

# Student Research Seminar

## *Software Debugging through Dynamic Analysis of Program Structures*

**Mr. Zhang Zhenyu**

**Date & Time:**  
**December 23, 2009**  
**Wednesday**  
**3:30 pm**

**Venue:**  
**Room 313**  
**Chow Yei Ching Building**  
**The University of Hong Kong**

### **Abstract:**

Software debugging is difficult and time-consuming because developers do not know the locations of the faults in advance and they may require many failed executions to be produced in addition to the very first observed one. Modern software often provides error reporting for users to feedback the failure information to developers, opening a new door of using many low-cost failed executions statistically. Many existing fault-localization techniques correlate failures to fault positions. They not only ignore how infected program states propagate among statements, but also require passed executions to be available. Such unnecessary constraints limit the power of statistical approach to fault localization. Zhenyu Zhang's thesis work presents the results of my investigation to tackle these problems.

Zhenyu Zhang addresses the first problem, namely strong correlation not necessarily the root cause of the observed failures, both by developing a model to compute how infected program states statistically propagate along edges inside a control-flow graph and by distinguishing different short-circuiting evaluation that coarsely written statically as individual program statements. Specifically, he models the propagation traffic through each edge and apportion proportionally the probability of a block being faulty to its directly connected blocks, resulting in a set of homogenous equations, in which the probabilities of individual blocks being faulty after propagation of infected program states are the unknown. By solving the equation set, the model can determine the fault suspiciousness for individual statements accordingly. Empirical studies on real-life medium-size programs show that this technique is more effective than the state-of-the-art techniques. He also conducts an empirical study, which shows that differentiating the short-circuit evaluations of individual program predicates incurs relatively small performance overhead and significantly improves existing techniques.

Zhenyu Zhang addresses the second problem, naming the mandate of passed execution being available for statistical fault localization, by developing a technique that uses the trends of failures of individual program statements. The technique uses the execution count of each statement over each execution to categorize the execution, calculates the fraction of failed executions for each category. It treats every pair of such an execution count and the corresponding fraction as a point in a two-dimensional space. It then performs linear regression on these points, treats the slope of the obtained line as the signal and the fitting error as the noise to measure the suspiciousness of each statement. To make the technique independent to passed executions, for each statement, it approximates each above-mentioned fraction by the average of such fractions. Empirical study shows that this technique is both effective with and without passed executions.

Moreover, in his thesis work, Zhenyu Zhang empirically validates that using standard non-parametric hypothesis testing methods can achieve better effectiveness than the state-of-the-art predicate-based fault-localization techniques.

In conclusion, Zhenyu Zhang's thesis work contributes to software debugging by developing a family of effective statistical fault-localization techniques. They are particularly useful when the faulty statements are generally not strongly correlated to failures, program statements have many evaluation alternatives, passed executions can be unavailable, and distribution of program spectra cannot be assumed.

### **About the Speaker:**

Zhenyu Zhang is a Ph.D candidate in department of computer science at The University of Hong Kong, Pokfula, Hong Kong. He received his BEng and MSc degrees in computer science at Tsinghua University, Beijing, China. His current research interests are software debugging, software testing and analysis, continuous integration of software engineering activities, and engineering web services and service compositions. He has published his research results in many top-notch international journals and conferences. He received two best paper awards for his COMPSAC'08 paper and COMPSAC'09 paper, and a best paper nominee from his WWW'09 paper. He is in the program committee of ICSOFT'09, SEKE'10, SOSE'10, is a refereed reviewer of Journal of Systems and Software, and serves as external reviewers of many international journals and conferences.

**All are welcome!**  
**For enquiries, please call 2859-2180 or email [enquiry@cs.hku.hk](mailto:enquiry@cs.hku.hk)**  
**Department of Computer Science**  
**The University of Hong Kong**

