FEROSH JACOB

https://feroshjacob.github.io
feroshjacob@gmail.com

Summary

My primary research area of interest lies in software engineering. I have worked in many interdisciplinary projects involving software engineering and several fields (e.g., High-Performance computing, Machine learning, Data mining, Mathematics, Cloud computing) to improve design, development, and maintenance of software. I carefully 1) analyzed existing tool support 2) proposed new solutions, and 3) determined advantages and disadvantages of the proposed approach compared to the existing tools in the relevant area.

Education

Doctor of Philosophy *2009-2013*
Department of Computer Science, University of Alabama, Tuscaloosa, AL
Advisor: Dr. Jeff Gray
Thesis: *Modeling Computation Intensive Problems for Heterogeneous Computing and Improved Source Code Maintenance*

Master of Science *2007-2009*
Department of Computer Science, Clarkson University, Potsdam, NY
Advisor: Dr. Daqing Hou
Thesis: *CSeR - A Code Editor for Tracking & Visualizing Detailed Clone Differences*

Bachelor of Technology *2000-2004*
Department of Electronics and Communication, National Institute of Technology, Calicut, India

Teaching Experience

Assistant Professor (part-time) *May 2017 - Present*
Kennesaw State University, Kennesaw, GA

- CS 7263 Text Mining (2019 Spring) (Online and In-Class)
- CS 6045 Advanced Algorithms (2018 Fall) (Online and In-Class)
- CS 4242 Artificial Intelligence (2018 Summer) (Online)
- CS 7263 Text Mining (2018 Spring) (Online and In-Class)
- CS 1301 Programming Principles I (2017 Fall)
- CS 6021 Advanced Computer Architecture (2017 Summer) (Online and In-Class)

Teaching Assistant (Instructor) *Aug 2008 - May 2009*
Clarkson University, Potsdam, NY

- MA 131 Calculus I (Fall 2008)
- MA 132 Calculus II (Spring 2009)

Teaching Assistant (Grader) *January 2010 - April 2010*
University of Alabama at Birmingham, Birmingham, AL

- CS 402 Compiler Design (Spring 2008)

Staff Software Engineer *Sep 2015 - Present*
The Home Depot, Smyrna, GA

Sorting and Ranking (Role: Lead) The project focuses on improving the sorting and ranking of the products in HD.com. Our two-stage strategy involves: 1) Using customer engagement and merchants input for ranking the products and 2) Learning from the customer engaged products on how to rank the remaining products.

Search Evaluation (Role: Lead) One of our main search evaluation goals is to estimate the overall impact of a new feature to the HD.COM search without actually deploying the feature into production. Our current approach is to estimate the overall impact by comparing the search results (products) for the past searches (last month, last year) with and without the feature. Our future goal is to develop and deploy an absolute search evaluation system.

HD Ontology (Role: Engineer) We developed an ontology of products that are relevant to the home improvement store. The ontology was designed to improve the precision and recall of the HD.com search. My focus involved determining the synonyms and the hypernym/hyponym relationships between the products. I have used The HomeDepot search log and customer behavior data along with the publicly available data sets (e.g., Wikipedia, Wordnet) for the data extraction.

Senior Software Engineer *Sep 2013 - Sep 2015*
Careerbuilder, Norcross, GA

School Normalization (Role: Lead) I led and completed the Careerbuilder initiative on a system that can recognize and normalize school and university names from text documents (e.g., resumes, job postings). This system is currently in use and we leveraged several machine learning and data mining techniques for the entity recognition and normalization.

Recruitment Edge (Role: Engineer) We created a master profile of all the job seekers using the publicly available information in social media and Careerbuilder resume database. I worked on how to match the profile of one candidate for different sources using state-of-art instance matching techniques from literature.

Graduate Research Assistant *June 2010 - May 2013*
Department of Computer Science, University of Alabama, Tuscaloosa, AL

Software Modeling in HPC: I identified, analyzed, and developed three modeling approaches (Code-level, Program-level, and Domain-level) for computation intensive problems to leverage portability, improved source code maintenance, and heterogeneous computing. I also collaborated with researchers working on cloud computing to understand and evaluate the execution of such problems in a cloud. I was supported by a grant from Air Force/AFRL STTR, which I helped Dr. Jeff Gray in preparing and submitting.

Ph.D. Intern *May 2012 - Aug 2012*
Pacific Northwest National Laboratory, Richland, WA

HPC Pipeline DSLs (Role: Engineer) I designed and developed two Domain-Specific Languages (DSLs) for scientists to describe and deploy signature discovery workflows.

