# Min Xu


**Contact Information**

1100 East 58th Street,
University of Chicago,
Chicago, IL, 60637,

312-874-2708
xum@cs.uchicago.edu
Webpage


**Research Interests**

My current research interest is cloud security & privacy. Previously, I have done research in distributed systems, especially storage systems.

**Education**

**University of Chicago**, Chicago, IL

Ph.D. Candidate, Computer Science — *Expected:* 2020

- Advisor: Ariel Feldman

**The Chinese University of Hong Kong**, Shatin, N.T., Hong Kong

M.Phil., Computer Science and Engineering — Aug 2015

- Thesis: *Even Data Placement for Load Balance in Distributed Storage Systems with Deduplication and Erasure Coding*
- Advisor: Patrick, P.C. Lee

B.S., Mathematics and B.Eng., Information Engineering — Jul 2013

**Work Experience**

**Research Assistant** — Sep 2015 to present
Department of Computer Science, University of Chicago — *Chicago, IL*
Supervisor: Ariel Feldman

**Research Assistant** — Jul 2013 to Jul 2015
Department of Computer Science and Engineering, The Chinese University of Hong Kong — *Hong Kong*
Supervisor: Patrick, P.C. Lee

**Summer Research Intern** — Jun 2012 to Aug 2012
Department of Computer Science and Engineering, The Chinese University of Hong Kong — *Hong Kong*
Supervisor: Patrick, P.C. Lee

**Software Engineer** — Jul 2011 to Jun 2012
WisdomOne Ltd. — *Hong Kong*

**Summer Research Intern** — Jun 2010 to Aug 2010
Oak Ridge National Laboratory — *Knoxville, TN*
Supervisor: Prof. Yvonne M.J. Ou
Partner: Mr. T.K. Mak

**Current Research**

**Secure computation outsourcing to the untrusted cloud** A lot of computations on sensitive data tend to be outsourced onto the untrusted cloud for efficiency and reliability. How to completely secure the data owners' privacy against the untrusted software stack at the cloud, however, is still an ongoing pursuit. In particular, various side channels exposed from these computations have been demonstrated to be exploitable to break the privacy guarantees of existing cloud vendors. In this work, we plan to implement expressive development APIs using existing oblivious algorithms to eliminate crucial side channels, and, furthermore, we plan to re-design the software stack at the untrusted cloud infrastructure to accelerate these oblivious algorithms without compromising their security guarantees.

**Cloud privacy and integrity against untrusted app and OS** This project aims to protect cloud users' private data from the potentially untrusted cloud applications and OS. In our threat model, the untrusted OS could eavesdrop or tamper users' private

data in memory. Furthermore, the untrusted applications have access to all the memory pages in their address space with specified privileges, and can cooperate with unstrusted OS to leak sensitive information via possible covert channels. There is a limit to how much security a cloud provider can guarantee. It would be better if users didnt have to trust them as much. Our goal is to redesign data analytics outsourced to third party cloud providers to reduce the degree to which users have to trust them. We combine trusted hardware, e.g. Intel Software Guard eXtension (SGX), and software fault isolation technique, e.g. Google Native Client (NaCl), to protect application memory from untrusted OS and to sandbox untrusted application, respectively. We also extend the NaCl sandbox to shutdown all feasible covert channels exposed to the untrusted application, including abundant syscalls, timing channel, size channel, etc.

## Past Researches

**Distributed Storage Systems with Deduplication** We designed and implemented a distributed deduplication system to tackle the read imbalance problem induced by deduplication in a distribution storage system. We formulated a combinatorial optimization problem, and proposed a greedy, polynomial-time Even Data Placement (EDP) algorithm, which identifies a data placement that effectively achieves read balance while maintaining write balance. We further extended our EDP algorithm to heterogeneous environments. Through extensive simulations and prototype testbed experiments, we showed that our EDP algorithm could reduce the file read time by 37.41% compared to the baseline round-robin placement, and the reduction could further reach 52.11% in a heterogeneous setting.

