Meiyappan Nagappan

Software Engineering Department, Rochester Institute of Technology Phone: 585-475-2949
134 Lomb Memorial Drive Rochester, NY 14623-5608
http://www.se.rit.edu/~mei/ Email: mei@se.rit.edu

Education

- **Assistant Professor, Software Engineering Department, Rochester Institute of Technology, Rochester, NY (Current)**
- **Post Doctoral Fellow, Software Analysis and Intelligence Laboratory, Queen's University, Kingston, ON (Oct 2014)**

  Research Area: *Empirical Software Engineering using Ultra Large Repositories*

  Supervisor: Dr. Ahmed E. Hassan
- **Doctor of Philosophy (PhD), Computer Science, North Carolina State University, Raleigh, NC (May 2011).**

  Research Area: *Framework for Analyzing Log Files from a Cloud Computing Infrastructure*

  Advisor: Dr. Mladen A. Vouk
- **Master of Science (M.S.), Computer Science, North Carolina State University, Raleigh, NC (May 2008).**

  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
- **Bachelor of Engineering (B.E.), Computer Science, Anna University, Madras, India (May 2006).**

  Undergraduate Project: *Dynamic Scheduling of Tasks in a Multiprocessor Environment*

  Graduated *First Class with Distinction.*

Research Experience

Software Analysis and Intelligence Laboratory May 2011 - Oct 2014
Queen's University, Kingston, ON

Empirical Software Engineering using Ultra Large Repositories - Worked on mining, and analyzing ultra large software repositories (*100K - 1M projects, or TBs of data*), in order to identify patterns and relationships across entire ecosystems of projects. In this work, the concerns of various Software (S/W) stakeholders - S/W operators, build engineers, and project managers, in addition to S/W developers, testers and Software Engineering (SE) researchers were addressed.

Department of Computer Science May 2008 - May 2011
North Carolina State University, Raleigh, NC

Framework for Analyzing Log Files from a Cloud Computing Infrastructure - Proposed a framework with linearly scaling solutions for filtering, abstracting, and analyzing execution logs. Besides theoretical evaluation of the solutions, empirical and qualitative evaluation was carried out by determining the operational profile of the Virtual Computing Lab, a *cloud computing infrastructure* used by more than 30,000 users.

Scientific Data Management Center May 2007 - May 2011
North Carolina State University, Raleigh, NC

Provenance Management in Big Data Scientific Workflows - Built the provenance management system for the workflows of *Big Data scientific applications* like Plasma Edge Simulation for Nuclear Fusion, Climate Modelling, and Underground Waterflow Modelling (funded by the United States-Department of Energy).

Research Intern - Microsoft Research Jun 2010 - Aug 2010
Cambridge, UK

Analyzed the data describing the development of Microsoft Products to identify the root causes in development activities that lead to defects in the code.

Research Intern - ABB US Corporate Research Center Jan 2010 - May 2010
Raleigh, NC

Built tools for the operational profiling of ABB systems by mining log files.

Research Intern - Lawrence Berkeley National Labs May 2008 - Aug 2008
Berkeley, CA

Used the suffix array and the longest common prefix array data-structures on execution logs of software systems to calculate the operational profile.

Department of Computer Science and Engineering Jan 2006 - May 2006
Anna University, Chennai, India

Dynamic Scheduling of Tasks in a Multiprocessor Environment - Investigated various algorithms for scheduling in a multiprocessor environment, and then proposed a *Parallel Genetic Algorithm* to better optimize the schedule of tasks. Conducted Java simulation experiments to compare the proposed algorithm and the other existing algorithms.

Summer Project - Chennai Container Terminal Limited Apr 2004 - Jun 2004
Chennai, India

Designed an algorithm and implemented a C++ based application to pack a cargo container, taking constraints such as center of gravity, strength and stability of various containers into consideration.

Teaching Experience

Instructor

- Empirical Software Engineering using Ultra Large Repositories (CISC 835 - Queen's University - Graduate Higher Level Course) - Fall 2012. (6 students - two projects accepted as publications and three projects under submission).
- Software Engineering (CSC 326L-00X - NCSU - Undergraduate Core Course Lab) - Fall 2006, Spring 2007, Fall 2007. (20-25 students each semester).