On Campus Ambassador *Oct 2007 - May 2008*
Sun MicroSystems, Clarkson University, Potsdam, NY

I worked as a technological evangelist for Sun (Oracle) technologies in Clarkson University. My duties included creating and supporting user groups and conducting seminars about the latest tools and technological advances.

Software Engineer *Jun 2004 - Jul 2007*

Stabilix Private Solutions Ltd, Trivandrum, India

Felix (Role: Engineer) I worked in the security module of a generic application/claim processing product for insurance companies.

Refereed Book Chapters

Faizan Javed and **Ferosh Jacob**, Data Science and Big Data Analytics at Career Builder, *Big-Data Analytics and Cloud Computing: Theory, Algorithms and Applications(Marcello Trovati, Richard Hill, Ashiq Anjum, Ying Shao Zhu, and Lu Liueds.)*, Springer International Publishing, ISBN: 978-3-319-25313-8, 2016, pp. 83-96.

Refereed Journal Articles

Ferosh Jacob, Adam Wynne, Yan Liu, Nathan Baker, and Jeff Gray, Domain-Specific Languages For Developing and Deploying Signature Discovery Workflows, *IEEE Computing in Science and Engineering*, vol. 16, no. 1, 2014, pp. 52-64.

Weihua Geng and **Ferosh Jacob**, A GPU Accelerated Direct-sum Boundary Integral Poisson Boltzmann Solver, *Computer Physics Communications*, vol. 184, no. 6, 2013, pp. 1490-1496.

Ferosh Jacob, Jeff Gray, Jeffrey C. Carver, Marjan Mernik, and Purushotham Bangalore, "PPModel: A Modeling Tool for Source Code Maintenance and Optimization of Parallel Programs," *The Journal of Supercomputing*, vol. 62, no 3, 2012, pp. 1560-1582.

Ferosh Jacob, Sonqging Yue, Jeff Gray, and Nicholas Kraft, "SRLF: A Simple Refactoring Language for High Performance FORTRAN," *Journal of Convergence Information Technology*, vol. 7, no. 12, 2012, pp. 256-263.

Ferosh Jacob, Amber Wagner, Prateek Bahri, Susan Vrbsky, and Jeff Gray, "Simplifying the Development and Deployment of MapReduce Algorithms," *International Journal of Next-Generation Computing* (Special Issue on Cloud Computing Yugyung Lee and Praveen Rao, eds.), vol. 2, no. 2, 2011, pp. 123-142.

Ferosh Jacob, Ashfakul Islam, Weihua Geng, Jeff Gray, Brandon Dixon, Susan Vrbsky, and Purushotham Bangalore, "PNBsolver: A Case Study on Modeling Parallel N-body Programs," *Journal of Parallel and Distributed Computing*, 26 pages (in submission).

Refereed Conference Publications

Ferosh Jacob, Ilamgumaran Karunanithi, Pramod Salian, and Ravi Sambhu, BBC: A DSL for Designing Cloud-based Heterogeneous Bigdata Pipelines *Proceedings of the International Conference on Big Data*, Boston, MA, December 2017.

Mayank Kejriwal, Qiaoling Liu, **Ferosh Jacob**, and Faizan Javed, A Pipeline for Extracting and Deduplicating Domain-Specific Knowledge Bases in *Proceedings of the International Conference on Big Data*, Santa Clara, CA, November 2015, pp. 1144-1153.

Qinlong Luo, Meng Zhao, Faizan Javed, and **Ferosh Jacob**, Macau: Large-Scale Skill Sense Disambiguation in the Online Recruitment Domain in *Proceedings of the International Conference on Big Data*, Santa Clara, CA, November 2015, pp. 1324-1329.

Ferosh Jacob, Aaron Johnson, Faizan Javed, Meng Zhao, and Matt McNair, WebScalding: A Framework for Big Data Web Services in *Proceedings of the International Conference of Big Data Computing Service and Applications*, San Francisco Bay, CA, April 2015, pp. 493-498.

Faizan Javed, Qinlong Luo, Matt McNair, **Ferosh Jacob**, Meng Zhao, Tae Seung Kang, Carotene: A Job Title Classification System for the Online Recruitment Domain in *Proceedings of the International Conference of Big Data Computing Service and Applications*, San Francisco Bay, CA, April 2015, pp. 286-293.

Meng Zhao, Faizan Javed, **Ferosh Jacob**, and Matt Mcnair, SKILL: A System for Skill Identification and Normalization in *Proceedings of the Association for the Advancement of Artificial Intelligence (Industry track)*, Austin, TX, January 2015, pp. 4012-4017.

