



Montreal, March 2-6, 2015

SANER 2015

22nd IEEE International Conference
on Software Analysis, Evolution, and
Reengineering

génie :: informatique
génie :: logiciel

POLYTECHNIQUE MONTRÉAL

Google

IEEE
computer
society



REENGINEERING
FORUM

SAP

tcse

IEEE Technical Council
on Software Engineering

à la
Montréal

2015 IEEE
22nd International Conference
on Software Analysis,
Evolution, and Reengineering
(SANER)

Proceedings

Yann-Gaël Guéhéneuc, Bram Adams, and Alexander Serebrenik

March 2-6, 2015
Montréal, Canada

2015 IEEE

22nd International Conference on Software Analysis, Evolution, and Reengineering (SANER)

IEEE Catalog Number: CFP15102-ART

ISBN: 978-1-4799-8469-5

Copyright and Reprint Permission:

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Operations Center, 445 Hoes Lane, P.O. Box 1331, Piscataway, NJ 08855-1331.

All rights reserved.

Copyright © 2015 by the Institute of Electrical and Electronics Engineers

Contents

Frontmatter

Message from the Chairs	iii
Committees	viii

Keynotes

On Whose Shoulders? (Keynote)	
Jane Cleland-Huang — <i>DePaul University, USA</i>	1
Checkpoint Alpha (Keynote)	
Boris Debić — <i>Google, USA</i>	2

Main Research

Information Retrieval

Modeling the Evolution of Development Topics using Dynamic Topic Models	
Jiajun Hu, Xiaobing Sun, David Lo, and Bin Li — <i>Yangzhou University, China; Nanjing University, China; Singapore Management University, Singapore</i>	3
Understanding Developers' Natural Language Queries with Interactive Clarification	
Shihai Jiang, Liwei Shen, Xin Peng, Zhaojin Lv, and Wenyun Zhao — <i>Fudan University, China</i>	13

APIs and Patterns

Mining Multi-level API Usage Patterns	
Mohamed Aymen Saied, Omar Benomar, Hani Abdeen, and Houari Sahraoui — <i>Université de Montréal, Canada</i>	23
An Observational Study on API Usage Constraints and Their Documentation	
Mohamed Aymen Saied, Houari Sahraoui, and Bruno Dufour — <i>Université de Montréal, Canada</i>	33
Improving Pattern Tracking with a Language-Aware Tree Differencing Algorithm	
Nicolas Palix, Jean-Rémy Falleri, and Julia Lawall — <i>University of Grenoble, France; LaBRI, France; University of Bordeaux, France; INRIA, France</i>	43
Measuring the Quality of Design Pattern Detection Results	
Shouzheng Yang, Ayesha Manzer, and Vassilios Tzerpos — <i>York University, Canada</i>	53

Analysis of Programming Languages

Are PHP Applications Ready for Hack?	
Laleh Eshkevari, Fabien Dos Santos, James R. Cordy, and Giuliano Antoniol — <i>Polytechnique Montréal, Canada; Polytech Montpellier, France; Queen's University, Canada</i>	63
Does JavaScript Software Embrace Classes?	
Leonardo Humberto Silva, Miguel Ramos, Marco Tulio Valente, Alexandre Bergel, and Nicolas Anquetil — <i>Federal Institute of Northern Minas Gerais, Brazil; Federal University of Minas Gerais, Brazil; University of Chile, Chile; INRIA, France</i>	73
Evolution Analysis for Accessibility Excessiveness in Java	
Kazuo Kobori, Makoto Matsushita, and Katsuro Inoue — <i>NTT DATA, Japan; Osaka University, Japan</i>	83

A Software Quality Model for RPG Gergely Ladányi, Zoltán Tóth, Rudolf Ferenc, and Tibor Keresztesi — <i>University of Szeged, Hungary; R&R Software, Hungary</i>	91
---	----

