

Yuanfang Hu

Email: yuanfang@gmail.com , Phone: 15011299068

Objective

Seek for a senior software engineer position in an internet, e-commerce, social mobile, or online video company

Skills

- Extensive industrial experience on data mining, machine learning, recommender system, collaborative filtering, web search, large scale distributed computing platform Hadoop map/reduce, ads, etc
- Research experience on optimization problems, VLSI, computer architecture, and compiler optimizations, etc
- Proficient in C/C++, Python, PIG. Knowledge of Hive, Perl, Java, Shell, etc
- Strong problem solving skills, good in data structures and algorithms
- Quick learner, good team player

Education

Ph.D. Computer Science, University of California, San Diego, 2007
M.S., Computer Science, Tsinghua University, Beijing, China, 2000
B.S., Computer Science, Tsinghua University, Beijing, China, 1998

Work Experience

03/2011 – present, Senior Applied Researcher / Senior Software Engineer, eBay Inc., San Jose, USA

- Extensive experience on processing terabyte data on Hadoop Map/Reduce platform using PIG/Hive/Python/Java.
- Worked on a similar-item clustering flow, which uses K-means clustering algorithm to generate similar-item clusters, and recommend these similar items to users when they lose their biddings or seek for alternatives.
- Developed a popularity flow, which recommends the most popular/trending items on eBay site.
- Developed a merchandising metrics dashboard tool, which tracks the eBay site events, such as impression, click, bid, watch, purchase etc., and analyzes how recommendation algorithms help to improve site CTR and revenue.
- Developed a personalized gift recommendation flow to recommend eBay inventories to users based on their gender, age, and personal profile.
- These recommendation algorithms (including similar items clustering, related item clustering, and popular items/watches) lifts 45+ million dollars revenue to eBay site in 2011.

04/2007 – 03/2011, Senior Software Engineer, Mentor Graphics Inc., San Jose, USA

- Developed a flow to fit machine-learned linear regression models to predict and correct optical distortion in mask process during chip-making. The linear regression is based upon a cost function representing the norm of the difference between the simulated mask contours and the measured mask contours.
- Implemented optimizations to improve performance of large scale optical mask data preparation, such as a fast indexing mechanism to speed up the access of large scale data on disk, various hashing, caching mechanisms to speed up performance.
- Filed a US patent on the machine learning model work.

11/2003 – 01/2007, Software Programmer, San Diego Supercomputer Center, San Diego, USA

- Developed a computational high-resolution earthquake simulation on terabyte scale of data for earthquake ground motion prediction using MPI.
- Achieved high performance parallelization to scale onto 100,000+ cores through combined optimizations on data localization, cache optimization, I/O mapping/optimization, etc.
- The work is later scaled onto petabyte scale of data and received multiple NSF(national science foundation) grants, and is one of the largest earthquake simulations in US and worldwide.
- Published 4 papers on this work.

06/2003 – 09/2003, Summer Intern, Programming Systems Lab, Intel Corp

- Analyzed and tested the auto parallelization optimizations in Intel's Open Research Compiler, including software pipelining, inter-procedural analysis, etc

Research Experience

09/2000 – 02/2007, Research Assistant, University of California, San Diego

- Performance Driven Low Power Network-on-Chip Design (Ph.D. Thesis)
 - o Proposed and implemented a machine learned multi-commodity flow model to improve communication power and latency of Network-on-Chip given on-chip routing area constraints. The model is based on an approximation algorithm and is learned through guidance of a prime/dual optimization scheme.
 - o Achieved 36-52% power/latency improvement by optimizing on-chip interconnections.
 - o Published 5 papers at top-tier journal and international conferences.
- Dynamic Optimizer Design for VLIW Architecture
 - o Implemented an hardware optimizer simulation that dynamically optimizes loop instructions through software pipelining for VLIW architecture
 - o Achieved an average of 12-16% speedup on the SPEC2000 benchmarks.
 - o Published 2 papers at UCSD technical report and a top-tier international conference.

09/1998 – 06/2000, Research Assistant, Tsinghua University, China

- Developed an object oriented integrated configuration software for distributed control systems.
- Published 1 paper at an international conference.

Publications

- Y. Zhu, **Y. Hu**, M. Taylor, C.K. Cheng, "Energy and Switch Area Optimizations for FPGA Global Routing Architectures", *ACM Transactions on Design Automation of Electronic Systems*, 2009
- **Y. Hu**, Y. Zhu, M. Taylor, C.K. Cheng, "FPGA Global Routing Architecture Optimization Using A Multi-Commodity Flow Approach", *International Conference on Computer Design*, 2007
- Y. Cui, R. Moore, K. Olsen, A. Chourasia, **Y. Hu**, J. Zhu, "Enabling Very Large Scale Earthquake Simulations on Parallel Machines", *International Conference on Computational Science*, 2007
- M. Faerman, R. Moore, Y. Cui, **Y. Hu**, J. Zhu, "Managing Large Scale Data for Earthquake Simulations", *Journal of Grid Computing*, 2007
- Y. Cui, R. Moore, K. Olsen, A. Chourasia, **Y. Hu**, J. Zhu, B. Minster, P. Maechling, "Enabling a Large-Scale Application on the TeraGrid", *TeraGrid Conference*, Indianapolis, June, 2006
- S. Zhou, Y. Zhu, **Y. Hu**, R. Graham, M. Hutton, C.K. Cheng, "Timing Model Reduction for Hierarchical Timing Analysis", *International Conference on Computer Aided Design*, 2006
- **Y. Hu**, Y. Zhu, H. Chen, C.K. Cheng, "Communication Latency Aware Low Power NoC Synthesis", *Design Automation Conference*, 2006
- **Y. Hu**, H. Chen, Y. Zhu, A. A. Chien, C.K. Cheng, "Physical Synthesis of Energy-Efficient NoCs Through Topology Exploration and Wire Style Optimization", *International Conference on Computer Design*, 2005
- J. Weinberg, A. Jagatheesan, A. Ding, M. Faerman, **Y. Hu**, "Gridflow Description, Query, and Execution at SCEC using the SDSC Matrix", *International Symposium on High-Performance Distributed Computing*, 2004
- S. Narayanasamy, **Y. Hu**, S. Sair, B. Calder, "An Instruction Scheduling Co-Processor for Adaptive VLIW Schedules", *International Conference on High Performance Computer Architecture*, 2004
- S. Sair, **Y. Hu**, T. Sherwood, B. Calder, "Optimized Trace Binaries for Architectural Evaluation", *UCSD Technical Report CS2002-0711*, 2002
- **Y. Hu**, S. Zhang, W. Jiang, "Applying Object-Oriented Method to CSIE System", *International Conference on Technology on Object-Oriented Languages and Systems*, 1999

US Patents

- E. Sahouria, **Y. Hu**. "Joint Calibration For Mask Process Models", *U.S. Patent Application No.61/116,203*, 2010

References

Available upon request