

Chengliang Zhang

Department of Computer Science
University of Rochester
Rochester, NY 14627
<http://www.cs.rochester.edu/u/zhangchl>

Email: zhangchl@cs.rochester.edu
Phone: (585)275-1192 (office)
(585)503-3358 (cell)
Fax: (585)273-4556

Objective: Full-time researcher or software engineer

Summary

- Excellent research background in areas:
 - Large-scale data analysis
 - Performance optimization
 - Data mining and machine learning
- Rich industrial internship experience at the IBM Toronto Lab, the IBM T. J. Watson Research Center, the NEC American Lab, and Xerox Corporation
- Languages: C, C++, Java, Perl, SQL, and MATLAB
- OS: Linux and MSWindows
- Others: Compiler construction, binary instrumentation, data mining tools
- Active in social activities

Education

- Ph.D. 2002-2007, Computer Science Department, University of Rochester.
Thesis Title: *Behavior Models for Computer Programs*
Advisors: Mitsunori Ogihara and Chen Ding
GPA: 3.9/4.0
- M.S. 1999-2002, Department of Computer Science, Beijing University, China.
Thesis Title: *Research and Implementation on Storing and Retrieving Multidimensional Data*
Advisor: Dongqing Yang
GPA: 3.8/4.0
- B.S. 1995-1999, Department of Computer Science, Shandong University, China.
GPA: 3.9/4.0

Honors

- 2006 IBM CAS fellowship
- 2002 - Present Conference travel awards for PLDI&ISMM'06, POPL'06, PLDI'05, ASPLOS'04
- 1995 - 2002 President Scholarship, Excellent Student Scholarship
Jinan Government Scholarship, Konica Scholarship, Panasonic Scholarship
Excellent Student Leader, Excellent Bachelor Degree Dissertation
Special honor of entering university without admission test
- 1994 First prize in Chinese National Olympic Mathematics Competition

Research Experience

- 09/2002 - Present: Research Assistant University of Rochester
My research mainly focuses on modeling and adapting large-scale program behavior. I have worked on both theoretical analysis and practical applications. On the theoretical analysis side, I have proved the hardness of finding reference affinity groups and have proposed an approximate algorithm [POPL'06,TDM'04]. I have redefined the spatial locality to identify components that have poor locality [TR'06d]. On the practical application side, I have used the behavior models to direct data placement [TR'04], code placement [POPL'06], adaptive memory management [ISMM'06,TR'06b], memory monitoring, memory leak detection [MSP'05], transactional memory scheduling [TR'06a], behavior-oriented program parallelization [PLDI'07] and support for unsafe code optimizations [TR'06c].
In addition, I have done research on text mining [ICMLA'06], multilabel classification [ICTAI'06], and multi-class classification on gene expression datasets [BIOINFORMATICS'04].
- 09/1999-07/2002: Research Assistant Beijing University
I participated in the state key research project “Content-Oriented Massive Information Integration, Analyzing, Processing and Services”. I was part of the team responsible for the theory and techniques of massive information analyzing and processing [M.S.Thesis'00]. I also participated in developing Cyber Magic, a data warehousing software tool suite, and the Oversea Travel System, an information management system. I tested and compared the performance of IBM db2, Oracle, Terradata, Sybase, CA, and Cognos for the Agricultural Bank of China with the other team members.

Work Experience

- 09/2006 - 12/2006: Visiting Student IBM - Toronto Software Lab
The new IBM PowerPC chip has provided us an opportunity of flushing a cache line into a lower level of storage. This research investigates how to generate flush instructions based on the reuse signature of arrays and small objects. This work has been implemented in the IBM Toronto Portable Optimizer. This is joint work with Dr. Yaoqing Gao and Dr. Roch Archambault.
- 05/2006 - 08/2006: Summer Intern IBM - T. J. Watson
Poor data locality causes poor performance, but unfortunately finding an optimal data placement is NP-hard. Choosing the best garbage collection method a priori for all programs running on different hardware configurations is thus a mission impossible. We formulate the locality optimization problem as a well-known reinforcement learning problem—the multi-armed bandit machine. We find the best layout on-line by trial and errors and then exploit it to improve the overall performance. This work has been implemented in the IBM J9 Java Virtual Machine with Dr. Martin Hirzel and Dr. Michael J. Hind [PATENT'07].
- 05/2005 - 08/2005: Summer Intern NEC Labs America
More and more powerful computer technology inspires us to investigate information hidden under huge amounts of documents. We use a time sensitive Dirichlet process

mixture model with different decay function to find the events among the document streams. We use a theme river to provide a visualization of the events along the time axis. Experiments on the TDT dataset have demonstrated the effectiveness. This is a joint work with Dr. Shenghuo Zhu and Dr. Yihong Gong [ICMLA'06].

- 06/2003 - 08/2003: Summer Intern Xerox Webster Lab
I worked on the Phaser usage profile database with other team members. We used statistical and Oracle data mining tools to analyze the data and find useful information to reduce the cost of delivering service to customers.

Publications

Journal Articles

- [BIOINFORMATICS'04] "A Comparative Study of Feature Selection and Multiclass Classification Methods for Tissue Classification Based on Gene Expression." Tao Li, **Chengliang Zhang** and Mitsunori Ogihara. *Bioinformatics* 20(15), October 12 2004.
- [SC'00] "Web-based Pattern Data Browsing and Query System in Cooperative Environment." Gang Yang, Xiangxu Meng, Yanning Xu, Chenglei Yang and **Chengliang Zhang**. *Shandong Science* 13(14A), 2000.