On Crashes and Traces

JCHARMING: A Bug Reproduction Approach using Crash Traces and Directed Model Checking Mathieu Nayrolles, Abdelwahab Hamou-Lhadj, Sofiène Tahar, and Alf Larsson — <i>Concordia University, Canada; Ericsson, Sweden</i>	101
Towards a Common Metamodel for Traces of High Performance Computing Systems to Enable Software Analysis Tasks Luay Alawneh, Abdelwahab Hamou-Lhadj, and Jameleddine Hassine — <i>Jordan University of Science and Technology, Jordan; Concordia University, Canada; KFUPM, Saudi Arabia</i>	111
Automated Extraction of Failure Reproduction Steps from User Interaction Traces Tobias Roehm, Stefan Nosovic, and Bernd Bruegge — <i>TU München, Germany</i>	121
Misery Loves Company: CrowdStacking Traces to Aid Problem Detection Tommaso Dal Sasso, Andrea Mocci, and Michele Lanza — <i>University of Lugano, Switzerland</i>	131

Code Reviews

Who Should Review My Code? A File Location-Based Code-Reviewer Recommendation Approach for Modern Code Review Patanamon Thongtanunam, Chakkrit Tantithamthavorn, Raula Gaikovina Kula, Norihiro Yoshida, Hajimu Iida, and Kenichi Matsumoto — <i>NAIST, Japan; Osaka University, Japan; Nagoya University, Japan</i>	141
Code Review: Veni, ViDI, Vici Yuriy Tymchuk, Andrea Mocci, and Michele Lanza — <i>University of Lugano, Switzerland</i>	151
Would Static Analysis Tools Help Developers with Code Reviews? Sebastiano Panichella, Venera Arnaoudova, Massimiliano Di Penta, and Giuliano Antoniol — <i>University of Zurich, Switzerland; Polytechnique Montréal, Canada; University of Sannio, Italy</i>	161
Do Code Review Practices Impact Design Quality? A Case Study of the Qt, VTK, and ITK Projects Rodrigo Morales, Shane McIntosh, and Foutse Khomh — <i>Polytechnique Montréal, Canada; Queen's University, Canada</i>	171

Searching and Cloning

Scaling up Evaluation of Code Search Tools through Developer Usage Metrics Kostadin Damevski, David C. Shepherd, and Lori Pollock — <i>Virginia State University, USA; ABB, USA; University of Delaware, USA</i>	181
Optimized Feature Selection towards Functional and Non-functional Requirements in Software Product Lines Xiaoli Lian and Li Zhang — <i>Beihang University, China</i>	191
Threshold-Free Code Clone Detection for a Large-Scale Heterogeneous Java Repository Iman Keivanloo, Feng Zhang, and Ying Zou — <i>Queen's University, Canada</i>	201
Detecting Duplicate Bug Reports with Software Engineering Domain Knowledge Karan Aggarwal, Tanner Rutgers, Finbarr Timbers, Abram Hindle, Russ Greiner, and Eleni Stroulia — <i>University of Alberta, Canada</i>	211

Change Impact Analysis

Impact Analysis Based on a Global Hierarchical Object Graph Marwan Abi-Antoun, Yibin Wang, Ebrahim Khalaj, Andrew Giang, and Václav Rajlich — <i>Wayne State University, USA</i>	221
A Framework for Cost-Effective Dependence-Based Dynamic Impact Analysis Haipeng Cai and Raul Santelices — <i>University of Notre Dame, USA</i>	231
Circular Dependencies and Change-Proneness: An Empirical Study Tosin Daniel Oyetoyan, Jens Dietrich, Jean-Rémy Falleri, and Kamil Jezek — <i>NTNU, Norway; Massey University, New Zealand; LaBRI, France; University of Bordeaux, France; University of West Bohemia, Czech Republic</i>	241
An Empirical Study of Work Fragmentation in Software Evolution Tasks Heider Sanchez, Romain Robbes, and Victor M. Gonzalez — <i>University of Chile, Chile; ITAM, Mexico</i>	251

