Web工学 (Web Engineering) http://tinyurl.com/webeng2013

豊田正史 (Masashi Toyoda)

E-mail: mtoyoda@acm.org

目標

- Webに関する研究分野の基礎、および、最近の動向について学ぶ

Goal

- Learn basics of Web related research and its recent trends

進め方

- 論文の輪講形式
- 毎回3,4本の論文を担当者が発表する
  (15-20分発表、5-10分質疑）
- 論文は、以下に上げる基礎文献か、メジャーな国際会議から自由に選択

Format

- Seminar style
- Three or four students present research papers at each session
  (15-20 min. presentation, 5-10 min. Q&A)
- Choose papers from the following basic papers, or from measure international conferences

単位

- 出席は取らない
- 条件：
  - 発表を1回行う
    OR
  - レポートを2回提出

Credits

- No roll call
- Requirement:
  - 1 presentation,
    OR
  - 2 reports

Organization of sessions

- First half: Basic papers from 1998-2007
  - Google, PageRank, HITS,…
- Second half: Recent research papers
  - WWW Conference, WSDM
  - SIGMOD, VLDB, ICDE
  - SIGIR
  - KDD, ICDM
  - AAAI, ICWSM
  - ACL, EMNLP
  - Etc.

スケジュール

- 10/ 7(月）ガイダンス
- 10/15(火）休講
  - 発表論文候補提出
- 10/21(月）休講
- 10/28(月）Search engine, PageRank
- 11/5(火）Hub and Authority
- 11/11(月）Crawler
- 11/18(月）Structure and Evolution
- 11/25(月）Social Network
- 12/2(月）Recent papers 1
- 12/9(月）Recent papers 2
- 12/16(月）Recent papers 3
- 1/15(水）Recent papers 4
- 1/20(月）Recent papers 5

Schedule

- 10/ 7 (Mon）Guidance
- 10/15(Tue）Cancel,
  - Submit list of papers to read
- 10/21(Mon）Cancel
- 10/28(Mon）
- 11/5(Tue）
- 11/11(Mon）
- 11/18(Mon）
- 11/25(Mon）
- 12/2(Mon）
- 12/9(Mon）
- 12/16(Mon）
- 1/15(Wed）
- 1/20(Mon）

Papers for Session 1 (10/28)

- Search engine and PageRank
  - Sergey Brin and Lawrence Page. The Anatomy of a Large-Scale Hypertextual Web Search Engine. WWW7, 1998. (平川)
  - Page, Lawrence; Brin, Sergey; Motwani, Rajeev; Winograd, Terry. The PageRank Citation Ranking: Bringing Order to the Web. 1999. (槇)

Papers for Session 2 (11/5)

- Hub and Authority Analysis
  - J. Kleinberg. Authoritative Sources in a Hyperlinked Environment, Proc. 9th ACM-SIAM Symposium on Discrete Algorithms, 1998. Extended version in Journal of the ACM 46(1999). Also appears as IBM Research Report RJ 10076, May 1997. (井上)
  - Krishna Bharat and Monika R. Henzinger. Improved algorithms for topic distillation in a hyperlinked environment, SIGIR '98, 1998. （河村）
  - Ravi Kumar, Prabhakar Raghavan, Sridhar Rajagopalan, Andrew Tomkins. Trawling the Web for Emerging Cyber-Communities. WWW8, 1999

Papers for Session 3 (11/11)

- Crawler
  - A. Heydon and M. Najork. Mercator: A scalable, extensible web crawler. WWW8, 1999. （浅見）
  - J. Cho and H. Garcia-Molina. The evolution of the web and implications for an incremental crawler. VLDB2000, 2000. （碓井）
  - Junghoo Cho, Hector Garcia-Molina. Parallel crawlers. WWW2002, 2002. (中島)

Papers for Session 4 (11/18)

- Structure and evolution of the Web
  - Andrei Broder, Ravi Kumar, Farzin Maghoul, Prabhakar Raghavan, Sridhar Rajagopalan, Raymie Stata, Andrew Tomkins, Janet Wiener. Graph structure in the web. WWW9, 2000. (高木)
  - Dennis Fetterly, Mark Manasse, Marc Najork, and Janet Wiener. A Large-Scale Study of the Evolution of Web Pages. WWW2003, 2003. (栗原）
  - Alexandros Ntoulas, Junghoo Cho, Christopher Olston. What's new on the web?: the evolution of the web from a search engine perspective. WWW2004, 2004.

Papers for Session 5 (11/25)

- Social network
  - R. Kumar, J. Novak, P. Raghavan, and A. Tomkins. On the bursty evolution of blogspace. WWW2003, 2003. (菊地)
  - D. Gruhl, R. Guha, D. Liben-Nowell, A. Tomkins. Information Diffusion Through Blogspace. WWW2004, 2004.
  - Y. Y. Ahn, S. Han, H. Kwak, S. Moon, H. Jeong. Analysis of Topological Characteristics of Huge Online Social Network Services. WWW2007, 2007.（鈴木）

割り当て

- 最初の日に発表する2,3名は今日決定
- 全員メールで以下を知らせること (10/15〆切)
  - 学籍番号、氏名
  - 発表希望日（3日）
  - 発表希望論文3本 (基礎論文、最近の論文を両方含めること）
- 10/18に割り当てとスケジュールを発表する
- 注意：
  - 希望日が早いほど、論文選択は優先される
  - 論文を変更したい場合、速やかに申し出ること

Assignment

- Today, we decide the first 2 or 3 students
- All students should e-mail following info. by 15 Oct.
  - Student No, Name
  - 3 candidate dates, and
  - 3 papers you want to present (include both basic and recent papers)
- Assignment & Schedule will be released on 18 Oct.
- Note:
  - Students who select earlier day have more chance to present a preferred paper
  - Let me know, if you change the paper to present