Vivek Nair

Contact Information

3240 EBII, 890 Oval Drive
Raleigh, NC, 27606
Github: `https://github.com/vivekaxl`

Phone: (919) 523-9920
E-mail: vivekaxl@gmail.com
www: `http://www.vivekaxl.com`

Objective

My primary research interest involves the development of analytical tools that increase the strength of inferences made from scientific data. The application of my work focus on the application of next generation of AI and Machine Learning techniques to help improve the performance of software (or any) system or processes. My research question is always: how can we save resources (or money)?

Education

North Carolina State University — Raleigh, NC
Ph.D. candidate in Computer Science — May 2013 - Dec 2018
Advisor: Dr. Tim Menzies

National Institute of Technology, Durgapur — Durgapur, India
M.Tech in Information Technology — *2011*
Advisor: Dr. Subhrabrata Choudhury, Associate Professor

West Bengal University of Technology — Kolkata, West Bengal
B.Tech in Computer Science — *2009*

Professional Experience

Microsoft Research — Software Engineering Research Intern
Telemetry, Big Data, Analytics, C#, Azure Kusto, Azure CosmosDB — *May, 2018 - Aug, 2018*

- **Optimizing Distributed Testing Platform.** Optimized the performance (makespan) of Cloud Test, an internal tool used for testing software projects across Microsoft. The optimization strategy uses historical data to general models to predict an optimal schedule for the test jobs. The strategy saved 380 hours in 7 days (3%) and is currently the default scheduling strategy in CloudTest.

LexisNexis - Risk Solutions — Intern
Big Data, Usabulity, FUSE plugins, ECL, Python, Apache Spark — *June,2015 - August, 2017*

- **Enhanced ML Capabilities of HPCC (a big data cluster)**
  - Developed a FUSE plugin for HPCC to connect with Apache Spark. This plugin decreases data query time (up to 20%) as well as the overhead necessary to download the files to local machines.
  - Implemented an automated testing suite for the ML library and ensured that the testing time was < 24 hours.
  - Developed ML plugins for the Data Science Portal (an internal tool), which required massive refactoring of the codebase.
  - One of the largest contributors to the Machine Learning codebase (addition).

Samsung Software Engineering Labs, India — Software Engineer
Performance Analysis, Embedded System, Flash Memory — *June 2011 - May 2013*

- **Developed file-system and memory solutions for cellular phones.** Analyzed data from projects based on NOR Flash for Ultra Low-Cost cell phones, to reduce latency of applications. For example, reduced the boot time of the E1200 from 30 seconds to < 10 seconds.

Projects

Faster Discovery of Configuration Options of Software System

- Used optimization and ML techniques to discover (near) optimal system configurations.
- Ran benchmarks to collect data over nine months and found performance metrics corresponding to configurations of the system under analysis.

- Spent one month analyzing data to build a model used for performance optimization.

Sampling to discover optimal product configurations

- Explored various alternatives to expensive evolutionary searches by intelligent sampling.
- Approximated Principal Component Analysis to quickly prune the space of products.

Storm

- Storm is a Multi-Objective Optimization framework containing the latest state of the art optimization algorithms like NSGA-III and MOEA/D.
- It is implemented on Python and graphic rendering via MatPlotLib. → `http://tiny.cc/so8ixy`

Publications

- **Vivek Nair**, Zhe Yu, Tim Menzies, Norbert Siegmund, and Sven Apel. *Finding faster configurations using FLASH.* TSE (2018) → `http://tiny.cc/vivek_2018g`
- Chin-Jung Hsu, **Vivek Nair**, Vincent W. Freeh, Tim Menzies. *Micky: A Cheaper Alternative for Selecting Cloud Instances.* IEEE CLOUD (2018) → `http://tiny.cc/vivek_2018e`
- Chin-Jung Hsu, **Vivek Nair**, Vincent W. Freeh, Tim Menzies. *Low-Level Augmented Bayesian Optimization for Finding the Best Cloud VM.* ICDCS (2018) → `http://tiny.cc/vivek_2018a`
- **Vivek Nair**, Amritanshu Agrawal, Jianfeng Chen, Wei Fu, George Mathew, Tim Menzies, Leandro Minku, Markus Wagner, and Zhe Yu. *Data-Driven Search-based Software Engineering.* MSR (2018). → `http://tiny.cc/vivek_2018c`
- Jianfeng Chen, **Vivek Nair**, Rahul Krishna, Tim Menzies. *"Sampling" as a Baseline Optimizer for Search-based Software Engineering.* IEEE TSE 2018. → `http://tiny.cc/vivek_2016b`
- Jianfeng Chen, **Vivek Nair**, Tim Menzies. *Beyond Evolutionary Algorithms for Search-based Software Engineering* in IST 2017. → `http://tiny.cc/vivek_2017c`
- **Vivek Nair**, Tim Menzies, Norbert Seigmund, Sven Apel. *Using Bad Learners to find Good Configurations* in FSE 2017. → `http://tiny.cc/vivek_2017b`
- **Vivek Nair**, Tim Menzies, Norbert Seigmund, Sven Apel. *Faster Discovery of Faster System Configurations with Spectral Learning* in ASE Journal 2017. → `http://tiny.cc/vivek_2017a`
- **Vivek Nair**, Tim Menzies, Jianfeng Chen. *"An (Accidental) Exploration of Alternatives to Evolutionary Algorithms for SBSE"* in SSBSE 2016. → `http://tiny.cc/vivek_2016a`

Technical Skills

- Language: Python (4+ years), C (6+ years), Java (2+ years), ECL (2 years), C# (3 months)
- Tools: scikit-learn, pandas, numpy, auto-ml
- OS: Linux, Windows
- Version Control: git, Perforce

Honors and Awards

- Awarded for innovative use of HPCC Systems, 2017
- Awarded the 2^{nd} prize in HPCC System Poster Competition, 2016.
- Awarded the Employee of the Month in January 2012 at Samsung Engineering Lab
- Awarded Scholarship by the HRD Ministry for pursuing M.Tech