

Bee-Chung Chen

525 Tahoe Terrace
Mountain View, CA 94041
<http://www.cs.wisc.edu/~beechung>

Phone: 408-230-5982
Email: beechun@yahoo-inc.com

EDUCATION

- University of Wisconsin – Madison** Feb. 2008
Ph.D. in Computer Sciences, Minor in Statistics (**GPA: 4.0**)
Dissertation: Cube-Space Data Mining
Advisor: Professor Raghu Ramakrishnan
- National Taiwan University**, Taipei, Taiwan Jun. 2000
M.S. in Computer Science and Information Engineering (**GPA: 4.0, Rank: 1/54**)
Thesis: Content-Based Image Retrieval of Butterflies
Advisor: Professor Jieh Hsiang
- National Taiwan University**, Taipei, Taiwan Jun. 1998
B.S. in Computer Science and Information Engineering (**GPA: 3.9, Rank: 1/45**)

HONORS AND AWARDS

- Outstanding Graduate Student Research Award**, CS, UW-Madison 2007
This award is given to one student in the department of Computer Sciences, UW-Madison each year to give recognition to excellence in research by the student
- Microsoft Research Graduate Fellowship** 2006 – 2007
- Master Thesis Award of Inalways Foundation**, Taiwan 2000
- Presidential Award of National Taiwan University**, Taiwan 1995 – 1998
Seven times, the **best record** at National Taiwan University
This award is given each semester to students ranking within the top 5% of their class
- Pan Wen-Yuan Scholarship**, Taiwan 1997

RESEARCH EXPERIENCE

- Web Content and Ad Optimization** 2008 – Present
Research Scientist, Yahoo! Research
Working with my colleagues at Yahoo!, I am developing machine-learning and statistical models [15, 19, 21, P2, P3, P5, P6] and schemes for better exploration [20, P4] for Web content and ad optimization.
- Exploratory Data Analysis and Monitoring** 2004 – 2007
Research Assistant, University of Wisconsin – Madison
Under the supervision of Professor Raghu Ramakrishnan, I defined and investigated several novel data mining problems [6, 10, 17] that combine predictive modeling and OLAP multidimensional exploration, which result in a promising research direction, cube-space data mining [11, 14]. I was also involved in the design of an aggregation workflow framework [9], research on network coordinated attack data analysis [7] and research on atmospheric aerosol monitoring.
- Goal-Oriented Privacy Preservation** 2005 – 2007
Research Assistant, University of Wisconsin – Madison
Under the supervision of Professor Raghu Ramakrishnan, I investigated the tradeoff between data utility and privacy preservation in data publishing. Specifically, I defined and studied the problem of learning classification models from (privacy-preserved) SQL aggregate views [8], and proposed a novel framework that allows publishing organizations to investigate privacy threats and enforce privacy requirements in the presence of various types and amounts of adversaries' external knowledge [13, 16].

Data Cleaning Summer 2005, Summer 2006
 Intern, DMX Group, Microsoft Research
 I worked with Dr. Surajit Chaudhuri, Dr. Venkatesh Ganti and Dr. Raghav Kaushik on example-driven design of record matching queries for data cleaning [12, P1].

Fuzzy Logic-Based Knowledge Retrieval 2002 – 2003
 Research Assistant, National Taiwan University, Taipei, Taiwan
 Under the supervision of Professor Jieh Hsiang, I developed a novel knowledge retrieval framework and its fuzzy logic inference theory [5].

Content-Based Image Retrieval in Digital Libraries 1998 – 2000
 Research Assistant, National Taiwan University, Taipei, Taiwan
 Under the supervision of Professor Jieh Hsiang, I built a content-based image retrieval (CBIR) system for butterflies [1, 3, 4]. Because of the success of my methodology, in the following years, there were **five M.S.** and **one Ph.D.** studies based on my CBIR architecture in Professor Hsiang’s research group.

