

# **2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER 2018)**

**Campobasso, Italy  
20-23 March 2018**



**IEEE Catalog Number: CFP18102-POD  
ISBN: 978-1-5386-4970-1**

**Copyright © 2017 by the Institute of Electrical and Electronics Engineers, Inc.  
All Rights Reserved**

*Copyright and Reprint Permissions:* 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 Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP18102-POD
ISBN (Print-On-Demand):	978-1-5386-4970-1
ISBN (Online):	978-1-5386-4969-5
ISSN:	1534-5351

**Additional Copies of This Publication Are Available From:**

Curran Associates, Inc  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: (845) 758-0400  
Fax: (845) 758-2633  
E-mail: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# Contents

## Frontmatter

---

Message from the Chairs . . . . .	iii
SANER 2018 Organization . . . . .	v
SANER 2018 Sponsors and Supporters . . . . .	xii

## Keynotes

---

<b>A Decade of Software Quality Analysis in Practice: Surprises, Anecdotes, and Lessons Learned (Keynote)</b> Elmar Juergens — <i>CQSE, Germany</i> . . . . .	1
<b>Towards a New Digital Business Operating System: Speed, Data, Ecosystems, and Empowerment (Keynote)</b> Jan Bosch — <i>Chalmers University of Technology, Sweden</i> . . . . .	2
<b>Compilers Are Sprinters – IDEs Are Marathoners (Keynote)</b> Peter Gromov — <i>JetBrains, Germany</i> . . . . .	3

## Retrospective Papers

---

<b>Ten Years of JDeodorant: Lessons Learned from the Hunt for Smells</b> Nikolaos Tsantalis, Theodoros Chaikalas, and Alexander Chatzigeorgiou — <i>Concordia University, Canada; University of Macedonia, Greece</i> . . . . .	4
<b>Design Patterns Impact on Software Quality: Where Are the Theories?</b> Foutse Khomh and Yann-Gaël Guéhéneuc — <i>Polytechnique Montréal, Canada; Concordia University, Canada</i> . . . . .	15
<b>Benchmarks for Software Clone Detection: A Ten-Year Retrospective</b> Chanchal K. Roy and James R. Cordy — <i>University of Saskatchewan, Canada; Queen's University, Canada</i> . . . . .	26

## Technical Research Papers

---

### Program Analysis

<b>Context Is King: The Developer Perspective on the Usage of Static Analysis Tools</b> Carmine Vassallo, Sebastiano Panichella, Fabio Palomba, Sebastian Proksch, Andy Zaidman, and Harald C. Gall — <i>University of Zurich, Switzerland; University of Zurich, Netherlands; Delft University of Technology, Netherlands</i> . . . . .	38
<b>Micro-clones in Evolving Software</b> Manishankar Mondal, Chanchal K. Roy, and Kevin A. Schneider — <i>University of Saskatchewan, Canada</i> . . . . .	50

### Software Logging

<b>SMARTLOG: Place Error Log Statement by Deep Understanding of Log Intention</b> Zhouyang Jia, Shanshan Li, Xiaodong Liu, Xiangke Liao, and Yunhuai Liu — <i>National University of Defense Technology, China; Peking University, China</i> . . . . .	61
---	----

### Testing

<b>Exploring the Integration of User Feedback in Automated Testing of Android Applications</b> Giovanni Grano, Adelina Ciurumelea, Sebastiano Panichella, Fabio Palomba, and Harald C. Gall — <i>University of Zurich, Switzerland</i> . . . . .	72
<b>Structured Random Differential Testing of Instruction Decoders</b> Nathan Jay and Barton P. Miller — <i>University of Wisconsin-Madison, USA</i> . . . . .	84
<b>Clustering Support for Inadequate Test Suite Reduction</b> Carmen Coviello, Simone Romano, Giuseppe Scanniello, Alessandro Marchetto, Giuliano Antoniol, and Anna Corazza — <i>University of Basilicata, Italy; Polytechnique Montréal, Canada; Federico II University of Naples, Italy</i> . . . . .	95

