Ye Yu

PhD Candidate (ABD)
Department of Computer Science
University of Kentucky
Lexington, KY, 40506-0495, USA

Cell phone: (859)559-2918
Email: ye.yu@uky.edu
Homepage: http://cs.uky.edu/∼yyu233/

Research Interests

Cloud Computing and Data Centers, Data Science and Bioinformatics, Large Data Storage and High-speed Queries, Computer Networks, Distributed Systems, and Internet of Things

Education

08/2013 - 06/2018 **University of Kentucky**, Ph.D in Computer Science, GPA 4.0/4.0
Advisor: **Chen Qian** (Now faculty at University of California Santa Cruz)
Dissertation: *Ultra-Fast and Memory-Efficient Lookups for Cloud, Massive Data Management, and High-Speed Networks*
Best Computer Sciece Ph.D. student award recipient

08/2009 - 06/2013 **Beihang University**, B.E. in Computer Science, GPA 3.8/4.0
Outstanding graduate of the Honors College of Beihang University

Referred Journal Publications

Remark: *Bioinformatics* is the leading top journal on bioinformatics, *ToN* is the leading top journal on networking, and *TPDS* is the leading top journal on distributed systems.

1. Memory-efficient and Ultra-fast Network Lookup and Forwarding using Othello Hashing
   Ye Yu, Djamal Belazzougui, Chen Qian, and Qin Zhang
   accepted, to appear in **IEEE/ACM Transactions on Networking (ToN)**, 2018.
2. A Novel Data Structure to Support Ultra-fast Taxonomic Classification of Metagenomic Sequences with k-mer Signatures
   Ye Yu*, Xinan Liu* (Co-first authorship), James N. MacLeod, Chen Qian and Jinze Liu
   in **Bioinformatics**, Oxford Academic, 2017. (**Impact Factor 7.307**)
3. Practical Network-wide Packet Behavior Identification by AP Classifier
   Huazhe Wang, Chen Qian, Ye Yu, Hongkun Yang, and Simon S. Lam
   in **IEEE/ACM Transactions on Networking (ToN)**, 2017.
4. Space Shuffle: A Scalable, Flexible, and High-Bandwidth Data Center Network
   Ye Yu and Chen Qian
   in **IEEE Transactions on Parallel and Distributed Systems (TPDS)**, 2016.
5. DiFS: Distributed Flow Scheduling for adaptive switching in FatTree data center networks
   Wenzhi Cui, Ye Yu, and Chen Qian
   In **Computer Networks**, 2016.

Referred Conference and Workshop Publications (with at least three reviewers)

Remark: *ICNP*, *CoNEXT*, and *SIGMETRICS* are all single-session top conferences on computer systems and networking (including cloud, data center networks, and distributed systems), *CloudCom* is one of the good conferences on cloud computing research, *IoTDI* is a relatively new top conference in IoT research.

1. A Concise Forwarding Information Base for Scalable and Fast Name Switching
   Ye Yu, Djamal Belazzougui, Chen Qian, and Qin Zhang
   in **IEEE International Conference on Network Protocols** (**ICNP**), 2017.
   Acceptance Rate: 18%.
2. A Fast, Small, and Dynamic Forwarding Information Base
   Ye Yu, Djamal Belazzougui, Chen Qian, and Qin Zhang

in **ACM SIGMETRICS**, 2017. Extended abstract.
Acceptance Rate: 17%.

