CURRICULUM VITAE

NAME	POSITION TITLE
Judy Fox (Legal) Judy Qiu (Professional on papers) Email: xqiu@indiana.edu Research Website: https://www.judyfox.online Cell: (812)219-4719 Work: (812)856-7977	Associate Professor of Intelligent Systems Engineering and Computer Science Director, Graduate Studies of Data Science

INSTITUTION AND LOCATION	DEGREE	YEAR(s)	FIELD OF STUDY
Beihang University, Beijing	BS	1987-1991	Computer Science & Engineering
Beihang University, Beijing	MS	1995-1998	Computer Science & Engineering
Syracuse University, Syracuse, NY	MS	1998-2000	Computer Science
Syracuse University, Syracuse, NY	PhD	2000-2005	Computer Science
Indiana University, Bloomington, IN	Postdoc	2007-2009	Computer Science

RESEARCH SUMMARY

My current interdisciplinary research focuses on designing experimental systems that enable Data Science, Machine Learning, and the Internet of Things applications to harness the computational resources of Cloud and HPC platforms effectively. My specific contributions to this community have been in the area of programming models and performance optimization for Data Analytics and Machine Learning, where I have designed and implemented novel high performance Map-Collective runtimes and fundamental computation models for batch (past data), streaming data-models (present data), and learning and prediction (future data) scenarios. High performance graph algorithms are a particular focus. More broadly, my interests include parallel and distributed systems, including HPC and scientific computing. I have strong, successful collaborations with industry, including leadership of an Intel Parallel Computing Center, and work with the IndyCar Company on real-time analysis of racecar data. In the past, I worked extensively with Microsoft on parallel computing and Azure and developed online Courses with Google Funding.

My research has an integrating theme of Data-Enabled Discovery Environments for Science and Engineering, using novel technologies driven by applications. I was one of the first to demonstrate the effectiveness of Iterative MapReduce on both HPC and commercial cloud environments, and I established the importance of collective communication in Big Data systems. Our paper "Twister: A Runtime for Iterative MapReduce" was published in 2010 *and has been cited 1080 times*. The open source Harp framework generalized this to include support of four distinct computation models (Locking, Rotation, Allreduce, and Asynchronous) found in different parallel patterns (kernels) of machine learning and included an extended collective communication library, for high-performance graphs, and machine learning. I am currently extending it to support streaming data analysis, including indexing extensions for HBase and parallel clustering using Apache Storm. Through several collaborations with application scientists, I have built an interdisciplinary program to motivate and validate my computer science systems work. The applications include Exascale Simulation, Data analytics, Bioinformatics, Streaming Time series, Nanoengineering, Neural Networks, and Network Science.

According to Google Scholar, my h-index is 28, with a total of 4093 citations (February, 2020). They verify the outside interest in my recent research. I was honored by giving an invited talk on our Harp-DAAL research project at the SC17 Supercomputing conference, a leading international venue on High-Performance Computing. My earlier paper from 2009 (CW29) with a follow-up journal article (J6) was cited as paper with the greatest impact in the cloud area, and I gave another invited keynote talk. In 2019, I've been working on large-scale graph analytics, machine learning (e.g. Gradient Boost Decision Tree) and deep learning on Anomaly Detection and Prediction over streaming data. This research has been presented at SC and Big Data conferences in collaboration with the University of Virginia and industry partners Intel, NEC and IndyCar.

EDUCATION SUMMARY

I have received three educational funding awards, including those from Google, NIH, and NSF, to develop new curricula in emerging areas to address national big data challenges and workforce development. At Indiana University, I have developed and taught four distinct new courses: Distributed Systems for graduates and undergraduates, Cloud Computing for residential and online/MOOC versions, High-Performance Big Data Systems, and Data Structure courses. Most classes were given an overall rating of outstanding by the majority (60%-80%) of the students. I co-founded the IU Data Science program in 2014 with the Cloud MOOC course and continue to teach big data and systems courses offering to Computer Science, Engineering, and Data Science students. Currently, I am the Director of Graduate Studies and oversee both the residential and online Data Science programs. In this role, I work with the Data Science Curriculum committee and proposed a new "Big Data Systems" track proposal that establishes a new MS degree concentration oriented to "Data Engineering" and "Intelligent Systems Engineering." I supervise students and outreach to industrial partners and employers, especially related to their Curriculum Practical Training (CPT) and integration with the existing curriculum development.

SIGNIFICANT SOFTWARE

As part of my research that involves building large software systems, major open source software releases are listed below. Harp has been used as an innovative High-Performance Big Data Computing framework by over 370 IU graduate students for their class projects, where Harp-DAAL is a collaboration of NSF and Intel projects.

- Judy Qiu, Chathura Widanage, IndyCar Streaming predicative Analytics, October 2019

- Judy Qiu, Bingjing Zhang, Harp (production release with tutorial), February 2017 (initial porotype on February 2014; updated on February 2019 for new software libraries such as parallel graph algorithms)

- Judy Qiu and Xiaoming Gao, IndexedHBase, January 2014

- Judy Qiu and Thilina Gunarathne, Twister4Azure, September 2012

- Judy Qiu and Jong Young Choi, Plotvz, November 2010

- Judy Qiu and Jaliya Ekanayake, Twister, March 2010

POSITIONS AND HONORS

Positions

2019-	Director, Graduate Studies of Data Science, School of Informatics, Computing, and Engineering
2016-	Associate Professor (tenured), Intelligent Systems Engineering, School of Informatics, Computing, and Engineering
2016	Associate Professor (tenured), Computer Science, School of Informatics and Computing, Indiana University
2010-2015	Assistant Professor, Computer Science, School of Informatics and Computing, Indiana University
2009-	Assistant Director, Digital Science Center, Indiana University
2007-2009	Postdoctoral Researcher, Pervasive Technology Labs, Indiana University
2006	Senior Software Project Lead, Anabas Inc., Bloomington, Indiana
2001-2005	Graduate Research Assistant, Pervasive Technology Labs, Indiana University
1999	Graduate Teaching Assistant, Syracuse University, Syracuse, NY, fall semester 1999
1995-1998	Graduate Research Assistant, National Lab of Software Development Environment, Beihang University
1991-1998	Assistant Lecturer and Lecturer, Beihang University

Honors

2014-2015	Indiana University Outstanding Junior Faculty Award

2013-2014   Indiana University Trustees Award for Teaching Excellence

2012   National Science Foundation CAREER Award

2005   Outstanding Graduate Student, College of Engineering and Computer Science, Syracuse University

1998-2001   Fellowship, Syracuse University

Honors or Research and Publication Highlights

2014   NSF Student Travel Grant (13 out of 70 students) for Ph.D. student Xiaoming Gao for his paper "Supporting Queries and Analyses of Large-Scale Social Media Data with Customizable and Scalable Indexing Techniques over NoSQL Databases" at IEEE/ACM CCGrid conference, March 26-29, 2014.

2012   Cloud Challenge Award at 5th Workshop on Many-Task Computing on Grids and Supercomputers (MTAGS) of SC12 for the paper "Cloud Technologies for Bioinformatics Applications."

2012   Best student paper runner-up of IEEE Utility and Cloud Computing 2011 conference for the paper "Portable Parallel Programming on Cloud and HPC: Scientific Applications of Twister4Azure".

2010   First place at the 2010 STARS Celebration Conference from her HBCU student Keenan Black's poster presentation on "Data Point Visualization and Clustering Analysis."

Other Experience and Professional Memberships

- Professional Activities

I have been actively contributing to my profession. I have organized several workshops and journal special issues in multicore computing, cloud computing, and data-enabled science. For instance, I am currently organizing the Streaming Systems and Realtime Machine Learning (STREAM ML) workshop at the IEEE Big Data Conference 2019. This follows my organization of the Big Data, and Extreme-Scale Computing 2 (BDEC2) workshop in Bloomington, and the High Performance Systems, and Analytics for Big Data Workshop HPSA2018 also in Bloomington. I am on the editorial board of the International Journal of Cloud Computing (IJCC) and an organizer or member of many top-tier activities, program committees, and panels.

- Conference and Workshop Leadership

2019   Real-Time Machine Learning Workshop (STREAM-ML) of the IEEE Big Data Conference, December 9-12, 2019, Las Angeles, CA.

