2013 20th Working Conference on Reverse Engineering

(WCRE 2013)

Koblenz, Germany
14-17 October 2013
## Preface

Message from the Chairs ......................................................... iii
Committees ................................................................. v

## Invited Papers

Genetic Programming for Reverse Engineering (Invited Paper)
Mark Harman, William B. Langdon, and Westley Weimer — University College London, UK; University of Virginia, USA . 1

The First Decade of GUI Ripping: Extensions, Applications, and Broader Impacts (Invited Paper)
Atif Memon, Ishan Banerjee, Bao N. Nguyen, and Bryan Robbins — University of Maryland at College Park, USA . 11

Reverse Engineering in Industry (Panel Paper)
Ralf Lämmel — University of Koblenz-Landau, Germany . 21

## Main Research Track

### Binary Reverse Engineering

Who Allocated My Memory? Detecting Custom Memory Allocators in C Binaries
Xi Chen, Asia Slowinska, and Herbert Bos — VU University Amsterdam, Netherlands . 22

MemPick: High-Level Data Structure Detection in C/C++ Binaries
Istvan Haller, Asia Slowinska, and Herbert Bos — VU University Amsterdam, Netherlands . 32

Reconstructing Program Memory State from Multi-gigabyte Instruction Traces to Support Interactive Analysis
Brendan Cleary, Patrick Gorman, Eric Verbeek, Margaret-Anne Storey, Martin Salois, and Frederic Painchaud — University of Victoria, Canada; Defence R&D Canada, Canada . 42

Static Binary Rewriting without Supplemental Information: Overcoming the Tradeoff between Coverage and Correctness
Matthew Smithson, Khaled ElWazeer, Kapil Anand, Aparna Kotha, and Rajeev Barua — University of Maryland at College Park, USA . 52

### Bug Management

An Incremental Update Framework for Efficient Retrieval from Software Libraries for Bug Localization
Shivani Rao, Henry Medeiros, and Avinash Kak — Purdue University, USA . 62

Accurate Developer Recommendation for Bug Resolution
Xin Xia, David Lo, Xinyu Wang, and Bo Zhou — Zhejiang University, China; Singapore Management University, Singapore . 72

Has This Bug Been Reported?
Kaiping Liu, Hee Beng Kuan Tan, and Hongyu Zhang — Nanyang Technological University, Singapore; Tsinghua University, China . 82

Automatic Recovery of Root Causes from Bug-Fixing Changes
Ferdian Thung, David Lo, and Lingxiao Jiang — Singapore Management University, Singapore . 92
Clones

Distilling Useful Clones by Contextual Differencing
Zhenchang Xing, Yinxing Xue, and Stanislaw Jarzabek — Nanyang Technological University, Singapore; National University of Singapore, Singapore .................................................................................................................. 102

Effects of Cloned Code on Software Maintainability: A Replicated Developer Study
Debarshi Chatterji, Jeffrey C. Carver, Nicholas A. Kraft, and Jan Harder — University of Alabama, USA; University of Bremen, Germany .................................................................................................................. 112

Human Studies

The Influence of Non-technical Factors on Code Review
Olga Baysal, Oleksii Kononenko, Reid Holmes, and Michael W. Godfrey — University of Waterloo, Canada .................................................................................................................. 122

Understanding Project Dissemination on a Social Coding Site
Jing Jiang, Li Zhang, and Lei Li — Beihang University, China .................................................................................................................. 132

What Help Do Developers Seek, When and How?
Hongwei Li, Zhenchang Xing, Xin Peng, and Wenyun Zhao — Fudan University, China; Nanyang Technological University, Singapore .................................................................................................................. 142

Towards Understanding How Developers Spend Their Effort during Maintenance Activities
Zéphyrin Soh, Foutse Khomh, Yann-Gaël Guéhéneuc, and Giuliano Antoniol — Polytechnique Montréal, Canada .................................................................................................................. 152

Re-documenting Legacy Code

Leveraging Specifications of Subcomponents to Mine Precise Specifications of Composite Components
Ziying Dai, Xiaoguang Mao, Yan Lei, and Liqian Chen — National University of Defense Technology, China .................................................................................................................. 162

A Model-Driven Graph-Matching Approach for Design Pattern Detection
Mario Luca Bernardi, Marta Cimitile, and Giuseppe Antonio Di Luca — University of Sannio, Italy; Unitelma Sapienza University, Italy .................................................................................................................. 172

Recommendation Systems

Automated Library Recommendation
Ferdian Thung, David Lo, and Julia Lawall — Singapore Management University, Singapore; INRIA, France .................................................................................................................. 182