Teaching Assistant

- Software Testing and Reliability (CSC 712 - NCSU - Graduate Level Course) - Fall 2006. (12 students).
- Computer Architecture and Multiprocessors (CSC 456 - NCSU - Undergraduate Course) - Spring 2007. (25-30 students).

Publications

Journal Publications (5 year Impact factors in brackets)

1. Israel J. Mojica Ruiz, **Meiyappan Nagappan**, Bram Adams, Thorsten Berger, Steffen Dienst, Ahmed E. Hassan, "An Examination of the Current Rating System used in Mobile App Stores", *IEEE Software (1.408)*, Accepted Jan 2015.
2. Hammad Khalid, **Meiyappan Nagappan**, Ahmed E. Hassan, "Examining the Relationship between FindBugs Warnings and End User Ratings: A Case Study On 10,000 Android Apps", *IEEE Software (1.408)*, Accepted Nov 2014.
3. Hadi Hemmati, **Meiyappan Nagappan**, Ahmed E. Hassan, "Investigating the Effect of 'Defect Co-fix' on Quality Assurance Resource Allocation: A Search-based Approach", *Journal of Systems and Software*, Accepted November 2014.
4. Mark D. Syer, **Meiyappan Nagappan**, Bram Adams and Ahmed E. Hassan, "Replicating and Re-evaluating the Theory of Relative Defect-Proneness", *IEEE Transactions on Software Engineering (2.786)*, Accepted September 2014.
5. Shane McIntosh, **Meiyappan Nagappan**, Bram Adams, Audris Mockus, Ahmed E. Hassan, "Large-Scale Empirical Study of the Relationship Between Build Technology and Build Maintenance ", *Empirical Software Engineering (2.610)*, Accepted May 2014.
6. Israel J. Mojica Ruiz, **Meiyappan Nagappan**, Thorsten Berger, Bram Adams, Steffen Dienst, Ahmed E. Hassan, "On Ad Library Updates in Android Apps", *IEEE Software (1.408)*, Accepted May 2014.
7. Israel J. Mojica Ruiz, **Meiyappan Nagappan**, Bram Adams, Thorsten Berger, Steffen Dienst, Ahmed E. Hassan, "An Empirical Study on the Relationship between the Number of Ad Libraries in an App and its Rating", *IEEE Software (1.408)*, Accepted May 2014.
8. Mark D. Syer, **Meiyappan Nagappan**, Bram Adams, Ahmed E. Hassan, "Studying the Relationship Between Source Code Quality and Mobile Platform Dependence", *Springer Software Quality Journal (0.889)*, Accepted April 2014.
9. Hammad Khalid, Emad Shihab, **Meiyappan Nagappan**, Ahmed E. Hassan, "What Do Mobile App Users Complain About? A Study on Free iOS Apps", *IEEE Software (1.408)*, Accepted March 2014.
10. Nicolas Bettenburg, **Meiyappan Nagappan**, Ahmed E. Hassan, "Towards Improving Statistical Modelling of Software Engineering Data: Think Locally, Act Globally!", *Empirical Software Engineering (2.610)*, Accepted November 2013.
11. Israel J. Mojica, Bram Adams, **Meiyappan Nagappan**, Steffen Dienst, Thorsten Berger, Ahmed E. Hassan, "A Large Scale Empirical Study on Software Reuse in Mobile Apps", *IEEE Software Special Issue on Next Generation Mobile Computing (1.408)*, Accepted November 2013.
12. Stephen W. Thomas, **Meiyappan Nagappan**, Dorothea Blostein, Ahmed E. Hassan, "The Impact of Classifier Configuration and Classifier Combination on Bug Localization", *IEEE Transactions on Software Engineering (2.786)*, Vol 39, Issue 10, October 2013. Pages 1427-1443.

13. Weiyi Shang, **Meiyappan Nagappan**, Ahmed E. Hassan, "Studying the Relationship between Logging Characteristics and the Code Quality of Platform Software", *Empirical Software Engineering (2.610)*, Published Online September 08 2013.

