

Avani P. Gadani/Wildani

CONTACT INFORMATION 1010 Pacific Ave #501 *Voice:* +1 909.437.8626
Santa Cruz, CA, 95060 *E-mail:* avani@soe.ucsc.edu

Objective: Summer Internship: June 2010 through September 2010

EDUCATION **University of California**, Santa Cruz, CA USA
PhD Computer Science (ongoing: started September 2007)

University of New Mexico, Albuquerque, NM USA
MS Computer Science : Machine Learning (May 2007)

Harvey Mudd College, Claremont, CA USA
B.S Joint Math / Computer Science (May 2003)

Relevant Classes: Advanced Math and Computer Science Courses including -

Advanced Algorithms, **Scientific Computing**, Graph Theory, Computational Geometry, Probability, Differential Geometry, **Machine Learning**, Geometric and Probabilistic Methods in CS, **Artificial Intelligence**, Software Engineering, **Computational Biology**, Programming Languages Theory, Computational Linguistics, **Networks**, **Storage Systems**, Adv. Operating Systems, **Information Theory**, and Databases

PROJECTS **UCSC Storage Systems Research Center**

I currently work in Distributed Storage. Specifically, I am in the early stages of developing a system that uses machine learning and probabilistic modeling to create a distributed storage network that is resilient to fluctuations in node reliability and trust. A secondary goal is to apply this research to create a global, censorship-resistant storage system that is more scalable and decentralized than current projects.

I am also working in power-aware coding to explore the trade-offs between power consumption, performance, and availability when reconstructing data using irregular erasure codes.

MIND Institute

This research involved using fMRI images to detect interconnections within the brains of both healthy and schizophrenic patients. From this, we worked on deriving an efficient technique to perform clustering and hidden variable **Bayesian Structure Search** analyses. I have also used **Support Vector Machines** to obtain a baseline for our other classification methods, and worked extensively with **Spectral Graph Clustering** methods to handle the patients' categorical data.

SLC Project

Research Assistant in the proposal stage of a large scale learning project involving gene expression data from rodents. I modeled the data using **Bayesian Networks** after doing basic clustering for manual parameter selection.

Sandia National Laboratories Clinic

Program Manager of a team working with Dr. Kevin Boyack at Sandia to cluster discrete data points and produce a visualization tool to help with large data-set analysis. I **implemented density clustering** and also researched validity metrics to integrate with the tool we wrote in MATLAB.

Implemented a PID-based control system in Python to automatically balance customers for Bigtable (a key/value store). Extensively researched reward functions and SLA/SLO management techniques and wrote out a detailed experiment plan to investigate the tradeoffs between economic and standard scheduling techniques. Worked with the Census team to expand their SLO framework to include SLAs.

IBM Almaden Research Center: *Research Intern* **Summer 2008 (continued through fall)**

Worked with a team of 4 researchers to develop a highly accurate method to automatically classify log excerpts from commercially deployed storage systems to speed up problem isolation. This involved coding in Java, particularly the Lucene search engine, and Python for the backend and data-manipulation respectively. I also did a significant portion of the system design and analysis. This is ongoing work to be submitted to FAST in September.

Google: *Intern*

Summer 2006, Summer 2007

Documented Internal Build System Components using HTML/CSS and Wiki. Studied and documented Code Dependencies in Python, UNIX Shell, and C++ scripts. Used Python to implement statistical techniques to analyze web data.

CENIC: *Systems Administrator*

Oct 2003 - May 2004

Configured and supported a variety of **internal solutions** for CENIC's network management. This included limited exposure to **Cisco optical equipment** (15808 and 15540), and the deployment of a complete Cisco VoIP solution for the office. Also **administered 5-7 OS X and Solaris servers** including mail, news, Request Tracker, and Ciscoworks.

University of California, Santa Cruz, CA USA:

Research Assistant

March 2008 - present

Storage Systems research under Dr. Ethan Miller. (see Projects)

University of New Mexico, Albuquerque, NM USA: *Research Assistant*

Jan 2005 - May 2007

Machine Learning research under Dr. Terran Lane. (see Projects)

Teaching Assistant

Aug 2004 - Dec 2004

Lead teaching assistant for **Introduction to Java Programming**. Wrote and graded assignments, Taught Labs, Managed and coordinated with 5 TAs, and maintained the class web site.

Harvey Mudd College, Claremont, CA USA *Research Assistant*

May 2002 - Jan 2008

IDXP/BEEP **Intrusion Detection** project at Harvey Mudd College, with Prof. Mike Erlinger, Head of the IDWG. Over Summer 2002, we **implemented a functional IDS** using Perl and Java with help from Aerospace and Silicon Defense. We also monitored traffic using a modification of the snort tool and NFR and send these alerts into a **MySQL database**.

PUBLICATIONS

Avani Wildani, Thomas Schwarz, Ethan L. Miller, Darrell D. E. Long, Protecting Against Rare Event Failures in Archival Systems, Proceedings of the 17th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS 2009), September 2009.

Kevin Greenan, Darrell D. E. Long, Ethan L. Miller, Thomas Schwarz, Avani Wildani, Building Flexible, Fault-Tolerant Flash-based Storage Systems, Proceedings of the Fifth Workshop on Hot Topics in System Dependability (HotDep 2009), June 2009.

STRENGTHS

Programming and Markup Languages I've used for large projects:

Python (including NumPy/Numeric and SimPy), Java, Prolog, MATLAB, \LaTeX , Shell Scripting, XML, HTML.

Programming Languages I have used in a classroom setting:

C, C++ (STL), SML, Lisp, Ciao, Rex, ns2 simulator Tcl

Software and Operating Systems I can use proficiently:

UNIX Utilities, MATLAB, Microsoft Office
Solaris 7-9, Mac OS X, Windows 3.1 through XP, Linux (Ubuntu, Redhat, Debian, etc.)