3. SDLB: A Scalable and Dynamic Software Load Balancer for Fog and Mobile Edge Computing
Ye Yu, Xin Li, and Chen Qian
in **ACM SIGCOMM Workshop on Mobile Edge Communications** (**MECOMM**), 2017
4. A Novel Data Structure to Support Ultra-fast Taxonomic Classification of Metagenomic Sequences with k-mer Signatures
Ye Yu*, Xinan Liu*(Co-first authorship), James N. MacLeod, Chen Qian and Jinze Liu
in **RECOMB Workshop on Massively Parallel Sequencing** (**RECOMB-Seq**), 2017.
5. An IoT Data Communication Framework for Authenticity and Integrity
Xin Li, Huazhe Wang, Ye Yu, and Chen Qian
in **ACM/IEEE International Conference on Internet-of-Things Design and Implementation** (**IoTDI**), 2017.
Acceptance Rate: 29%.
6. Failure-Resilient Routing for Server-Centric Data Center Networks with Random Topologies
Ye Yu and Chen Qian
in **IEEE International Conference on Cloud Computing Technology and Science** (**CloudCom**), 2016.
Acceptance Rate: 26%.
7. Garlic Cast: Lightweight and Decentralized Anonymous Content Sharing
Chen Qian, Junjie Shi, Zihao Yu, Ye Yu, and Sheng Zhong
in **IEEE International Conference on Parallel and Distributed Systems (ICPADS)**, 2016
8. FTDC: A Fault-Tolerant Server-Centric Data Center Network
Ye Yu and Chen Qian
in **IEEE/ACM International Symposium on Quality of Service** (**IWQoS**), Poster, 2016.
9. Practical Network-wide Packet Behavior Identification by AP Classifier
Huazhe Wang, Chen Qian, Ye Yu, Hongkun Yang, and Simon S. Lam
in **ACM Conference on emerging Networking Experiments and Technologies** (**CoNEXT**), 2015.
Acceptance Rate: 21%.
10. Distributed Collaborative Monitoring in Software Defined Networks
Ye Yu, Chen Qian, and Xin Li
in **ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking** (**HotSDN**), 2014.
Acceptance rate: 28.9%
11. Space Shuffle: A Scalable, Flexible, and High-Bandwidth Data Center Network
Ye Yu and Chen Qian
in **IEEE International Conference on Network Protocols** (**ICNP**), 2014.
Acceptance Rate: 20%.
12. NetCP: Multi-level Checkpointing and Rollback of Software Defined Non-interrupted Networks
Ye Yu, Chen Qian, Ying Zhang, Ravi Manghirmalani
in **IEEE International Conference on Network Protocols** (**ICNP**), Poster, 2014.

Publications in submission

Enabling deBruijn-informed online sequence query against thousands of RNA-seq datasets
Ye Yu, Jinpeng Liu, Xinan Liu, Eamonn Magner, Chen Qian, and Jinze Liu
under review by **Genome Biology**

Concury: Consistent, Fast, and Scalable Layer-4 Load Balancing
Shouqian Shi, Ye Yu, Ying Zhang, Xin Li, and Chen Qian

submitted to **flagship annual conference of the ACM Special Interest Group on Data Communication (SIGCOMM), 2018** .

A Secure and Efficient Communication Framework for Internet of Things Xin Li, Minmei Wang, Huazhe Wang, Ye Yu, and Chen Qian
submitted to **IEEE Transactions on Dependable and Secure Computing (TDSC)**

Distributed Collaborative Monitoring in Software Defined Networks
Ye Yu, Xin Li, and Chen Qian
submitted to **Computer Networks**

SICS: Secure and Dynamic Middlebox Outsourcing
Huazhe Wang, Xin Li, Ye Yu, Yu Zhao, Hongkun Yang, Chen Qian, Yang Wang, and Jinsong Han
conference submission

NetCP: Multi-level Checkpointing and Rollback of SDN
Ye Yu, Ying Zhang, Wenfei Wu, and Chen Qian
conference Submission

Memory Chord: A Scalable, Flexible Memory Network Architecture
Matheus Ogleari, Ye Yu, Chen Qian, Ethan Miller, and Jishen Zhao
submitted to **International Symposium on Computer Architecture (ISCA), 2018**

Proposal Writing Participation

"CRII: NeTS: A Coordinate System based Network for Data Plane Scalability and High Throughput", 2015-2018, $175,000 funded by National Science Foundation. (PI Chen Qian)

"CAREER: Othello Hashing and Its Applications to Scalable and Dynamic Network Forwarding and Functions", $500,000 pending by National Science Foundation. (PI Chen Qian)

"BigData: IA: Collaborative Research: Finding Needles in a Haystack: A Cloud-based Query System to Empower Search in Large Scale Sequencing Data", $800,000 pending by National Science Foundation. (PIs Jinze Liu and Chen Qian)