2018   Big Data and Extreme-Scale Computing 2 Workshop (BDEC2), November 28-30, 2018.

2018   High-Performance Systems and Analytics for International Big Data Workshop (HPSA18), April 11, 2018, Indiana

2017   Big Simulation and Big Data Workshop (BSBD), Indiana University, January 9, 2017.

2016   Doctoral Symposium Chair, the IEEE International Conference on Cloud Engineering (IC2E2016)

2014   Doctoral Symposium, 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid2014), May 26-29, 2014

2013   Summer workshop on "Algorithms and Cyberinfrastructure for large scale optimization/AI", August 9, 2013

2012   Emerging Computational Methods for the Life Sciences Workshop (ECMLS 2012) at the 21th ACM High Performance Distributed Computing Conference (HPDC2012), Delft, The Netherlands

2012   Grid and Cloud Computing (GCC) of International Conference on Communication Networks (ICCCN2012), München, Germany

2012   Instructor and organizer, NCSA Science Cloud Summer School, July 30~August 3, 2012

2011   Organizing Committee and Chair, Emerging Computational Methods for the Life Sciences Workshop (ECMLS 2011) at the 20th ACM High Performance Distributed Computing Conference

2010   2nd IEEE International Conference on Cloud Computing Technology and Science Conference (CloudCom2010)

2010   Big Data for Science Workshop at NCSA virtual summer school of Computational Science and Engineering (VSCSE 2010)

2010   Organizing Committee and Chair, Frontiers of GPU, Multi- and Many-Core Systems Workshop (FGMMS2010) at the 10th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing Conference

2010   Emerging Computational Methods for the Life Sciences Workshop (ECMLS 2010) at the 19th ACM High Performance Distributed Computing Conference

- Other Service Activities

2020   Program Committees (Clouds & Distributed Computing of SC20, IPDPS 2020, ISC 2020)

2019   Program Committees (Poster SC 2019, Clouds & Distributed Computing SC19, IEEE Big Data 2019, IEEE Clustering 2019, HiPC 2019)

2018   Program Committees (Poster SC 2018, Clouds & Distributed Computing SC18, HPDC 2018, IEEE Big Data 2018)

2017   Invited talk, SC17, November 15, 2017, Denver, Colorado

2017   Program Committees (IC2E2017, ISC2017, Poster SC 2017, Systems SC 2017, ICPP2017)

2016   Keynote talk, 9th Workshop on Many-Task Computing on Clouds, Grids, and Supercomputers (MTAGS2016)

2016   Invited talk, NSF Workshop on Applications of Analytics and Machine Learning in Energy Industry-Academia Workshop (Energy), Georgia Institute of Technology, Atlanta, September 6, 2016

2016   Panelist and invited talk, NSF Workshop on Geospatial Data Science in the Era of Big Data and CyberGIS, Urbana, Illinois, July 25-26, 2016

2016   Instructor, ADMI Cloud Computing Workshop (ADMICloud2016), Elizabeth City, North Carolina, June 10-12, 2016

2016   Organizing Committee and Co-chair, Applications Track of the 44th Annual International Conference on Parallel Processing (ICPP2016)

2016   Program Committees (CCGrid2016, XSEDE2016, MTAGS2016)

2016   National Science Foundation (NSF) review panelist, May 22-23, 2016

2015   National Science Foundation (NSF) review panelist, December 16-17, 2015

2015   External reviewer of University of Missouri Research grant proposal, May 2015

2015   Panelist, the 2nd IEEE Workshop on Cloud Analytics (IWCA) of IEEE IC2E Conference, March 12, 2015

2015   Panelist, Internet of Things Panel of IEEE IC2E Conference, March 11, 2015

2015   Panelist, DOE ASCR Machine Learning workshop, January 5-7, 2015

2015   Program Committees (CCGrid2015, IC2E2015, BDC2015, XSEDE2015, DSDIS2015)

2014   Book Co-Editor, Cloud Computing for Data Intensive Applications, Springer Publisher, December 8, 2014

2014   Exhibitor, SOIC/IU booth, SC'14 conference, November 17-19, 2014

2014   Program Committees (SuperComputing2014, IPDPS2014, CCGrid2014, ICCCN2014, CSWC2014, BigSystem2014, UCC2014, DataCloud2014, MTAGS2014, CASK2014)

2013 Program Committees ([SuperComputing2013](), [HPDC2013](), [XSEDE2013](), [CCGrid2013](), ClOUD2013, CGC2013, CAC2013, [DataCloud13]())

2013 Panelist, Birds of a Feather at SC13 on "Community MOOCs for Computing Technologies and Applications", November 20, 2013

2013 Panelist, 6th Workshop on Many-Task Computing on Clouds, Grids, and Supercomputers ([MTAGS]()) at SC'13, November 17, 2013

2013 Book Reviewer, Cloud Computing: Technologies, Principles, and Paradigms. MIT Press, 2014

2013 Guest Editor, Special Issue on Emerging Computational Methods for the Life Sciences for Journal of Concurrency and Computation: Practice and Experience ([CCPE]()), Wiley Publishers

2013 National Science Foundation (NSF) review panelist, May 28, 2013

2012 National Science Foundation (NSF) CAREER review panelist, October 25-26, 2012

2012 National Science Foundation (NSF) review panelist, April 19-20, 2012

2012 [Guest Editor](), Special Issue on Data Intensive eScience for Journal of Parallel and Distributed Database ([DAPD]()), Volume 30, Issue 5-6, Springer Publishers, August 25, 2012

2012 Guest Editor, Special Issue on Emerging Computational Methods for the Life Sciences for Journal of Concurrency and Computation: Practice and Experience ([CCPE]()), Wiley Publishers

2012 Program Committees ([SuperComputing2012](), [SOCC2012](), [eScience2012](), [CCGrid2012](), [CGC2012](), MapReduce2012, ISDP2012, [CloudCom2012]())

2011- Member, ACM SIGHPC

2011 National Science Foundation (NSF) review panelist, March 2011

2011 [Guest Editor](), Special Issue on Emerging Computational Methods for the Life Sciences for Journal of Concurrency and Computation: Practice and Experience ([CCPE]()), Wiley Publishers, October 17, 2011

2011 Program Committees ([SuperComputing2011](), [MTAGS 2011, DataCloud-SC11](), [UCC 2011](), [CGC2011](), [Grid 2011](), [PBC 2011](), IEA/AIE 2011, MapReduce2011, HiPC 2011, [AusPDC 2011]())

2010 Editorial board, International Journal of Cloud Computing ([IJCC]()), InderScience Publishers

2009 Mentor of 2019, 2013, 2012, 2011, 2010 and 2009 HBCU Summer [STEM]() Initiative program at Indiana University

RESEARCH SUPPORT

My research has been funded by NSF, NIH, Microsoft, Google, Intel, NEC, and Indiana University. I have been on the leadership teams of 21 funded projects since 2009, of which I was PI for 10 ($1.76 million research grants including $400K for educational grants). I'm Co-PI and project lead (with $2.21 million as my portion) of $18.5 million funded support. My total personal funding is $3.97 million. Eighteen of the projects relate directly to my research, and the remaining three focus on education and training.

Funded Research

C1. **National Science Foundation: Judy Qiu (Role: Co-PI 33%; PI: Xiaofeng Wang)**

BIGDATA: IA: Enabling Large-Scale, Privacy-Preserving Genomic Computing with a Hardware-Assisted Secure Big-Data Analytics Framework

Total Award Amount: $1,000,000 Awarded: 1/1/2019 – 12/31/2022

This project focuses on studying a big-data analytics framework on Intel SGX and applying it to support privacy-preserving, large-scale genomic data analyses, and other computing tasks. The new MPI-based cluster computing framework is built to automatically optimize the deployment of computing nodes across enclaves and CPU packages

under resource constraints, supporting iterative computing and a set of fundamental genomic computing tasks while effectively controlling side-channel leaks.

C2. **National Science Foundation: Judy Qiu (Role: Co-PI 75% of IU component)**

Collaborative Research: Framework: Software: CINES: A Scalable Cyberinfrastructure for Sustained Innovation in Network Engineering and Science (Geoffrey Fox and Judy Qiu)

Total Award Amount: $500,000 (at IU) Awarded: 11/01/2018 - 10/31/2023

This project is to build self-sustaining cyberinfrastructure (CI) named CINES (Cyberinfrastructure for Sustained Innovation in Network Engineering and Science) that will be a community resource for network science. CINES will be an extensible and sustainable platform for producers and consumers of network science data, information, and software. CINES will enable fundamental changes in the way researchers (e.g. social scientists) study and teach complex networks.

C3. **NEC Corporation and Indiana University Research Collaboration: Judy Qiu (Role: PI)**

High-Performance Systems and Analytics for Big Data

Total Award Amount: $64,175 Awarded: 8/31/2019-3/31/2020

This project utilizes High-Performance Computing (HPC) clusters to optimize data analytics, and on graph and/or machine learning algorithms that explore the NEC hardware architectures. The Machine Learning and Graph Analytics Algorithms and Applications leverage linear algebra kernels for the systems that advance the programmability and computing capabilities through NEC's vector architectures.

C4. **Intel Parallel Computing Center at Indiana University (Renewed IPCC@IU): Judy Qiu (Role: PI)**

Address grand challenges in High Performance Data Analytics on HPC-Cloud.

Total Award Amount: $200,000 Awarded: 7/1/2018 – 6/30/2020

The goal is to advance the research of the participants by taking advantage of the Intel® architecture, including addressing programmer productivity and performance portability. We focus on Big Data applications, including analysis of IndyCar streaming data.

C5. **National Science Foundation: Judy Qiu (Role: Senior Personnel 7%)**

Network for Computational Nanotechnology (NCN) – Engineered nanoBio node (Geoffrey Fox, James Alexander Glazier, Vikram Jadhao, Paul Macklin, Trevor Douglas, Judy Qiu, Eleftherios Garyfallidis, and Suresh Marru)

Total Award Amount: $4,000,000 Awarded: 7/1/2017 – 8/31/2022

The node aims to create computational tools designed for cutting-edge research to develop biocompatible, safe, and efficient nanoscale devices. The node is contributing tools that: 1) design functional nanoparticles and self-assembled nanostructures with user-selected physicochemical, mechanical, and biocompatible properties, 2) evaluate and control nanodevice-cell interactions and establish nanodevice-cell phenotype links, and 3) enable the engineering of multicellular systems using the nanoscale design elements and the nanodevice-cell phenotype links.

C6. **National Science Foundation: Judy Qiu (Role: Co-PI 33% IU)**

CIF21 DIBBS: Middleware and High Performance Analytics Libraries for Scalable Data Science (Geoffrey Fox, Judy Qiu, Shantenu Jha, Madhav Marathe, Fusheng Wang, John Paden, Oliver Beckstein and Thomas Cheatham)

Total Award Amount: $5,250,000 (33% IU) Awarded: 10/01/2014 – 09/30/2020

The project is designing and developing software and algorithm components to support large scale parallel data analytics on cloud or HPC platforms. Applications in Polar Science, Computer Vision, Pathology Informatics, and Network Science motivate algorithms and software and are used to evaluate new middleware and parallel algorithms.

