Hao Jiang

Mobile: +1 (315) 267-6271 | Email: hajiang@clarkson.edu | Webpage: http://www.clarkson.edu/~hajiang
Current Address: Department of Computer Science, Clarkson University, Potsdam, NY, 13676

PROFILE

Master graduate seeks a Ph.D. degree in distributed system and machine learning research.

Main strength is the proven ability to do independent research supported by both academic acumen and system design programming skill to implement them. Have gained research experience and studied broad topics with excellent credits during master study together with 7 years of industry experience before that.

Have self-motivated enthusiasm in research. Proposed multiple research topics to advisor and kept working on them. With solid background in programming and distributed system design, have leaded design and development of large distributed system and online data mining system.

EDUCATION

Clarkson University, Potsdam, NY
M.S in Computer Science   expected Jan 2015   GPA 4.0
* *Research Direction* : Application of machine learning with big data and distributed system

Fudan University, Shanghai, China
B.S in Computer Science   2005   Major GPA 3.5

RESEARCH EXPERIENCE

Master Research   Department of Computer Science, Clarkson University
Advisor Prof. Jeanna N. Matthews   Sep. 2012 – Now
* *Distributed platform for Internet Topology research*
    Independently design a distributed system that allows Internet Topology Dataset (huge graph representation of internet connectivity, contains 54 millions of vertices and 2 billion of edges) to be processed in memory efficiently. This work enables more in-depth data mining on Internet structure.
* *IP Geolocation with two-tiered neural network*
    Propose an innovative method of locating the physical location of arbitrary IP address with a two tier neural network. Independently design and implemented the system. Paper is submitted to NSDI'15 as first author.
* *Phishing website detection with logo recognition*
    Use SVM to extract the embedded logos of known websites from a large image and identify phishing websites based on these logos. Independently design and implement the system.
* *Distributed System Simulator*
    Design and implement a simulator that enables the comparison of performance between different distributed system and consistency models on multiple metrics.

Research Assistant in DGDC Project   Clarkson University
Advisor Prof. Pier Marzocca   Dec. 2012 – Now
        Prof. Jeanna N. Matthews

Prof. Stephen Bird

- *Control Plane for Distributed Green Data Center*
  Implement a control plane for virtual machines migration between data centers powered on pure green power. Use optimization method to design a consistency model for the system. Use Internet Topology platform to estimate network latencies between prospective data center positions.

Independent Research   Architect Team, Baidu.Inc
                                                            Mar.2011 – Aug.2012

- *MySQL master-slave replication optimization*
  Rewrite MySQL replication module to support writing to different databases/tables in parallel and allow customized code to be interposed. This increases system throughput by 15% in average.

Undergraduate Research Assistant   Database Lab,
Advisor Prof. Liang Zhang   Department of Computer Science, Fudan University
                                                Jul 2003 – Jun 2005

- *Grid Service Computing*
  Use IBM Globus to implement a grid system providing uninterrupted bioinfomatic database query service to research community

RESEARCH INTEREST

- Theory and application of machine learning
- Distributed system design and implementation
- Big data analysis and data mining

SKILLS

- Application experience with common machine learning algorithms
- Distributed system design and implementation
- Big data processing with Apache Hadoop and customized system
- Proficient programming skills with Java / Scala / C++ / Python / Shell

TEACHING EXPERIENCE

Clarkson University   Department of Math & Computer Science
Teaching Assistant   Introduction to Computer Science   Fall 2014
- Labs, lectures and recitation classes
Teaching Assistant   Calculus I   Fall 2013
- Recitation and discussion classes
Teaching Assistant   Linear Algebra and Differential Equation   Fall 2012
- Proctoring exam and grading.

PUBLICATIONS

Hao Jiang, Jeanna N. Matthews
IP Geolocation with Two-tiered Neural Networks
Submitted to USENIX Symposium on Network System Designs and Implementation (NSDI'15), April 2015

Wenjin Hu, Long Zhang, **Hao Jiang,** Jeanna N. Matthews

A Quantitative Study of Virtual Machine Live Migration
The ACM Cloud and Autonomic Computing Conference (CAC2013), August 2013

RELATED PROFESSIONAL EXPERIENCE

Baidu Inc. – Shanghai, China   Senior System Architect
                                                    Mar 2011 – Aug 2012

- *MySQL Replication Optimization*
    As system architect, I rewrite MySQL's replication module to allow replications operation to be processed in parallel. This system also allows users to interpose customized code to transpose data before it is replicated to slave nodes. The system performance is improved by 15% in average.
- *Data mining with Apache Hadoop*
    Baidu is a search engine with online ad system similar to Google. I, as technical leader, use Apache Hadoop to process Baidu's query log, estimate users' daily expense on ad, and give them personalized recommendation on ad keywords. Our system works on a cluster with more than 1000 job nodes and the data volume in total is over 20 terabytes daily.
- *JVM memory leaking monitoring*
    I design and develop a JVM plugin to automatically log object creation and releasing to file system. The log is then processed with Apache Hadoop to monitor production memory leak in semi real time. This system is deployed on a production system with millions of requests daily. It is capable of finding possible memory leak within minutes.

OOCL Co. Ltd. – Shanghai, China   System Architect
                                                    Jul 2005 – Mar 2011

- *Company Financial System*
    I lead the design and implementation of company's internal financial management system as system architect and technical leader. The system is implemented with Javascript, J2EE and Oracle database. It handled 50,000 requests daily.

REFERENCE

Professor Jeanna. N. Matthews   Professor Christino Tamon
Department of Computer Science   Department of Computer Science
Clarkson University   Clarkson University
Potsdam, NY, 13676   Potsdam, NY, 13676
jnm@clarkson.edu   tino@clarkson.edu

Professor Christopher Lynch   Professor Liang Zhang
Department of Computer Science   Department of Computer Science
Clarkson University   Fudan University
Potsdam, NY, 13676   Shanghai, China, 200433
clynch@clarkson.edu   lzhang@fudan.edu.cn