Honors and Awards

- Best Ph.D. Student, Department of Computer Sciece, University of Kentucky, 2018.
- Student Travel Awards: IEEE ICNP 2014, 2017. IEEE CloudCom, 2016. ACM SIGCOMM, 2017. ACM SIGMETRICS 2017.
- MSRA Undergraduate Scholarship, Microsoft Research Asia, 2012.
- Scholarship by Qian Changzhao and Shen Xingyuan, Beihang University, 2011.
- Scholarship for disciplinary contest winners, Beihang Univeristy, 2010.
- Scholarship for outstanding new students, Beihang University, 2009.
- Silver Medal in the ACM International Collegiate Programming Contest (ICPC) Asia Regional Contest, Tianjin, China, 2010.
- First place prize in the Olympiad in Informatics of Shandong Province, Jinan, China, 2008.
- Second prize in the Shandong Province Middle School Mathematics Olympics, Jinan, China, 2007.

Research Experience

Research Assistant Aug. 2013 - present
University of Kentucky Advisor: Chen Qian. Lexington, KY

- **Cloud, Distributed Systems, and Networking research theme**

Fast and Scalable Forwarding Information Base (Concise)
Designed and implemented a unique data structure (Othello) that supports efficient classification queries. Othello uses significantly less (10% $\sim$ 30) memory to achieve 2x $\sim$ 4x higher query speed compared to other Forwarding Information Base (FIB) designs. Implemented and evaluated the fast FIB prototype on Click Modular Router and Intel DPDK.

Software Load Balancer
Designed and implemented a software load balancer with Othello Hashing techniques named SDLB, which enables the load balancer to perform the functionalities for stateful and stateless flows simultaneously using one query on the table. Evaluation results show that the system is faster by 4x to 10x and uses much less ($< 50\%$) memory than the current widely-used load balancer designs.

Distributed Collaborative Monitoring (DCM):
Proposed a scalable software-defined measurement solution for SDN, namely DCM. DCM distributes and balances the measurement load among switches in the network. DCM is able to perform fine-grained per-flow measurement. It costs very little memory by using the novel design of two-stage Bloom filters. DCM achieves better measurement accuracy than existing solutions given the same memory space on switches.

Data Center Networks with Random Topologies (SpaceShuffle and FTDC):
Designed a flexible data center network architecture (SpaceShuffle) that achieves high-bandwidth, flexibility and scalability. Designed the greedy routing protocol. This is the first greedy routing protocol to enable high-throughput multi-path routing on random topologies. Designed, implemented, and evaluated the fault-tolerant routing protocol for random topologies.

•Bioinformatics and Big Data research theme Supervisor: Jinze Liu.
Metagenomic Sequencing Reads Classification
Designed and implemented the algorithm for a hashing classifier for metagenomics sequencing reads (RNA sequences). This algorithm is an order-of-magnitude faster than other state-of-art metagenomics classification algorithms.

SeqOthello Sequence Search Engine
Designed and implemented a system that supports online query of transcripts from thousands of RNA-seq datasets. For the sequence containment query, SeqOthello achieves a two orders of magnitude improvement over previous approaches while retaining comparable accuracy. Furthermore, it is the first approach capable of answering the sequence coverage query and retrieving the relative expression levels of the query transcript. SeqOthello achieves a 700:1 compression ratio relative to the original sequencing data. We built the first sequence search index constructed on the scale of TCGA data (54TB in compressed fastq format).

Software Development Engineer Intern May. 2017 - Aug. 2017
Facebook, Inc. Menlo Park, CA
Implemented the new framework for load-balancing and bandwidth allocation for Facebook's internal inter-data center backbone SDN. Designed and implemented internal tools for SDN management. Worked on BGP protocols, SDN controller programs, etc.

Research Intern Jun. 2014 - Aug. 2014
Silicon Valley Lab, Ericsson Research Host: Ying Zhang. San Jose, CA
Checkpointing and Rollback in SDN (NetCP):
Designed a multi-level checkpointing/rollback framework for SDN systems. Designed and implemented the controller checkpointing/rollback mechanisms for OpenDaylight and POX, and the OpenFlow protocol extension on OpenVSwitch for NetCP data plane. NetCP significantly improves the network performance during failure recovery.