C7. **Indiana University Faculty Research Support Grant: Judy Qiu (Role: Co-PI 50%; PI: Lei Jiang)**

Constructing a High Performance and Low Power Campus Research Cyberinfrastructure at IU

Total Award Amount: $28,500 Awarded: 06/01/2018 – 03/31/2019

The project was a collaboration of campus cyberinfrastructure researchers to advance Science Applications with innovative modern computer systems. It uses IU facilities as an example in this FRSP project to prepare for a federal proposal next year, which investigates a general approach.

C8. **Indiana University Precision Health Initiative: Judy Qiu (Role: data analytics framework lead)**

Data Analytics software environment for Precision Heath (Director of SICE PHI: Kay Connelly)

Total Award Amount: $756,000 (50%) Awarded: 07/01/2016 – 06/30/2020

The project developed High-Performance Big Data Software environment aimed at Precision Health, including social media and medical diagnostic applications.

C9. **National Science Foundation XPS: Judy Qiu (Role: Co-PI 33% IU)**

Collaborative Research: Rapid Prototyping HPC Environment for Deep Learning (Andrew Ng, Jack Dongarra, Geoffrey Fox)

Total Award Amount: $900,000 (33% IU) Awarded: 10/01/2014 – 09/30/2017

This research investigated and developed a high-performance software run time for Deep Learning with a Python front end, benchmarks, and supported with NoSQL storage.

C10. **Intel Parallel Computing Center at Indiana University (IPCC@IU): Judy Qiu (Role: PI)**

Address grand challenges in High Performance simulation and data analytics with innovative solutions and software development using Intel architecture (Judy Qiu and Steve Gottlieb)

Total Award Amount: $330,000 Awarded: 9/1/2015 – 7/31/2017

The goal was to advance the research of the participants by taking advantage of the Intel® architecture, including the addressing of programmer productivity and performance portability. We developed and benchmarked on physics MILC simulations as well as Large-Scale Data Analysis and Applications on systems built around the Intel® Xeon Phi™ Processor family.

C11. **National Science Foundation CAREER Award: Judy Qiu (Role: PI)**

Programming Environments and Runtime for Data Enabled Science [OCI-1149432]

Total Award Amount: $499,994 Awarded: 03/01/2012 – 02/28/2017

This research was at the nexus of the data deluge in science and business and two major computing thrusts – clouds and Exascale scientific systems that are unified with an interoperable runtime system. The project aimed to transform the approach to applications that varies from data mining of genomic and proteomic data for science to data analytics for business. Computer science areas at the heart of the research – namely Iterative Map Collective runtime, fault tolerance, data-computing co-location, and high-level languages – were advanced. Furthermore, the new applications enabled, and new software paradigms were fed back into the architecture of cloud and Exascale systems suggesting particular storage and communication choices and new directions for the national infrastructure.

C12. **Microsoft Azure Resource Allocation Grant: Judy Qiu (Role: PI)**

Map-Collective on Azure Cloud

Total Award Amount: $40,000 Awarded: 3/15/2014 – 3/15/2015

Total Award Amount: Azure compute hours approx. 200,000 and 200GB of "out" data transfer per month, one terabyte of stored data, and 50,000 Storage Transactions (Queue, Table and Blob operations) measured in units of 100K.

This research focused on a generalization of MapReduce to a Map-Collective model where the "reduce" phase in MapReduce is supported by a library of powerful optimized collective communication routines covering operations like (all)reduce, scatter, gather, broadcast, regroup, combine, and merge, which cover the key primitives in MapReduce and MPI.

C13. **Indiana University Faculty Research Support Grant: Judy Qiu (Role: PI)**

Research in Data Intensive Cloud Programming (Twister)

Total Award Amount: $50,501 Awarded: 09/01/2011 – 06/30/2013

This project researched Iterative MapReduce as a user-friendly, efficient cloud programming model with a focus on Computer Science research issues. In particular, we addressed communication performance, fault-tolerance, and storage issues, and combined new research with extending the open-source software Twister environment on both Azure and private Linux clouds. We used bioinformatics and cheminformatics applications to motivate and validate our research results.

C14. **Indiana University Faculty Research Support Grant: Judy Qiu (Role: Co-PI 50%; PI Raquel Hill)**

Childhood Obesity Studies with Secure Cloud Data Mining (Twister Application)

Total Award Amount: $ 36,501 Awarded: 09/01/2011 – 06/30/2013

The collaborative project applied cloud analysis techniques to health data linked to rich geographical location data coming from GPS sensors and accelerometers. This project was needed to show that the richer time series information can be analyzed effectively and securely. Further, it designed on-demand, secure computing services for the processing of health data.

C15. **Microsoft Azure Grant: Judy Qiu (Role: PI)**

Microsoft Azure Cloud Academic Resources Allocation

Total Award Amount: 50 core computing instances and 5 TB of storage (approximate value of $123,000) Awarded: 03/30/2011 – 03/30/2013

The Windows Azure platform included a cloud services operating system and a set of developer services, which provided the functionality to build data analysis applications on demand. We researched Cloud Programming Models and their Runtime and used Azure as a testbed to develop MapReduce and iterative MapReduce (Twister) framework for data intensive applications on public Clouds.

C16. **Microsoft Foundation Grant: Judy Qiu (Role: PI)**

MapReduce and Twister on Windows Clusters

Total Award Amount: $125,000 Awarded: 11/01/2010

This project investigated novel parallel and distributed systems on multicore clusters with an emphasis on Expectation-Maximization (EM) data mining algorithms. We researched many important aspects of MapReduce and iterative MapReduce (Twister) approach to data intensive applications in cheminformatics and bioinformatics. Dimension reduction, clustering, and Latent Dirichlet Allocation were focus algorithms.

C17. **National Science Foundation: Judy Qiu (Role: Co-PI; PI: Madhav Marathe)**

From Desktops to Clouds – A Middleware for Next Generation Network Science [OCI-1032677]

Total Award Amount to IU: $255,000 (Total Award $1,281,562) Awarded: 8/01/2010 – 08/30/2014

This project involved five different biology projects, including gene assembly and metagenomics where applications are using Cloud systems and technologies. Performance and ease of use were evaluated compared to traditional approaches. Methods that scale to large problems were developed.

C18. **National Institutes of Health: Judy Qiu (Role: Co-PI 20%; PI: Geoffrey Fox)**

Investigation of Cloud Computing to Support Data-Parallel Health Research [1RC2HG005806-01]

Total Award Amount: $1,506,926.00 Awarded: 09/30/2009 – 09/29/2011

This project formed a multidisciplinary team of Indiana University computer scientists, biologists, and bioinformaticians to develop and deploy new large-scale computing infrastructure and tools that enabled fundamental health research. Our research investigated the impact of Cloud computing architectures on large-scale computational biology, and covered widely encountered, "data-parallel" problems including but not limited to DNA sequence analysis.

Educational Grants

C19. **National Science Foundation: Judy Qiu (Role: PI)**

EAGER: Remote Sensing Curriculum Enhancement using Cloud Computing (with Linda Hayden)

Total Award Amount: $299,781 (IU $189,780) Awarded: 09/01/2015 – 2/28/2017

This project was a collaboration between Prof. Linda Hayden of Elizabeth City State University and Prof. Judy Qiu of Indiana University in environmental applications of Microwave Remote Sensing using Cloud Computing technology. Our prototype demonstrated the concept that Data and Computational Science (remote sensing) curriculum can drive new workforce and research opportunities at Minority Serving Institutions (MSI) by exploiting enhancements using Cloud Computing technology.

C20. **National Institute of Heath: Judy Qiu (Role: Senior Personnel; PI: Ilkay Altintas)**

R25 HG008292-01 entitled An Open Resource for Collaborative Biomedical Big Data Training

Total Award Amount to IU: $40,000 (Total Award $600,000) Awarded: 12/01/2014 – 11/30/2017

This project built a Biomedical Big Data Training Collaborative repository and created lecture content and example courses using MOOC technologies, with hands-on application use cases for biomedical big data training. It developed and communicated best practices for content development and deployment to a wide range of educators and trainees along with associated adaptive learning, assessment metrics, and testing practices.

C21. **Google MOOC Grant: Judy Qiu (Role: PI)**

Customizable MOOC for Cloud Computing

Total Award Amount: $49,695 Awarded: 12/20/2012

This project used MOOCs to produce custom modules and courses optimized in cloud computing applications for particular communities and topics. We explored the use of Google Course Builder, emphasizing the expansion beyond current MOOC formats and how social interactions can be used to enhance the student experience.

PUBLICATIONS

I have over 90 peer-reviewed book, journal, conference, and workshop publications. I contributed several sections to a published textbook by Morgan Kaufmann Publishers (October 2011) on "Distributed and Cloud Computing" with Kai Hwang as lead author. The book achieved "Academic Library Community and American Library Association for Best Scholarly Title" in the Elsevier 2012 Choice Outstanding Academic Titles List. I'm co-editor of a book on "Cloud Computing for Data Intensive Applications" published by Springer in 2014. I contributed to another book on "Data Intensive Computing for

Bioinformatics" in Data Intensive Distributed Computing by IGI Global Publishers (January 2012), and to Book Series of "HPC and Big Data: Convergence and Ecosystem", "New Frontiers in High Performance Computing", "Advances in Parallel Computing", "Big Data in Complex and Social Networks Handbook" between 2016 and 2018.

