Jinyang Li

715 Broadway Rm 708
New York, NY
10003

Phone: (212) 998-3098
jinyang@cs.nyu.edu
http://cs.nyu.edu/~jinyang

Education

MASSACHUSETTS INSTITUTE OF TECHNOLOGY — Cambridge, MA

Ph.D., Computer Science, 2006
Thesis: Tradeoffs of routing in dynamic peer-to-peer networks
Minor in Applied Mathematics

S.M. in Computer Science, 2001
Thesis: A scalable location service for geographic ad hoc routing

NATIONAL UNIVERSITY OF SINGAPORE — Singapore

B.S. (first class honors) in Computer Science, 1998

Research Interests

Distributed systems and computer networks

Professional Experience

2006-present — **Assistant Professor** — New York University Department of Computer Science

Current projects include cooperative distributed storage (Friendstore), censorship circumvention (Kaleidoscope), wide area distributed file systems (WheelFS) and high throughput multi-radio wireless mesh networks.

2005-2006 — **Postdoctoral Researcher** — UC Berkeley

Designed and built a multichannel MAC protocol for mesh networks.

1999–2005 — **Research Assistant** — CSAIL, MIT

Major projects at MIT include cooperative digital library (OverCite), robust routing protocols for peer-to-peer systems (Accordion) and scalable wireless mesh networks (Grid).

6/2001-9/2001 — **Summer Intern** — International Computer Science Institute, Berkeley

Worked on understanding the characteristics of observed IP addresses on the Internet.

Teaching at New York University

Fall 2007 — Networks and Distributed Systems (G22.2620.001, graduate level).

Spring 2007 — Distributed Storage Systems (G22.3033.006, graduate level).

Fall 2006 — Networks and Distributed Systems (G22.3033.010, graduate level).

Honors and Awards

2008 — NSF CAREER Award

2004–2005 Microsoft Research Graduate Fellowship

1998 IEEE Singapore Information Technology Gold Medal

Professional Services

2008 Program Committee, First International Workshop on Social Network Systems, 2008

2007 Program Committee and Poster Co-chair, USENIX Networked Systems Design and Implementation (NSDI)

2007 Program Committee, International Workshop on Peer-to-Peer Systems (IPTPS), Feb 2007

2006 Panel Reviewer, NSF Future Internet Network Design (FIND)

2000-2005 External Reviewer, Symposium on Operating Systems Principles(SOSP), Symposium on Networked System Design and Implementation(NSDI), Symposium on Operating System Design and Implementation(OSDI), ACM SIGCOMM

Journal Publications

[1] Eddie Kohler, Jinyang Li, Vern Paxson, and Scott Shenker. Observed structure of addresses in ip traffic. *IEEE/ACM Transactions on Networking (TON)*, 16:1207–1218, December 2006.

Refereed Conference and Workshop Publications

[2] Yair Sovran, Alana Libonati, and Jinyang Li. Pass it on: Social networks stymie censors. In *Proc. of the 7th International Workshop on Peer-to-Peer Systems (IPTPS)*, Feb 2008.

[3] Jeremy Stribling, Emil Sit, M. Frans Kaashoek, Jinyang Li, and Robert Morris. Don't give up on distributed file systems. In *Proc. of the 6th International Workshop on Peer-to-Peer Systems (IPTPS)*, Feb 2007.

[4] Jeremy Stribling, Jinyang Li, Isaac G. Councill, M. Frans Kaashoek, and Robert Morris. Overcite: A distributed, cooperative citeseer. In *the 3rd Symposium on Networked Systems Design and Implementation (NSDI'06)*, May 2006.

[5] Jinyang Li and Frank Dabek. F2F: reliable storage in open networks. In *Proceedings of the 5th International Workshop on Peer-to-Peer Systems (IPTPS)*, February 2006.

[6] Jinyang Li, Jeremy Stribling, Robert Morris, M. Frans Kaashoek, and Thomer Gil. A performance vs. cost framework for evaluating DHT design tradeoffs under churn. In *Proceedings of 24th IEEE Infocom*, March 2005.

[7] Jinyang Li, Jeremy Stribling, Robert Morris, and M. Frans Kaashoek. Bandwidth efficient management of DHT routing tables. In *the 2nd Symposium on Networked Systems Design and Implementation (NSDI'05)*, May 2005.

[8] Jeremy Stribling, Isaac G. Councill, Jinyang Li, M. Frans Kaashoek, David R. Karger, Robert Morris, and Scott Shenker. OverCite: A cooperative digital research library. In *Proceedings of the 4rd International Workshop on Peer-to-Peer Systems (IPTPS'05)*, February 2005.

[9] Jinyang Li, Jeremy Stribling, Thomer Gil, Robert Morris, and M. Frans Kaashoek. Comparing the performance of distributed hash tables under churn. In *Proceedings of the 3rd International Workshop on Peer-to-Peer Systems (IPTPS'04)*, February 2004.

[10] Frank Dabek, James Robertson, Jinyang Li, Emil Sit, M. Frans Kaashoek, and Robert Morris. Designing a DHT for low latency and high throughput. In *Proceedings of the 1st Symposium on Networked Systems Design and Implementation (NSDI'04)*, 2004.

[11] Russ Cox, Frank Dabek, M. Frans Kaashoek, Jinyang Li, and Robert Morris. Practical, distributed network coordinates. In *2nd Workshop on Hot Topics in Networks (HotNets-II)*, 2004.

