DEEPAYAN CHAKRABARTI

701 1st Ave, Sunnyvale, CA 94089 *(412) 855 1186; deepay@cs.cmu.edu*

Personal Information

- Work address: 701 1st Ave, Sunnyvale, CA 94089. (412) 855 − 1186
- Email: deepay@cs.cmu.edu
- Citizenship: Indian; Visa status: Permanent Resident

Professional Preparation

Employment

- Senior Research Scientist, Yahoo! Inc. (08/2005 to present)

Education

- Ph.D. in Computational and Statistical Learning, (08/2002 to 06/2005)
  School of Computer Science, CMU, Pittsburgh PA.
  Thesis Title: Tools for Large Graph Mining (*Advisor: Dr. Christos Faloutsos*)
- M.S. in Knowledge Discovery and Data Mining, (09/2000 to 06/2002)
  School of Computer Science, CMU, Pittsburgh PA. GPA: 4.0
- B.Tech. in Computer Science and Engineering, (07/1996 to 05/2000)
  Indian Institute of Technology (IIT), Kanpur, India. GPA: 9.7 (out of 10)

Research Experience

I have worked on algorithmic and statistical challenges arising in a broad range of problems primarily derived from Web Search, Computational Advertising, and Graph Mining.

Web Search: The Web is characterized by multiple sources of information: searches, anchortexts, user tags, etc. Combining these to extract the "true signal" has been the underlying theme of my work, which includes webpage segmentation, template detection, and title generation.

Computational Advertising: In the same spirit, we combined IR scores with click feedback, & merged data from coarser to finer scales, for CTR estimation as well as explore-exploit strategies.

Graph Mining: My thesis, titled **Tools for Large Graph Mining**, under Dr. Christos Faloutsos, looks at realistic graph generators, analysis of network epidemics, and community detection.

Awards and Honors

- Our R-MAT generator is the basis for the new Graph500 supercomputer benchmark.
- Our COLT 2010 paper received the student best paper award.
- One of only five *"Siebel Scholars"* in 2002 from the CMU School of Computer Science.
- "Certificate of Merit" for 1996-97 and 1997-98 in IIT-Kanpur, India.
- National Talent Search Scholarship in 1994 from the Govt. of India.

Professional Service

- Tutorial on *Statistical Challenges in Computational Advertising*, with D. Agarwal, in CIKM 2008 and KDD 2009.
- Posters chair for WWW 2010.
- Senior PC on the KDD 2010 research track.
- Served as a PC member and reviewer on several conferences, journals, and NSF Panels.
- Local arrangements co-chair for KDD 2007.
- Student member of CMU departmental Ph.D. admissions committee for 2001-2003.

Released Software (at http://www.cs.cmu.edu/∼deepay/index.html#Sw)

- The NetMine system extracts many patterns given a large graph as input, and has been used by the Northrop Grumman Corp. (Mark Hoy and Jayshree Ranka).
- The CrossAssociations system automatically "groups" nodes in a large graph.
- The F4 system performs automatic time-series prediction using chaotic time series methods.

Selected Refereed Conference Papers (full list at http://www.cs.cmu.edu/∼deepay/#Pubs)

1. P. Sarkar, D. Chakrabarti, and A. W. Moore: *Theoretical Justification of Popular Link Prediction Heuristics*, in COLT 2010 (**best student paper**) and invited to IJCAI 2011.
2. D. Chakrabarti, and K. Punera: *Event Summarization using Tweets*, in ICWSM 2011.
3. A. Vattani, M. Gurevich, and D. Chakrabarti: *Preserving Pairwise Relationships in Subgraphs*, in ICML 2011.
4. D. Chakrabarti, and R. Mehta: *The Paths More Taken: Matching DOM Trees to Search Logs for Accurate Webpage Clustering*, in WWW 2010.
5. D. Chakrabarti, R. Kumar, and K. Punera: *Quicklink Selection for Navigational Query Results*, in WWW 2009.
6. X. Wang, D. Chakrabarti, and K. Punera: *Mining Broad Latent Query Aspects from Search Sessions*, in KDD 2009.
7. D. Chakrabarti, R. Kumar, F. Radlinski, and E. Upfal: *Mortal Multi-Armed Bandits*, in NIPS 2008.
8. D. Chakrabarti, R. Kumar, and K. Punera: *Generating Succinct Titles for Web URLs*, in KDD 2008.
9. D. Chakrabarti, R. Kumar, and K. Punera: *A Graph-Theoretic Approach to Webpage Segmentation*, in WWW 2008, Beijing, China.
10. D. Chakrabarti, D. Agarwal, and V. Josifovski: *Contextual Advertising by Combining Relevance with Click Feedback*, in WWW 2008, Beijing, China.
11. D. Agarwal, A. Broder, D. Chakrabarti, D. Diklic, V. Josifovski, and M. Sayyadian: *Estimating Rates of Rare Events at Multiple Resolutions*, in KDD 2007, San Jose, CA.
12. S. Pandey, D. Chakrabarti, and D. Agarwal: *Multi-armed Bandit Problems with Dependent Arms*, in ICML 2007, Corvallis, OR.
13. D. Chakrabarti, R. Kumar, and K. Punera: *Page-level Template Detection via Isotonic Smoothing*, in WWW 2007, Banff, Canada.
14. S. Pandey, D. Agarwal, D. Chakrabarti, and V. Josifovski: *Bandits for Taxonomies: A Model-based Approach*, in SDM 2007, Minneapolis, MN.
15. J. Leskovec, D. Chakrabarti, C. Faloutsos, S. Madden, C. Guestrin and M. Faloutsos: *Information Survival Threshold in Sensor and P2P Networks*, in IEEE INFOCOM 2007, Anchorage, Alaska.
16. D. Chakrabarti, R. Kumar and A. Tomkins: *Evolutionary Clustering*, in KDD 2006.
17. D. Chakrabarti, S. Papadimitriou, D. Modha and C. Faloutsos: *Fully Automatic Cross-Associations*, in KDD 2004, Washington, USA; also a CMU Tech Report.
18. D. Chakrabarti, Y. Zhan, C. Faloutsos: *R-MAT: A Recursive Model for Graph Mining*, in SDM '04.
19. Y. Wang, D. Chakrabarti, C. Wang and C. Faloutsos: *Epidemic Spreading in Real Networks: An Eigenvalue Viewpoint*, in SRDS 2003, Florence, Italy.

Selected Refereed Journal Papers (full list at http://www.cs.cmu.edu/∼deepay/#Pubs)

1. J. Leskovec, D. Chakrabarti, J. Kleinberg, C. Faloutsos, and Z. Ghahramani: *Kronecker Graphs: An Approach to Modeling Networks*, in JMLR 2010, 11 (Feb).
2. D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos: *Epidemic Thresholds in Real Networks*, in ACM TISSEC, 10(4), 2008.
3. D. Chakrabarti and C.Faloutsos: *Graph Mining: Laws, Generators and Algorithms*, in ACM Computing Surveys, 38(1), 2006.

Selected Book Chapters

- D. Chakrabarti and C. Faloutsos: *Graph Patterns and the R-MAT Generator*, in *Mining Graph Data*, editors L. Holder and D. Cook, published by Wiley in 2006.

Patents

- Issued 6 patents in the USA: 7987417, 7974934, 7921073, 7870474, 7734629, and 6611834.
- Named as a co-inventor in 18 patents (applied and issued) and 2 defensive publications.