Refereed Conference and Workshop Proceedings

CW1. L. Chen, J. Li, C. Sahinalp, M. Marathe, A. Vullikanti, A. Nikolaev, E. Smirnov, R. Israfilov, and J. Qiu, "Subgraph2Vec: Highly-vectorized tree-like subgraph counting," in 2019 IEEE International Conference on Big Data, IEEE, 2019.

CW2. S. Akkas and S. S. Maini and J, Qiu, "A fast video image detection using tensorflow mobile networks for racing cars," in Stream Systems and Real-time Machine Learning (STREAM-ML) Workshop of IEEE Big Data Conference, IEEE, 2019.

CW3. J. Li, F. Wang, T. Araki and J. Qiu, "Generalized Sparse Matrix-Matrix Multiplication for Vector Engines and Graph Applications," in MCHPC'19: Workshop on Memory Centric High Performance Computing of SC'19 Conference.

CW4. C. Widanage, J. Li, S. Tyagi, R. Teja, B. Peng, S. Kamburugamuve, D. Baum, D. Smith, J. Qiu, and J. Koskey, "Anomaly detection over streaming data: Indy500 case study," in 2019 IEEE 12th International Conference on Cloud Computing (CLOUD), pp. 9–16, IEEE, 2019.

CW5. B. Peng, L. Chen, J. Li, M. Jiang, S. Akkas, E. Smirnov, R. Israfilov, S. Khekhnev, A. Nikolaev, and J. Qiu, "HarpGBDT: Optimizing gradient boosting decision tree for parallel efficiency," in 2019 IEEE International Conference on Cluster Computing (CLUSTER), pp. 1–11, IEEE, 2019.

CW6. G. Fox, J. A. Glazier, J. Kadupitiya, V. Jadhao, M. Kim, J. Qiu, J. P.Sluka, E. Somogyi, M. Marathe, A. Adiga, et al., "Learning everywhere : Pervasive machine learning for effective high-performance computation," in 2019 HPCDC workshop of IPDPS conference, pp. 422–429, 2019.

CW7. L. Jiang, L. Chen, and J. Qiu, "Performance characterization of multi-threaded graph processing applications on many-integrated-core architecture," in 2018 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), pp. 199–208, IEEE, 2018.

CW8. B. Peng, B. Zhang, L. Chen, M. Avram, R. Henschel, C. Stewart, S. Zhu, E. Mccallum, L. Smith, T. Zahniser, et al., "Harplda+: Optimizing Latent Dirichlet Allocation for Parallel Efficiency," in 2017 IEEE International Conference on Big Data (Big Data), pp. 243–252, IEEE, 2017.

CW9. J. Qiu, S. Kamburugamuve, H. Lee, J. Mitchell, R. Caldwell, G. Bullock, and L. Hayden, "Teaching, learning and collaborating through cloud computing online classes," in the proceedings of the Workshop on Education for High-Performance Computing (EduHPC-17), Denver, Colorado. November 13, 2017.

CW10. L. Chen, B. Peng, B. Zhang, T. Liu, Y. Zou, L. Jiang, R. Henschel, C. Stewart, Z. Zhang, E. Mccallum, et al., "Benchmarking Harp-DAAL: High Performance Hadoop on KNL clusters," in 2017 IEEE 10th International Conference on Cloud Computing (CLOUD), pp. 82–89, IEEE, 2017.

CW11. B. Zhang, B. Peng, and J. Qiu, "Model-Centric Computation Abstractions in Machine Learning Applications," in Proceedings of the 3rd ACM SIGMOD Workshop on Algorithms and Systems for MapReduce and Beyond, p. 3, ACM, 2016.

CW12. B. Zhang, B. Peng, and J. Qiu, "High performance LDA through Collective Model Communication Optimization," Procedia Computer Science, vol. 80, pp. 86–97, 2016.

CW13. B. Zhang, Y. Ruan, and J. Qiu, "Harp: Collective Communication on Hadoop," in 2015 IEEE International Conference on Cloud Engineering, pp. 228–233, IEEE, 2015.

CW14. G. Fox, J. Qiu, S. Kamburugamuve, S. Jha, A. Luckow, "HPC-ABDS High Performance Computing Enhanced Apache Big Data Stack". In CCGRID Conference, pp. 1057-1066, IEEE, 2015.

CW15. X. Gao, E. Ferrara, and J. Qiu, "Parallel Clustering of High-Dimensional Social Media Data Streams," in 2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, pp. 323–332, IEEE, 2015.

CW16. J. Qiu, S. Jha, A. Luckow, and G. C. Fox, "Towards HPC-ABDS: an initial high-performance big data stack," Building

Robust Big Data Ecosystem ISO/IEC JTC, vol. 1, pp. 18–21, 2014.

CW17. T.-L. Wu, A. Koppula, and J. Qiu, "Integrating pig with harp to support iterative applications with fast cache and customized communication," in Proceedings of the 5th International Workshop on Data-Intensive Computing in the Clouds, pp. 33–39, IEEE Press, 2014.

CW18. S. Jha, J. Qiu, A. Luckow, P. Mantha, and G. C. Fox, "A tale of two data-intensive paradigms: Applications, Abstractions, and Architectures," in 2014 IEEE International Congress on Big Data, pp. 645–652, IEEE, 2014.

CW19. T. Gunarathne, J. Qiu, and D. Gannon, "Towards a collective layer in the big data stack," in 2014 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, pp. 236–245, IEEE, 2014.

CW20. X. Gao and J. Qiu, "Supporting queries and analyses of large-scale social media data with customizable and scalable indexing techniques over NOSQL databases," in 2014 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, pp. 587–590, IEEE, 2014.

CW21. G. C. Fox, S. Jha, J. Qiu, and A. Luckow, "Towards an understanding of facets and exemplars of big data applications," in Proceedings of the 20 Years of Beowulf Workshop on Honor of Thomas Sterling's 65th Birthday, pp. 7–16, ACM, 2014.

CW22. X. Gao and J. Qiu, "Supporting end-to-end social media data analysis with the IndexedHbase platform," in Proceedings of the 6th workshop on many- task computing on clouds, grids, and supercomputers (MTAGS) at SC13, Citeseer, 2013.

CW23. B. Zhang and J. Qiu, "High performance clustering of social images in a map-collective programming model," in Proceedings of the 4th annual Symposium on Cloud Computing, p. 44, ACM, 2013.

CW24. S. E. Abdelhamid, R. Alo, S. Arifuzzaman, P. Beckman, M. H. Bhuiyan, K. Bisset, E. A. Fox, G. C. Fox, K. Hall, S. S. Hasan, J. Qiu, et al., "CINET: A cyberinfrastructure for Network Science," in 2012 IEEE 8th International Conference on E-Science, pp. 1–8, IEEE, 2012.

CW25. L. Stanberry, R. Higdon, W. Haynes, N. Kolker, W. Broomall, S. Ekanayake, A. Hughes, Y. Ruan, J. Qiu, E. Kolker, et al., "Visualizing the protein sequence universe," in Proceedings of the 3rd international workshop on Emerging computational methods for the life sciences, pp. 13– 22, ACM, 2012.

CW26. A. Hughes, Y. Ruan, S. Ekanayake, S.-H. Bae, Q. Dong, M. Rho, J. Qiu, and G. Fox, "Interpolative multidimensional scaling techniques for the identification of clusters in very large sequence sets," in BMC bioinformatics, vol. 13, p. S9, BioMed Central, 2012.

CW27. Y. Ruan, Z. Guo, Y. Zhou, J. Qiu, and G. Fox, "Hymr: a hybrid mapreduce workflow system," in Proceedings of the 3rd international workshop on Emerging computational methods for the life sciences, pp. 39–48, ACM, 2012.

CW28. J. Y. Choi, H. Abbasi, D. Pugmire, N. Podhorszki, S. Klasky, C. Capdevila, M. Parashar, M. Wolf, J. Qiu, and G. Fox, "Mining hidden mixture context with ADIOS-p to improve predictive pre-fetcher accuracy," in 2012 IEEE 8th International Conference on E-Science, pp. 1–8, IEEE, 2012.

CW29. S.-H. Bae, J. Qiu, and G. Fox, "Adaptive interpolation of multidimensional scaling," Procedia Computer Science, vol. 9, pp. 393–402, 2012.

CW30. H. Li, G. Fox, and J. Qiu, "Performance model for parallel matrix multiplication with dryad: Dataflow graph runtime," in 2012 Second International Conference on Cloud and Green Computing, pp. 675–683, IEEE, 2012.

CW31. Y. Ruan, S. Ekanayake, M. Rho, H. Tang, S.-H. Bae, J. Qiu, and G. Fox, "Dacidr: deterministic annealed clustering with interpolative dimension reduction using a large collection of 16s RNA sequences," in Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine, pp. 329–336, ACM, 2012.

CW32. T. Gunarathne, B. Zhang, T.-L. Wu, and J. Qiu, "Portable parallel programming on cloud and HPC: Scientific applications of twister4azure," in 2011 Fourth IEEE International Conference on Utility and Cloud Computing, pp. 97–104, IEEE, 2011.

CW33. H. Li, Y. Ruan, Y. Zhou, J. Qiu, and G. Fox, "Design patterns for scientific applications in DryadLinq CTP," in

Proceedings of the second international workshop on Data intensive computing in the clouds, pp. 61–70, ACM, 2011.

CW34. A. J. Younge, R. Henschel, J. T. Brown, G. Von Laszewski, J. Qiu, and G. C. Fox, "Analysis of virtualization technologies for high performance computing environments," in 2011 IEEE 4th International Conference on Cloud Computing, pp. 9–16, IEEE, 2011.