Automatic Discovery of Function Mappings between Similar Libraries
Cédric Teyton, Jean-Rémy Falleri, and Xavier Blanc — University of Bordeaux, France .................................................................................................................. 192

Find Your Library Experts
Cédric Teyton, Jean-Rémy Falleri, Floréal Morandat, and Xavier Blanc — University of Bordeaux, France .................................................................................................................. 202

Refactoring and Re-modularization

Towards Automatically Improving Package Structure while Respecting Original Design Decisions
Hani Abdeen, Houari Sahraoui, Osama Shata, Nicolas Anquetil, and Stéphane Ducasse — Qatar University, Qatar; Université de Montréal, Canada; INRIA, France .................................................................................................................. 212

Heuristics for Discovering Architectural Violations
Cristiano Maffort, Marco Tulio Valente, Mariza Bigonha, Nicolas Anquetil, and André Hora — UFMG, Brazil; INRIA, France .................................................................................................................. 222

Recommending Move Method Refactorings using Dependency Sets
Vitor Sales, Ricardo Terra, Luis Fernando Miranda, and Marco Tulio Valente — UFMG, Brazil; UFSJ, Brazil .................................................................................................................. 232

Do Developers Care about Code Smells? An Exploratory Survey
Aiko Yamashita and Leon Moonen — Mesan, Norway; Simula Research Laboratory, Norway .................................................................................................................. 242

Security and Testing

LigRE: Reverse-Engineering of Control and Data Flow Models for Black-Box XSS Detection
Fabien Duchêne, Sanjay Rawat, Jean-Luc Richier, and Roland Groz — LIG, France; Ensimag, France .................................................................................................................. 252

Circe: A Grammar-Based Oracle for Testing Cross-Site Scripting in Web Applications
Andrea Avancini and Mariano Ceccato — Fondazione Bruno Kessler, Italy .................................................................................................................. 262
## Capture-Replay vs. Programmable Web Testing: An Empirical Assessment during Test Case Evolution
Maurizio Leotta, Diego Clerissi, Filippo Ricca, and Paolo Tonella — University of Genova, Italy; Fondazione Bruno Kessler, Italy

## Software Maintenance

### Clustering Static Analysis Defect Reports to Reduce Maintenance Costs
Zachary P. Fry and Westley Weimer — University of Virginia, USA

### Lehman's Laws in Agile and Non-agile Projects
Kelley Durán, Gabbie Burns, and Paul Snell — Rochester Institute of Technology, USA

### Inferring Extended Finite State Machine Models from Software Executions
Neil Walkinshaw, Ramsay Taylor, and John Derrick — University of Leicester, UK; University of Sheffield, UK

### Comparing and Combining Evolutionary Couplings from Interactions and Commits
Fasil Bantelay, Motahareh Bahrani Zanjani, and Huzefa Kagdi — Wichita State University, USA

## Software Quality

### Improving SOA Antipatterns Detection in Service Based Systems by Mining Execution Traces
Mathieu Nayrolles, Naouel Moha, and Petko Valtchev — Université du Québec à Montréal, Canada

### Mining System Specific Rules from Change Patterns
André Hora, Nicolas Anquetil, Stéphane Ducasse, and Marco Tulio Valente — INRIA, France; University of Lille, France; UFMG, Brazil

### Empirical Evidence of Code Decay: A Systematic Mapping Study
Ajay Bandi, Byron J. Williams, and Edward B. Allen — Mississippi State University, USA

### Mining the Relationship between Anti-patterns Dependencies and Fault-Proneness
Fehmi Jaafar, Yann-Gaël Guéhéneuc, Sylvie Hamel, and Foutse Khomh — Polytechnique Montréal, Canada; Université de Montréal, Canada

## Traceability and Feature Location

### Leveraging Historical Co-change Information for Requirements Traceability
Nasir Ali, Fehmi Jaafar, and Ahmed E. Hassan — Queen's University, Canada; Université de Montréal, Canada

### Using Relationships for Matching Textual Domain Models with Existing Code
Raghavan Komondoor, Indrajit Bhattacharya, Deepak D’Souza, and Sachin Kale — Indian Institute of Science, India; IBM Research, India

### On the Effectiveness of Accuracy of Automated Feature Location Technique
Takashi Ishio, Shinpei Hayashi, Hiroshi Kazato, and Tsuyoshi Oshima — Osaka University, Japan; Tokyo Institute of Technology, Japan; NTT Data Intellilink, Japan; NTT, Japan

