

Proceedings

First International Symposium on Empirical Software Engineering and Measurement

ESEM 2007

*Madrid, Spain
20-21 September 2007*



Los Alamitos, California
Washington • Tokyo



Table of Contents

ESEM 2007

General Chair's Message
Program Chair's Message
Organizing Committee
Program Committee
Reviewers
Supporting and Sponsoring Organisations
Local Arrangements Committee
Steering Committee
Program Chair Short Papers Message
Program Committee — Short Papers
Program Chair Posters Message
Program Committee — Posters

Keynote Addresses

A Critical Analysis of Empirical Research in Software Testing..... 1
Lionel C. Briand

Helping End-User Programmers “Engineer” Software: An Opportunity for Empirical Researchers.....9
Gregg Rothermel

Full Papers

Experience Management

Proposal of a Complete Life Cycle In-Process Measurement Model Based on Evaluation of an In-Process Measurement Experiment Using a Standardized Requirement Definition Process	11
<i>Yoshiki Mitani, Tomoko Matsumura, Mike Barker, Seishiro Tsuruho, Katsuro Inoue, and Ken-Ichi Matsumoto</i>	
Observing Software Testing Practice from the Viewpoint of Organizations and Knowledge Management.....	21
<i>Ossi Taipale, Katja Karhu, and Kari Smolander</i>	
Relevant Information Sources for Successful Technology Transfer: A Survey Using Inspections as an Example.....	31
<i>Andreas Jedlitschka, Marcus Ciolkowski, Christian Denger, Bernd Freimut, and Andreas Schlichting</i>	

Session: Empirical Studies of Testing and Validation

Assessing, Comparing, and Combining Statechart-based Testing and Structural Testing: An Experiment.....	41
<i>Samar Mouchawrab, Lionel C. Briand, and Yvan Labiche</i>	
Test Inspected Unit or Inspect Unit Tested Code?	51
<i>Atul Gupta and Pankaj Jalote</i>	
Defect Detection Efficiency: Test Case Based vs. Exploratory Testing.....	61
<i>Juha Itkonen, Mika V. Mäntylä, and Casper Lassenius</i>	
Comparing Model Generated with Expert Generated IV&V Activity Plans.....	71
<i>Haruka Nakao, Daniel Port, Yoshikawa Shigeo, Yuko Miyamoto, and Masa Katahira</i>	

Session: Experimental Techniques and Infrastructure

Requirement and Design Trade-offs in Hackstat: An In-Process Software Engineering Measurement and Analysis System	81
<i>Philip M. Johnson</i>	
Tuning Anonymity Level for Assuring High Data Quality: An Empirical Study	91
<i>Garardo Canfora, and Corrado Aaron Visaggio</i>	
Filtering, Robust Filtering, Polishing: Techniques for Addressing Quality in Software Data.....	99
<i>Gernot Liebchen, Martin Shepperd, Michelle Cartwright, Bheki Twala, and Mark Stephens</i>	

Session: Effort Estimation

An Estimation Model for Test Execution Effort.....	107
<i>Eduardo Henrique da Silva Aranha, and Paulo Henrique Monteiro Borba</i>	
Cognitive Limits of Software Cost Estimation.....	117
<i>Ricardo Valerdi</i>	
Impact Analysis of Missing Values on the Prediction Accuracy of Analogy-Based Software Effort Estimation Method AQUA	126
<i>Jingzhou Li, Ahmed Al-Emran, and Guenther Ruhe</i>	

Session: Empirical Studies of Software Products

Static Members and Cycles in Java Software	136
<i>Hayden Melton and Ewan Tempero</i>	
Usability Evaluation Based on Web Design Perspectives	146
<i>Tayana Conte, Jobson Massolar, Emilia Mendes, and Guilherme Horta Travassos</i>	
An Empirical Study of the Effects of Gestalt Principles on Diagram Understandability	156
<i>Krystle D. Lemon, Edward B. Allen, Jeffrey C. Carver, and Gary L. Bradshaw</i>	

Session: Empirical Studies of Software Processes

A Comparative Study of Aspect-Oriented Requirements Engineering Approaches	166
<i>Américo Sampaio, Phil Greenwood, Alessandro F. Garcia, and Awais Rashid</i>	
Toward Reducing Fault Fix Time: Understanding Developer Behavior for the Design of Automated Fault Detection Tools	176
<i>Lucas Layman, Laurie Williams, and Robert St. Amant</i>	