CW35. Y. Luo, Z. Guo, Y. Sun, B. Plale, J. Qiu, and W. W. Li, "A hierarchical framework for cross-domain Mapreduce execution," in Proceedings of the second international workshop on Emerging computational methods for the life sciences, pp. 15–22, ACM, 2011.

CW36. T. Gunarathne, T.-L. Wu, J. Y. Choi, S.-H. Bae, and J. Qiu, "Cloud Computing paradigms for pleasingly parallel biomedical applications," Concurrency and Computation: Practice and Experience, vol. 23, no. 17, pp. 2338– 2354, 2011.

CW37. J. Y. Choi, S.-H. Bae, J. Qiu, B. Chen, and D. Wild, "Browsing large-scale cheminformatics data with dimension reduction," Concurrency and Computation: Practice and Experience, vol. 23, no. 17, pp. 2315–2325, 2011.

CW38. J. Ekanayake, H. Li, B. Zhang, T. Gunarathne, S.-H. Bae, J. Qiu, and G. Fox, "Twister: a runtime for iterative Mapreduce," in Proceedings of the 19th ACM international symposium on high performance distributed computing, pp. 810–818, ACM, 2010.

CW39. J. Qiu, S. Beason, S.-H. Bae, S. Ekanayake, and G. Fox, "Performance of windows multicore systems on threading and MPI," in 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, pp. 814–819, IEEE, 2010.

CW40. J. Y. Choi, S.-H. Bae, X. Qiu, and G. Fox, "High performance dimension reduction and visualization for large high-dimensional data analysis," in Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, pp. 331–340, IEEE Computer Society, 2010.

CW41. J. Y. Choi, J. Qiu, M. Pierce, and G. Fox, "Generative topographic map- ping by deterministic annealing," Procedia Computer Science, vol. 1, no. 1, pp. 47–56, 2010.

CW42. S.-H. Bae, J. Qiu, and G. C. Fox, "Multidimensional scaling by deterministic annealing with iterative majorization algorithm," in 2010 IEEE sixth international conference on e-Science, pp. 222–229, IEEE, 2010.

CW43. B. Zhang, Y. Ruan, T.-L. Wu, J. Qiu, A. Hughes, and G. Fox, "Applying twister to scientific applications," in 2010 IEEE Second International Conference on Cloud Computing Technology and Science, pp. 25–32, IEEE, 2010.

CW44. T. Gunarathne, T.-L. Wu, J. Qiu, and G. Fox, "Mapreduce in the clouds for science," in 2010 IEEE second international conference on cloud computing technology and science, pp. 565–572, IEEE, 2010.

CW45. S.-H. Bae, J. Y. Choi, J. Qiu, and G. C. Fox, "Dimension reduction and visualization of large high-dimensional data via interpolation," in Proceedings of the 19th ACM international symposium on high performance distributed computing, pp. 203–214, ACM, 2010.

CW46. X. Qiu, J. Ekanayake, S. Beason, T. Gunarathne, G. Fox, R. Barga, and D. Gannon, "Cloud technologies for bioinformatics applications," in Proceedings of the 2nd Workshop on Many-Task Computing on Grids and Supercomputers, p. 6, ACM, 2009.

CW47. G. Fox, X. Qiu, S. Beason, J. Choi, J. Ekanayake, T. Gunarathne, M. Rho, H. Tang, N. Devadasan, and G. Liu, "Biomedical case studies in data intensive computing," in IEEE International Conference on Cloud Computing, pp. 2–18, Springer, 2009.

CW48. X. Qiu, G. Fox, H. Yuan, S.-H. Bae, G. Chrysanthakopoulos, and H. Nielsen, "Parallel data mining on multicore clusters," in 2008 Seventh International Conference on Grid and Cooperative Computing, pp. 41–49, IEEE, 2008.

CW49. X. Qiu, G. C. Fox, H. Yuan, S.-H. Bae, G. Chrysanthakopoulos, and H. F. Nielsen, "Performance of multicore systems on parallel data clustering with deterministic annealing," in International Conference on Computational Science, pp. 407–416, Springer, 2008.

CW50. X. Qiu, G. C. Fox, H. Yuan, S.-H. Bae, G. Chrysanthakopoulos, and H. F. Nielsen, "Parallel clustering and dimensional

scaling on multicore systems," HIGH PERFORMANCE COMPUTING & SIMULATION (HPCS 2008), p. 67, 2008.

CW51. X. Qiu, G. C. Fox, H. Yuan, S.-H. Bae, G. Chrysanthakopoulos, and H. F. Nielsen, "High performance multi-paradigm messaging runtime integrating grids and multicore systems," in Third IEEE International Conference on e-Science and Grid Computing (e-Science 2007), pp. 407–414, IEEE, 2007.

CW52. X. Qiu and A. Jooloor, "Web service architecture for e-learning," Journal of Systemics, Cybernetics and Informatics, vol. 3, no. 5, pp. 92–101, 2006.

CW53. X. Qiu, S. Pallickara, and A. Uyar, "Making SVG a Web Service in a Message-based MVC Architecture," 2004.

CW54. X. Qiu, "Building desktop applications with web services in a Message-based MVC paradigm," in Proceedings. IEEE International Conference on Web Services, pp. 765–768, IEEE, 2004.

CW55. X. Qiu, B. Carpenter, and G. C. Fox, "Collaborative SVG as a web service," in SVG Open 2003 Conference and Exhibition, Vancouver, Canada, 2003.

CW56. X. Qiu, B. Carpenter, G. C. Fox, et al., "Internet collaboration using the w3c document object model.," in International Conference on Internet Computing, pp. 643–647, Citeseer, 2003.

CW57. G. Fox, H. Bulut, K. Kim, S.-H. Ko, S. Lee, S. Oh, S. Pallickara, X. Qiu, A. Uyar, M. Wang, et al., "Collaborative Web Services and Peer-to-Peer Grids," SIMULATION SERIES, vol. 35, no. 1, pp. 3–12, 2003.

Book Contributions

B1. M. Marathe, L. Jiang, and J. Qiu, "High-performance massive subgraph counting using pipelined adaptive-group communication," Big Data and HPC: Ecosystem and Convergence, vol. 33, p. 173, 2018.

B2. B. Zhang, B. Peng, and J. Qiu, "Parallelizing big data machine learning applications with model rotation," New Frontiers in High Performance Computing and Big Data, vol. 30, p. 199, 2017.

B3. T. Wu, B. Zhang, C. Davis, E. Ferrara, A. Flammini, F. Menczer, J. Qiu, M. Thai, H. Xiong, and W. Wu, "Scalable query and analysis for social networks: an integrated high-level dataflow system with Pig and Harp," Big data in complex and social networks, 2016.

B4. G. Fox, J. Qiu, S. Jha, S. Ekanayake, and S. Kamburugamuve, "Big data, Simulations and HPC Convergence," in Big Data Benchmarking, pp. 3–17, Springer, 2015.

B5 G. Fox, S. Jha, J. Qiu, S Ekanazake, A. Luckow, "Towards a comprehensive set of big data benchmarks", in Big Data and High Performance Computing, pp. 47-66, IOS Press, 2015.

B6. X. Li and J. Qiu, Cloud computing for data-intensive applications, vol. 1. Springer, 2014.

B7. X. Gao, E. Roth, K. McKelvey, C. Davis, A. Younge, E. Ferrara, F. Menczer, and J. Qiu, "Supporting a social media observatory with customizable index structures: architecture and performance," in Cloud Computing for Data-Intensive Applications, pp. 401–427, Springer, 2014.

B8. J. Q. B. Zhang, "Mammoth data in the cloud: clustering social images," Cloud computing and big data, vol. 23, p. 231, 2013.

B9. K. Hwang, J. Dongarra, and G. C. Fox, Distributed and cloud computing: from parallel processing to the internet of things. Morgan Kaufmann, 2013.

B10. J. Qiu, J. Ekanayake, T. Gunarathne, J. Y. Choi, S.-H. Bae, Y. Ruan, S. Ekanayake, S. Wu, S. Beason, G. Fox, et al., "Data intensive computing for bioinformatics," in Data Intensive Distributed Computing: Challenges and Solutions for Large-scale Information Management, pp. 207–241, IGI Global, 2012.

B11. J. Ekanayake and G. Fox, "High performance parallel computing with clouds and cloud technologies," in International Conference on Cloud Computing, pp. 20–38, Springer, 2009.

B12. G. Fox, S.-H. Bae, J. Ekanayake, X. Qiu, and H. Yuan, "Parallel Data Mining from Multicore to Cloudy Grids," in High Performance Computing Workshop, vol. 18, pp. 311–340, 2009.

B13. G. Fox, D. Gannon, S.-H. Ko, S. Pallickara, X. Qiu, and A. Uyar, "Peer- to-Peer Grids," 2003.

Journal Papers

J1. Z. Zhao, L. Chen, M. Avram, M. Li, G. Wang, A. Butt, M. Khan, M. Marathe, J. Qiu, and A. Vullikanti, "Finding and counting tree-like subgraphs using mapreduce," IEEE Transactions on Multi-Scale Computing Systems, vol. 4, no. 3, pp. 217–230, 2017.

J2. C. A. Davis, G. L. Ciampaglia, L. M. Aiello, K. Chung, M. D. Conover, E. Ferrara, A. Flammini, G. C. Fox, X. Gao, B. Gon¸calves, et al., "Osome: the IUNI observatory on social media," PeerJ Computer Science, vol. 2, p. e87, 2016.

J3. T. Gunarathne, B. Zhang, T.-L. Wu, and J. Qiu, "Scalable parallel computing on clouds using twister4azure iterative Mapreduce," Future Generation Computer Systems, vol. 29, no. 4, pp. 1035–1048, 2013.

