Maheswaran Sathiamoorthy

CONTACT
INFORMATION

RTH 419, 3710 McClintock Ave
University of Southern California
Los Angeles, CA-90089

E-mail: msathiam at usc.edu
WWW: http://smahesh.com

RESEARCH
INTERESTS

Data Center Networks: Erasure coding techniques for distributed storage, reliable block storage.
Traditional Networks: Analysis and design of content distribution strategies.
General: Distributed Storage, Distributed Systems, Big Data, Data Science, and Social Networks.

EDUCATION

University of Southern California, Los Angeles, California, USA
Doctor of Philosophy (PhD), Electrical Engineering   **Aug 2008 – Nov 2013**
- GPA: 3.93/4
- Advisors: Prof. Bhaskar Krishnamachari & Prof. Alexandros G. Dimakis
- Thesis title: Optimizing Distributed Storage for Cloud Environments

Indian Institute of Technology (IIT), Kharagpur, West Bengal, India
B.Tech(H), Electronics and Electrical Communication Engineering   **July 2004 – May 2008**
- GPA: 9.27 out of 10.00
- Ranked 3/50 in the department and 7/650 in the Institute

PUBLICATIONS

- M. Sathiamoorthy, M. Asteris, D. Papailiopoulos, A. G. Dimakis, R. Vadali, S. Chen, D. Borthakur, "XORing Elephants: Novel Erasure Codes for Big Data", *VLDB 2013*.
  Work featured in famous storage industry blogs StorageMojo and High Scalability.
- M. Sathiamoorthy, A. G. Dimakis, B. Krishnamachari, F. Bai, "Distributed Storage Codes Reduce Latency in Vehicular Networks", *Accepted for publication, Transactions on Mobile Computing 2013.*
- J. Ahn, M. Sathiamoorthy, B. Krishnamachari, F. Bai, L. Zhang, "Optimizing Content Dissemination in Vehicular Networks with Radio Heterogeneity", *Accepted for publication, Transactions on Mobile Computing 2013.*
- M. Sathiamoorthy, W. Gao, B. Krishnamachari, G. Cao, "Minimum Latency Data Diffusion in Intermittently Connected Mobile Networks", in *2012 IEEE 75th Vehicular Technology Conference: VTC2012-Spring, 6-9 May 2012, Yokohama, Japan.*
- M. Sathiamoorthy, A. G. Dimakis, B. Krishnamachari, F. Bai, "Distributed Storage Codes Reduce Latency in Vehicular Networks", in *Proceedings of the IEEE INFOCOM Mini-conference, 2012.*
- M. Alresaini, M. Sathiamoorthy, B. Krishnamachari, M. J. Neely, "Backpressure with Adaptive Redundancy (BWAR)", in *Proceedings of the IEEE INFOCOM, 2012.*
- S. Lee, S. Pattem, M. Sathiamoorthy, B. Krishnamachari, A. Ortega, "Spatially-Localized Compressed Sensing and Routing in Multi-Hop Sensor Networks", in *3rd International Conference on Geosensor Networks*, July 2009, Pages 11-20.

INTERNSHIP
EXPERIENCE

TURN Inc., Redwood City, CA
Big Data Intern   Sept 2013 – Nov 2013
- Working on Column-storage for Hadoop.

Symantec Research Labs, Culver City, CA
Research Intern   May 2013 – Aug 2013
- Designed and implemented a new RAID scheme that can protect from two disk failures and can reduce single disk rebuilding time by a few factors than RAID-6 (by utilizing a fraction of extra storage).
- A software RAID implementation over 12 disks demonstrated $3\times$ reduction in disk rebuild time.
- To be filed for a Patent.

General Motors R&D, Warren, MI
Visiting Scholar   May 2011 – Aug 2011
- Developed a new inter-vehicular video sharing application based on GM's existing Wavecast system for vehicular communication.
- Integrated erasure coding into the application for faster distributed file downloads.
- Additionally, developed an Android application to act as the front end (which connects to and controls the Linux based video sharing application wirelessly).

University of Southern California, Los Angeles, CA

Summer Intern May 2007 – July 2007
- Worked on the energy reduction of Wireless Sensor Networks using Compressed Sensing
- Compressed Sensing is used to integrate compression and sensing to achieve energy gains as high as 90% in ideal conditions.
- The work involved using wavelet and DCT domains to perform compressed sensing to sense temperature. Sparse random projections were used to effectively reduce the number of communications to the sink. Methods for optimizing the projections were also investigated.

Nanyang Technological University, Singapore

Summer Intern May 2006 – July 2006
- Developed an English Continuous Speech Recognizer based on TIMIT Database using HTK software.
- Worked on Variable Frame Rate Algorithms and on using Spectral Entropy based features and tested them on the CENSREC-3 database.

RESEARCH
PROJECTS

Reliable Block Placement for Cold Data in Data Centers
- In this ongoing work, I am investigating how best to store cold data in a data center with maximum efficiency while maintaining high reliability. To facilitate the study, I have designed a data center block storage simulator, and utilized Akka distributed computation framework to scale-out processing.
- Future work involves implementing a Block Placement Policy in Hadoop.