### On the Effect of Program Exploration on Maintenance Tasks
Zéphyrin Soh, Foutse Khomh, Yann-Gaël Guéhéneuc, Giuliano Antoniol, and Bram Adams — Polytechnique Montréal, Canada

## Practice Track

### Practice Papers I

#### Documenting APIs with Examples: Lessons Learned with the API Miner Platform
João Eduardo Montandon, Hudson Borges, Daniel Felix, and Marco Tulio Valente — UFMG, Brazil

#### Extracting Business Rules from COBOL: A Model-Based Framework
Valerio Cosentino, Jordi Cabot, Patrick Albert, Philippe Baquiel, and Jacques Perronnet — AtlanMod, France; IBM, France

#### Evaluating Architecture Stability of Software Projects
Lerina Aversano, Marco Molfetta, and Maria Tortorella — University of Sannio, Italy

#### Migrating a Large Scale Legacy Application to SOA: Challenges and Lessons Learned
Ravi Khadka, Amir Saeidi, Slinger Jansen, Jurriaan Hage, and Geer P. Haas — Utrecht University, Netherlands; IBM, Netherlands
Practice Papers II
Assessing the Complexity of Upgrading Software Modules
Bram Schoenmakers, Niels van Den Broek, Istvan Nagy, Bogdan Vasilescu, and Alexander Serebrenik — ASML, Netherlands; Eindhoven University of Technology, Netherlands ............................. 433

Analyzing PL/1 Legacy Ecosystems: An Experience Report
Erik Aeschlimann, Mircea Lungu, Oscar Nierstrasz, and Carl Worms — University of Bern, Switzerland; Credit Suisse, Switzerland ................................. 441

Psybo0t Malware: A Step-by-Step Decompilation Case Study
Lukáš Durfina, Jakub Kroustek, and Petr Zemek — Brno University of Technology, Czech Republic ................................. 449

ERA Track
Reusing Reused Code
Tomoya Ishihara, Keisuke Hotta, Yoshiki Higo, and Shinji Kusumoto — Osaka University, Japan ................................. 457

Specification Extraction by Symbolic Execution
Josef Pichler — Software Competence Center Hagenberg, Austria ................................. 462

An IDE-Based Context-Aware Meta Search Engine
Mohammad Masudur Rahman, Shamima Yeasmin, and Chanchal K. Roy — University of Saskatchewan, Canada ................................. 467

An Approach to Clone Detection in Behavioural Models
Elizabeth P. Antony, Manar H. Alalfi, and James R. Cordy — Queen’s University, Canada ................................. 472

Tool Demonstrations
MemBrush: A Practical Tool to Detect Custom Memory Allocators in C Binaries
Xi Chen, Asia Slowinska, and Herbert Bos — VU University Amsterdam, Netherlands ................................. 477

MemPick: A Tool for Data Structure Detection
Istvan Haller, Asia Slowinska, and Herbert Bos — VU University Amsterdam, Netherlands ................................. 479

Gelato: GEneric LAnguage TOols for Model-Driven Analysis of Legacy Software Systems
Amir Saeidi, Jurriaan Hage, Ravi Khadka, and Slinger Jansen — Utrecht University, Netherlands ................................. 481

Extracting Business Rules from COBOL: A Model-Based Tool
Valerio Cosentino, Jordi Cabot, Patrick Albert, Philippe Bauquel, and Jacques Perronnet — AtlanMod, France; IBM, France ................................. 483

Detecting Dependencies in Enterprise JavaBeans with SQuAVisiT
Alexandru Sutii, Serguei Roubtsov, and Alexander Serebrenik — Eindhoven University of Technology, Netherlands ................................. 485

REdiffs: Refactoring-Aware Difference Viewer for Java
Shinpei Hayashi, Sirinut Thangthumachit, and Motoshi Saeki — Tokyo Institute of Technology, Japan ................................. 487

CCCD: Concolic Code Clone Detection
Daniel E. Krutz and Emad Shihab — Rochester Institute of Technology, USA ................................. 489

Workshop Summaries
3rd Workshop on Mining Unstructured Data
Alberto Bacchelli, Nicolas Bettenburg, Latifa Guerrouj, and Sonia Haiduc — Delft University of Technology, Netherlands; Queen’s University, Canada; Polytechnique Montréal, Canada; Florida State University, USA ................................. 491

Workshop on Open and Original Problems in Software Language Engineering
Anya Helene Bagge and Vadim Zaytsev — University of Bergen, Norway; CWI, Netherlands ................................. 493

Author Index