J4. A. Hughes, Y. Ruan, S. Ekanayake, S.-H. Bae, Q. Dong, M. Rho, J. Qiu, and G. Fox, "Interpolative multidimensional scaling techniques for the identification of clusters in very large sequence sets," in BMC bioinformatics, vol. 13, p. S9, BioMed Central, 2012.

J5. J. Qiu, S. Beason, S.-H. Bae, S. Ekanayake, and G. Fox, "Performance of windows multicore systems on threading and MPI," in 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, pp. 814–819, IEEE, 2010.

J6. T. Gunarathne, T.-L. Wu, J. Y. Choi, S.-H. Bae, and J. Qiu, "Cloud Computing paradigms for pleasingly parallel biomedical applications," Concurrency and Computation: Practice and Experience, vol. 23, no. 17, pp. 2338– 2354, 2011.

J7. J. Y. Choi, S.-H. Bae, J. Qiu, B. Chen, and D. Wild, "Browsing large-scale cheminformatics data with dimension reduction," Concurrency and Computation: Practice and Experience, vol. 23, no. 17, pp. 2315–2325, 2011.

J8. J. Qiu, J. Ekanayake, T. Gunarathne, J. Y. Choi, S.-H. Bae, H. Li, B. Zhang, T.-L. Wu, Y. Ruan, S. Ekanayake, et al., "Hybrid cloud and cluster computing paradigms for life science applications," in BMC Bioinformatics, vol. 11, p. S3, BioMed Central, 2010.

J9. X. Qiu, J. Ekanayake, S. Beason, T. Gunarathne, G. Fox, R. Barga, and D. Gannon, "Cloud technologies for Bioinformatics applications," in Proceedings of the 2nd Workshop on Many-Task Computing on Grids and Supercomputers, p. 6, ACM, 2009.

J10. X. Qiu and A. Jooloor, "Web service architecture for e-learning," Journal of Systemics, Cybernetics and Informatics, vol. 3, no. 5, pp. 92–101, 2006.

J11. G. Fox, S.-H. Ko, M. Pierce, O. Balsoy, J. Kim, S. Lee, K. Kim, S. Oh, X. Rao, M. Varank, X. Qiu, et al., "Grid services for earthquake science," Concurrency and Computation: Practice and Experience, vol. 14, no. 6-7, pp. 371–393, 2002.

Other Publications

- Magazine

M1. J. Qiu, Harp-DAAL for High Performance Big Data Computing, Intel Magazine, March 17, 2018.

- Editorials

O1. J. Qiu, I. Foster, C. A. Goble, "Editorial for Special Issue for Emerging Computational Methods for the Life Sciences Workshop 2012," Journal of Concurrency and Computation: Practice and Experience 26(6): 1231-1233 (2014).

O2. J. Qiu, I. Foster and R. Taylor "Editorial for Special Issue for Emerging Computational Methods for the Life Sciences Workshop 2011," Journal of Concurrency & Computation: Practice & Experience 26(4): 851-853, published in February 4, 2013. DOI: 10.1002/cpe.2998.

O3. J. Qiu, D. Gannon , "Editorial for Special issue for data intensive eScience," Journal of Distributed and Parallel Databases, August 25, 2012 (online).

O4. J. Qiu, I. Foster and G. Fox, "Editorial for Special Issue for Emerging Computational Methods for the Life Sciences Workshop 2010," Journal of Concurrency & Computation: Practice & Experience, August 23, 2011.

O5. S. Zhou, J. Qiu and K. Hawick, "Editorial for Special Section on Challenges and Solutions in Multicore and Many-Core Computing," Journal of Concurrency & Computation: Practice & Experience, published November 9, 2011. DOI:

10.1002/cpe.1861.

- Position Papers

O6. J. Qiu, "Generalizing MapReduce as a Unified Cloud and HPC Runtime," position paper in the proceedings of ACM Petascale Data Analytics Workshop (PDACO'11) at Supercomputing11, December 6, 2011.

O7. X.Gao, V. Nachankar and J. Qiu, "Experimenting Lucene Index on HBase in an HPC Environment," position paper in the proceedings of ACM High Performance Computing meets Databases workshop (HPCDB'11) at SuperComputing 11, December 6, 2011.

08. S.-H Bae, J. Qiu, G. C. Fox, "Scalable Dimension Reduction for Large Abstract Data Visualization," Poster at IEEE Cluster 2011 at the Hilton Hotel in Austin, TX, and hosted by the Texas Advanced Computing Center, September 26-30 2011.

- Technical Reports

R1. B. Zhang, B. Peng, J. Qiu, "Parallelizing Big Data Machine Learning Applications with Dynamic Model Rotation," Technical report, August 7, 2016

R2. L. Chen, J. Qiu, "Accelerating Machine Learning with Model-Centric Approach on Emerging Architectures," Technical report, June 29, 2016

R3. T.-L. Wu, B. Zhang, C. Davis, E. Ferrara, S. Flammini, F. Menczer, J. Qiu, "Scalable Query and Analysis for Social Networks: An Integrated High-Level Dataflow System with Pig and Harp," Technical report, 2015.

R4. B. Zhang, J. Qiu, S. Lee, D. Crandall, "Large-Scale Image Classification using High Performance Clustering," Technical report, 2014.

R5. T. Gunarathne, J. Qiu, D. Gannon, "Towards a Collective Layer in the Big Data Stack," submitted to the 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing conference (CCGrid) 2014.

R6. J. Qiu and B. Zhang, "Mammoth Data in the Cloud: Clustering Social Images," Technical report, January 2013.

R7. B. Zhang and J. Qiu, "Clustering Social Images with MapReduce and High Performance Collective Communication," Technical report, January 2013.

R8. X. Gao and J. Qiu, "Scalable Inverted Indexing on NoSQL Table Storage," Technical report, January, 2013

R9. S.-H. Bae, J. Qiu, G. Fox, "Visualization of Large High-Dimensional Data via Interpolation Approach of Multidimensional Scaling," Technical Report, July 14, 2012.

R10. J.Klinginsmith, J. Qiu "Using Cloud Computing for Scalable, Reproducible Experimentation," Technical report, August 22, 2012.

R11. T. Gunarathne, X. Gao and J. Qiu, "Twister4Azure: Data Analytics in the Cloud," Technical report, December 31, 2011.

R12. H. Li, R. Yang, Y. Zhou, J. Qiu, "Draft Report: DRYADLINQ CTP EVALUATION," Technical report, December 13, 2011.

R13. J.Ekanayake, T. Gunarathne, J. Qiu, G. Fox, S. Beason, J.Y. Choi, Y. Ruan, S.-H. Bae, H. Li, "Draft Report: Applicability of DryadLINQ to Scientific Applications," Technical report, October 16, 2009.

R14. X. Qiu and G. Fox, "Parallel Data Mining for Medical Informatics," Technical report, January 13, 2009.

R15. X. Qiu, G. Fox and A. Ho, "Analysis of Concurrency and Coordination Runtime CCR and DSS for Parallel and Distributed Computing," Technical report, January 21, 2007.

Media Coverage

Harp, Harp-DAAL, Twister and Twister4Azure are our research projects that support large-scale iterative computations for HPC and Cloud.

- Harp-DAAL Raises Hadoop Platform to Knights Landing Height, an article on The Next Platform, August 15, 2017.

- Twister4Azure is first project listed in Cloud Research Projects of Microsoft Research, 2012.

- Twister is listed in Microsoft's Research Project <u>Daytona</u> (released July 2011).
- IU "Twister" software improves Google's MapReduce for large-scale scientific data analysis, <u>IU News Room</u>, Released March 5, 2010.

Invited Talks

- SpGEMM for Vector Engines and Graph Applications, NEC Booth of SC'19, Denver, Colorado. Nov. 19th, 2019.
- Real-Time Anomaly Detection from Edge to HPC-Cloud, Big Data and Extreme-Scale Computing 2 Workshop (BDEC2), Bloomington, Indiana. November 28-30, 2018.
- High-Performance AI: A View from Systems and Frameworks, Our Deep Learning on Supercomputers workshop of SC18, Dallas, Texas. November 16, 2018.
- Real-Time Anomaly Detection from Edge to HPC-Cloud, Workshop on Clusters and Computational Data for Scientific Computing (CCDSC), Lyon, France. September 4-7, 2018
- High-Performance Big Data Computing with Harp-DAAL, HPC 2018 Workshop, July 2-6, 2018, Cetraro, Italy.
- SC17 Invited Talk by Judy Qiu on "Harp-DAAL: A Next Generation Platform for High Performance Machine Learning on HPC-Cloud", November 15, 2017, Denver, Colorado.
- Convergence of HPC and Clouds for Large-Scale Data Enabled Science, July 26-31, 2016, Cetraro, Italy.
- Convergence of HPC and Clouds for Large-Scale Data Enabled Science, UIUC/NCSA, March 13, 2016.
- Date-enabled Science and Engineering: Scalable High Performance Data Analytics, CS Colloquium at Indiana University, August 28, 2015
- Towards HPC-ABDS: An Initial High-Performance Big Data Stack, Seminar talk at Virginia Tech, July 9, 2015
- Harp: Collective Communication on Hadoop, invited talk at Computer and Information Science Department Seminar, IUPUI, November 7, 2014.
- Harp: Collective Communication on Hadoop. Invited talk at High Performance Computing – From Clouds and Big Data to Exascale and Beyond Workshop (HPC 2014), July 7~11, 2014, Cetraro, Italy.
- Supporting End-to-End Social Media Data Analysis with the IndexedHBase Platform, invited talk at 6th Workshop on Many-Task Computing on Clouds, Grids, and Supercomputers (MTAGS) 2013, November 17, 2013 at SC13.
- Generalizing MapReduce as a Unified Cloud and HPC Runtime, ECE Colloquium at RUTGERS University, October 10, 2013.
- Clustering Social Images in the Cloud, invited talk, Trends in High Performance Distributed Computing Workshop (Trends-HPDC-2013), Rutgers University, March 12, 2013.
- Portable Data Mining on Azure and HPC Platforms, invited talk (Cloud Challenge Award), 5th Workshop on Many-Task Computing on Grids and Supercomputers (MTAGS) of SC12, Salt Lake City, November 12, 2012.
- High Performance Computing, GRIDS and clouds, invited talk, An International Advanced Workshop (HPC2012), Cetraro – Italy, June 25 – 29, 2012.
- Twister4Azure: Parallel Data Analytics on Azure, invited talk, Microsoft Research Cloud Future Workshop 2012, UC Berkeley, California, May 7-8, 2012.
- Iterative MapReduce enabling HPC-Cloud Interoperability, invited talk, Department of Computer Science, Illinois Institute of Technology, Chicago, Illinois, November 4, 2011.
- Generalizing MapReduce as a Unified Cloud and HPC Runtime, invited talk, Petascale Data Analytics Workshop (PDAC-11) at The International Conference for High Performance Computing, Networking, Storage and Analysis (SuperComputing2011), Seattle, WA, November 12-18, 2011.
- Beyond MapReduce: Twister enabling HPC-Cloud Interoperability, invited talk, Tsinghua University, Beijing, China,

