# Table of Contents

**Message from the General Chair** ........................................................................... x
**Message from the Program Chairs** .......................................................................... xi
**Organizing Committee** ............................................................................................ xiii
**Steering Committee** ................................................................................................ xiv
**Program Committee** ................................................................................................ xv
**Additional Reviewers** .............................................................................................. xvii

**Keynote I: Evolving Critical Systems** .................................................................... xviii
**Keynote II: Architecture-Driven Modernization of the European Air Traffic Management System** ........................................................................................................... xix

**WCRE 2001 Most Influential Paper: Robust Parsing using Island Grammars Revisited** .......................................................... xxvi

---

**Session I – Clones**

Incremental Code Clone Detection: A PDG-based Approach .................................................. 3
_Yoshiki Higo, Ueda Yasushi, Minoru Nishino, and Shinji Kusumoto_

On the Effectiveness of Simhash for Detecting Near-Miss Clones in Large Scale Software Systems ............................................................................................................. 13
_Md. Sharif Uddin, Chanchal K. Roy, Kevin A. Schneider, and Abram Hindle_

Internet-scale Real-time Code Clone Search Via Multi-level Indexing ................................... 23
_Iman Keivanloo, Juergen Rilling, and Philippe Charland_

An Investigation into the Impact of Software Licenses on Copy-and-paste Reuse among OSS Projects ................................................................................................................. 28
_Yu Kashima, Yasuhiro Hayase, Northiro Yoshida, Yuki Manabe, and Katsuro Inoue_

**Session II – Comprehension I**

Automatic Segmentation of Method Code into Meaningful Blocks to Improve Readability ................................................................................................................................. 35
_Xiaoran Wang, Lori Pollock, and K. Vijay-Shanker_

Requirements Traceability for Object Oriented Systems by Partitioning Source Code .......... 45
_Nasir Ali, Yann-Gaël Guéhéneuc, and Giuliano Antoniol_
Can We Predict Dependencies Using Domain information? .................................................................55
Amir Aryani, Fabrizio Perin, Mircea Lungu, Abduh Naser Mahmood, and Oscar Nierstrasz

Session III – Comprehension II
Does the Documentation of Design Pattern Instances Impact on Source Code Comprehension? Results from Two Controlled Experiments ...............................................................67
Carmine Gravino, Michele Risi, Giuseppe Scanniello, and Genoveffa Tortora
Towards the Extraction of Domain Concepts from the Identifiers .........................................................77
Surafel Lemma Abebe and Paolo Tonella
Reverse Engineering Co-maintenance Relationships Using Conceptual Analysis of Source Code .................................................................................................................................87
Scott Grant, James R. Cordy, and David B. Skillicorn
Concern Localization using Information Retrieval: An Empirical Study on Linux Kernel .................................................................92
Shaowei Wang, David Lo, Zhemchang Xing, and Lingxiao Jiang

Session IV – Code Search
Locating the Meaning of Terms in Source Code Research on “Term Introduction” ................................99
Jan Nonnen, Daniel Speicher, and Paul Imhoff
Approximate Code Search in Program Histories ..................................................................................109
Saman Bazrgeshan, Rainer Koschke, and Nils Göde
Code Search via Topic-Enriched Dependence Graph Matching ................................................................119
Shaowei Wang, David Lo, and Lingxiao Jiang
Measuring the Accuracy of Information Retrieval Based Bug Localization Techniques ..................124
Matthew Beard, Nicholas Kraft, Letha Etzkorn, and Stacy Lukins

Session V – APIs
Exploring the Intent behind API Evolution: A Case Study ......................................................................131
Daqing Hou and Xiaojia Yao
A Framework Profile of .NET ..............................................................................................................141
Ralf Lömmel, Rufus Linke, Ekaterina Pek, and Andrei Varanovich
Useful, But Usable? Factors Affecting the Usability of APIs ................................................................151
Minhas F. Zibran, Farjana Z. Eishita, and Chanchal K. Roy

Session VI – Security
Automatic Extraction of Secrets from Malware .....................................................................................159
Ziming Zhao, Gail-Joon Ahn, and Hongxin Hu
Reverse Engineering of Protocols from Network Traces ........................................................................169
João Antunes, Nuno Neves, and Paulo Verissimo
An Android Security Case Study with Bauhaus

Bernhard J. Berger, Michaela Bunke, and Karsten Sohr

An Exploratory Study of Software Reverse Engineering in a Security Context

Christoph Treude, Fernando Figueira Filho, Margaret-Anne Storey, and Martin Salois

Session VII – Bugs

How Long Does a Bug Survive? An Empirical Study

Gerardo Canfora, Michele Ceccarelli, Luigi Cerulo, and Massimiliano Di Penta

A Preliminary Evaluation of Text-based and Dependency-based Techniques for Determining the Origins of Bugs

Steven Davies, Marc Roper, and Murray Wood

Got Issues? Do New Features and Code Improvements Affect Defects?

