

Meiyappan Nagappan

890 Oval Drive, EB II Room 3231
Raleigh, NC 27695-8206
Phone: (919) -513-5082, Fax: (919) -515 -7896
Email: mnagapp@ncsu.edu
<http://www4.ncsu.edu/~mnagapp/>

RESEARCH INTERESTS

My research interests are in the field of Software Engineering Systems focusing on Software System Log File Analysis, Software Fault Identification and Operational Profiling of Software Systems. I am particularly interested in the analysis of log files to identify abnormal behavior and help the developer locate the source of the problem. I am also interested in the application of algorithms that are traditionally used in the pure sciences domain, for solving day to day software engineering tasks.

EDUCATION

North Carolina State University, Raleigh Ph.D. in Computer Science, Area: <i>Software Fault Identification by Log File Analysis</i> Advisor: Dr. Mladen A. Vouk	2008 - Present
North Carolina State University, Raleigh M.S. in Computer Science Qualifiers Title: " <i>A Model for Sharing of Confidential Provenance Information in a Query Based System</i> " Advisor: Dr. Mladen A. Vouk GPA: 4.0/4.0	2008
Anna University, Madras B.E. in Computer Science Undergraduate Project Topic: " <i>Dynamic Scheduling of tasks in a Multiprocessor Environment</i> " First class with Distinction	2006

RESEARCH EXPERIENCE

North Carolina State University - Research Assistant (May 07 – Present)

Research Assistant to Dr. Mladen A. Vouk in the Scientific Data Management (SDM) Center working on the Scientific Process Automation (SPA) group's Provenance project.

Software Provenance (Jan 08 –May 08)

For my Qualifier's Exam at NCSU, I wrote a paper titled "A Model for Sharing of Confidential Provenance Information in a Query Based System". This was published in the 2nd International Provenance and Annotation Workshop. This research involved the identification of the confidentiality issues in a Collaborative query based system where the provenance data was stored. Then we present a model for sharing provenance information when the confidentiality level is decided by the user dynamically.

SDM Center Research (Jan 07-Present)

Investigating and constructing a provenance framework for scientific workflow systems. I am currently building the provenance collection system for DOE scientific applications in the Kepler workflow system. This is a collaborative project between NCSU and the Scientific Workflow Automation Team (SWAT) in San Diego Supercomputing Center (SDSC).

Building a dashboard solution for DOE scientists in Oak Ridge National Laboratories (ORNL). This is aimed at easing the daily tasks of the Application Scientist. This is under the "End to End Computing" project at ORNL. The system and data provenance codes are operational in ORNL and can be viewed on the Dashboard.

Did a literature survey of the past and present provenance techniques in scientific workflow space and other computer science areas.

Lawrence Berkeley National Laboratories - Intern

(May 08 - Aug 08)

Did a literature survey on the current state of research in Suffix Array data structures. Used these to speed the searching of all occurrences a particular text pattern in a text file that has a given gene sequence.

Used the suffix array and the longest common prefix array on execution logs of software systems to calculate the operational profile. A student paper titled "Efficient Operational Profiling of Systems using Suffix Arrays on Execution Logs" was selected for the International Symposium on Software Reliability Engineering (ISSRE) 2008. Wrote a research paper titled "Efficiently Extracting Operational Profiles from Execution Logs using Suffix Arrays" with validation data from 2 different systems. Currently in the process of addressing the issue of parsing the log files for message types, in an efficient and automated manner.

Did a literature survey on using execution logs for fault identification in software systems. Built the first version of a tool that will mine Non Finite Automaton Models from execution logs. Waiting on the data set to carry on the research.

Anna University, Chennai, India – Senior Thesis

(Jan 06 - May 06)

Investigated various algorithms for scheduling in a multiprocessor environment. We then proposed a Parallel Genetic Algorithm to better optimize the schedule of tasks without the time delays involved in usual genetic algorithms. We conducted java simulation experiments on our algorithm and the other algorithms. Published the work in a paper titled "Task Scheduling Using Parallel Genetic Algorithms For Heterogeneous Distributed Computing" at Grid Computing Applications (GCA) 2006.

Chennai Container Terminal Limited (CCTL) - Summer Project

(Apr 04 - Jun 04)

Identified the various processes involved in packing a cargo container with goods and analyzed key areas of center of gravity, strength and stability of various containers for cargo ships. Designed an algorithm and implemented C++ simulation software to pack the container taking into effect all the constraints.

TEACHING EXPERIENCE

CSC 326 - Software Engineering (NCSU - Undergraduate Core Course):– Teaching Assistant
(Fall 2006, Spring 2007, Fall 2007)

- Held weekly lab sections where students participated in hands-on software engineering exercises.
- Create multiple homework assignments and detailed programming projects, and grade them.
- Awarded the Outstanding Teaching Assistant Award for the year 2007 by the University Graduate Student Association.

CSC 712 - Software Testing and Reliability (NCSU - Graduate Higher Level Course):– Teaching Assistant
(Fall 2006)

CSC 456 - Computer Architecture and Multiprocessors (NCSU - Undergraduate Course):– Teaching Assistant
(Spring 2007)

- Assisted in grading and general course administration and held weekly student sessions for all the three courses.

PUBLICATIONS

CONFERENCE PUBLICATIONS

Nagappan, M., Vouk, M.A., Wu, K., Sim, A., Shoshani, A.. "Efficient Operational Profiling of Systems using Suffix Arrays on Execution Logs." Accepted as Student Paper in the 19th International Symposium on Software Reliability Engineering, 11-14 Nov, 2008, Redmond, WA.

Nagappan, M., Vouk, M.A., "A Model for Sharing of Confidential Provenance Information in a Query Based System" 2nd International Provenance and Annotation Workshop, 17 - 18 Jun, 2008, Salt Lake City, Utah.