Conference Publications

1. Jiaping Gui, Stuart Mcilroy, **Meiyappan Nagappan**, William G.J. Halfond, "Truth in Advertising: The Hidden Cost of Mobile Ads for Software Developers", Accepted to be published in the proceedings of the *37th International Conference on Software Engineering (ICSE 2015)*, Florence, Italy, May 16 - 24, 2015. (Acceptance Rate: 19%).

2. (**ACM SIGSOFT Distinguished Paper Award and Best Paper Award**) Felivel Camilo, Andrew Meneely and **Meiyappan Nagappan**, "Do Bugs Foreshadow Vulnerabilities? A Study of the Chromium Project", *In the proceedings of the 12th ACM/IEEE Working Conference on Mining Software Repositories (MSR 2015)*, Florence, Italy, May 16 - 17, 2015. (Acceptance Rate: 30%).

3. Baishakhi Ray, **Meiyappan Nagappan**, Christian Bird, Nachiappan Nagappan and Thomas Zimmermann, "The Uniqueness of Changes: Characteristics and Applications", *In the proceedings of the 12th ACM/IEEE Working Conference on Mining Software Repositories (MSR 2015)*, Florence, Italy, May 16 - 17, 2015. (Acceptance Rate: 30%).

4. Hammad Khalid, **Meiyappan Nagappan**, Emad Shihab, Ahmed E. Hassan, "Prioritizing The Devices To Test Your App On: A Case Study Of Android Game Apps", Published in the proceedings of the *22nd ACM SIGSOFT International Symposium on the Foundations of Software Engineering (FSE 2014)*, Hong Kong, China, November 16 - 21, 2014. (Acceptance Rate: 22%).

5. Weiyi Shang, **Meiyappan Nagappan**, Ahmed E. Hassan, Zhen Ming Jiang, "Understanding Log Lines Using Development Knowledge", Published in the proceedings of the *30th IEEE International Conference on Software Maintenance and Evolution (ICSME 2014)*, Victoria, Canada, October 1 - 3, 2014. (Acceptance Rate: 19%).

6. Shane Mcintosh, Bram Adams, **Meiyappan Nagappan**, Ahmed E. Hassan, "Mining Co-Change Information to Understand when Build Changes are Necessary", Published in the proceedings of the *30th IEEE International Conference on Software Maintenance and Evolution (ICSME 2014)*, Victoria, Canada, October 1 - 3, 2014. (Acceptance Rate: 19%).

7. Tse-Hsun Chen, **Meiyappan Nagappan**, Emad Shihab, Ahmed E. Hassan, "An Empirical Study of Dormant Bugs", In the proceedings of the *11th ACM/IEEE Working Conference on Mining Software Repositories (MSR 2014)*, Hyderabad, India, May 31 - June 1, 2014. (Acceptance Rate: 34%).

8. Thanh H. D. Nguyen, **Meiyappan Nagappan**, Ahmed E. Hassan, Mohamed Nasser, Parminder Flora, "An Industrial Case Study of Automatically Identifying Performance Regression-Causes", *In the proceedings of the Practice Track at the 11th ACM/IEEE Working Conference on Mining Software Repositories (MSR 2014)*, Hyderabad, India, May 31 - June 1, 2014. (Acceptance Rate: 34%).

9. Mark D. Syer, Zhen Ming Jiang, **Meiyappan Nagappan**, Ahmed E. Hassan, Mohamed Nasser, Parminder Flora, "Continuous Validation of Load Test Suites", In the proceedings of the *5th ACM/SPEC International Conference on Performance Engineering (ICPE 2014)*, Dublin, Ireland, March 23 - 26, 2014. (Acceptance Rate: 30%).

10. Mark D. Syer, **Meiyappan Nagappan**, Bram Adams, Ahmed E. Hassan, "Revisiting Prior Empirical Findings For Mobile Apps: An Empirical Case Study on the 15 Most Popular Open Source Android Apps", In the proceedings of the *23rd International Conference hosted by the Centre for Advanced Studies Research, IBM Canada Software Laboratory (CASCON 2013)*, Markham, Canada, November 18-20, 2013. (Acceptance Rate: 36%).

