Ph.D. Preliminary Oral Defense

Tuesday, December 13th, 2016
223 Atanasoff Hall at 12:00 p.m.

Ganesha Upadhyaya

On Accelerating Big Code Analysis

Big Code analysis is about analyzing massive codebases. Big Code analysis has found numerous applications in defect prediction, discovery of programming patterns, specification inference, etc. Analyzing thousands of source code artifacts contained in Big Code, can be very expensive and needs to be accelerated. The current infrastructure support for Big Code analysis leverages parallelization techniques to accelerate. We believe understanding the interaction between the analysis and the source code artifact, and leveraging it, can further accelerate Big Code analysis. In this work, we present the notion of analysis-graph interactions and leverage it to perform a pre-analysis, that identifies the analysis relevant portions of the source code and performs program compaction to retain only the analysis relevant portions. In this way, running analysis on the reduced source code is much faster than running analysis on the original source code. We applied our technique to analyzing over 7 million control flow graphs (CFGs) and we achieved over 40% reduction in the analysis time. As part of our future work, we propose to apply the concept of analysis-graph interactions for program understanding and program characterization.
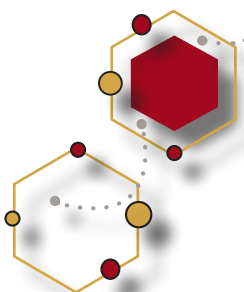
Major Professor:
Hridesh Rajan
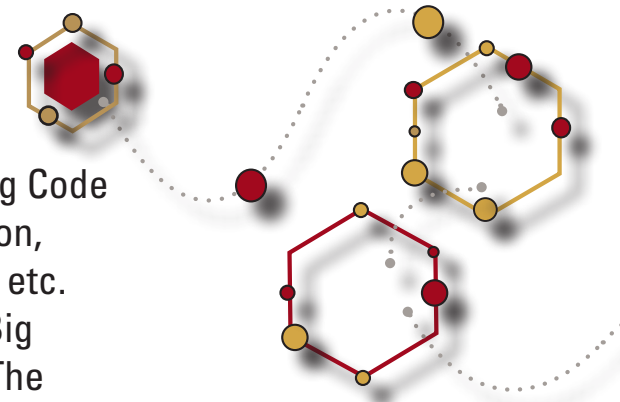
Committee Members:
Pavankumar Aduri

Wei Le

Robyn Lutz

Arun Somani, E CPE

IOWA STATE UNIVERSITY
Department of Computer Science

www.cs.iastate.edu