Daryl Posnett, Abram Hindle, and Prem Devanbu

Session VIII – Quality

Impact of Installation Counts on Perceived Quality: A Case Study on Debian

Israel Herraiz, Emad Shihab, Thanh H.D. Nguyen, and Ahmed E. Hassan

Monitoring Software Quality Evolution by Analyzing Deviation Trends of Modularity Views

Tianmei Zhu, Yijian Wu, Xin Peng, Zhenchang Xing, and Wenyun Zhao

An Empirical Validation of the Benefits of Adhering to the Law of Demeter

Yi Guo, Michael Wuersch, Emanuel Giger, and Harold C. Gall

Assessing Software Quality by Program Clustering and Defect Prediction

Xi Tan, Xin Peng, Sen Pan, and Wenyun Zhao

Session IX – Dynamic Analysis

Object-Based Dynamic Protocol Recovery for Multi-Threading Programs

Bernhard Scholz and Rainer Koschke

An Entropy Evaluation Approach for Triaging Field Crashes: A Case Study of Mozilla Firefox

Foutse Khomh, Brian Chan, Ying Zou, and Ahmed E. Hassan

Meta-Level Runtime Feature Awareness for Java

Andrzej Olszak, Martin Rytter, and Bo Norregaard Jørgensen

Using Dynamic Analysis and Clustering for Implementing Services by Reusing Legacy Code

Andreas Fuhr, Tassilo Horn, and Volker Riediger
Session X – Reverse Engineering

Reverse Engineering of Mobile Application Lifecycles ................................................... 283
Dominik Franks, Corinna Elsemann, Stefan Kowalewski, and Carsten Weise

Reverse Engineering of Event Handlers of RAD-Based Applications .......................... 293
Óscar Sánchez Ramón, Jesús Sánchez Cuadrado, and Jesús García Molina

Analyzing the Source Code of Multiple Software Variants for Reuse Potential ............. 303
Sławomir Duszynski, Jens Knodel, and Martin Becker

Reverse Engineering Feature Models from Programs’ Feature Sets ............................... 308
Evelyn Nicole Haslinger, Roberto E. Lopez-Herrejon, and Alexander Egyed

Session XI – Evolution

Make it or Break it: Mining Anomalies from Linux Kbuild ........................................ 315
Sarah Nadi and Ric Holt

An Exploratory Study of Macro Co-changes ................................................................. 325
Fehmi Jaafar, Yann-Gaët Guédénaux, Sylvie Hamal, and Giuliano Antoniol

An Exploratory Study of the Evolution of Communicated Information about the Execution of Large Software Systems ......................................................... 335
Wei Yi Shang, Zhen Ming Jiang, Bram Adams, Ahmed E. Hassan, Michael W. Godfrey, Mohamed Nasser, and Parminder Flora

Session XII – Binary Analysis and Refactoring

SmartDec: Approaching C++ Decompilation ............................................................... 347
Alexander Fokin, Egor Derevenetc, Alexander Chernov, and Katerina Troshina

Precise Static Analysis of Binaries by Extracting Relational Information ....................... 357
Alexander Sepp, Bogdan Mihaila, and Axel Simon

Refactoring Traditional Forms into Ajax-enabled Forms ............................................... 367
Ming Ying and James Miller

An Empirical Study of Refactoring in the Context of FanIn and FanOut Coupling ........... 372
Alessandro Murgia, Roberto Tonelli, Steve Counsell, Giulio Concas, and Michele Marchesi

Session XIII – Developers and Modularization

Recommending People in Developers’ Collaboration Network ....................................... 379
Didi Surian, Nian Liu, David Lo, Hanghang Tong, Ee-Peng Lim, and Christos Faloutsos

Reasoning over the Evolution of Source Code Using Quantified Regular Path Expressions ...................................................................................................................... 389
Andy Kellens, Coen De Roover, Carlos Noguera, Reinout Stevens, and Viviane Jonckers

Modularization Metrics: Assessing Package Organization in Legacy Large Object-Oriented Software ................................................................. 394
Hani Abdeen, Stéphane Ducasse, and Houari Sahraoui
Industrial Experience Reports

Application Architecture Discovery - Towards Domain-driven, Easily-Extensible Code Structure

Hitesh Sajnani, Ravindra Naik, and Cristina Lopes

Renovation by Machine-Assisted Program Transformation in Production Reporting and Integration

Sava Mintchev

Reconstructing Traceability between Bugs and Test Cases: An Experimental Study

Nilam Kaushik, Ladan Tahvildari, and Mark Moore

Doctoral Symposium

Recovering Test-to-Code Traceability Via Slicing and Conceptual Coupling

Abdallah Qusef

In-Memory Database Support for Source Code Search and Analytics

Oleksandr Panchenko

Extending Bug Localization Using Information Retrieval and Code Clone Location Techniques

Matthew Beard

Tool Demonstrations

Stealthy Profiling and Debugging of Malware Trampolining from User to Kernel Space

Jason Raber

Function Insight: Highlighting Suspicious Sections in Binary Run Traces

Michelle Cheatham and Jason Raber

Understanding Legacy Features with Featureous

Andrzej Olszak and Bo Nørregaard Jørgensen

iProblems - An Integrated Instrument for Reporting Design Flaws, Vulnerabilities and Defects

Mihai Codoban, Cristina Marinescu, and Radu Marinescu

Focusing on Execution Traces Using Diver

Del Myers and Margaret-Anne Storey

jPET: An Automatic Test-Case Generator for Java

Elvira Albert, Israel Cabañas, Antonio Flores-Montoya, Miguel Gómez-Zamalloa, and Sergio Gutiérrez

SCOTCH: Slicing and Coupling Based Test to Code Trace Hunter

Abdallah Qusef, Gabriele Bavota, Rocco Oliveto, Andrea De Lucia, and David Binkley

Author Index