Research Assistant 2011–2013
Beihang University Beijing, China
Research Topics: Max clique problem on model RB and combinational optimization problems.
Designed and implemented an Android application that collects data on how human solves NP-Hard combinational problems. Discovered a heuristic function for the max clique problem using the user-generated data.

Teaching and Advising Experience

Programming Competition Coaching/Organizing

- Assistant Coach, University of Kentucky ACM/ICPC Team, 2017.
- Organizing member, the University of Kentucky Student Programming Contests, 2014 - 2017.
- Organizing member, Beijing High School Programming Contest, Beijing, China, 2012.

Teaching Assistant

- CS101 - Introduction to Computers, University of Kentucky, Spring 2015, Fall 2016.
- the C programming language, Beihang University, Fall 2013.

Guest Lecturer

- CS 315 - Algorithm Design and Analysis. University of Kentucky, Multiple semesters, 2014 - 2016.

Student Advising. Helped my advisors to advise other students in the group:

- Wenzhi Cui, now PhD student at University of Texas at Austin, co-authored one paper.
- Huazhe Wang, now PhD student at University of California Santa Cruz, co-authored two papers.
- Eamonn Magner, now Master student at University of Kentucky, co-authored one paper.

Service and Event Organizing

Student member, the Higher Degrees Committee of Computer Science, University of Kentucky, 2014 - 2015.

Demo presenter, "Face morphing with delaunay triangulation", Univeristy of Kentucky Enginering Day, Lexington, KY, 2015, 2016, and 2017

Technical volunteer, Kentucky Department of Education's Student Technology Leadership Program (STLP). Lexington, KY, 2015 and 2016.

Talks & Presentations

1. Othello Hashing in Bioinfomatics

   in Markey Cancer Center, Lexington, KY, 2017.

2. A Concise Forwarding Information Base for Scalable and Fast Name Switching,

   on IEEE ICNP conference, Toronto, ON, Canada, 2017.
   poster presentation ACM SIGMETRICS conference, Champain, IN, 2017.

3. SDLB: A Scalable and Dynamic Software Load Balancer for Fog and Mobile Edge Computing

   on ACM SIGCOMM MECOMM workshop, Los Angles, CA, 2017.

4. Distributed Collaborative Monitoring in Software Defined Networks

   on ACM ACM HotSDN workshop, Chicago, IL, 2014.

5. Failure-Resilient Routing for Server-Centric Data Center Networks with Random Topologies

   on IEEE CloudCom conference, Luxembourg, 2016.

6. FTDC: A Fault-Tolerant Server-Centric Data Center Network

   poster presentation on IEEE IWQOS, Beijing, China, 2016.

7. NetCP: Multi-level Checkpointing and Rollback of Software Defined Non-interrupted Networks

   in Ericsson Research, San Jose, CA, 2014.

poster presentation on IEEE ICNP conference, Raleigh, NC, 2014.

8. Micro Unmanned Aerial Vehicle Embedded Systems: Design & Implementation,

   National Cheng Kung University, Tainan, 2011.

   Microsoft Research Asia, Beijing, 2011.

Open Source Software Projects

Othello Hashing. `https://github.com/sdyy1990/Othello`

MetaOthello Taxonomic classifier. `https://zenodo.org/record/808941`

SeqOthello. `https://github.com/sdyy1990/SeqOthello`

S2 Network Simulator. `https://github.com/sdyy1990/S2Simulator`

Professional Activities

• Conference Reviewers:
ICCCN 2015,2016; ICPADS 2014; ICPP 2015; CoolSDN 2015; SDNNFVSEC 2016; INFOCOM 2016, 2017, 2018, Globecom 2016; ICNC 2018;
• Journal Reviewers:
IEEE/ACM International Journal of Distributed Sensor Networks, IEEE/ACM Transaction on Networking

Extracurricular Projects/Activities

The Unmanned Aerial Vehicle Project, School of Advanced Engineering, Beihang University.

- Designed the structure of the system of a commodity standard drone.
- Worked on navigation and flight-control algorithms in ARM/Linux embedded systems.
- This system won the fourth place prize of the 2012 Taiwan International Innovative UAV Design Competition.

Membership

IEEE student member
ACM student member