## Program Repair

- Automatically Repairing Dependency-Related Build Breakage**  
Christian Macho, Shane McIntosh, and Martin Pinzger — *University of Klagenfurt, Austria; McGill University, Canada* . . . . . 106
- Mining StackOverflow for Program Repair**  
Xuliang Liu and Hao Zhong — *Shanghai Jiao Tong University, China* . . . . . 118
- Dissection of a Bug Dataset: Anatomy of 395 Patches from Defects4J**  
Victor Sobreira, Thomas Durieux, Fernanda Madeiral, Martin Monperrus, and Marcelo de Almeida Maia — *Federal University of Uberlândia, Brazil; University of Lille, France; KTH, Sweden* . . . . . 130

## Mobile Development

- Detecting Third-Party Libraries in Android Applications with High Precision and Recall**  
Yuan Zhang, Jiarun Dai, Xiaohan Zhang, Sirong Huang, Zhemin Yang, Min Yang, and Hao Chen — *Fudan University, China; University of California at Davis, USA* . . . . . 141

## Software Quality

- How Do Developers Fix Issues and Pay Back Technical Debt in the Apache Ecosystem?**  
Georgios Digkas, Mircea Lungu, Paris Avgeriou, Alexander Chatzigeorgiou, and Apostolos Ampatzoglou — *University of Groningen, Netherlands; University of Macedonia, Greece* . . . . . 153
- How Good Is Your Puppet? An Empirically Defined and Validated Quality Model for Puppet**  
Eduard van der Bent, Jurriaan Hage, Joost Visser, and Georgios Gousios — *Utrecht University, Netherlands; Software Improvement Group, Netherlands; Delft University of Technology, Netherlands* . . . . . 164

## Behavior and Runtime Analysis

- Maintaining Behaviour Driven Development Specifications: Challenges and Opportunities**  
Leonard Peter Binamungu, Suzanne M. Embury, and Nikolaos Konstantinou — *University of Manchester, UK* . . . . . 175
- Recursion Aware Modeling and Discovery for Hierarchical Software Event Log Analysis**  
Maikel Leemans, Wil M. P. van der Aalst, and Mark G. J. van den Brand — *Eindhoven University of Technology, Netherlands* . . . . . 185

## Design Analysis

- Automatically Exploiting Implicit Design Knowledge When Solving the Class Responsibility Assignment Problem**  
Yongrui Xu, Peng Liang, and Muhammad Ali Babar — *Wuhan University, China; University of Adelaide, Australia* . . . . . 197

## Defect Prediction

- Cross-Version Defect Prediction via Hybrid Active Learning with Kernel Principal Component Analysis**  
Zhou Xu, Jin Liu, Xiapu Luo, and Tao Zhang — *Wuhan University, China; Hong Kong Polytechnic University, China; Harbin Engineering University, China* . . . . . 209
- Using a Probabilistic Model to Predict Bug Fixes**  
Mauricio Soto and Claire Le Goues — *Carnegie Mellon University, USA* . . . . . 221
- Connecting Software Metrics across Versions to Predict Defects**  
Yibin Liu, Yanhui Li, Jianbo Guo, Yuming Zhou, and Baowen Xu — *Nanjing University, China; Tsinghua University, China* . . . . . 232

## APIs

- Classifying Stack Overflow Posts on API Issues**  
Md Ahasanuzzaman, Muhammad Asaduzzaman, Chanchal K. Roy, and Kevin A. Schneider — *Queen's University, Canada; University of Saskatchewan, Canada* . . . . . 244
- Why and How Java Developers Break APIs**  
Aline Brito, Laerte Xavier, Andre Hora, and Marco Tulio Valente — *Federal University of Minas Gerais, Brazil* . . . . . 255
- Mining Accurate Message Formats for Service APIs**  
Md Arafat Hossain, Steve Versteeg, Jun Han, Muhammad Ashad Kabir, Jiaojiao Jiang, and Jean-Guy Schneider — *Swinburne University of Technology, Australia; CA Technologies, Australia* . . . . . 266

