

Research Scientist/Engineer
Facebook
Palo Alto, CA

U.S. Permanent Resident
Email: dr.dingzhou@gmail.com
Homepage: <http://dzhou.net/>

Background Summary

Ding is a researcher and engineer experienced in large-scale machine learning and text mining. He currently focuses on optimizing ad ranking and targeting for Facebook, where he led the development of behavioral and keyword advertisement targeting systems, and played key roles in the company's many early machine learning products that were applied for advertising, data analytics and recommendation. Prior to that, he took positions at Google Labs, Yahoo! Search, and NEC Research Labs. He was a co-developer of the world famous open source search engine: CiteSeer^x. He invented five U.S. patents, authored research papers that have been cited hundreds of times, and served on program committees of influential academic conferences.

Industry Experience

Facebook

Technical Lead/Senior Research Scientist

Jan 2008 -

Machine learning products for Internet advertising and social recommendation:

- *development of a real-time trained CTR estimation system (engineering lead)*
- *led development of a behavioral and contextual ad targeting system (engineering lead, 1 patent)*
- *led development of keyword-based ad targeting (1/3 ad revenue) (engineering lead)*
- *key developer of first company-wide data analytics systems based on Hadoop (system architect)*
- *key developer a/b testing frameworks based on Hadoop (over 2T/day log)*
- *optimizing ad targeting and ranking features, delivering double-digit CTR improvement (2 U.S. patents)*
- *developer of recommendation and reputation systems for social applications (2 U.S. patents)*
- *mentored 5 Ph.D. interns; published 3 research papers; interviewed 100+ engineer candidates;*
- *experienced the company's fast growth (users: 50M to 850M; employees: 300 to 4000);*

NEC Research Labs

Research Intern

Summer 2007

Prototyped a document recommendation system based on a new graph merging technique. Offline experiments showed 35% recommendation accuracy improvement on CiteSeer data.

Google Labs

Engineering Intern

Summer 2006

Developed in C++ an automatic Optical Character Recognition (OCR) correction system for Google Books based on probabilistic language modeling.

Yahoo! Search

Engineering Intern

Summer 2005

Prototyped in Java and Python an automatic realtime Web search result clustering system.

Porsche Asia-Pacific

Software Engineer

Feb 2004 - Aug 2004

Developed in C# a Client/Server based enterprise resource management (ERP) system.

U.S. Patents

Cognitive Relevance Targeting in a Social Network System, U.S. Patent pending, 2011.

Inferring User Profile Attributes from Social Information, U.S. Patent pending, 2010.

Leveraging Information in a Social Network for Inferential Targeting of Advertisements, U.S. Patent pending, 2009.

Determining User Affinity Towards Applications on a Social Networking Website, U.S. Patent pending, 2008.

Resource Management of Social Network Applications, U.S. Patent pending, 2008.

Selected Publications

Learning Relevance in a Heterogeneous Social Network and Its Application in Online Targeting, Chi Wang, Rajat Raina, David Fong, Ding Zhou, Jiawei Han, Greg Badros, In *SIGIR'11; The 34th Annual ACM SIGIR Conference on Information Retrieval*, Beijing, China, 2011 (SIGIR 2011).

Keyword Extraction for Social Snippets, Zhenhui Li, Ding Zhou, Yun-Fang Juan, Jiawei Han, In *WWW'10: The 19th International World Wide Web Conference (Poster)*, North Carolina, USA, 2010 (WWW 2010).

Learning to Recognize Reliable Users and Content in Social Media with Coupled Mutual Reinforcement, Jiang Bian, Yandong Liu, Ding Zhou, Eugene Agichtein and Hongyuan Zha, In *WWW'09: The 18th International World Wide Web Conference*, Spain, 2009 (WWW 2009).

Learning Multiple Graphs for Document Recommendations, Ding Zhou, Shenghuo Zhu, Kai Yu, Xiaodan Song, Belle Tseng, Hongyuan Zha, C. Lee Giles In proceedings of *WWW '08: The 17th international conference on World Wide Web*, Beijing, China, 2008 (WWW 2008).

