Risk-Based Quantifiable Quality Improvement

Jeff Tian (tian@ engr. smu.edu)
Southern Methodist University
Dallas, Texas, USA

Contents

• Quality, Quantification, and Risk

• Quality↑ ⇐ Risk↓

• Quality↑ for Customers/Users

• Quality↑ for Software Organizations
Risk: Defect/Metrics Data

Defects (failures) by impact areas

0 50 100 150 200

NA Usability Standards Service Security Reliability Perforce Migration Maintain'ly Install'n Document'n Install'n Maintain'ly

0.0 0.2 0.4 0.6 0.8 1.0 HLSC HLDC MLSC MLDC MLC LOC Com CSI SMI McCabe RDecl UDecl Stmt Proc preDF postDF

March 30, 2007

Jeff Tian, SMU
Risk: Usage/Workload Data

Test runs in execution sequence

March 30, 2007

Jeff Tian, SMU
Quality, Quantification, and Risk

- **Risk:** Highly uneven distribution of cost, usage, quality, defect, performance, etc.
  - “80:20” rule or Pareto’s principle.
  - Units: component, owner, feature, etc.
  - Focus: high-risk/high-leverage units.

- Quantifiable quality improvement:
  - Quality: reliability, low-defects, etc.
  - Quantified: data and measurement
  - Improvement: risk-id/management
Risk Identification and Management

- Risk identification:
  - Qualitative: Causal analysis, etc.
  - Quantitative:
    - Old: correlation, regression, etc.
    - New: PCA, DA, TBM, etc.
    - AI/learning: NN, OSR, etc.

- Risk management:
  - Current project: Remedial actions
  - Similar projects: Corrective actions
  - Future projects: Preventive actions
In-/Ex-ternal Perspectives

- External risk to customers/users
  - Reliability: probability of failure-free operations for a time period or input set
  - Usage affects reliability
    - usage: OP (operational profile) for usage-based statistical testing (UBST)
  - Risk-based reliability improvement

- Internal risk to software organizations:
  - Defect risk identification & reduction.
  - Metrics-defect predictive modeling.
  - ODC and extensions.
Application Domains

- NASA/SEL: metrics in modeling
  - early work (pre-1992) at U. Maryland
  - axiomatic framework on measurement
  - selection procedure as optimization
  - used for better effort prediction

- Commercial software:
  - UBST and OP for compilers
  - Reliability modeling for DBMS etc.
  - ODC and extensions
  - Defect/metrics analysis and modeling
  - Reliability improvement using TBRM
Application Domains

- Telecommunications software:
  - Nortel (pre-2002):
    - similar to IBM work +
    - formal hypothesis testing HC-HD?
    - characterization of HD modules
  - Verizon (2004+):
    - e-commerce quality improvement

- Other applications: quality improvement
  - LM: aeronautic/embedded QI
    - reliability, DEA, HP
  - web testing and QI (later)
  - open source software QI
  - safety-critical systems
Risk Focus: Important Usage

- Focusing on functions/modules with:
  - High usage frequency and importance
  - Non-uniform testing effort
    ⇒ usage-based statistical testing (UBST)
  - Other focused quality assurance

- UBST = OP-guided testing
  - Capture user/usage information
  - Usage model = Operational profile (OP)
  - SRMs: Testing results ⇒ reliability
  - New applications in web, etc.
Web Testing Example

- High-level usage: Musa’s flat OP
- Other usage/testing: UMMs, etc
Web Testing Example

- UMMs (Unified Markov Models):
  - Navigation patterns/probabilities
  - Expansion to lower-level models
  - Driving existing white-box testing
Risk Focus: Reliability Growth

- Focused/accelerated reliability improvement via tree-based reliability models (TBRMs)
  - Measure: Purification level $\rho = \frac{\lambda_0 - \lambda_T}{\lambda_0}$
  - A/B/C: $0.35 \sim 0.72$ vs. D: $0.94 \sim 0.99$

March 30, 2007
Jeff Tian, SMU
Risk Focus: Web Reliability

- Different web sites: SMU/KDE/AMS
  - workload measurement:
    - hits, bytes, users, sessions
  - operational reliability: Nelson model
  - reliability growth simulation:
    - \( \approx \frac{2}{3} \) defect reduction in 1 month
  - accelerated reliability growth via (ODC inspired) risk identification (below)
Risk Focus: Safety

- TFM: Two-Frame-Model
  - Physical and logical frame
  - Sensors: physical $\Rightarrow$ logical
  - Actuators: logical $\Rightarrow$ physical
  - Focus: interface/interaction problems
Risk Focus: Safety

- Prescriptive specification checking:
  - Analyze sources of hazard
    - frame inconsistencies, sub-types
  - Derive systematic assertions
  - Dynamically check the assertions
Risk Focus: Defect-Prediction

- Analyzing defect-metrics relations
  - Correlation/regression (example above)
  - Impact: Behavior modification
Risk Focus: Defect-Reduction

• Early successes ⇒ Behavior modification
  ▶ Validation ⇒ hypothesis testing (HT)
  ▶ Need more sophisticated methods

• HT in Koru and Tian, IEEE-TSE 8/2005:
  ▶ High-defect (HD) modules vs. high-complexity (HC) modules
  ▶ HD and HC statistically different
  ▶ Complexity ranking of HD: 60 ∼ 80%
Risk Focus: Defect-Reduction

Legend:
- \( d \) = mean DF
- \( sd \) = std.dev
- \( n \) = #modules
- attr\(<\)cutoff
- attr\(>\)cutoff

- TBDM for defect↓ and quality↑:
  - Tian and Troster, JSS 12/1998
  - Tian/Nguyen/Allen/Appan, JSS 9/2001

March 30, 2007
Jeff Tian, SMU
Summary and Perspectives

- Existing work and successes:
  - Size/complexity $\uparrow \Rightarrow$ selective effort
  - 80:20 $\Rightarrow$ risk focus
  - Risk identification/management:
    - usage-based statistical testing
    - defect-prone module characterization
    - risk-based reliability improvement

- Positive impact on different systems:
  - Commercial software: defect/reliability
  - Web-based: heterogeneity/quality
  - Embedded: safety/performance
  - Other types too