N. Nedunchezian, K. Koushik, N. Meiyappan, V. Raghu, "Dynamic Task Scheduling Using Parallel Genetic Algorithms for Heterogeneous Distributed Computing", International Conference on Grid Computing and Applications (GCA'06), Las Vegas, USA, pp. 82-88

THESES

Nagappan, M, "A Model for Sharing of Confidential Provenance Information in a Query Based System", PhD. Qualifiers Exam, North Carolina State University, 2008.

Nagappan, M, "Dynamic Scheduling of tasks in a Grid Environment", Senior Thesis, SVCE, Anna University, May 2006.

REFEREED POSTERS

Ilkay Altintas, George Chin, Daniel Crawl, Terence Critchlow, David Koop, Jeff Ligon, Bertram Ludaescher, Pierre Mouallem, Meiyappan Nagappan, Norbert Podhorszki, Claudio Silva, Mladen Vouk, "Provenance in Kepler-based Scientific Workflow Systems," Poster # 41, at Microsoft eScience Workshop Friday Center, University of North Carolina, Chapel Hill, NC, October 13 - 15, 2007, pp. 82.

Roselyne Barreto, Terence Critchlow, Ayla Khan, Scott Klasky, Leena Kora, Jeffrey Ligon, Pierre Mouallem, Meiyappan Nagappan, Norbert Podhorszki, Mladen Vouk, "Managing and Monitoring Scientific Workflows through Dashboards," Poster # 93, at Microsoft eScience Workshop Friday Center, University of North Carolina, Chapel Hill, NC, October 13 - 15, 2007, pp. 108.

TUTORIALS PRESENTED

SuperComputing 08, "Introduction to Scientific Workflow Management and the Kepler System" Austin, TX (11/08)

SuperComputing 07, "Introduction to Scientific Workflow Management and the Kepler System" Reno, NV (11/07)

ORNL All Hands Meeting, "Introduction to Scientific Workflow Management and the Kepler System", Oak Ridge, TN (03/07)

PROFESSIONAL ACTIVITIES

Organizing activities

- Web co-chair, IEEE International Symposium on Software Reliability Engineering (ISSRE) 2008, Redmond, WA, USA.

RESEARCH PROJECTS

Operational Profiler

(<http://www4.ncsu.edu/~mnagapp/code.html>)

Using Suffix Arrays and Longest Common Prefix Array on execution logs, we construct the operational profile of a given software system.

Kepler

(<http://kepler-project.org/>)

Kepler is an open source cross-project, cross-institution collaboration to build and evolve a scientific workflow system on top of the (also evolving) Ptolemy II system.

Dashboard

(<https://dashboard.ccs.ornl.gov>)

Dashboard at ORNL is to provide application scientists with an easy-to-use tool for machine and simulation monitoring.

SCHOOL PROJECTS

iTrust	A medical records application developed in java, with MySql as backend and jsp as front end for the Software Engineering and Software Testing and Reliability course.
Bus-based cache coherence protocols	Edited the MSI protocol simulator to behave as a MESI protocol simulator and collected the performance statistics for it as part of the Architecture of Parallel Computers course.
Decaf	A programming language that is a subset of java. I implemented the compiler for it as part of the Compiler Construction course.
Spaghetti Resolver	Designed a system to manage the spaghetti of wires in a server room for the Human Computer Interaction course.
Wolflnn	A hotel management system written in java with Oracle backend for the Database Management Systems course.
CirclePack	A Software design project to write an algorithm to accept or reject a potential packing set of circles $S = \{ R(i) : i = 1,2,\dots,n \}$ in the given target circle with radius R, as a packable set.
Virtual Stock Market	Developed stock market simulation software for a Kapital FY'05, a National level technical symposium at SVCE.

UNIVERSITY SERVICE

Committee Member - University Library Committee (North Carolina State University) - 2008-2009

Committee Member- FORESE (Forum for Economic Studies by Engineers, SVCE, India) – 2004-2006.

Committee member at “KAPITAL '04”, and “KAPITAL '05” a National level technical symposium at SVCE.

Committee member at “Interrupt '04”, a National level technical symposium at SVCE.

Member of the organizing committee for Nanosphere 2004, a National Conference on Nanotechnology hosted by SVCE, Chennai, India.

Committee Member- ACE (Association of Computer Engineers, SVCE, India) – 2002-2006.

AWARDS AND HONORS

Member of Phi Kappa Phi, The Honor society

Recipient, IEEE ISSRE 2008 Student paper Travel award.

Research Assistantship funded by DOE Grant for SDM center, 2007 – Present.

Outstanding Teaching Assistant Award for the year 2007, University Graduate Student Association, North Carolina State University, 2008.

N.C. State Teaching Assistantship funded by the Graduate Student Support Plan, 2006 – 2007.

Elected Departmental representative in the academic council, SVCE, 2004-2005.

PERSONAL INFORMATION

Indian Citizen. US Student Visa (F-1).

RESEARCH COLLABORATORS

Lawrence Berkeley National Laboratories
Oak Ridge National Laboratories
San Diego Supercomputing Center

Pacific Northwest National Laboratories
University of California Davis
University of Utah

REFERENCES

Dr. Mladen A, Vouk
Professor and Chair
Department of Computer Science
North Carolina State University
Raleigh, NC 27695
vouk@csc.ncsu.edu

Dr. Arie Shoshani
Senior Staff Scientist
Lawrence Berkeley Laboratory
Berkeley, CA 94720
shoshani@lbl.gov

Dr. Kesheng Wu
Staff Scientist
Lawrence Berkeley Laboratory
Berkeley, CA 94720
kwu@lbl.gov