[12] Jinyang Li, Boon Thau Loo, Joseph M. Hellerstein, M. Frans Kaashoek, David Karger, and Robert Morris. On the feasibility of peer-to-peer web indexing and search. In *Proceedings of the 2nd International Workshop on Peer-to-Peer Systems (IPTPS'03)*, February 2003.

[13] Eddie Kohler, Jinyang Li, Vern Paxson, and Scott Shenker. Observed structure of addresses in IP traffic. In *Internet Measurement Workshop 2002*, 2002.

[14] Jinyang Li, Charles Blake, Douglas De Couto, Hu Imm Lee, and Robert Morris. Capacity of ad hoc wireless networks. In *Proceedings of the 7th ACM Mobicom*, August 2001.

[15] Robert Morris, M. Frans Kaashoek, John Jonnotti, Jinyang Li, and Douglas S. J. De Couto. CarNet: A scalable ad hoc wireless network system. In *9th ACM SIGOPS European Workshop*, 2000.

[16] Jinyang Li, John Jannotti, Douglas De Couto, David R. Karger, and Robert Morris. A scalable location service for geographic ad hoc routing. In *Proceedings of the 6th ACM Mobicom*, August 2000.

Graduate students under supervision

9/2006-8/2011 (Ph.D. expected) Dinh Nguyen Tran, New York University

9/2007-8/2012 (Ph.D. expected) Yair Sovran, New York University

9/2006-8/2011 (Ph.D. expected) Arthur Meacham, New York University

9/2006-8/2008 (M.S. expected) Alana Libonati, New York University

9/2006-12/2007 (M.S. expected) Frank Chiang, New York University

Research - Distributed systems

2006–now **Censorship circumvention** NYU

Kaleidoscope circumvents Internet censorship using volunteer proxies outside the censored domain to relay traffic to blocked websites[2]. It disseminates proxy information along a trust graph whose links reflect the real world trust relationships among users. Kaleidoscope helps users discover proxies while minimizing the chances of the censor discovering and blocking them.

2006–now **Cooperative distributed storage** NYU

Cooperative storage systems promise cheap and easy online storage. However, such systems are traditionally considered too unreliable to use because of high churn and the presence of selfish and malicious participants. Our system, Friendstore, gives users explicit control over what nodes to contribute their storage and network resources to and what nodes to entrust their replicated data with. Friendstore stores coded data to maximize storage utilization despite the constraint that each node can only backup its data on a small subset of other nodes.

2006–now **Wide area distributed file systems** NYU

WheelFS is a wide area file system for running distributed applications. WheelFS[3] gives applications explicit control over the tradeoff between performance, failure resilience and data consistency. We expect WheelFS to dramatically simplify the constructions of applications such as content distribution networks, distributed digital library and traditional Grid applications. This project is being done in collaboration with MIT.

2004–2006 **Distributed digital library** MIT

With Jeremy Stribling et. al., I am building and deploying OverCite [4][8], a distributed digital library of published research papers. OverCite aggregates the storage, computing and network resources of more than twenty donated machines distributed over the wide area network to achieve good end-to-end capacity in terms of keyword searching and paper downloading.

2003–2005 **Lookup protocols for Distributed Hash Tables** MIT

My dissertation research examines the tradeoffs of DHT lookup protocols in dynamic networks where nodes constantly join and leave. All DHTs incur communication costs to keep up with membership changes in order to route lookups efficiently. Fundamentally, we are interested in the *efficiencies* of different DHTs in their abilities to use extra communication bits for low latency lookups.

I designed PVC [6], a performance vs. cost framework and evaluation methodologies that allow designers to quantify and compare the efficiencies of different DHT design choices. With Jeremy Stribling et. al., I conducted extensive simulations of existing DHTs and used PVC to study different design choices embedded in the protocols. PVC analysis reveals that the key to efficiently use additional bandwidth is to adjust a node's routing table size according to available bandwidth and churn rate.

Based on lessons learnt from PVC study, I designed *Accordion* [7], a new DHT that adjusts its routing table size automatically according to the churn rate and user specified protocol bandwidth budget. Accordion's bandwidth efficiency approximates the best performance/cost tradeoffs of existing protocols with manually tuned best parameter values.

2002–2003 **Distributed keyword search** MIT

With Boon Thau Loo et. al., I investigated various algorithms for performing scalable keyword search in peer-to-peer networks. Specifically, we evaluated two designs of distributing the global index on different nodes: partition by document and partition by

keyword with its various optimizations [12]. We analyzed the bandwidth cost of a full-text distributed web search engine using such designs.

Research - Wireless networks

2007–now **Multi-radio mesh networks** NYU

Mesh networks that consist of nodes with multiple inexpensive radios can achieve much higher potential performance than single radio meshes. I am building a joint routing and channel assignment protocol to assign optimal channels to nodes with multiple radios. Our protocol exploits the fact that the majority of traffic in mesh networks is to access a few gateway nodes. This project is in collaboration with David Bindel, Lakshmi Subramanian.

2001 **Capacity analysis of ad hoc networks** MIT

Multi-hop ad hoc networks have surprisingly low data capacity. We performed extensive simulation studies of end-to-end throughput with different workloads and simple experiments on hardware. I also showed analytically that the capacity available to each node in ad hoc networks depends critically on the traffic pattern [14].

2000 **Grid ad hoc routing protocol** MIT

I helped design and evaluate the Grid Ad hoc Routing Protocol. Grid uses geographic forwarding and a fully distributed location service (GLS) [16] to achieve scalable ad hoc routing. GLS requires each node to store the current geographic location of only $O(\log n)$ other nodes. A GLS query to a geographically close-by destination node can be resolved without contacting far-away nodes in the networks.