11. Mark D. Syer, Zhen Ming Jiang, **Meiyappan Nagappan**, Ahmed E. Hassan, Mohamed Nasser, Parminder Flora, "Leveraging Performance Counters and Execution Logs to Diagnose Memory-Related Performance Issues", In the proceedings of the *29th IEEE International Conference on Software Maintenance (ICSM 2013)*, Eindhoven, The Netherlands, September 24-26, 2013. (Acceptance Rate: 22%).

12. Seyyed Ehsan Salamati Taba, Foutse Khomh, Ying Zou, Ahmed E. Hassan, **Meiyappan Nagappan**, "Predicting Bugs Using Antipatterns", In the proceedings of the *29th IEEE International Conference on Software Maintenance (ICSM 2013)*, Eindhoven, The Netherlands, September 24-26, 2013. (Acceptance Rate: 22%).

13. **Meiyappan Nagappan**, Thomas Zimmermann, Christian Bird, "Diversity in Software Engineering Research", In the proceedings of the *9th joint meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering (ESEC/FSE 2013)*, Saint Petersburg, Russia, August 21-23, 2013. Pages 466-476. (Acceptance Rate: 20%).

14. Masateru Tsunoda, Koji Toda, Kyohei Fushida, Yasutaka Kamei, **Meiyappan Nagappan**, Naoyasu Ubayashi, "Revisiting Software Development Effort Estimation Based on Early Phase Development Activities", In the proceedings of the *10th ACM/IEEE Working Conference on Mining Software Repositories (MSR 2013)*, San Francisco, CA, USA, May 18-19, 2013. Pages 429-438. (Acceptance Rate: 38%).

15. Roman Suvorov, Bram Adams, **Meiyappan Nagappan**, Ahmed E. Hassan, Ying Zou, "An Empirical Study of Build System Migrations in Practice: Case Studies on KDE and the Linux Kernel", In the proceedings of the *28th IEEE International Conference on Software Maintenance (ICSM 2012)*, Riva del Garda, Trento, Italy, September 25-27, 2012. Pages 160-169. (Acceptance Rate: 25%).

16. Israel J. Mojica Ruiz, **Meiyappan Nagappan**, Bram Adams, Ahmed E. Hassan, "Understanding Reuse in the Android Market", In the proceedings of the *20th IEEE International Conference on Program Comprehension (ICPC 2012)*, Passau, Germany, June 11-13, 2012. Pages 113-122. (Acceptance Rate: 41%).

17. (**Best Paper Award**) Nicolas Bettenburg, **Meiyappan Nagappan**, Ahmed E. Hassan, "Think Locally, Act Globally: Improving Defect and Effort Prediction Models", In the proceedings of the *9th ACM/IEEE Working Conference on Mining Software Repositories (MSR 2012)*, Zurich, Switzerland. June 2-3, 2012. Pages 60-69. (Acceptance Rate: 28%).

18. Tse-Hsun Chen, Stephen W. Thomas, **Meiyappan Nagappan**, Ahmed E. Hassan, "Explaining Software Defects Using Topic Models", In the proceedings of the *9th ACM/IEEE Working Conference on Mining Software Repositories (MSR 2012)*, Zurich, Switzerland. June 2-3, 2012. Pages 189-198. (Acceptance Rate: 28%).

19. **Meiyappan Nagappan**, Kesheng Wu, Mladen A. Vouk, "Efficiently Extracting Operational Profiles from Execution Logs using Suffix Arrays", In the proceedings of the *20th IEEE International Symposium on Software Reliability Engineering (ISSRE 2009)*, Mysuru, India. November 16-19, 2009. Pages 41-50. (Acceptance Rate: 25%).

20. Nedunchelian Ramanujam, Koushik Kalyanaraman, **Meiyappan Nagappan**, Raghu Viswanathan, "Dynamic Task Scheduling Using Parallel Genetic Algorithms for Heterogeneous Distributed Computing", In the proceedings of the *International Conference on Grid Computing and Applications (GCA'06)*, Las Vegas, USA. June 26-29, 2006.