SCAM at SANER

Library Functions Identification in Binary Code by Using Graph Isomorphism Testings Jing Qiu, Xiaohong Su, and Peijun Ma — <i>Harbin Institute of Technology, China</i>	261
A Non-convex Abstract Domain for the Value Analysis of Binaries Sven Mattsen, Arne Wichmann, and Sibylle Schupp — <i>TU Hamburg, Germany</i>	271
Precision vs. Scalability: Context Sensitive Analysis with Prefix Approximation Raveendra Kumar Medicherla and Raghavan Komondoor — <i>Tata Consultancy Services, India; Indian Institute of Science, India</i>	281
MG++: Memory Graphs for Analyzing Dynamic Data Structures Vineet Singh, Rajiv Gupta, and Iulian Neamtiu — <i>University of California at Riverside, USA</i>	291

Mining Software Repositories

SQA-Profiles: Rule-Based Activity Profiles for Continuous Integration Environments Martin Brandtner, Sebastian C. Müller, Philipp Leitner, and Harald C. Gall — <i>University of Zurich, Switzerland</i>	301
Cross-Project Build Co-change Prediction Xin Xia, David Lo, Shane McIntosh, Emad Shihab, and Ahmed E. Hassan — <i>Zhejiang University, China; Singapore Management University, Singapore; Queen's University, Canada; Concordia University, Canada</i>	311
The Influence of App Churn on App Success and StackOverflow Discussions Latifa Guerrouj, Shams Azad, and Peter C. Rigby — <i>Concordia University, Canada</i>	321
Beyond Support and Confidence: Exploring Interestingness Measures for Rule-Based Specification Mining Tien-Duy B. Le and David Lo — <i>Singapore Management University, Singapore</i>	331

On Code Changes

Untangling Fine-Grained Code Changes Martín Dias, Alberto Bacchelli, Georgios Gousios, Damien Cassou, and Stéphane Ducasse — <i>INRIA, France; University of Lille, France; Delft University of Technology, Netherlands; Radboud University Nijmegen, Netherlands</i>	341
A Comprehensive and Scalable Method for Analyzing Fine-Grained Source Code Change Patterns Masatomo Hashimoto, Akira Mori, and Tomonori Izumida — <i>RIKEN Advanced Institute for Computational Science, Japan; National Institute of Advanced Industrial Science and Technology, Japan</i>	351
Summarizing Evolutionary Trajectory by Grouping and Aggregating Relevant Code Changes Qingtao Jiang, Xin Peng, Hai Wang, Zhenchang Xing, and Wenyun Zhao — <i>Fudan University, China; Nanyang Technological University, Singapore</i>	361
Identifying the Exact Fixing Actions of Static Rule Violation Hayatou Oumarou, Nicolas Anquetil, Anne Etien, Stéphane Ducasse, and Kolyang Dina Taiwe — <i>University of Maroua, Cameroon; INRIA, France; University of Lille, France</i>	371

The Human Within

CloCom: Mining Existing Source Code for Automatic Comment Generation Edmund Wong, Taiyue Liu, and Lin Tan — <i>University of Waterloo, Canada</i>	380
amAssist: In-IDE Ambient Search of Online Programming Resources Hongwei Li, Xuejiao Zhao, Zhenchang Xing, Lingfeng Bao, Xin Peng, Dongjing Gao, and Wenyun Zhao — <i>Fudan University, China; Jiangxi Normal University, China; Nanyang Technological University, Singapore; Zhejiang University, China</i>	390
Reverse Engineering Time-Series Interaction Data from Screen-Captured Videos Lingfeng Bao, Jing Li, Zhenchang Xing, Xinyu Wang, and Bo Zhou — <i>Zhejiang University, China; Nanyang Technological University, Singapore</i>	399
Niche vs. Breadth: Calculating Expertise over Time through a Fine-Grained Analysis Jose Ricardo da Silva Junior, Esteban Clua, Leonardo Murta, and Anita Sarma — <i>Federal Fluminense University, Brazil; University of Nebraska-Lincoln, USA</i>	409

Search, Touch, Tweet

Protecting Web Applications via Unicode Extension Boze Zekan, Mark Shtern, and Vassilios Tzerpos — <i>York University, Canada</i>	419
A Search-Based Approach to Multi-view Clustering of Software Systems Amir M. Saeidi, Jurriaan Hage, Ravi Khadka, and Slinger Jansen — <i>Utrecht University, Netherlands</i>	429

