

Wei Xu

Assistant Professor, IIS, Tsinghua University

WORKING EXPERIENCE

Tsinghua University

Assistant Professor at Institute for Interdisciplinary Information Sciences IIIS

Research Area: Distributed Systems + Machine Learning

Administrative work:

- Assistant Dean of IIIS
- Director, International Partnership, Tsinghua Online Education Office
- Supervisory Committee Member, Tsinghua Data Science Institute
- Director, Open Compute Lab at Tsinghua

Beijing, China

Feb, 2013 – Now

Jun, 2014 - Now

Jun, 2013 – Now

Sept, 2015 – Now

Jan, 2015 - Now

Google

Software Engineer

-Design and implementation of Google's next generation debug logging and cluster monitoring infrastructure.

Mountain View, CA

Aug, 2010 to Feb, 2013

Netease Beijing R&D Center

Software Research Engineer

- Designer and lead engineer for a large scale, distributed storage system
- Designer and lead engineer for image search (<http://image.yodao.com>)

Beijing, China

Jun – Dec, 2006

HP Labs

Research Intern

- Research on predictive control in virtual machine resource allocations

Palo Alto, CA

Jun - Aug, 2005

Institute of Software, Chinese Academy of Science (ISCAS)

Research Assistant

Beijing, China

Jul-Aug, 2002

EDUCATION

University of California at Berkeley

Ph.D. in Computer Science

Advisors: Prof. David Patterson and Prof. Armando Fox

Dissertation: Detecting Large Scale System Problems by Mining Console Logs

Certificate in Management of Technology (MOT)

Berkeley, CA

Dec. 2005 – July, 2010

University of California at Berkeley

M.S. in Computer Science, GPA: 3.95 / 4.00

Research Advisor: Prof. David Patterson

Berkeley, CA

Aug, 2003- Dec. 2005

University of Pennsylvania

B.S. Eng. in Computer Science, Summa Cum Laude, GPA 4.00 / 4.00

with Electrical Engineering Minor

Philadelphia, PA

Sept. 2001- May. 2003

Tsinghua University

Major in Computer Science

Beijing, China

Sept. 1999 - Aug. 2001

PUBLICATIONS

Yang Zhang, Fangzhou Xu, Erwin Frise, Siqu Wu, Bin Yu and Wei Xu. **DataLab: A Version Data Management and Analytics System**, in proceedings of *ICSE first International Workshop on BIG Data Software Engineering*, Austin, USA, May, 2016

Hu Li and Wei Xu. **Improving Spark Performance with Zero-copy Buffer Management and RDMA**, to appear in proceedings of *IEEE INFOCOM First International Workshop on Big Data Sciences, Technologies and Applications (BDSTA)*

2016), San Francisco, USA, Apr, 2016

Yuchao Pan, Yuxi Dong, Jingtian Zhou, Mark Hallen, Bruce R. Donald, Jianyang Zeng and Wei Xu. **cOSPREY: A Cloud-Based Distributed Algorithm for Large-Scale Computational Protein Design**, to appear in *Journal of computational biology (JCB)*, 2016

Guosai Wang, Shuhao Wang, Bing Luo, Xin Jin, Yinghang Zhu, Wenjun Yang, Longbo Huang, Weisong Shi, Dianming Hu and Wei Xu. **Increasing Large-Scale Data Center Capacity by Statistical Power Control**. To appear in proceedings of *the European Conference on Computer Systems (EuroSys 2016)*, London, UK, Apr, 2016

Xiaofeng Lin, Yu Chen, Xiaodong Li, Junjie Mao, Wei Xu, Jiaquan He, and Yuanchun Shi, **Scalable Kernel TCP Design and Implementation for Short-Lived Connections**, to appear in proceedings of *the 21th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS 2016)*, Atlanta, USA, Apr, 2016

Yi Li, Hong Liu, Wenjun Yang, Dianming Hu and Wei Xu. Inter-Data-Center Network Traffic Prediction with Elephant Flows, to appear in proceedings of *IEEE/IFIP Network Operations and Management Symposium (NOMS 2016)*, Istanbul, Turkey, Apr, 2016

Quanquan Zhi, Wei Xu. **MED: The Monitor-Emulator-Debugger for Software-Defined Networks**, to appear in proceedings of *IEEE International Conference on Computer Communications (INFOCOM 2016)*, San Francisco, USA, Apr, 2016

Yichao Zhou, Wei Xu, Bruce R. Donald, Jianyang Zeng. **An Efficient Parallel Algorithm for Accelerating Computational Protein Design** in proceedings of *ISMB 2014*. Boston, Massachusetts, USA, July 2014. *Bioinformatics*.

