WEBKDD'06: The 8th International Workshop on Knowledge Discovery on the Web

August 20, 2006, Philadelphia, Pennsylvania

Organizers:

Olfa Nasraoui, University of Louisville
Myra Spiliopoulou, University of Magdeburg
Jaideep Srivastava, University of Minnesota
Bamshad Mobasher, DePaul University
Brij Masand, Data Miners, Inc.
2006: Goals

• Research results in all areas of Web mining and Semantic Web mining, with an interest in a seven years' update:

  – What are the lessons learned on algorithms, semantics, data preparation, data integration and applications of the Web?

  – How are new technologies, like adaptive mining methods, stream mining algorithms and techniques for the Grid apply to Web mining?

  – What new challenges are posed by new forms of data, especially flat texts, documents, pictures, and streams?

  – Which lessons have we learned about usability, e-commerce applications, personalization, and recommendation engines?
Solicited Topics

- Web usage mining and Web analytics
- Web content and structure mining
- Modeling users and their interaction with the Web
- Adaptability and evolution in the Web
- Enabling technologies for Web mining, including technologies for
  - Web warehousing,
  - Data preparation,
  - Management of multi-relational data and streams,
  - Visualization,
  - Evaluation
- Application areas, including
  - Recommendation engines,
  - Web communities,
  - Web marketplaces,
  - Web search,
  - Web security and misuse and abuse of the Web and its services
Submitted Topics (25 papers, 17 accepted)

- Web usage mining and Web analytics
- Web content and structure mining
- Modeling users and their interaction with the Web
- Adaptability and evolution in the Web
- Enabling technologies for Web mining, including technologies for
  - Web warehousing,
  - Data preparation,
  - Management of multi-relational data and streams,
  - Visualization, and
  - Evaluation
- Application areas, including
  - Recommendation engines,
  - Web communities,
  - Web marketplaces,
  - Web search,
  - Web security and misuse and abuse of the Web and its services
Program

8:30 - 10:00

- 08:30-08:45 Introduction & Welcome

- **Session 1: Adaptive Web Navigation & Collaborative Filtering**
  - 08:45-09:10 Adaptive Website Design using Caching Algorithms
    Justin Brickell, Inderjit S. Dhillon, and Dharmendra S. Modha
  - 09:10-09:35 Usage-Aware Average Clicks
    Kalyan Beemanapalli, Ramya Rangarajan, and Jaideep Srivastava
  - 09:35-10:00 Nearest-Biclusters Collaborative Filtering
    Panagiotis Symeonidis, Alexandros Nanopoulos, Apostolos Papadopolos, and Yannis Manolopoulos

- 10:00-10:30 Coffee Break
Program

10:35 - 12:05

**Session 2: Web Content, Search, and Queries**

- 10:30-10:55 **Model-Based Classification of Web Documents Represented by Graphs**
  Alex Markov, Mark Last, and Abraham Kandel
- 10:55-11:20 **Keyword Weight Propagation for Indexing Structured Web Content**
  Jong Wook Kim and K. Selcuk Candan
- 11:20-11:35 **How to Define Searching Sessions on Web Search Engines**
  Bernard J. Jansen, Amanda Spink, Vinish Kathuria, and Sherry Koshman
- 11:35-11:50 **The Evolution of Web Content and Search Engines**
  Ricardo Baeza-Yates, Alvaro Pereira Jr and Nivio Ziviani
- 11:50-12:05 **Analysis of Web Search Engine Query Sessions**
  David Nettleton, Liliana Calderon, and Ricardo Baeza-Yates
- 12:05-1:30 **Lunch** (on your own)
Program

1:30 - 3:05

• **Session 3: Recommender Systems**
  • 1:30-1:55  *Incorporating Concept Hierarchies into Usage Mining Based Recommendations*
    Amit Bose, Kalyan Beemanapalli, Jaideep Srivastava, and Sigal Sahar
  • 1:55-2:20  *A Random-Walk Based Scoring Algorithm with Application to Recommender Systems for Large-Scale E-Commerce*
    Augusto Pucci and Marco Gori
  • 2:20-2:35  *ClustKNN: A Highly Scalable Hybrid Model- & Memory-Based CF Algorithm*
    Al Mamunur Rashid, Shyong K. Lam, George Karypis, and John Riedl
  • 2:35-2:50  *Recommendation Based on Influence Sets*
    Jian Chen and Jian Yin
  • 2:50-3:05  *Relevance of Time Spent on Web Pages*
    Peter I. Hofgesang

• 3:05-3:30  **Coffee Break**
Session 4: Web Mining Emerging Trends

- 3:30-3:55  The Impact of Attack Profile Classification on the Robustness of Collaborative Recommendation
  Chad Williams, Runa Bhaumik, Robin Burke, and Bamshad Mobasher
- 3:55-4:20  Client-side Web Mining for Community Formation in Peer-to-Peer Environments
  Kun Liu, Kanishka Bhaduri, Kamalika Das, Phuong Nguyen and Hillol Kargupta
- 4:20-4:45  Using Rank Propagation and Probabilistic Counting for Link-Based Spam Detection
  Luca Becchetti, Carlos Castillo, Debora Donato, Stefano Leonardi and Ricardo Baeza-Yates
- 4:45-5:00  Mining Sentiment Classification from Political Web Logs
  Kathleen T. Durant and Michael D. Smith
- 5:00-5:30  Joint Discussion and Closing
Guidelines for Speakers

• **Before their session, speakers should**
  – introduce themselves to the organizers
  – **Transfer their slides onto the presentation laptop if they did not e-mail it**

• **Time:**
  – 12 minutes talk + 3 minutes for questions/answers (CHECK THE PROGRAM)
  – 20 minutes talk + 5 minutes for questions/answers (CHECK THE PROGRAM)
Time for Audience Introductions

• Please fill out your information on the circulating page

• Tell us who you are?
  – Name?
  – Affiliation?
  – E-mail?
  – Area of work / research interests
Welcome and Enjoy the Sessions!!!