

# DEEPAYAN CHAKRABARTI

701 1st Ave, Sunnyvale, CA 94089

(408) 349 2950; [deepay@yahoo-inc.com](mailto:deepay@yahoo-inc.com)

## Personal Information

- Work address: 701 1st Ave, Sunnyvale, CA 94089. (408) 349 – 2950
- Home address: 1604 Hope Dr., Apt 123, Santa Clara, CA 95054. (408) 235 – 8917
- Email: [deepay@yahoo-inc.com](mailto:deepay@yahoo-inc.com)
- Citizenship: Indian; Visa status: H1-B
- Birth year: 1979

## Professional Preparation

### Employment

- 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, anchor-texts, 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

- 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 *Algorithmic Challenges in Online Advertising*, with D. Agarwal, in CIKM 2008.
- Local arrangements co-chair for KDD 2007.
- Student member of CMU departmental Ph.D. admissions committee for 2001-2003.
- Program Committee member of KDD 2008, WWW 2008 and 2009, ICDE 2008, PKDD 2008, ECDM 2008, MLG 2007, and LinkKDD 2006.
- Reviewer for ICDE 2006, JMLR 2005 and 2006, DMKD 2005, INFOCOM 2004, SIGMOD 2003, IEEE Communication Letters 2003, and VLDB 2002.

### 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.

## Publications

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

1. D. Chakrabarti, R. Kumar, F. Radlinski, and E. Upfal: *Mortal Multi-Armed Bandits*, in NIPS 2008.
2. D. Chakrabarti, R. Kumar, and K. Punera: *Generating Succinct Titles for Web URLs*, in KDD 2008.
3. D. Chakrabarti, R. Kumar, and K. Punera: *A Graph-Theoretic Approach to Webpage Segmentation*, in WWW 2008, Beijing, China.
4. D. Chakrabarti, D. Agarwal, and V. Josifovski: *Contextual Advertising by Combining Relevance with Click Feedback*, in WWW 2008, Beijing, China.
5. 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.
6. S. Pandey, D. Chakrabarti, and D. Agarwal: *Multi-armed Bandit Problems with Dependent Arms*, in ICML 2007, Corvallis, OR.
7. D. Chakrabarti, R. Kumar, and K. Punera: *Page-level Template Detection via Isotonic Smoothing*, in WWW 2007, Banff, Canada.
8. S. Pandey, D. Agarwal, D. Chakrabarti, and V. Josifovski: *Bandits for Taxonomies: A Model-based Approach*, in SDM 2007, Minneapolis, MN.
9. 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.
10. D. Chakrabarti, R. Kumar and A. Tomkins: *Evolutionary Clustering*, in KDD 2006.
11. J. Leskovec, D. Chakrabarti, J. Kleinberg and C. Faloutsos: *Realistic, Mathematically Tractable Graph Generation and Evolution, Using Kronecker Multiplication*, in PKDD 2005, Porto, Portugal.
12. D. Chakrabarti: *AutoPart: Parameter-Free Graph Partitioning and Outlier Detection*, in PKDD 2004.
13. D. Chakrabarti, S. Papadimitriou, D. Modha and C. Faloutsos: *Fully Automatic Cross-Associations*, in KDD 2004, Washington, USA; also a CMU Tech Report.
14. D. Chakrabarti, Y. Zhan, C. Faloutsos: *R-MAT: A Recursive Model for Graph Mining*, in SDM '04.
15. Y. Wang, D. Chakrabarti, C. Wang and C. Faloutsos: *Epidemic Spreading in Real Networks: An Eigenvalue Viewpoint*, in SRDS 2003, Florence, Italy.
16. Y. Liu, R. Emery, D. Chakrabarti, W. Burgard and S. Thrun: *Using EM to Learn 3D Models of Indoor Environments with Mobile Robots*, in ICML 2001, Williamstown, USA.

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

1. D. Chakrabarti, Y. Wang, C. Wang, J. Leskovec, and C. Faloutsos: *Epidemic Thresholds in Real Networks*, in ACM TISSEC, 10(4), 2008.
2. D. Chakrabarti and C. Faloutsos: *Graph Mining: Laws, Generators and Algorithms*, in ACM Computing Surveys, 38(1), 2006.
3. S. Thrun, C. Martin, Y. Liu, D. Hahnel, R. Emery-Montemerlo, D. Chakrabarti, and W. Burgard: *A Real-Time Expectation Maximization Algorithm for Acquiring Multi-Planar Maps of Indoor Environments with Mobile Robots*, in IEEE Transactions on Robotics and Automation, 20 (3), 2003.

## 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

- Granted: *Customization of information retrieval through user-supplied code*, Patent number 6,611,834, by G. Aggarwal, D. Chakrabarti, P. K. Dubey, N. P. Garg, S. Ghosal, A. K. Gupta, A. Kulshreshtha, Ashutosh and S. K. V. Murthy; assignee IBM Corp.
- Filed: Nine patent applications, all in the USA, by Yahoo! Inc.