EMPLOYMENT

Yahoo! Research, Santa Clara, CA 2008 – Present
 Research Scientist

University of Wisconsin – Madison 2004 – 2007
 Research Assistant

Yahoo! Research, Santa Clara, CA Summer 2007
 Research Intern

Microsoft Research, Redmond, WA Summer 2005, Summer 2006
 Research Intern

University of Wisconsin – Madison 2003 – 2004
 Teaching Assistant

National Taiwan University, Taipei, Taiwan 2002 – 2003
 Research Assistant

Combat System Bureau, Headquarters of ROC Navy, Taiwan 2000 – 2002
 Software Design Officer

EpicStream Inc, Taipei, Taiwan 2000
 System Engineer

National Taiwan University, Taipei, Taiwan 1998-2000
 Research Assistant, Teaching Assistant

PUBLICATIONS

- [21] Deepak Agarwal and Bee-Chung Chen. **Regression based Latent Factor Models**. ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2009. (peer-reviewed)
- [20] Deepak Agarwal, Bee-Chung Chen and Pradheep Elango. **Multi-armed Bandit Schemes for Dynamic Web Applications**. 2009. (under review)
- [19] Deepak Agarwal, Bee-Chung Chen and Pradheep Elango. **Spatio-Temporal Models for Estimating Click-through Rate**. In Proceedings of the 18th International World Wide Web Conference (WWW), 2009. (peer-reviewed, 12% acceptance)
- [18] Bee-Chung Chen, Daniel Kifer, Kristen LeFevre and Ashwin Machanavajjhala. **A Survey on Privacy-Preserving Data Publishing**. To appear in Foundations and Trends in Databases, invited in 2008.
- [17] Bee-Chung Chen, Raghu Ramakrishnan, Jude W. Shavlik and Pradeep Tamma. **Bellwether Analysis: Searching for Cost-Effective Query-defined Predictors in Large Databases**. To appear in ACM Transactions on Knowledge Discovery from Data (TKDD), Vol. 3, Issue 1, 2009.
- [16] Bee-Chung Chen, Kristen LeFevre and Raghu Ramakrishnan. **Adversarial-Knowledge Dimensions in Data Privacy**. International Journal on Very Large Data Bases (VLDB Journal), Vol. 18, Issue 2, 2009.
- [15] Deepak Agarwal, Bee-Chung Chen, Pradheep Elango, Raghu Ramakrishnan, Nitin Motgi, Scott Roy and Joe Zachariah. **Online Models for Content Optimization**. In Proceedings of the 22nd Annual Conference on Neural Information Processing Systems (NIPS), 2008. (peer-reviewed, 24.5% acceptance)
- [14] Bee-Chung Chen. **Cube-Space Data Mining**. Ph.D. Dissertation, Computer Sciences, University of Wisconsin – Madison, 2008.
- [13] Bee-Chung Chen, Kristen LeFevre and Raghu Ramakrishnan. **Privacy Skyline: Privacy with Multidimensional Adversarial Knowledge**. In Proceedings of the International Conference on Very Large Data Bases (VLDB), 2007. (peer-reviewed, 16.9% acceptance)
- [12] Surajit Chaudhuri, Bee-Chung Chen, Venkatesh Ganti and Raghav Kaushik. **Example-Driven Design of Efficient Record Matching Queries**. In Proceedings of the International Conference on Very Large Data Bases (VLDB), 2007. (peer-reviewed, 16.9% acceptance)
- [11] Raghu Ramakrishnan and Bee-Chung Chen. **Exploratory Mining in Cube Space**. Data Mining and Knowledge Discovery, 10th Anniversary Issue, 2007. (invited article)
- [10] Bee-Chung Chen, Raghu Ramakrishnan, Jude W. Shavlik and Pradeep Tamma. **Bellwether Analysis: Predicting Global Aggregates from Local Regions**. In Proceedings of the International Conference on Very Large Data Bases (VLDB), 2006. (peer-reviewed, 13.2% acceptance)
- [9] Lei Chen, Raghu Ramakrishnan, Paul Barford, Bee-Chung Chen and Vinod Yegneswaran. **Composite Subset Measures**. In Proceedings of the International Conference on Very Large Data Bases (VLDB), 2006. (peer-reviewed, 13.2% acceptance)
- [8] Bee-Chung Chen, Lei Chen, Raghu Ramakrishnan and David R. Musicant. **Learning from Aggregate Views**. In Proceedings of the IEEE International Conference of Data Engineering (ICDE), 2006. (peer-reviewed, 19% acceptance)
- [7] Bee-Chung Chen, Vinod Yegneswaran, Paul Barford, Raghu Ramakrishnan. **Toward a Query Language for Network Attack Data**. In Proceedings of the IEEE International Workshop on Networking Meets Databases (NetDB), 2006. (peer-reviewed)
- [6] Bee-Chung Chen, Lei Chen, Yi Lin and Raghu Ramakrishnan. **Prediction Cubes**. In Proceedings of the International Conference on Very Large Data Bases (VLDB), 2005. (peer-reviewed, 16.4% acceptance)

