

Meiyappan Nagappan

890 Oval Drive, EB II Room 3231
Raleigh, NC 27695-8206
Phone: (919) -607-2110, 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: *Software Fault Identification by Log File Analysis*
Advisor: Dr. Mladen A. Vouk
GPA: 4.0/4.0

North Carolina State University, Raleigh

M.S. in Computer Science
Qualifiers Title: *"A Model for Sharing of Confidential Provenance Information in a Query Based System"*
Advisor: Dr. Mladen A. Vouk
GPA: 4.0/4.0

Anna University, Madras

B.E. in Computer Science
Undergraduate Project Topic: *"Dynamic Scheduling of tasks in a Multiprocessor Environment"*
First class with Distinction

RESEARCH EXPERIENCE

Microsoft Research, Cambridge - Intern (May 10 - Aug 10)
Analyzed the data collected from the development of Microsoft Products to identify the root developmental causes of injecting bugs in the code.

ABB-US Corporate Research Center - Intern (Jan 10 - May 10)
Built tools for the operational profiling of ABB systems, by using their log files.

Lawrence Berkeley National Laboratories - Intern (May 08 - Aug 08)
Used the suffix array and the longest common prefix array on execution logs of software systems to calculate the operational profile.

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.

Log File Analysis (Aug 08-Present)
Analyzed log files from the Virtual Computing Lab at NCSU to build the operational profile for the use of the system, which could be used in prioritizing test cases, dynamic logging of information to help the developers of VCL debug the system, and abstracting the log lines in the log file to log events.

Software Provenance**(Jan 08 –May 08)**

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.

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.

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) - Fall 2006, Spring 2007, Fall 2007

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

CSC 456 - Computer Architecture and Multiprocessors (NCSU - Undergraduate Course) - Spring 2007

PUBLICATIONS

CONFERENCE PUBLICATIONS

Nagappan, M., Vouk, M.A., "Adaptive Logging: A Case Study of Logs from a Cloud Computing Environment". In the Fast Abstracts track of International Symposium on Software Reliability Engineering, 1-4 Nov, 2010, San Jose, California.

Nagappan, M., Vouk, M.A., "Abstracting Log Lines to Log Event Types for Mining Software System Logs". In the proceedings of Mining Software Repositories (Co-Located with ICSE 2010), 2-3 May, 2010, Cape Town, South Africa.

Nagappan, M., "Analysis of Execution Log Files". Published in the Doctoral Symposium track of the 32th International Conference on Software Engineering, 2-8 May, 2010, Cape Town, South Africa.

Nagappan, M., Wu, K., Vouk, M.A., "Efficiently Extracting Operational Profiles from Execution Logs using Suffix Arrays." In the proceedings of the 20th International Symposium on Software Reliability Engineering, 16-19 Nov, 2009, Mysuru, India.

Nagappan, M., Vouk, M.A., Wu, K., Sim, A., Shoshani, A.. "Efficient Operational Profiling of Systems using Suffix Arrays on Execution Logs." 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 Moullem, *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 Moullem, *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

Log File Abstraction

Used the empirical data on the frequency of words in a log file, we built a tool that would abstract the log lines to log events automatically.

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.

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.

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.

PERSONAL INFORMATION

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

REFERENCES

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

Dr. Brian Robinson
Senior Principal Scientist
ABB - US Corporate Research Center
Raleigh, NC 27695
brian.p.robinson@us.abb.com

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

Dr. Brendan Murphy
Principal Researcher
Microsoft Research
Cambridge, UK CB3 0FB
bmurphy@microsoft.com

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

Dr. Laurie Williams
Associate Professor
Department of Computer Science
North Carolina State University
Raleigh, NC 27695
williams@csc.ncsu.edu