Conference Papers

- [PLDI'07] "Behavior-oriented Parallelization." Chen Ding, Xipeng Shen, Kirk Kelsey, Chris Tice, Ruke Huang, **Chengliang Zhang**. *the Conference on Programming Language Design and Implementation, San Diego, CA, June 2007 (Acceptance Rate: 25%)*.
- [ICMLA'06] "Trend Analysis for Large Document Streams." **Chengliang Zhang**, Shenghuo Zhu, Yihong Gong. *Fifth International Conference on Machine Learning and Applications*, Orlando, Florida, December 2006 (Acceptance Rate: 22.8%).
- [ICTAI'06] "Empirical Studies on Multilabel Classification." Tao Li, **Chengliang Zhang** and Shenghuo Zhu. *Proceedings of the 18th IEEE International Conference on Tools with Artificial Intelligence*, Washington D.C, November 2006 (Acceptance Rate: 35%).
- [ISMM'06] "Program-level Adaptive Memory Management." **Chengliang Zhang**, Kirk Kelsey, Xipeng Shen, Chen Ding, Matthew Hertz, and Mitsunori Ogihara. *International Symposium on Memory Management*, Ottawa, Canada, June 2006 (Acceptance Rate: 38%).
- [POPL'06] "A Hierarchical Model of Data Locality." **Chengliang Zhang**, Chen Ding, Mitsunori Ogihara, Yutao Zhong and Youfeng Wu. *Symposium on Principles of Programming Languages*, Charleston, South Carolina, January 2006 (Acceptance Rate: 20%).

Workshop Papers

- [ExpCS'07] "Analysis of Input-Dependent Program Behavior Using Active Profiling." Xipeng Shen, **Chengliang Zhang**, Chen Ding, Michael Scott, Sandhya Dwarkadas, Mitsunori Ogihara, *The Workshop on Experimental Computer Science*, to appear.

- [MSP'05] "Gated Memory Control for Memory Monitoring, Leak Detection and Garbage Collection." Chen Ding, **Chengliang Zhang**, Xipeng Shen, and Mitsunori Ogihara. *The Third ACM SIGPLAN Workshop on Memory System Performance*, Chicago, Illinois, June 2005.
- [TDM'04] "Finding Reference Affinity Groups in Traces Using Sampling Method." **Chengliang Zhang**, Yutao Zhong, Chen Ding, Mitsunori Ogihara. *The third International Workshop on Mining temporal and Sequential Data(TDM-04) in conjunction with SIGKDD-04*, Seattle, Washington, August 2004.

Patent

- [PATENT'07] "Method and System for Determining Optimal Data Layout Using Blind Justice." Martin Hirzel, **Chengliang Zhang**. US Patent filed by IBM.

Other Publications

- [TR'06a] "A Key-based Adaptive Transactional Memory Executor." Tongxin Bai, Xipeng Shen, **Chengliang Zhang**, William N. Scherer III, Chen Ding, Michael L. Scott. *TR909, Computer Science Department, University Rochester*, December 2006.
- [TR'06b] "Waste Not, Want Not: Adaptive Garbage Collection in a Shared Environment." **Chengliang Zhang**, Kirk Kelsey, Xipeng Shen, Chen Ding, Matthew Hertz, and Mitsunori Ogihara. *TR908, Computer Science Department, University Rochester*, December 2006.
- [TR'06c] "Fast Track: Supporting Unsafe Optimizations with Software Speculation ." Kirk Kelsey, **Chengliang Zhang**, and Chen Ding. *TR907, Computer Science Department, University Rochester*, December 2006.
- [TR'06d] "A Component-based Definition of Spatial Locality." **Chengliang Zhang** and Chen Ding. *TR906, Computer Science Department, University Rochester*, December 2006.
- [TR'04] "A Method for Hierarchical Data Placement." **Chengliang Zhang**, Yutao Zhong, Mitsunori Ogihara and Chen Ding. *TR845, Computer Science Department, University Rochester*, August 2004.
- [M.S.Thesis'02] "Research and Implementation on Storing and Retrieving Multidimensional Data." **Chengliang Zhang**. *Master's thesis, Department of Computer Science, Beijing University*, 2002.

Professional Activities

- ACM, SIGPLAN student member
- Reviewer for conferences: ISPASS'07, PLDI'05, ICPP'04, and NPC'04
- Comprehensive exams committee member, Computer Science Department, 2005-2007
My responsibilities include reviewing the comprehensive exam questions from multiple areas and providing feedback on difficulty and relevance
- Graduate admission committee member, Computer Science Department, 2004-2005
My responsibilities include reviewing applications, participating in face-to-face interviews during applicant weekend, and participating in phone interviews with foreign candidates

- Lab committee member, Computer Science Department, 2002-2004

References

Mitsunori Ogihara
Department of Computer Science
University of Rochester
Rochester, NY 14627
(585)275-7263
ogihara@cs.rochester.edu

Chen Ding
Department of Computer Science
University of Rochester
Rochester, NY 14627
(585)275-1373
cding@cs.rochester.edu

Robert A. Jacobs
Meliora 416
Brain and Cognitive Sciences
University of Rochester
Rochester, NY 14627-0268
(585)275-0753
robbie@bcs.rochester.edu

Michael Hind
IBM T.J. Watson Research Center
19 Skyline Drive, 1N-B18
Hawthorne, NY 10532
(914)784-7589
hind@watson.ibm.com