. 2008

2. Sang Kim

Week 15 (11/27/08): No class due to Thanksgiving Break

Week 16 (12/4/08): No class: Data Mining Group Summary Meeting