Other Refereed Publications

1. Catherine Ramirez, **Meiyappan Nagappan**, Mehdi Mirakhorli, "Studying the Impact of Evolution in R Libraries on Software Engineering Research", *In the proceedings of The first International Workshop on Software Analytics (SWAN 2015)*, Montreal, Canada, March 2, 2015. (Acceptance Rate: 73%).

2. Lucas Layman, Madeline Diep, **Meiyappan Nagappan**, Janice Singer, Robert Deline, Gina Venolia, "Debugging Revisited: Toward Understanding the Debugging Needs of Contemporary Software Developers", *In the proceedings of the 7th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM 2013) - Industrial Track*, Baltimore, Maryland, October 10-11, 2013. (Acceptance Rate: 45%).

3. **Meiyappan Nagappan**, Brendan Murphy, Mladen A. Vouk, "Which Code Construct Metrics are Symptoms of Post Release Failures?", In the proceedings of the *2nd International Workshop on Emerging Trends in Software Metrics (WETSoM 2011)*, Waikiki, Honolulu, HI, May 24, 2011. Pages 65-68. (Acceptance Rate: 57%).

4. **Meiyappan Nagappan**, Brian Robinson, "Creating Operational Profiles of Software Systems by Transforming their Log Files to Directed Cyclic Graphs", In the proceedings of the *6th International Workshop on Traceability in Emerging Forms of Software Engineering (TEFSE 2011)*, Waikiki, Honolulu, HI, May 23, 2011. Pages 54-57. (Acceptance Rate: 62%).

5. **Meiyappan Nagappan**, Aaron Peeler, Mladen A. Vouk, "Modeling Cloud Failure Data: A Case Study of the Virtual Computing Lab", In the proceedings of the *2nd International Workshop on Software Engineering for Cloud Computing (SECLOUD 2011)*, Waikiki, Honolulu, HI, May 22, 2011. Pages 8-14. (Acceptance Rate: 56%).

6. **Meiyappan Nagappan**, Mladen A. Vouk, "Abstracting Log Lines to Log Event Types for Mining Software System Logs", In the proceedings of the *7th Working Conference on Mining Software Repositories (MSR 2010)*, Cape Town, South Africa, May 2-3, 2010. Pages 14-17. (Acceptance Rate: 31%).

7. **Meiyappan Nagappan**, "Analysis of Execution Log Files", In the proceedings of the *Doctoral Symposium track of the 32th International Conference on Software Engineering (ICSE 2010)*, Cape Town, South Africa, May 4, 2010. Pages 409-412. (Acceptance Rate: 67%).

8. **Meiyappan Nagappan**, Mladen A. Vouk, Kesheng Wu, Alex Sim, Arie Shoshani, "Efficient Operational Profiling of Systems using Suffix Arrays on Execution Logs", In the proceedings of the *Student Paper track in the 19th International Symposium on Software Reliability Engineering (ISSRE 2008)*, Redmond, WA, November 11-14, 2008. Pages 313-314.

9. **Meiyappan Nagappan**, Mladen A. Vouk, "A Model for Sharing of Confidential Provenance Information in a Query Based System", In the proceedings of the *2nd International Provenance and Annotation Workshop (IPAW 2008)*, Salt Lake City, Utah, June 17-18, 2008. Pages 62-69.

Theses

1. **Meiyappan Nagappan**, "A Framework for Analyzing Software System Log Files", PhD Dissertation, North Carolina State University, 2011.

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

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

Invited Talks and Tutorials Presented

1. "Big(ger) Data in Software Engineering", Invited talk at *MSR Asia*, Tokyo, Japan, November 2014. (25-30 attendees).

2. "Analytics on Ad Library Maintenance in Android Apps", Invited talk at *Dagstuhl Seminar 14261 - Software Development Analytics*, Dagstuhl, Germany, June 2014. (30 - 40 attendees).

3. "Big(ger) Data in Software Engineering", Invited talk at *IIT Madras*, Chennai, India, June 2014. (25-30 attendees).

4. "Leveraging Performance Counters and Execution Logs to Diagnose Memory-Related Performance Issues", Invited talk at *Tata Research Development & Design Centre*, Pune, India, October 2013. (15-20 attendees).