September 30, 2011.

- Iterative MapReduce enabling HPC-Cloud Interoperability, invited poster presentation, NSF-NSFC U.S.-China Software Workshop, Peking University, Beijing, China, September 26-29, 2011.

- Data Analysis for Scientific Applications in Cloud, invited talk in Broadband Technologies and Applications, Cyber Security, Cloud Computing, and Soc (System-on-a-Chip) Workshop (C4I and SoC) of The 11 Emerging Information and Technology Conference (EITC-2011), University of Chicago, Illinois, July 28-29, 2011.

- Hybrid Cloud and Cluster Computing Paradigms for Scalable Data Intensive Applications, invited talk for Computer Science Department Colloquium, University of Alabama, April 15, 2011.

- Applying Twister for Scientific Applications, invited talk in NSF Cloud PI Workshop, Waterview Conference Center, Arlington, Virginia, March 17, 2011.

- Analyzing large-scale cheminformatics and chemogenomics datasets through dimension reduction, invited talk in the workshop of Charting Chemical Space: Challenges and Opportunities for AI and Machine Learning at Neural Information Processing Systems Foundation Conference (NIPS 2010), Vancouver, B.C., Canada, December 10-11, 2010.

- Performance of MapReduce on Multicore Clusters, invited talk in the 2nd Workshop on "Frontiers of Multicore Computing" (FMC II) at University of Maryland, Baltimore County. September 22-23, 2010.

- Cloud Technologies and Their Applications, keynote talk in the 5th International Workshop on Content Delivery Networks (CDN) at the 10th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing Conference (CCGrid), Melbourne, Australia. May 17, 2010.

- Using MapReduce Technologies in Bioinformatics and Medical Informatics, invited talk in Using clouds for parallel computations in systems biology Workshop (part of DoE Knowledgebase in Systems Biology) at SuperComputing 09 Conference. Portland, Oregon. November 16, 2009.

- Computational Methods for Large Scale DNA Data Analysis, invited talk in Microsoft eScience Workshop. Gates Center, Carnegie Mellon University. Pittsburgh, PA. October 16, 2009.

- High Performance Biomedical Applications Using Cloud Technologies, invited talk in HPC and Grid Computing in the Cloud Workshop (Summit 09 of Open Grid Forum 27). Banff, Canada. October 13, 2009.

- Data Intensive Biomedical Computing System, invited talk in Statewide IT Conference 2009. Indianapolis, Indiana.

- September 31 and October 1, 2009.

- Large Scale DNA Sequence Analysis and Biomedical Computing using MapReduce, MPI and Threading, invited talk in Enabling Data-Intensive Computing: from Systems to Applications (NSF Data Intensive Computing Workshop). University of Pittsburgh, Pittsburgh, PA. July 30-31, 2009.

TEACHING

Graduate level courses

- Fall 2019, EGNR 599/CSCI-B649 **High Performance Big Data Systems Course** (size: 6)
- Spring 2019, EGNR 599/CSCI-B649 **High Performance Big Data Systems Course** (size: 21)
- Fall 2018, EGNR 599 **High Performance Big Data Systems Course** (size: 25)
- Spring 2017, CSCI-B649/EGNR 599 **Cloud Computing Residential/Online Course** (size: 102)
- Fall 2016, CSCI-B534/EGNR 599 **Distributed Systems** (size: 44)
- Spring 2016, CSCI-B649 **Cloud Computing Residential/Online Course** (size: 137)
- Spring 2015, CSCI-B649 **Cloud Computing Residential/Online Course** (size: 89)

- Spring 2014, CSCI-B649 **Cloud Computing Residential/Online Course** (size: 73)

- Fall 2012, CSCI-B534 **Distributed Systems** (size: 34)

- Spring 2012, CSCI-B534 **Distributed Systems** (size: 50)

- Fall 2011, CSCI-B649 **Topics on Systems: Cloud Computing for Data Intensive Sciences** (size: 30)

- Spring 2011, CSCI-B534/CSCI-B490 **Distributed Systems** (size: 58)

- Fall 2010, CSCI-B649 **Topics on Systems: Cloud Computing for Data Intensive Sciences** (size: 24)

Undergraduate level courses

- Fall 2015, CSCI-P434 **Distributed Systems** (size: 21)

- Fall 2014, CSCI-P434 **Distributed Systems** (size: 43)

- Fall 2013, CSCI-P434 **Distributed Systems** (size: 28)

- Spring 2013, CSCI-C343 **Data Structures** (size: 60)

STUDENT SUPERVISION

I link my research to education with new courses Cloud computing and Distributed Systems. I have taught over graduate and undergraduate students in the School of Informatics, Computing and Engineering. I have been an advisor to 8 Ph.D. students, two of whom graduated, and served on the thesis committee of 30 Ph.D. students. I supervised over 50 Ph.D. and Master's students for independent study course. I am committed to involving students from under-represented communities and in my research. I have hosted 29 HBCU STEM students in my lab. In addition, I supervised 12 IU undergraduate students for research. I organized the weeklong NCSA Big Data for Science workshop 2010, which was offered to over 300 graduate students from 10 universities across the nation.

As advisor of PhD students

- Jiang Miao

- Selahattin Akkas

- Jiayu Li

- Meng Li

- Yiming Zou

- Bingjing Zhang (graduated)

- Xiaoming Gao (graduated)

As member of PhD research committees

- Lijiang Guo (2019)

- Udayanga Wickramasinghe (2018)

- Qian Luo (2018)

- Jiang Miao (2018)

- Pulasthi Supun Wickramasinghe (2017)

- Kushal Keshavamurthy Raviprakash (2017)

- Zong Peng (2017)

- Luyi Xing (2017)
- Lei Wang (2016)
- Quan Zhou(2016)
- Jesun Saharia Firoz (2015)
- Saliya Ekanayake (2015)
- Daniel Kogler (2015)
- Yongan Zhao (2015)
- Andrew Younge (2014)
- Yuan Luo (2014)
- Yangyi Chen (2014)
- Isuru Suriarachchi (2014)
- Milinda Pathirage (2014)
- Miao Zhang (2013)
- Supun Madhushanka Kamburugamuve (2013)
- Sumayah Alrwais (2013)
- Jerome Mitchell (2013)
- Zhenghua Guo (2012)
- Jiaan Zeng (2012)
- Hui Li (2012)
- Yuan Luo (2012)
- Thilina Gunarathne (2011)
- Tak-Lon Wu (2011)
- Yuan Yang (2011)

PhD, Master's and undergraduate students Qiu supervised via independent study courses and research internship

- 2019

- Jiayu Li (PhD)
- Miao Jian (PhD)
- Selahattin Akkas (PhD)
- Sahaj Sigh Maini (MS)
- Yining Wang (undergraduate, IU)
- La'Andrea Gates (undergraduate, Mississippi Valley State University)
- TaJuan Beckworth (undergraduate, Mississippi Valley State University)

- 2018

- Sahil Tyagi (PhD)
- Selahattin Akkas (PhD)
- Yining Wang (undergraduate, IU)
- Tiana Deckard (undergraduate, IU)
- Ravi Teja (undergraduate, IIT)

- 2017

- Sabra Ossen (PhD)
- Feng Bo (PhD)
- Yiming Zou (PhD)
- Meng Li (PhD)
- Mihai Avram (MS)
- Anchal Khandelwal (MS)
- Gowtham A R (MS)
- Yining Wang (undergraduate, IU)
- Anurag Sharma (undergraduate, IIT)
- Mayank Jindal (undergraduate, IIT)
- Prawal Gangwar (undergraduate, IIT)
- Ravi Teja (undergraduate, IIT)
- Manpreet Gulia (undergraduate, IIT)

- 2016

- Yiming Zou (PhD) Meng Li (PhD)
- Mayank Jaglan (MS)
- Yining Wang (undergraduate)

- 2015

- Mohsen Sayyadiharikandeh (PhD)
- Yiming Zou (PhD) Meng Li (PhD)
- Tessa Pilachowski (PhD) Rohit Ingle (MS)
- Zhenghao Gu (MS)
- Yining Wang (undergraduate)