Novel Erasure Codes for Big Data
- In collaboration with Facebook, we implemented regenerating codes specially designed for data centers over Hadoop HDFS. About 2x reduction in network utilization and disk I/O during rebuilds. Paper accepted for publication in VLDB 2013.
- This Hadoop version is available at https://github.com/madiator/HadoopUSC.

Distributed Storage Coding for Vehicular Networks
- Erasure Coding applied to distributed storage in Vehicular Networks to minimize the delay in content retrieval. I obtained theoretical upper bounds on delay and showed regions where Network Coding performs better than naive distribution strategy. Simulated on real taxi trace datasets to show the improvement. Work published at Infocom-Mini 2012 and TMC 2013 (accepted).
- Continued the work at General Motors to test on *real vehicles*, where I developed a new inter-vehicle video sharing application based on GM's existing Wavecast system for vehicular communication. Additionally, developed an Android application to act as the front end (which connects to and controls the Linux based video sharing application wirelessly).
- I have set up openvanet.org where the code and the datasets are available to be downloaded.

Twitter Retweet Dynamics
- In a collaborative work with USC Annenberg School of Communication and Journalism, we collected twitter graph and retweet data and showed an interesting trend between the number of retweets received and the probability of retweeting by Twitter users. Work was presented at the *2nd Annual Annenberg Symposium* and was invited to present again at the *3rd Annual Annenberg Reception*.

Implementation of MapReduce
- Implemented MapReduce (a framework for distributed processing) on USC's High-Performance Computing and Communications (HPCC) cluster consisting of thousands of nodes, as part of a course. Used it to study large network datasets.

Other projects
- Implementation of Nachos, a software that simulates a small OS, involving process management, memory management, interprocess communication, fault tolerance etc., for a course on Operating Systems.
- Flash Scheduling - proposed and analyzed a new scheduling algorithm in a multi-user communication system with varying number of users (term project for Computer Communications course).

	• Coded Distributed Storage for Cloud Environments. Pitch at *MHI Research Festival 2013*. **Best Pitch Award** (Honorable Mention). • Helper Node Allocation Strategies for Content Dissemination in Intermittently Connected Mobile Networks. E-Poster at *5th Annenberg Symposium 2013*. • Distributed Storage Codes Reduce Latency in Vehicular Networks. Presentation at *Infocom 2012, Orlando, FL*. Poster at *MHI Research Festival 2011*. • Minimum Latency Data Diffusion in Intermittently Connected Mobile Networks. Presentation at *VTC Spring 2012, Yokohoma, Japan*. • A Study of Twitter Retweet Dynamics. Presentation at *2nd Annenberg Symposium 2010*, invited again for *3rd Annenberg Fellows Reception*. • XORing Elephants: Novel Erasure Codes for Big Data. Poster at MHI Research Festival 2012 • Reliable Storage for Digital Media. *E-Poster at 4th Annenberg Symposium 2012*.

SKILLS	*Programming Languages*: C, C++, Java *Software Experiences*: Hadoop, Android, TCP/IP, Matlab, LaTeX, Processing

TEACHING	TA for **Wireless and Mobile Device Networks Design and Laboratory** (Spring 2012).

COURSES	Trends in Cloud Computing and Data Center Networking

	Distributed Storage Theory — Algebraic Coding Theory
	Analysis of Algorithms — Approximation Algorithms
	Computer Communications — Operating Systems
	Random Processes in Engineering — Wavelets
	Queueing Theory
	Design and Analysis of Computer Communication Networks
	Probabilistic Methods in Computer Systems Modeling

SOFTWARE RELEASES	• VANETSim: A Vehicular Network Content Dissemination Simulator. Java based tool that can talk with real-vehicle GPS trace datasets and can simulate content dissemination between vehicles. Code and datasets available at openvanet.org. • HadoopUSC: The raid-contrib project implements new regenerating (erasure) codes. Available at github.com/madiator/HadoopUSC.

HONORS AND AWARDS	• USC Annenberg Fellowship 2008-2012 - University of Southern California • InfoUSA Summer Fellowship 2007 - University of Southern California • One of the seven finalists of Trilogy's Pirates of the Corporate, a web 2.0 business plan contest held at Hong Kong. • Amateur Radio License (Grade II) from Ministry of Communications, Government of India (2003). • Ranked among top 0.5% of about 170,000 students appeared in JEE 2004.

REFERENCES	Prof. Bhaskar Krishnamachari Associate Professor Ming Hsieh Department of Electrical Engineering USC Viterbi School of Engineering, Los Angeles, CA Email: bkrishna at usc.edu http://ceng.usc.edu/~bkrishna Prof. Alexandros G. Dimakis Assistant Professor Department of Electrical Engineering University of Texas at Austin Email: dimakis at austin.utexas.edu http://users.ece.utexas.edu/~dimakis