- [5] Bee-Chung Chen and Jieh Hsiang. **A Logical Framework for Knowledge Retrieval with Fuzziness**. In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI), 2004. (peer-reviewed)
- [4] Jieh Hsiang, Wen-Jun Liu, Bee-Chung Chen, Hsieh-Chang Tu. **Multidimensional Interactive Retrieval on Fine-grained Images**. In Proceedings of the IEEE International Conference on Multimedia & Expo (ICME), 2003. (peer-reviewed)
- [3] Bee-Chung Chen and Jieh Hsiang. **Perception-Based Image Retrieval**, Technical Report NTUCSIE 02-04, National Taiwan University, 2002.
- [2] Bee-Chung Chen. **Intranet Data Access Control**, Journal of Communications, Electronics and Information, Ministry of National Defense, R.O.C. Taiwan, vol. 3, 2002. (Published in Chinese)
- [1] Bee-Chung Chen. **Content-Based Image Retrieval of Butterflies**, Master Thesis, Dept. of Computer Science and Information Engineering, National Taiwan University, 2000. (Published in Chinese)

PATENTS

- P6. Bee-Chung Chen, Pradheep Elango and Deepak Agarwal. **Dynamic Estimation of the Popularity of Web Content**. (Preparing for submission)
- P5. Wei Chu, Seung-Taek Park, Scott Roy, Pradheep Elango, Deepak Agarwal, Bee-Chung Chen, Raghu Ramakrishnan, Todd Beaupre. **Conjoint Analysis with Bilinear Regression Models for Segmented Predictive Content Ranking**. US Patent Filed at Yahoo!.
- P4. Bee-Chung Chen, Deepak Agarwal, Pradheep Elango, Wei Chu, Raghu Ramakrishnan, Scott Roy, Nitin Motgi. **Enhanced Matching through Explore/Exploit Schemes**. US Patent Filed at Yahoo!.
- P3. Pradheep Elango, Raghu Ramakrishnan, Seung-Taek Park, Bee-Chung Chen, Deepak Agarwal. **Framework to Evaluate Content Display Policies**. US Patent Filed at Yahoo!.
- P2. Bee-Chung Chen, Deepak Agarwal, Raghu Ramakrishnan, Vijay Narayanan, Pradheep Elango, Amit Seth, Vik Singh, Nitin Motgi, Joe Zachariah, Scott Roy. **Customized Today Module**. US Patent Filed at Yahoo!.
- P1. Chaudhuri, Bee-Chung Chen, Venkatesh Ganti and Raghav Kaushik. **Example-Driven Design of Efficient Record Matching Queries**. US Patent Filed at Microsoft.

PROFESSIONAL ACTIVITIES

Journal Reviewer

IEEE Transactions on Knowledge and Data Engineering, ACM Transactions on Knowledge Discovery from Data, ACM Transactions on Database Systems

Program Committee

2009 SIAM International Conference on Data Mining (SDM'09)

External Referee

ICDE 2009, VLDB 2008, SIGMOD 2006, EDBT 2006, ICDE 2005, VLDB 2004

CONFERENCE PRESENTATIONS

Privacy Skyline: Privacy with Multidimensional Adversarial Knowledge. Presented in the International Conference on Very Large Data Bases (VLDB), Vienna, Austria, 2007.

Learning from Aggregate Views. Presented in the IEEE International Conference of Data Engineering (ICDE), Atlanta, Georgia, USA, 2006.

Toward a Query Language for Network Attack Data. Presented in the IEEE International Workshop on Networking Meets Databases (NetDB), Atlanta, Georgia, USA, 2006.

Prediction Cubes. Presented in the International Conference on Very Large Data Bases (VLDB), Trondheim, Norway, 2005.