## Exploring Code Bases

- Mining Framework Usage Graphs from App Corpora**  
Sergio Mover, Sriram Sankaranarayanan, Rhys Braginton Pettee Olsen, and Bor-Yuh Evan Chang — *University of Colorado at Boulder, USA* . . . . . 277

<b>A Generalized Model for Visualizing Library Popularity, Adoption, and Diffusion within a Software Ecosystem</b>	
Raula Gaikovina Kula, Coen De Roover, Daniel M. German, Takashi Ishio, and Katsuro Inoue — <i>Osaka University, Japan; Vrije Universiteit Brussel, Belgium; University of Victoria, Canada; NAIST, Japan</i>	288
<b>Supporting Exploratory Code Search with Differencing and Visualization</b>	
Wenjian Liu, Xin Peng, Zhenchang Xing, Junyi Li, Bing Xie, and Wenyun Zhao — <i>Fudan University, China; Australian National University, Australia; Peking University, China</i>	300
<b>Language Models</b>	
<b>Syntax and Sensibility: Using Language Models to Detect and Correct Syntax Errors</b>	
Eddie Antonio Santos, Joshua Charles Campbell, Dhvani Patel, Abram Hindle, and José Nelson Amaral — <i>University of Alberta, Canada</i>	311
<b>A Deep Neural Network Language Model with Contexts for Source Code</b>	
Anh Tuan Nguyen, Trong Duc Nguyen, Hung Dang Phan, and Tien N. Nguyen — <i>Iowa State University, USA; University of Texas at Dallas, USA</i>	323
<b>Binary Analysis</b>	
<b>Efficient Features for Function Matching between Binary Executables</b>	
Chariton Karamitas and Athanasios Kehagias — <i>CENSUS, Greece; University of Thessaloniki, Greece</i>	335
<b>Using Recurrent Neural Networks for Decompilation</b>	
Deborah S. Katz, Jason Ruchti, and Eric Schulte — <i>Carnegie Mellon University, USA; GrammaTech, USA</i>	346
<b>Developers' Collaboration</b>	
<b>How Do Developers Discuss Rationale?</b>	
Rana Alkadhi, Manuel Nonnenmacher, Emitza Guzman, and Bernd Bruegge — <i>TU Munich, Germany; University of Zurich, Switzerland</i>	357
<b>Automated Quality Assessment for Crowdsourced Test Reports of Mobile Applications</b>	
Xin Chen, He Jiang, Xiaochen Li, Tieke He, and Zhenyu Chen — <i>Dalian University of Technology, China; Nanjing University, China</i>	368
<b>Refactoring</b>	
<b>The Impact of Refactoring Changes on the SZZ Algorithm: An Empirical Study</b>	
Edmilson Campos Neto, Daniel Alencar da Costa, and Uirá Kulesza — <i>Federal University of Rio Grande do Norte, Brazil; Queen's University, Canada</i>	380
<b>An Extensible Approach for Taming the Challenges of JavaScript Dead Code Elimination</b>	
Niels Groot Obbink, Ivano Malavolta, Gian Luca Scoccia, and Patricia Lago — <i>VU University Amsterdam, Netherlands; Gran Sasso Science Institute, Italy</i>	391
<b>Automated Refactoring of Client-Side JavaScript Code to ES6 Modules</b>	
Aikaterini Paltoglou, Vassilis E. Zafeiris, E. A. Giakoumakis, and N. A. Diamantidis — <i>Athens University of Economics and Business, Greece</i>	402
<b>Recommender Systems</b>	
<b>Improving Developers Awareness of the Exception Handling Policy</b>	
Taiza Montenegro, Hugo Melo, Roberta Coelho, and Eiji Barbosa — <i>Federal University of Rio Grande do Norte, Brazil</i>	413
<b>Detecting Faulty Empty Cells in Spreadsheets</b>	
Liang Xu, Shuo Wang, Wensheng Dou, Bo Yang, Chushu Gao, Jun Wei, and Tao Huang — <i>University at Chinese Academy of Sciences, China; Institute of Software at Chinese Academy of Sciences, China; North China University of Technology, China</i>	423
<b>Software Security</b>	
<b>Detection of Protection-Impacting Changes during Software Evolution</b>	
Marc-André Laverdière and Ettore Merlo — <i>Tata Consultancy Services, Canada; Polytechnique Montréal, Canada</i>	434
<b>Mining Sandboxes: Are We There Yet?</b>	
Lingfeng Bao, Tien-Duy B. Le, and David Lo — <i>Singapore Management University, Singapore</i>	445
<b>DeepWeak: Reasoning Common Software Weaknesses via Knowledge Graph Embedding</b>	
Zhuobing Han, Xiaohong Li, Hongtao Liu, Zhenchang Xing, and Zhiyong Feng — <i>Tianjin University, China; Australian National University, Australia</i>	456