CEL: Touching Software Modeling in Essence Remo Lemma, Michele Lanza, and Andrea Mocchi — <i>University of Lugano, Switzerland</i>	439
NIRMAL: Automatic Identification of Software Relevant Tweets Leveraging Language Model Abhishek Sharma, Yuan Tian, and David Lo — <i>Singapore Management University, Singapore</i>	449

Tool Demonstrations

A Static Code Analysis Tool for Control System Software Sreeja Nair, Raoul Jetley, Anil Nair, and Stefan Hauck-Stattelmann — <i>ABB Research, India; ABB Research, Germany</i>	459
RbG: A Documentation Generator for Scientific and Engineering Software Michael Moser, Josef Pichler, Günter Fleck, and Michael Wiltatschil — <i>Software Competence Center Hagenberg, Austria; Siemens Transformers Austria, Austria</i>	464
Historef: A Tool for Edit History Refactoring Shinpei Hayashi, Daiki Hoshino, Jumpei Matsuda, Motoshi Saeki, Takayuki Omori, and Katsuhisa Maruyama — <i>Tokyo Institute of Technology, Japan; Ritsumeikan University, Japan</i>	469
ClonePacker: A Tool for Clone Set Visualization Hiroaki Murakami, Yoshiki Higo, and Shinji Kusumoto — <i>Osaka University, Japan</i>	474
GiLA: GitHub Label Analyzer Javier Luis Cánovas Izquierdo, Valerio Cosentino, Belén Rolandi, Alexandre Bergel, and Jordi Cabot — <i>AtlanMod, France; University of Chile, Chile</i>	479
SPCP-Miner: A Tool for Mining Code Clones That Are Important for Refactoring or Tracking Manishankar Mondal, Chanchal K. Roy, and Kevin A. Schneider — <i>University of Saskatchewan, Canada</i>	484
TracerJD: Generic Trace-Based Dynamic Dependence Analysis with Fine-Grained Logging Haipeng Cai and Raul Santelices — <i>University of Notre Dame, USA</i>	489
Umple: A Framework for Model Driven Development of Object-Oriented Systems Miguel A. Garzón, Hamoud Aljamaan, and Timothy C. Lethbridge — <i>University of Ottawa, Canada</i>	494
Assessing the Bus Factor of Git Repositories Valerio Cosentino, Javier Luis Cánovas Izquierdo, and Jordi Cabot — <i>AtlanMod, France</i>	499

Industrial Research

Old Habits Die Hard: Why Refactoring for Understandability Does Not Give Immediate Benefits Erik Ammerlaan, Wim Veninga, and Andy Zaidman — <i>Exact International Development, Netherlands; Delft University of Technology, Netherlands</i>	504
Bash2py: A Bash to Python Translator Ian J. Davis, Mike Wexler, Cheng Zhang, Richard C. Holt, and Theresa Weber — <i>University of Waterloo, Canada; Owl Computing Technologies, USA</i>	508
On Implementational Variations in Static Analysis Tools Tukaram Muske and Prasad Bokil — <i>Tata Consultancy Services, India</i>	512
Tracking Known Security Vulnerabilities in Proprietary Software Systems Mircea Cadariu, Eric Bouwers, Joost Visser, and Arie van Deursen — <i>Software Improvement Group, Netherlands; Delft University of Technology, Netherlands; Radboud University Nijmegen, Netherlands</i>	516

Early Research Achievements

Evolution and Reuse

Trusting a Library: A Study of the Latency to Adopt the Latest Maven Release Raula Gaikovina Kula, Daniel M. German, Takashi Ishio, and Katsuro Inoue — <i>Osaka University, Japan; University of Victoria, Canada</i>	520
---	-----

Evolution of Dynamic Feature Usage in PHP Mark Hills — <i>East Carolina University, USA</i>	525
Towards Incremental Model Slicing for Delta-Oriented Software Product Lines Sascha Lity, Hauke Baller, and Ina Schaefer — <i>TU Braunschweig, Germany</i>	530
Understanding Software Performance Regressions using Differential Flame Graphs Cor-Paul Bezemer, Johan Pouwelse, and Brendan Gregg — <i>Delft University of Technology, Netherlands; Netflix, USA</i>	535

