NATIONAL UNIVERSITY OF SINGAPORE

School of Computing

CS SEMINAR

Title: **Mining Input Grammars for Security**

Speaker: Andreas Zeller
Professor for Software Engineering
Saarland University

Date/Time: 8 September 2016, Thursday, 04:00 PM to 05:00 PM

Venue: Video Conference Room, COM1-02-13

Chaired by: Dr Roychoudhury, Abhik, Professor, School of Computing
(abhik@comp.nus.edu.sg)

Abstract:

Knowing which part of a program processes which parts of an input can reveal the structure of the input as well as the structure of the program. In a URL "http://www.example.com/path/", for instance, the protocol "http", the host "www.example.com", and the path "path" would be handled by different functions and stored in different variables. Given a set of sample inputs, we use _dynamic tainting_ to trace the data flow of each input character, and aggregate those input fragments that would be handled by the same function into lexical and syntactical entities. The result is a _context-free grammar_ that accurately reflects valid input structure; as it draws on function and variable names, it can be as readable as textbook examples.

In my talk, I show how our AUTOGRAM prototype derives such grammars automatically, and point out their uses in software engineering and security:
* They facilitate reverse engineering of input formats as well as manually writing valid test inputs;
* They produce high numbers of varied and valid inputs, thus facilitating automated robustness testing and fuzzing;
* Integrated into a checking parser, they protect existing programs against invalid, unexpected, and malicious inputs and behaviors.

This work was conducted with Matthias Hoschele and Konrad Jamrozik, presented at ASE 2016 (https://www.st.cs.uni saarland.de/models/autogram/) and ICSE 2016 (http://www.boxmate.org). It is part of the ERC SPECMATE project, funded by an ERC Advanced Grant.

Biodata:

Andreas Zeller is a full professor for Software Engineering at Saarland University in Saarbrucken, Germany, since 2001. His research concerns the analysis of large software systems and their development process. In 2010, Zeller was inducted as Fellow of the ACM for his contributions to automated debugging and mining software archives, for which he also was awarded 10-year impact awards from ACM SIGSOFT and ICSE. In 2011, he received an ERC Advanced Grant, Europe's highest and most prestigious individual research grant, for work on specification mining and test case generation. In 2013, Zeller co-founded Testfabrik AG, a start-up on automatic testing of Web applications, where he chairs the supervisory board.