Session: Quality and Defect Prediction

Evaluating the Impact of Adaptive Maintenance Process on Open Source Software Quality	186
<i>Denis Kozlov, Jussi Koskinen, Jouni Markkula, and Markku Sakkinen</i>	
The Effects of Over and Under Sampling on Fault-prone Module Detection	196
<i>Yasutaka Kamei, Akito Monden, Shinsuke Matsumoto, Takeshi Kakimoto, and Ken-ichi Matsumoto</i>	
Generalizing Fault Contents from a Few Classes	205
<i>Hanna Scott and Philip M. Johnson</i>	

Session: Performing Systematic Reviews

Developing Search Strategies for Detecting Relevant Experiments for Systematic Reviews	215
<i>Oscar Dieste and Anna Grimán</i>	
Applying Systematic Reviews to Diverse Study Types: An Experience Report	225
<i>Tore Dybå, Torgeir Dingsøy, and Geir K. Hanssen</i>	
Using Context Distance Measurement to Analyze Results across Studies	235
<i>Daniela Cruzes, Manoel Mendonça, Victor Basili, Forrest Shull, and Mario Jino</i>	
A Visual Text Mining approach for Systematic Reviews	245
<i>Viviane Malheiros, Erika Höhn, Roberto Pinho, Manoel Mendonça, and José Carlos Maldonado</i>	

Session: Empirical Studies of Agile Methods

Usage and Perceptions of Agile Software Development in an Industrial Context: An Exploratory Study	255
<i>Andrew Begel and Nachiappan Nagappan</i>	
A Replicate Empirical Comparison between Pair Development and Software Development with Inspection	265
<i>Monvorath Phongpaibul and Barry Boehm</i>	
A Comparative Case Study on the Impact of Test-Driven Development on Program Design and Test Coverage	275
<i>Siniaalto Maria and Abrahamsson Pekka</i>	
An Experimental Evaluation of the Effectiveness and Efficiency of the Test-Driven Development	285
<i>Atul Gupta and Pankaj Jalote</i>	

Session: Software Metrics

Fine-Grained Software Metrics in Practice	295
<i>Michael English, Jim Buckley, and Tony Cahill</i>	
Conceptual Differences Among Functional Size Measurement Methods	305
<i>Cigdem Gencel and Onur Demirors</i>	
Evaluating Software Project Control Centers in Industrial Environments	314
<i>Marcus Ciolkowski, Jens Heidrich, Jürgen Münch, Frank Simon, and Mathias Radicke</i>	

Session: Effort Estimation in New Environments

Comparing Size Measures for Predicting Web Application Development Effort: A Case Study	324
<i>Carmine Gravino, Sergio Di Martino, Filomena Ferrucci, and Emilia Mendes</i>	
A Comparison of Techniques for Web Effort Estimation	334
<i>Emilia Mendes</i>	
Effort Prediction in Iterative Software Development Processes — Incremental Versus Global Prediction Models.....	344
<i>Pekka Abrahamsson, Raimund Moser, Witold Pedrycz, Alberto Sillitti, and Giancarlo Succi</i>	

Session: Software Data Mining

Mining Software Evolution to Predict Refactoring	354
<i>Jacek Ratzinger, Thomas Sigmund, Peter Vorburger, and Harald Gall</i>	
Using Software Dependencies and Churn Metrics to Predict Field Failures: An Empirical Case Study	364
<i>Nachiappan Nagappan and Thomas Ball</i>	
Fault-Prone Filtering: Detection of Fault-Prone Modules Using Spam Filtering Technique	374
<i>Osamu Mizuno, Shiro Ikami, Shuya Nakaichi, and Tohru Kikuno</i>	

Session: Homogeneity in Effort Estimation

Is This Cost Estimate Reliable? The Relationship between Homogeneity of Analogues and Estimation Reliability	384
<i>Naoki Ohsugi, Akito Monden, Nahomi Kikuchi, Michael D. Barker, Masateru Tsunoda, Takeshi Kakimoto, and Ken-ichi Matsumoto</i>	
Building Software Cost Estimation Models Using Homogenous Data	393
<i>Rahul Premraj and Thomas Zimmermann</i>	
Comparing Local and Global Software Effort Estimation Models — Reflections on a Systematic Review	401
<i>Stephen G. MacDonell, and Martin J. Shepperd</i>	