Exploring Social Annotations for Information Retrieval, Ding Zhou, Jiang Bian, Shuyi Zheng, Hongyuan Zha, C. Lee Giles, In proceedings of *WWW '08: The 17th international conference on World Wide Web*, Beijing, China, 2008 (WWW 2008).

Learning User Clicks in Web Search, Ding Zhou, Levent Bolelli, Jia Li, Lee Giles, Hongyuan Zha, In proceedings of *the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, 2007 (IJCAI 2007)*.

A Clustering Method for Web Data with Multi-Type Interrelated Components, Levent Bolelli, Seyda Ertekin, Ding Zhou, C. Lee Giles, In proceedings of *WWW '07: The 16th international conference on World Wide Web (Poster)*, Banff, Canada, (WWW 2007).

Discovering Temporal Communities from Social Network Documents, Ding Zhou, Isaac Council, Hongyuan Zha, C. Lee Giles, In proceedings of *the 7th IEEE International Conference on Data Mining*, Omaha, Nebraska, USA, 2007, (ICDM 2007).

Co-Ranking Authors and Documents in a Heterogeneous Network, Ding Zhou, Sergey Orshanskiy, Hongyuan Zha, C. Lee Giles, In proceedings of *the 7th IEEE International Conference on Data Mining*, Omaha, Nebraska, USA, 2007, (ICDM 2007).

Probabilistic Models for Discovering E-Communities, Ding Zhou, Eren Manavoglu, Jia Li, C. Lee Giles, Hongyuan Zha, In proceedings of *the 15th ACM International World Wide Web Conference*, Edinburgh, Scotland, 2006, (WWW 2006).

Rotational-Invariant L1-norm PCA for Robust Learning, Chris Ding, Ding Zhou, Xiaofeng He, Hongyuan Zha, In proceedings of *the 23rd International Conference on Machine Learning*, Pittsburgh, PA, 2006, (ICML 2006).

Topic Evolution and Social Interactions, How Authors Effect Research, Ding Zhou, Xiang Ji, Hongyuan Zha, Lee Giles, In proceedings of *the 15th ACM Conference on Information and Knowledge Management*, Arlington, VA, 2006 (CIKM 2006).

A New Mallows distance based Metric for Comparing Clusterings, Ding Zhou, Jia Li, Hongyuan Zha, In proceedings of the 22rd *International Conference on Machine Learning*, Bonn, Germany, 2005, (ICML 2005).

Recent Services

Co-Chair for the International Workshop on Large Scale Network Analysis (with WWW 2011)

Internet Technical Advisor for the Tourism Bureau of Changsha City, China (2010-2011).

Program Committee on ICDE, WWW, CIKM, SIGKDD, ICWSM, TOIS, SIGIR, etc (2007-2010).

Invited Talks at Hadoop Summit, JSM, UCSC, NEC Research, etc. (2008-2009)

Skills

Python/C++/Java/PHP : with 6/8/4/2 years of experience;

Expert experience with building large-scale data mining and analytics systems;

Experiences with Hadoop/Hive/HBase, Thrift, JSP, HTML/XML, R/Matlab, MySql, Linux Systems, Latex;

Technical leadership; Strong learning and communication; Self-motivated and diligent;

Education

Ph.D. Computer Science and Engineering, Pennsylvania State University, USA, 2007.

- Dissertation: mining social networks and documents
- Co-developed the open source search engine CiteSeer^x (<http://citeseerx.ist.psu.edu>, +1m users/day)
- Published 20+ papers that received over a hundred citations

B.S. Computer Science, Fudan University, China, 2004. *Summa Cum Laude*

- GPA ranked Top 2% and published two research papers
- Received Fred A. & Susan Breidenbach Fellowship of PSU
- City Mayor Fellowship, Government of Changsha City, China, 2000