# Journal-First Abstracts

---

<b>Towards Just-in-Time Suggestions for Log Changes (Journal-First Abstract)</b>	
Heng Li, Weiyi Shang, Ying Zou, and Ahmed E. Hassan — <i>Queen's University, Canada; Concordia University, Canada</i> . . . . .	467
<b>Which Log Level Should Developers Choose for a New Logging Statement? (Journal-First Abstract)</b>	
Heng Li, Weiyi Shang, and Ahmed E. Hassan — <i>Queen's University, Canada; Concordia University, Canada</i> . . . . .	468
<b>A Study of the Relation of Mobile Device Attributes with the User-Perceived Quality of Android Apps (Journal-First Abstract)</b>	
Ehsan Noei, Mark D. Syer, Ying Zou, Ahmed E. Hassan, and Iman Keivanloo — <i>Queen's University, Canada</i> . . . . .	469
<b>How Developers Micro-Optimize Android Apps (Journal-First Abstract)</b>	
Mario Linares-Vásquez, Christopher Vendome, Michele Tufano, and Denys Poshyvanyk — <i>Universidad de los Andes, Colombia; College of William and Mary, USA</i> . . . . .	470
<b>The Relationship between Evolutionary Coupling and Defects in Large Industrial Software (Journal-First Abstract)</b>	
Serkan Kirbas, Bora Caglayan, Tracy Hall, Steve Counsell, David Bowes, Alper Sen, and Ayse Bener — <i>Bloomberg, UK; Ryerson University, Canada; Brunel University London, UK; University of Hertfordshire, UK; Boğaziçi University, Turkey</i> . . . . .	471
<b>A Comparison Framework for Runtime Monitoring Approaches (Journal-First Abstract)</b>	
Rick Rabiser, Sam Guinea, Michael Vierhauser, Luciano Baresi, and Paul Grünbacher — <i>JKU Linz, Austria; Politecnico di Milano, Italy; University of Notre Dame, USA</i> . . . . .	472
<b>Modularity and Architecture of PLC-Based Software for Automated Production Systems: An Analysis in Industrial Companies (Journal-First Abstract)</b>	
Birgit Vogel-Heuser, Juliane Fischer, Stefan Feldmann, Sebastian Ulewicz, and Susanne Rösch — <i>TU Munich, Germany</i> . . . . .	473
<b>A Mapping Study on Design-Time Quality Attributes and Metrics (Journal-First Abstract)</b>	
Elvira Maria Arvanitou, Apostolos Ampatzoglou, Alexander Chatzigeorgiou, Matthias Galster, and Paris Avgeriou — <i>University of Groningen, Netherlands; University of Macedonia, Greece; University of Canterbury, New Zealand</i> . . . . .	474
<b>Review Participation in Modern Code Review: An Empirical Study of the Android, Qt, and OpenStack Projects (Journal-First Abstract)</b>	
Patanamon Thongtanunam, Shane McIntosh, Ahmed E. Hassan, and Hajimu Iida — <i>University of Adelaide, Australia; McGill University, Canada; Queen's University, Canada; NAIIST, Japan</i> . . . . .	475
<b>Spreadsheet Guardian: An Approach to Protecting Semantic Correctness throughout the Evolution of Spreadsheets (Journal-First Abstract)</b>	
Daniel Kulesz, Verena Käfer, and Stefan Wagner — <i>University of Stuttgart, Germany</i> . . . . .	476