- 2014

- Tessa Pilachowski (PhD)
- Abhilash Peddy Koppula (MS) Zhenghao Gu (MS) Tirtha Bhattacharjee (MS of Virginia Tech)

- 2013

- Fei Teng (PhD) Stephen Wu (PhD) Xiaoming Gao (PhD) Bingjing Zhang (PhD) Supun Madhushanka Kamburugamuve (PhD)
- Pavan Venkatramanachar (MS)
- Clayton Sheets (undergraduate) Jake Pusateri (undergraduate) Hayley Plageman (undergraduate)
- Steven Kutkoski (undergraduate) Fabricio Caixeta Andriani (undergraduate) Evan Roth (undergraduate)

- 2012
- Jonathan Klinginsmith (PhD) Fei Teng (PhD)
- Stephen Wu (PhD) Xiaoming Gao (PhD)
- Jerome Mitchell (PhD) Yuduo Zhou (PhD)
- Nabeel Ahamed Akheel (MS), Sandip Nandi (MS), Yuan Gao (MS), Vincy Joy (MS)

- 2011
- Xiaoming Gao (PhD), Jerome Mitchell (PhD), Yang Ruan (PhD)
- Fei Teng (PhD), Tak Lon Wu (PhD), Yuduo Zhou (PhD)

- 2010
- Madhav Remesh (MS), Yogendra Ghatpande (MS), Gagan Deep Arora (MS), Janhavi Virkar (MS), Geming Chen (MS)
- Jeremy Kasting (undergraduate), Zachary Adda (undergraduate)

27 Undergraduate students Qiu hosted via Indiana University HBCU STEM Summer Scholar Institute

- 2019
- La'Andrea Gates, Bachelor of Science in Computer Science, Mississippi Valley State University
- TaJuan Beckworth, Bachelor of Science in Computer Science, Mississippi Valley State University

- 2013
- Fabricio Caixeta Andriani
- Evan Roth

- 2012
- Brandi Smith
- Joseph Jackson

- 2011
- Chartese Jones, Franshetta Hibbler, Michael Austin, Jean Bevins, JerNettie Burney, Robyn Evans, Joyce Bevins, Autumn Luke, Lakesha Wells, Joseph Jackson

- 2010
- Michael Austin, Jean Bevins, Joyce Bevins, JerNettie Burney, Nadirah Cogbill, Robyn Evans, Keenan Black, Constance

Williams, Cornelius Myles, Cortnie Wright

- 2009

- Cedric Carter, Nadia Jones, Timothy Watson

UNIVERSITY SERVICE

2019   HPC Faculty Advisory Committee

2019   Selection Committee for IUB Outstanding Junior Faculty Awards

2019   ISE Colloquium Series Committee, School of Informatics, Computing and Engineering

2019   Data Science Curriculum Committee, School of Informatics, Computing and Engineering

2018   ISE Colloquium Series Committee, School of Informatics, Computing and Engineering

2018   ISE Curriculum Committee, School of Informatics, Computing and Engineering

2017   Data Science Curriculum Committee, School of Informatics and Computing

2017   School Curriculum Committee, School of Informatics and Computing

2017   ISE Colloquium Series Committee, Intelligent Systems Engineering Department, SOIC

2016   Data Science Curriculum Committee, School of Informatics and Computing

2015   Undergraduate Research Committee, School of Informatics and Computing

2015   Data Science Curriculum Committee, School of Informatics and Computing

2015   Computer Science Division of Admission and Award Committee, School of Informatics and Computing

2014   Data Science Curriculum Committee, School of Informatics and Computing

2014   Computer Science Division of Admission and Award Committee, School of Informatics and Computing

2014   Computer Science Division of Website (card sort) Design, School of Informatics and Computing

2014   Graduate Education Committee, Computer Science Department, School of Informatics and Computing

2013   Graduate Education Committee, Computer Science Department, School of Informatics and Computing

2012   Review Panelist, Faculty Research Support Program (FRSP), Indiana University

2012   Graduate Education Committee, Computer Science Department, School of Informatics and Computing

2012   Search Committee for Academic advisor, School of Informatics and Computing

2012   Undergraduate Scholarship Committee, School of Informatics and Computing

2011   Faculty Search Committee, Computer Science Department, School of Informatics and Computing

2010   Faculty Annual Review Committee (FAR), Computer Science Department, School of Informatics and Computing

- REFEREEING ACTIVITIES

Journal of Transactions on Parallel and Distributed Systems (TPDS), 2011, 2012, 2013, 2016

Journal of Future Generation Computer Systems (FGCS), 2011-2016

Computing in Science and Engineering Magazine (CISE), 2012, 2016

IEEE Transactions on Cloud Computing, 2015

Journal of ACM Computing Surveys, 2015

Journal of Clustering Computing (CLUS), 2014

Journal of Parallel and Distributed Computing (JPDC), 2013

Journal of Network and Systems Management (JNSM), 2013

Journal of Parallel and Distributed Database (DAPD), 2012

Journal of Concurrency & Computation: Practice & Experience (CCPE), 2012

IEEE Communications Magazine, Special Issue on QoS Architectures for Wireless Networks, 2011

RECOMMENDATION LETTERS I WROTE FOR STUDENTS

- Neha Maruti Nayak (Grace Hopper Scholarship), 2020

- Chathura Widanage (MS program of Intelligent System Engineering, Indian University), 2019

- Saliya Ekanayake (Postdoc, Lawrence Berkeley National Lab), 2018

- Mihai Avram (PhD program of IU, UIUC, Notre Dame, Chicago), 2017

- Sai Sri Vidya Chitta (Grace Hopper Scholarship), 2017

- Priyanka Cherukuri (Grace Hopper Scholarship), 2017

- Prateek Srivastava (PhD program of Intelligent Systems Engineering, Indiana University, Information Systems at Northeastern University), 2017

- Brahmendra Sravan Kumar Patibandla, (Data Science for Social Good program hosted by the University of Chicago), 2017

- Bingjing Zhang (Research Scientist position at Petuum Inc.), 2017

- Yueqi Tan (Electrical and Computer Engineering graduate program at Carnegie Mellon University), 2017

- Hsi-Yun(Karen) Cheng (Reference letter for Permanent Resident Card), 2017

- Vandana Kolli (Grace Hopper Scholarship), 2016

- Shwethambari Surendran (Grace Hopper Scholarship), 2016

- Akash Ram Gopal (reference check for Mathworks Inc.), 2016

- Zachary Sullivan (PhD program of Computer Science, IU and University of Oxford, CIS program in University of Oregon), 2015

- Patrick McChesney (PhD Minor representative of IU Physics Department), 2015

- Zachary King (MS program of Computer Science, Indiana University, MS program of ECE at IUPUI), 2015

- Mrinmoy Maity (Summer Fellowship of Data Science for Social Good at University of Chicago, Microsoft Research), 2015

- Supun Kamburugamuva (Microsoft Research Internship), 2015

- Seung-Hee Bae (faculty of Western Michigan University), 2015

- Ryan Chibana (MS program of Computer Science, Georgia Institute of Technology, Purdue University, Indiana University), 2014

- Andrew Sarkisian (MS program of Computer Science, Indiana University), 2014

- Qiuwei Shou (MS program of Computer Science, Indiana University), 2014

- Seung-Hee Bae (academic jobs online, Assistant Professor of University of Maryland-Baltimore County), 2014

- Jiaan Zeng (Yelp), 2014

- Bingjing Zhang (Internship at Microsoft), 2014

- Anesu Chaora (Data Science Program, School of Informatics and Computing, Indiana University), 2014

- Bina Bhaskar (MBA Program at Texas A&M University and Michigan State University), 2014

- Xiaoyang Chen (CS PhD program at University of Pittsburgh and University of Utah), 2014

- Chengzhi Wang (M.S. Administrative Studies at Boston University, MS Information Systems at the University of Washington, Teachers College Columbia University, Engineering Management program at John's Hopkins University), 2014

- Ila Jogaikar (IU experience letter for PERM card), 2013

- Jiaan Zeng (internship at Microsoft), 2013

- Fei Teng (internship at Microsoft), 2013

- Adarsh Yoga (CS PhD program at Rice University, Harvard, UC Riverside, University of Pittsburgh, University of Arizona, University of Virginia, Stony Brook University, Rutgers University, University of Rochester, Ohio State University, Indiana University), 2013, 2012, 2011

- Gordon Moon (CS PhD program at University of Washington, University of Minnesota, University of Maryland, Ohio State University, Boston University, University of Michigan, Yale University, Northwestern University, UC Santa Cruz, Rutgers University), 2013, 2012

- Hemanth Gokavarapu (reference check for VMware Inc.), 2012

- Kang Zhao (CS PhD program at University of Washington, Purdue University), 2012

- Alhanoof Althnian (CS PhD program at University of Southern California, Duke University, Penn State, University of Arizona), 2012

- Joyce Bevins (REU student applying for MS program at IU), 2012

- JerNettie Burney (REU student applying for MS program at IU), 2012

- Seung-Hee Bae (research associate at University of Washington), 2012

- Stephen Wu (Macau postgraduate scholarship), 2012, 2011, 2010

- Jong You Choi (postdoc fellowship program CIFP), 2011

- Jiang Wu (CS PhD program at IU), 2011

- Xiaoming Gao (internship at Microsoft), 2011

- Yuan Luo (internship at Microsoft), 2011

- Santhosh Kumar (internship at IU), 2011

- Pairoj Rattadilok (internship at IU), 2011

- Hemanth Gokavarapu (internship at IU), 2011

- Dhairya Gala (internship at IU), 2011

- Sonali Karwa (internship at IU), 2011

- Anand Mukundan (internship at IU), 2011