5. "Large Scale Studies of Mobile Apps in the Android and iOS Markets", Invited talk at *IBM Research, India*, Bangalore, India, October 2013. (25-30 attendees).

6. "Statistics in Software Engineering: Pitfalls and Good Practices", *ESEC/FSE 2013*, Saint Petersburg, Russia, August 2013. (25-30 attendees).

7. "Large Scale Studies of Mobile Apps in the Android and iOS Markets", Invited talk at *Microsoft Research*, Redmond, WA, May 2013. (25-30 attendees).

8. "Towards Improving Statistical Modelling of Software Engineering Data: Think Locally, Act Globally!", Talk at the *Symposium on Augmenting Software Developer Support to Improve Productivity*, Monte Verità, Ascona, Switzerland, March 2013. (50-60 attendees).

9. "Introduction to Scientific Workflow Management and the Kepler System", *SuperComputing 2008*, Austin, TX, November 2008. (25-30 attendees).

10. "Introduction to Scientific Workflow Management and the Kepler System", *SuperComputing 2007*, Reno, NV, November 2007. (25-30 attendees).

11. "Introduction to Scientific Workflow Management and the Kepler System", *Oak Ridge National Labs - All Hands Meeting*, Oak Ridge, TN, March 2007. (25-30 attendees).

Awards and Honors

- Best Presentation, Consortium for Software Engineering Research (CSER 2013 - Spring Meeting), 2013.

- *Best Paper Award* and journal extension invitation in Empirical Software Engineering, for the paper titled "Think Locally, Act Globally: Improving Defect and Effort Prediction Models" (Nicolas Bettenburg, **Meiyappan Nagappan**, and Ahmed E. Hassan), at *9th Working Conference on Mining Software Repositories (MSR 2012)*, Zurich, Switzerland. June 2-3, 2012.

- Best Poster, Consortium for Software Engineering Research (CSER 2011 - Winter Meeting), 2011.

- Member of Phi Kappa Phi, The Honor society, 2008.

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

- Recipient, IEEE ISSRE 2008 Student paper Travel award, 2008.

Professional Service

Chair - Conference/Workshop

- *Student Volunteer Chair*, 36th International Conference on Software Engineering (ICSE 2014).

- *PC Chair*, 4th IEEE International Workshop on Empirical Software Engineering in Practice (IWESEP 2012).
- *Co-Organizer*, Mining Software Repositories Vision 2020 (MSR Vision 2020 - August 2012).
- *Website Chair*, 19th IEEE International Symposium on Software Reliability Engineering (ISSRE 2008).

PC Member - Conference/Workshop

- 1st International Workshop on BIG Data Software Engineering (BIGDSE 2015).
- 22nd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER 2014).
- 37th International Conference on Software Engineering - Demo Track (ICSE 2015).
- 30th IEEE International Conference on Software Maintenance and Evolution (ICSME 2014).
- 11th ACM/IEEE Working Conference on Mining Software Repositories (MSR 2014).
- ERA track of the 22nd International Conference on Program Comprehension (ICPC 2014).
- ERA track of the 29th IEEE International Conference on Software Maintenance (ICSM 2013).
- Data Challenge track of the 10th ACM/IEEE Working Conference on Mining Software Repositories (MSR 2013).
- 6th ACM SIGSOFT India Software Engineering Conference (ISEC 2013).
- Tool Demo track of the 19th Working Conference on Reverse Engineering (WCRE 2012).
- 23rd IEEE International Symposium on Software Reliability Engineering (ISSRE 2012).

Reviewer - Journal

- IEEE Transactions on Software Engineering (TSE).
- Empirical Software Engineering (ESE).
- The Journal of Systems and Software (JSS).
- IEEE Software.
- Transactions on Software Engineering and Methodology (TOSEM).

Departmental Service

- Committee Member - GCCIS Outstanding Scholar Committee - 2014.
- Mentor in the STARS Program (North Carolina State University) - 2008-2011.
- Assisted in processing Graduate School applications for the Department of Computer Science (North Carolina State University) - 2010.
- Committee Member - University Library Committee (North Carolina State University) - 2008-2009.