**Efficient Hybrid Inline and Out-of-line Deduplication for Backup Storage** We designed and implemented RevDedup, an efficient hybrid inline and out-of-line deduplication system for backup storage. It applies coarse-grained inline deduplication to remove duplicates of the latest backup, and then fine-grained out-of-line reverse deduplication to remove duplicates from older backups. Our reverse deduplication design limits the I/O overhead and prepares for efficient deletion of expired backups. Through extensive testbed experiments using synthetic and real-world datasets, we showed that RevDedup can bring high performance to the backup, restore, and deletion operations, while maintaining high storage efficiency comparable to conventional inline deduplication.

## Referred Publications

1. **M. Xu**, Y.F. Zhu, P.P.C. Lee, Y.L. Xu, "Even Data Placement for Load Balance in Reliable Distributed Deduplication Storage Systems." Proceedings of the IEEE/ACM International Symposium on Quality of Service (IWQoS'15) (Full paper), Portland, Oregon, USA, Jun 2015.

2. Y.K. Li, **M. Xu**, C.H. Ng, P.P.C. Lee, "Efficient Hybrid Inline and Outofline Deduplication for Backup Storage." *ACM Transactions on Storage (TOS)*, 2014.

## Talks

**Isolation and Transparency for Outsourced Applications**
Center for Unstoppable Computing (CERES) Summit 2016 Poster Session, Chicago, IL — Jan 2016

**Even Data Placement for Load Balance in Reliable Distributed Deduplication Storage Systems**
IEEE/ACM International Symposium on Quality of Service (IWQoS'15), Portland, OR — Jun 2015

## Awards

Travel Grants

- CUHK CSE Department RPg Travel Grant, Portland, OR, USA — Jun 2015

- Summer Research on Applied Mathematics, Knoxville, TN, USA — Aug 2010

Student Awards

- Yasumoto Exchange Scholarship — Aug 2010
- HKSAR Government Admission Scholarship — 2008-2011,2013

## Teaching Experience

**Teaching Assistant**

University of Chicago

CMSC23200/33250 - Introduction to Computer Security — Autumn,Winter 2015
Instructor: Ariel Feldman

The Chinese University of Hong Kong

CSCI4180 - Introduction to Cloud Computing and Storage — Fall 2012,2013
Instructor: Patrick, P.C. Lee

ENGG5015/CSCI5470 - Computer and Network Security — Spring 2013,2014
Instructor: Patrick, P.C. Lee

## Courses

University of Chicago

CMSC33550 - Introduction to Databases — Winter 2015
Instructor: Aaron J. Elmore

CMSC37000 - Algorithms — Winter 2015
Instructor: Yury Makarychev

CMSC33250 - Introduction to Computer Security — Autumn 2015
Instructor: Ariel Feldman

CMSC33100 - Advanced Operating Systems — Autumn 2015
Instructor: Shan Lu

The Chinese University of Hong Kong

CSCI5010 - Practical Computational Geometry Algorithms — Spring 2014
Instructor: YuFei Tao

CSCI5060 - Semidefinite programming and approximation algorithms — Spring 2014
Instructor: Lap-Chi Lao

CSCI5120 - Advanced Topics in Database Systems — Fall 2013
Instructor: James Cheng

CSCI5510 - Big Data Analytics — Fall 2013
Instructor: Irwin, K.C. King and Michael, R.T. Lyu

## Skills

C/C++; Java; Python; shell scripting; Linux
Native speaker of Mandarin; Fluent in English and Cantonese.

## References

Ariel Feldman (Ph.D. advisor)
Assistant Professor — E-mail: arielfeldman@cs.uchicago.edu
Department of Computer Science
University of Chicago

Patrick P.C. Lee (M.Phil. advisor)
Assistant Professor — E-mail: pclee@cse.cuhk.edu.hk
Department of Computer Science and Engineering
The Chinse University of Hong Kong