Adam Oliner, Archana Ganapathi, Wei Xu. **Advances and Challenges in Log Analysis**. *Communications of ACM (CACM)*, Invited article, Feb, 2012

Wei Xu, Ling Huang, Armando Fox, David Patterson, and Michael Jordan. **Experience on Mining Google's Production Console Logs**. In the *Workshop on Managing Systems via Log Analysis and Machine Learning Techniques (SLAML '10)*, Vancouver, BC Oct, 2010

Ariel Rabkin, Avani Wildani, Randy Katz, Wei Xu, Armando Fox. **A graphical representation for identifier structure in logs**. In the *Workshop on Managing Systems via Log Analysis and Machine Learning Techniques (SLAML '10)*, Vancouver, BC Oct, 2010

Wei Xu. **Detecting Large Scale System Problems by Mining Console Logs**. PhD dissertation, UC Berkeley, July, 2010

Wei Xu, Ling Huang, Armando Fox, David Patterson, and Michael Jordan. **Using Machine Learning Techniques in Console Log Analysis**. Invited paper in *the 27th International Conference on Machine Learning (ICML'10)*, Haifa, Israel, June 2010

Wei Xu, Ling Huang, Armando Fox, David Patterson, and Michael Jordan. **Online System Problem Detection by Mining Patterns of Console Logs**. In Proc. of *the IEEE International Conference on Data Mining (ICDM'09)*, Miami, FL, December 2009

Wei Xu, Ling Huang, Armando Fox, David Patterson, and Michael Jordan. **Large-scale system problem detection by mining console logs**. In Proc. of *the 22nd ACM Symposium on Operating Systems Principles (SOSP'09)*, Big Sky, MT, October 2009

Wei Xu, Ling Huang, Armando Fox, David Patterson, and Michael Jordan. **Mining console logs for large-scale system problem detection**. In Proc. of *the 3rd workshop on Tackling Computer Systems Problems with Machine Learning Techniques (SysML'08)*, San Diego, CA, December 2008

Wei Xu, Zhangxi Tan, Armando Fox and David Patterson. **Regulating workload in J2EE application servers**. In Proc. of *the 1st International Workshop on Feedback Control Implementation and Design in Computing Systems and Networks (FeBID'06)*, Vancouver, Canada, April 2006

Wei Xu, Xiaoyun Zhu, Sharad Singhal, and Zhikui Wang. **Predictive control for dynamic resource allocation in enterprise data centers**. In Proc. of *the 10th IEEE/IFIP Network Operations & Management Symposium (NOMS'06)*, Vancouver, BC, Apr. 2006

Wei Xu. **Feedback control theory and processing system log streams**. *Master thesis, EECS Department, UC Berkeley*, December, 2005

Wei Xu, Joseph L. Hellerstein, Bill Kramer and David Patterson. **Control considerations for scaling event processing**. In Proc. of *the 16th IFIP/IEEE Distributed Systems: Operations and Management (DSOM'05)*, Barcelona, Spain, October 2005

Wei Xu, Peter Bodik and David Patterson. **A flexible framework for statistical learning and data mining from system log streams**. In Proc. of *Workshop on Temporal Data Mining: Algorithms, Theory and Applications at The Fourth IEEE International Conference on Data Mining (ICDM'04)*, Brighton, UK, Nov, 2004

Bjorn Knutsson, Honghui Lu, Wei Xu and Bryan Hopkins. **Peer-to-Peer support for massively multiplayer games**. In Proc.

of the 23rd Conference of the IEEE Communications Society (INFOCOM'04), Hong Kong, March 2004

PROFESSIONAL ACTIVITIES

Program Committee Chair / Organizer

Workshop "Hot Topics in Big Data and Networking (HotData-1), collocated with ICCCN 2014
IEEE Big Data Science Conference 2014
Workshop "Dependability and Security of System Operation" (DSSO), collocated with DSN 2014-2015

Guest Editor

Chinese Science Bulletin, Special issue on Data Science 2014-2015

Program Committee Member

IEEE International Conference on Parallel and Distributed Systems, ICPADS'14 2014
International Conference on Computer Communications and Networks, ICCCN'14 2014
Asia-Pacific Workshop on Systems 2014
19th IEEE International Conference on Parallel and Distributed Systems (ICPADS 2013) 2013
Workshop on Managing Systems via Log Analysis and Machine Learning Techniques (SLAML) and Workshop on Management Automation 2010-2012

Invited talks and panel discussions

CloudConnect, Shanghai 2015
Chinese Higher Education Development Network (CHED) annual conference 2014
OpenStack Summit - Paris 2014
China Future Networking and SDN Summit 2014
the 27th International Conference on Machine Learning (ICML'10) 2010
Improving Logging Practice, (SLAML'11) 2011
The 2nd Workshop on the Analysis of System Logs (WASL) 2009 2009

HONORS AND AWARDS

IBM Faculty Award, IBM 2015
Microsoft Research Asia Collaborative Research Awards, Microsoft 2014
Google Faculty Research Award, Google 2014
"Youth Talent 1000" Scholar Award, Chinese central government program 2013
Department Fellowship, EECS Department, UC Berkeley, 2004
Faculty Appreciation Award, Univ. of Penn 2003
Dean's List, School of Engineering and Applied Science, Univ. of Penn 2002-2003
Tau Beta Pai Honor Society and Golden Key Honor Society 2002