# ERA Track

---

<b>Extracting Features from Requirements: Achieving Accuracy and Automation with Neural Networks</b>	
Yang Li, Sandro Schulze, and Gunter Saake — <i>Otto von Guericke University Magdeburg, Germany; University of Magdeburg, Germany</i> . . . . .	477
<b>OctoBubbles: A Multi-view Interactive Environment for Concurrent Visualization and Synchronization of UML Models and Code</b>	
Rodi Jolak, Khanh-Duy Le, Kaan Burak Sener, and Michel R. V. Chaudron — <i>Chalmers University of Technology, Sweden; National Research University, Russia</i> . . . . .	482
<b>A Comparison of Software Engineering Domain Specific Sentiment Analysis Tools</b>	
Md. Rakibul Islam and Minhaz F. Zibran — <i>University of New Orleans, USA</i> . . . . .	487
<b>Generating Descriptions for Screenshots to Assist Crowdsourced Testing</b>	
Di Liu, Xiaofang Zhang, Yang Feng, and James A. Jones — <i>Soochow University, China; University of California at Irvine, USA</i> . . . . .	492
<b>Reconciling the Past and the Present: An Empirical Study on the Application of Source Code Transformations to Automatically Rejuvenate Java Programs</b>	
Reno Dantas, Antônio Carvalho Júnior, Diego Marcílio, Luísa Fantin, Uriel Silva, Walter Lucas, and Rodrigo Bonifácio — <i>University of Brasília, Brazil</i> . . . . .	497

# Tool Demos

---

## Mining

<b>The Statechart Workbench: Enabling Scalable Software Event Log Analysis using Process Mining</b>	
Maikel Leemans, Wil M. P. van der Aalst, and Mark G. J. van den Brand — <i>Eindhoven University of Technology, Netherlands</i> . . . . .	502

<b>APIDiff: Detecting API Breaking Changes</b>	
Aline Brito, Laerte Xavier, Andre Hora, and Marco Tulio Valente — <i>Federal University of Minas Gerais, Brazil; Federal University of Mato Grosso do Sul, Brazil</i> . . . . .	507
<b>LICCA: A Tool for Cross-Language Clone Detection</b>	
Tijana Vislavski, Gordana Rakić, Nicolás Cardozo, and Zoran Budimac — <i>University of Novi Sad, Serbia; Universidad de los Andes, Colombia</i> . . . . .	512
<b>GoldRusher: A Miner for Rapid Identification of Hidden Code</b>	
Aleiieldin Salem — <i>TU Munich, Germany</i> . . . . .	517