Ferosh Jacob, Faizan Javed, Meng Zhao, and Matt Mcnair, sCooL: A System for Academic Institution Name Normalization in *Proceedings of the Second International Symposium on Big Data and Data Analytics*, Minneapolis, MN, May 2014, pp. 86-93.

Ferosh Jacob, Yu Sun, Jeff Gray, and Purushotham Bangalore, "A platform-independent tool for modeling parallel programs," in *Proceedings of the 49th Southeast Regional Conference*, Kennesaw, GA, March 2011, pp. 138-143.

Ferosh Jacob, David Whittaker, Sagar Thapaliya, Purushotham Bangalore, Marjan Mernik, and Jeff Gray, "CUDACL: A tool for CUDA and OpenCL programmers," in *Proceedings of the 17th International Conference on High Performance Computing*, Goa, India, December 2010, pp. 1-11.

Ferosh Jacob, Ritu Arora, Purushotham Bangalore, Marjan Mernik, and Jeff Gray, "Raising the level of abstraction of GPU-programming," in *Proceedings of the 16th International Conference on Parallel and Distributed Processing*, Las Vegas, NV, July 2010, pp. 339-345.

Ferosh Jacob and Robert Tairas, "Template inference using language models," in *Proceedings of the 48th Southeast Regional Conference*, Oxford, MS, April 2010, pp. 104-109.

Daqing Hou, **Ferosh Jacob**, and Patricia Jablonski, "Exploring the design space of proactive tool support for copy-and-paste programming," in *Proceedings of the 19th International Conference of Center for Advanced Studies on Collaborative Research*, Ontario, Canada, November 2009, pp. 188-202.

Daqing Hou, **Ferosh Jacob**, and Patricia Jablonski, "Proactively managing copy-and-paste induced code clones," in *Proceedings of the 25th International Conference on Software Maintenance*, Alberta, Canada, September 2009, pp. 391-392.

Daqing Hou, **Ferosh Jacob**, and Patricia Jablonski, "CnP: Towards an environment for the proactive management of copy-and-paste programming," in *Proceedings of the 17th International Conference on Program Comprehension*, British Columbia, Canada, May 2009, pp. 238-242.

Refereed Workshop Publications

Ferosh Jacob, Adam Wynne, Yan Liu, Nathan Baker, and Jeff Gray, "Domain-specific languages for composing signature discovery Workflows," in *Proceeding of the 12th Workshop on Domain-Specific Modeling*, Tucson, AZ, October 2012, pp. 61-64.

Robert Tairas, **Ferosh Jacob**, and Jeff Gray, "Representing clones in a localized manner," in *Proceedings of the 5th International Workshop on Software Clones*, Honolulu, Hawai, May 2011, pp. 54-60.

Ferosh Jacob, Jeff Gray, Purushotham Bangalore, and Marjan Mernik, "Refining high performance FORTRAN code from programming model dependencies," in *Proceedings of the 17th International Conference on High Performance Computing Student Research Symposium*, Goa, India, December 2010, pp. 1-5.

Ferosh Jacob, Daqing Hou, and Patricia Jablonski, "Actively comparing clones inside the code editor," in *Proceedings of the 4th International Workshop on Software Clones*, Cape Town, South Africa, May 2010, pp. 9-16.

Presentations

"BBC: A DSL for Designing Cloud-based Heterogeneous Bigdata Pipelines" *International Conference on Big Data*, Boston, MA, December 2017.

"Developing an Ontology for the Home Improvement/Retail Domain" *ACM Southeast conference*, Kennesaw, GA, April 2017.

"Machine learning techniques in Java" *Atlanta Java Users Group meetup*, Atlanta, GA, May 2016.

"Modulo-X: A Simple Transformation Language for HPC Programs" *Southeast Regional Conference*, Tuscaloosa, AL, March 2012 (Poster).

"CUDACL+: A Framework for GPU Programs," *International Conference on Object-Oriented Programming, Systems, Languages, and Applications (SPLASH/OOPSLA)*, Portland, OR, October 2011 (Doctoral Symposium).

"Refining high performance FORTRAN code from programming model dependencies" *International Conference on High Performance Computing Student Research Symposium*, Goa, India, December 2010 (Poster). **Best Presentation Award**

"Extending abstract APIs to shared memory," *International Conference on Object-Oriented Programming, Systems, Languages, and Applications (SPLASH/OOPSLA)*, Reno, NV, October 2010 (Student Research Competition). **Bronze Medal**

"CSeR: A code editor for tracking and highlighting detailed clone differences," *International Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA) Refactoring Workshop*, Orlando, FL, October 2009 (Poster).