

2007 4th International Workshop on Mining Software Repositories

(MSR 2007)

**Minneapolis, Minnesota
20 – 26 May 2007**



IEEE Catalog Number: CFP0778C-PRT
ISBN: 978-0-7695-2950-9

TABLE OF CONTENTS

BUGS

How Long Will It Take to Fix This Bug?	1
<i>Cathrin Weiß, Rahul Premraj, Thomas Zimmermann, and Andreas Zeller</i>	
Determining Implementation Expertise from Bug Reports	9
<i>John Anvik and Gail C. Murphy</i>	
Defect Data Analysis Based on Extended Association Rule Mining	17
<i>Shuji Morisaki, Akito Monden, Tomoko Matsumura, Haruaki Tamada, and Ken-ichi Matsumoto</i>	
Spam Filter Based Approach for Finding Fault-Prone Software Modules	25
<i>Osamu Mizuno, Shiro Ikami, Shuya Nakaichi, and Tohru Kikuno</i>	

TEAMS

Recommending Emergent Teams	29
<i>Shawn Minto and Gail C. Murphy</i>	
Open Borders? Immigration in Open Source Projects	37
<i>Christian Bird, Alex Gourley, Prem Devanbu, Anand Swaminathan, and Greta Hsu</i>	
Correlating Social Interactions to Release History during Software Evolution	45
<i>Olga Baysal and Andrew J. Malton</i>	
Mining CVS Repositories to Understand Open-Source Project Developer Roles	53
<i>Liguo Yu and Srinini Ramaswamy</i>	

MINING

Visual Data Mining in Software Archives to Detect How Developers Work Together	57
<i>Peter Weißgerber, Mathias Pohl, and Michael Burch</i>	
Mining Software Repositories with iSPARQL and a Software Evolution Ontology	65
<i>Christoph Kiefer, Abraham Bernstein, and Jonas Tappelet</i>	
Mining Workspace Updates in CVS	73
<i>Thomas Zimmermann</i>	
Finding Relevant Applications for Prototyping	77
<i>Mark Grechanik, Kevin M. Conroy, and Katharina Probst</i>	
Lightweight Risk Mitigation for Software Development Projects Using Repository Mining	81
<i>Stephen P. Masticola</i>	

CHANGES

Identifying Changed Source Code Lines from Version Repositories	85
<i>Gerardo Canfora, Luigi Cerulo, and Massimiliano Di Penta</i>	

Mining a Change-Based Software Repository	93
<i>Romain Robbes</i>	
Studying Versioning Information to Understand Inheritance Hierarchy Changes	101
<i>Filip Van Rysselberghe and Serge Demeyer</i>	
Combining Single-Version and Evolutionary Dependencies for Software-Change Prediction	105
<i>Huzefa Kagdi and Jonathan I. Maletic</i>	
Evaluating the Harmfulness of Cloning: A Change Based Experiment	109
<i>Angela Lozano, Michel Wermelinger, and Bashar Nuseibeh</i>	

PATTERNS AND MODELS

Release Pattern Discovery via Partitioning: Methodology and Case Study	113
<i>Abram Hindle, Michael W. Godfrey, and Richard C. Holt</i>	
Comparing Approaches to Mining Source Code for Call-Usage Patterns	121
<i>Huzefa Kagdi, Michael L. Collard, and Jonathan I. Maletic</i>	
Towards a Theoretical Model for Software Growth	129
<i>Israel Herraiz, Jesus M. Gonzalez-Barahona, and Gregorio Robles</i>	
Analysis of the Linux Kernel Evolution Using Code Clone Coverage	137
<i>Simone Livieri, Yoshiki Higo, Makoto Matsushita, and Katsuro Inoue</i>	

DEVELOPERS, DEVELOPERS, DEVELOPERS

What Can OSS Mailing Lists Tell Us? A Preliminary Psychometric Text Analysis of the Apache Developer Mailing List	141
<i>Peter C. Rigby and Ahmed E. Hassan</i>	
Using Software Distributions to Understand the Relationship among Free and Open Source Software Projects	149
<i>Daniel M. German</i>	
Using Software Repositories to Investigate Socio-technical Congruence in Development Projects	157
<i>Giuseppe Valetto, Mary Helander, Kate Ehrlich, Sunita Chulani, Mark Wegman, and Clay Williams</i>	
Detecting Patch Submission and Acceptance in OSS Projects	161
<i>Christian Bird, Alex Gourley, and Prem Devanbu</i>	
Prioritizing Warning Categories by Analyzing Software History	165
<i>Sunghun Kim and Michael D. Ernst</i>	

MINING CHALLENGE

Impact of the Creation of the Mozilla Foundation in the Activity of Developers	169
<i>Jesus M. Gonzalez-Barahona, Gregorio Robles, and Israel Herraiz</i>	
Predicting Eclipse Bug Lifetimes	173
<i>Lucas D. Panjer</i>	
Mining Eclipse Developer Contributions via Author-Topic Models	177
<i>Erik Linstead, Paul Rigor, Sushil Bajracharya, Cristina Lopes, and Pierre Baldi</i>	

Predicting Defects and Changes with Import Relations181
Adrian Schröter

Forecasting the Number of Changes in Eclipse Using Time Series Analysis183
Israel Herraiz, Jesus M. Gonzalez-Barahona, and Gregorio Robles

Local and Global Recency Weighting Approach to Bug Prediction.....185
Hemant Joshi, Chuanlei Zhang, S. Ramaswamy, and Coskun Bayrak

Author Index