## Software Evolution

<b>BECLoMA: Augmenting Stack Traces with User Review Information</b>	
Lucas Pelloni, Giovanni Grano, Adelina Ciurumelea, Sebastiano Panichella, Fabio Palomba, and Harald C. Gall — <i>University of Zurich, Switzerland</i> . . . . .	522
<b>Bring Your Own Coding Style</b>	
Naoto Ogura, Shinsuke Matsumoto, Hideaki Hata, and Shinji Kusumoto — <i>Osaka University, Japan; NAIST, Japan</i> . . . . .	527
<b>FINALsT<sup>2</sup>: Feature Identification, Localization, and Tracing Tool</b>	
Andreas Burger and Sten Grüner — <i>ABB, Germany</i> . . . . .	532
<b>ChangeMacroRecorder: Recording Fine-Grained Textual Changes of Source Code</b>	
Katsuhisa Maruyama, Shinpei Hayashi, and Takayuki Omori — <i>Ritsumeikan University, Japan; Tokyo Institute of Technology, Japan</i> . . . . .	537
<b>RETICULA: Real-Time Code Quality Assessment</b>	
Luigi Frunzio, Bin Lin, Michele Lanza, and Gabriele Bavota — <i>University of Lugano, Switzerland</i> . . . . .	542

## Industry Track

---

### Reengineering

<b>Reengineering an Industrial HMI: Approach, Objectives, and Challenges</b>	
Bernhard Dorninger, Michael Moser, and Albin Kern — <i>Software Competence Center Hagenberg, Austria; ENGEL AUSTRIA, Austria</i> . . . . .	547
<b>Model-Based Software Restructuring: Lessons from Cleaning Up COM Interfaces in Industrial Legacy Code</b>	
Dennis Dams, Arjan Mooij, Pepijn Kramer, Andrei Rădulescu, and Jaromír Vaňhara — <i>ESI, Netherlands; Thermo Fisher Scientific, Netherlands</i> . . . . .	552
<b>Grammatical Inference from Data Exchange Files: An Experiment on Engineering Software</b>	
Markus Exler, Michael Moser, Josef Pichler, Günter Fleck, and Bernhard Dorninger — <i>Software Competence Center Hagenberg, Austria; Siemens, Austria</i> . . . . .	557

### Development and Testing

<b>Fuzz Testing in Practice: Obstacles and Solutions</b>	
Jie Liang, Mingzhe Wang, Yuanliang Chen, Yu Jiang, and Renwei Zhang — <i>Tsinghua University, China; University of Illinois at Urbana-Champaign, USA; Huawei, China</i> . . . . .	562
<b>Diggit: Automated Code Review via Software Repository Mining</b>	
Robert Chatley and Lawrence Jones — <i>Imperial College London, UK; GoCardless, UK</i> . . . . .	567

## RENE Track

---

### Examining Past Results

<b>Duplicate Question Detection in Stack Overflow: A Reproducibility Study</b>	
Rodrigo F. G. Silva, Klérisson Paixão, and Marcelo de Almeida Maia — <i>Federal University of Uberlândia, Brazil</i> . . . . .	572
<b>How Do Scientists Develop Scientific Software? An External Replication</b>	
Gustavo Pinto, Igor Wiese, and Luiz Felipe Dias — <i>Federal University of Pará, Brazil; Federal University of Technology Paraná, Brazil; University of São Paulo, Brazil</i> . . . . .	582
<b>Re-evaluating Method-Level Bug Prediction</b>	
Luca Pascarella, Fabio Palomba, and Alberto Bacchelli — <i>Delft University of Technology, Netherlands; University of Zurich, Switzerland</i> . . . . .	592

## Code Smells

### Keep It Simple: Is Deep Learning Good for Linguistic Smell Detection?

Sarah Fakhoury, Venera Arnaoudova, Cedric Noisieux, Foutse Khomh, and Giuliano Antoniol — *Washington State University, USA; Polytechnique Montréal, Canada* . . . . . 602

### Detecting Code Smells using Machine Learning Techniques: Are We There Yet?

Dario Di Nucci, Fabio Palomba, Damian A. Tamburri, Alexander Serebrenik, and Andrea De Lucia — *Vrije Universiteit Brussel, Belgium; University of Zurich, Switzerland; Eindhoven University of Technology, Netherlands; University of Salerno, Italy* . . . . . 612

**Author Index** . . . . . 622