Session: Empirical Studies of Architecture

Characterizing Software Architecture Changes: An Initial Study	410
<i>Byron J. Williams and Jeffrey C. Carver</i>	
The Impact of Group Size on Software Architecture Evaluation: A Controlled Experiment.....	420
<i>Muhammad Ali Babar and Barbara Kitchenham</i>	
Evaluating the Usefulness and Ease of Use of a Groupware Tool	430
<i>Muhammad Ali Babar, Dietmar Winkler, and Stefan Biffl</i>	

Short Papers

Software Quality

An Approach to Global Sensitivity Analysis: FAST on COCOMO.....	440
<i>Stefan Wagner</i>	
An Approach to Outlier Detection of Software Measurement Data Using the K-means Clustering Method.....	443
<i>Kyung-A Yoon, Oh-Sung Kwon, and Doo-Hwan Bae</i>	
A Cost-Effectiveness Indicator for Software Development	446
<i>Hakan Erdogmus</i>	
How Software Designs Decay: A Pilot Study of Pattern Evolution	449
<i>Clemente Izurieta and James M. Bieman</i>	
Estimating the Quality of Widely Used Software Products Using Software Reliability Growth Modeling: Case Study of an IBM Federated Database Project	452
<i>Paul Luo Li, Randy Nakagawa, and Rob Montroy</i>	
Correlation between Bug Notifications, Messages and Participants in Debian’s Bug Tracking System	455
<i>Miguel Pérez Francisco, Pablo Boronat Pérez, and Gregorio Robles</i>	
Investigating Test Teams’ Defect Detection in Function Test.....	458
<i>Carina Andersson and Per Runeson</i>	
Comparison of Outlier Detection Methods on Fault-proneness Models.....	461
<i>Shinsuke Matsumoto, Yasutaka Kamei, Akito Monden, and Ken-ichi Matsumoto</i>	
“Talking Tests”: A Preliminary Experimental Study on Fit User Acceptance Tests.....	464
<i>Marco Torchiano, Filippo Ricca, and Massimiliano Di Penta</i>	

Empirical Research, Case Study Design, and Surveys

A Survey of the Practice of Design — Code Correspondence amongst Professional Software Engineers	467
<i>Ariadi Nugroho and Michel R.V. Chaudron</i>	
Assessing the Quality Impact of Design Inspections.....	470
<i>Christopher Ackermann, Forrest Shull, Ralf Carbon, Christian Denger, and Mikael Lindvall</i>	
A Controlled Experiment on the Effects of Synchronicity in Remote Inspection Meetings	473
<i>Fabio Calefato, Filippo Lanubile, and Teresa Mallardo</i>	
A Survey of Enterprise Software Development Risks in a Flat World.....	476
<i>Jesal Bhuta, Sudeep Mallick, and S.V. Subrahmanya</i>	
Checklists for Software Engineering Case Study Research.....	479
<i>Martin Höst and Per Runeson</i>	
Evidence Relating to Object-Oriented Software Design: A Survey	482
<i>John Bailey, David Budgen, Mark Turner, Barbara Kitchenham, Pearl Brereton, and Stephen Linkman</i>	
A Snapshot of the State of Practice in Software Development for Medical Devices	485
<i>Christian Denger, Raimund L. Feldmann, Martin Höst, Christin Lindholm, and Forrest Shull</i>	
GQM+Strategies – Aligning Business Strategies with Software Measurement	488
<i>Victor Basili, Jürgen Münch, Jens Heidrich, Mikael Lindvall, and Myrna Regardie</i>	
Automated Information Extraction from Empirical Software Engineering Literature: Is that possible?	491
<i>Daniela Cruzes, Manoel Mendonça, Victor Basili, Forrest Shull, and Mario Jino</i>	

Posters

Value-Based Empirical Research Plan Evaluation	494
<i>Stefan Biffl and Dietmar Winkler</i>	
Decision Support with EMPEROR	495
<i>Michele A. Shaw, Raimund L. Feldmann, and Forrest Shull</i>	
An Empirical Study on MBASE and LeanMBASE	496
<i>Supannika Koolmanojwong and Barry Boehm</i>	
Evaluation of Feature Extraction Methods on Software Cost Estimation	497
<i>Burak Turhan, Onur Kutlubay, and Ayse Bener</i>	
A Controlled Experiment for Selecting Transformations Based on Quality Attributes in the Context of MDA	498
<i>Marcela Genero, Mario Piattini, Silvia Abrahao, Emilio Insfran, J.A. Carsi, and Isidro Ramos</i>	
Agile Software Assurance: An Empirical Study	499
<i>Noura Abbas</i>	
Approaching the ERP Project Cost Estimation Problem: An Experiment.....	500
<i>Maya Daneva</i>	