Text and Labeling

TextRank Based Search Term Identification for Software Change Tasks Mohammad Masudur Rahman and Chanchal K. Roy — <i>University of Saskatchewan, Canada</i>	540
Query Expansion via Wordnet for Effective Code Search Meili Lu, Xiaobing Sun, Shaowei Wang, David Lo, and Yucong Duan — <i>Yangzhou University, China; Nanjing University, China; Singapore Management University, Singapore; Hainan University, China</i>	545
Exploring the Use of Labels to Categorize Issues in Open-Source Software Projects Jordi Cabot, Javier Luis Cánovas Izquierdo, Valerio Cosentino, and Belén Rolandi — <i>AtlanMod, France</i>	550
Explore the Evolution of Development Topics via On-Line LDA Jiajun Hu, Xiaobing Sun, and Bin Li — <i>Yangzhou University, China; Nanjing University, China</i>	555

Bugs and Violations

Code Coverage and Test Suite Effectiveness: Empirical Study with Real Bugs in Large Systems Pavneet Singh Kochhar, Ferdian Thung, and David Lo — <i>Singapore Management University, Singapore</i>	560
Detection of Violation Causes in Reflexion Models Sebastian Herold, Michael English, Jim Buckley, Steve Counsell, and Mel Ó Cinnéide — <i>Lero, Ireland; University of Limerick, Ireland; Brunel University, UK; University College Dublin, Ireland</i>	565
A Comparative Study on the Effectiveness of Part-of-Speech Tagging Techniques on Bug Reports Yuan Tian and David Lo — <i>Singapore Management University, Singapore</i>	570

Static and Dynamic Analysis

Efficiently Identifying Object Production Sites Alejandro Infante and Alexandre Bergel — <i>University of Chile, Chile</i>	575
Where Was This SQL Query Executed? A Static Concept Location Approach Csaba Nagy, Loup Meurice, and Anthony Cleve — <i>University of Namur, Belgium</i>	580
Taint Analysis of Manual Service Compositions using Cross-Application Call Graphs Marc-André Laverdière, Bernhard J. Berger, and Ettore Merlo — <i>Tata Consultancy Services, Canada; Polytechnique Montréal, Canada; University of Bremen, Germany</i>	585

Tutorials and Briefings

TXL Source Transformation in Practice James R. Cordy — <i>Queen's University, Canada</i>	590
Software Risk Management in Practice: Shed Light on Your Software Product Jens Knodel, Matthias Naab, Eric Bouwers, and Joost Visser — <i>Fraunhofer IESE, Germany; Software Improvement Group, Netherlands; Radboud University Nijmegen, Netherlands</i>	592
Software Architecture Reconstruction: Why? What? How? Mehdi Mirakhorli — <i>Rochester Institute of Technology, USA</i>	595

Doctoral Symposium

SKilled Communication for Toolchains Timm Felden — <i>University of Stuttgart, Germany</i>	596
The Impact of Column-Orientation on the Quality of Class Inheritance Mapping Specifications Martin Lorenz — <i>HPI, Germany</i>	597

Improving the Integration Process of Large Software Systems	
Yujuan Jiang — <i>Polytechnique Montréal, Canada</i>	598
Handling the Differential Evolution of Software Artefacts: A Framework for Consistency Management	
Ildiko Pete and Dharini Balasubramaniam — <i>University of St. Andrews, UK</i>	599
Towards a Framework for Analysis, Transformation, and Manipulation of Makefiles	
Doug Martin — <i>Queen's University, Canada</i>	601
Towards a Framework for Automatic Correction of Anti-patterns	
Rodrigo Morales — <i>Polytechnique Montréal, Canada</i>	603
Towards an Ontology-Based Context-Aware Meta-Model for the Software Domain	
Mostafa Erfani — <i>Concordia University, Canada</i>	605
Investigating Modern Release Engineering Practices	
Md Tajmilur Rahman — <i>Concordia University, Canada</i>	607

Author Index