Software Reliability and Safety

CSE 8317 — Fall 2005

Prof. Jeff Tian, tian@engr.smu.edu
CSE, SMU, Dallas, TX 75275
(214) 768-2861; Fax: (214) 768-3085
www.engr.smu.edu/~tian/class/8317.05f

SRE.3: TBRMs & Integrated SRE

- Experience with existing approaches
- TBRMs: Tree-based reliability Models
- Integrated SRE using TBRMs & others

Overview

- Reliability: Prob(failure-free operations)
  - ▷ *Time domain*: for a specific period.
    ⇒ Reliability growth models.
  - ▷ *Input domain*: for a specific input set.
    ⇒ Repeated sampling models.

- A new integrated approach:
  - ▷ TBRMs: tree-based reliability models.
  - ▷ Both input/time domain information.
  - ▷ Data driven/sensitive partitions.
  - ▷ Method: Tree-based modeling (TBM).
  - ▷ Risk focusing and remedial actions.
  - ▷ Main info. source: AIC paper (Tian 1998)

Product Environment

- Large (medium-reliable) products:
  - ▷ Commercial: RDBMS, compilers, software tools and computing environments.
  - ▷ Later: Telecommunication products.
  - ▷ Size: Up to millions of LOC.
  - ▷ Widely distributed/large user population.
  - ▷ No precise operational profile.
  - ▷ Process: roughly waterfall.

- Overall testing:
  - ▷ Long testing period ($2 \sim 18$ months).
  - ▷ Different testing sub-phases.
  - ▷ System testing focuses on reliability.
  - ▷ Test-until-it-breaks commonly used.
  - ▷ Staffing level variations.
  - ▷ Code base stability.

Testing Environment

- Scenario-based testing.
  - ▷ Shifting focus: learning/dependency.
  - ▷ Structure: high level functions.
  - ▷ Within scenario class (SC):
    - randomized workload
    - progression: complexity & intensity ↑
    - defect fixing and related runs
    - division among testers.

- Specific reliability analysis issues:
  - ▷ Scenario-based $\sim$ random testing
    - parallelism and interleaving
  - ▷ Defect fixing effect:
    - no long-term dependency $\Rightarrow$ grouping
  - ▷ Uneven faults $\Rightarrow$ TBRMs

Needs and Constraints

- Need assessment and analysis:
  (current status & urgency of needs)
  - ▷ Track test effort, progress and defect.
  - ▷ Reliability assessment and prediction.
  - ▷ Effective defect detection and removal.
  - ▷ Process and quality improvement.

- Environmental constraints:
  - ▷ Minimize cost & schedule risks.
  - ▷ Data availability and affordability.
  - ▷ Process refinement.
  - ▷ Maximize data utilization.

- Recommendation:
  new, evolutionary approach.

Overall Solution

- Combine SRGMs and IDRMs into TBRMs.

- Analysis and control:
  - ▷ SRGMs (s/w rel. growth models).
  - ▷ TBRMs: tree-based reliability models.
  - ▷ Progress monitoring & exit criteria.

- Problem identification and correction:
  - ▷ Use of input domain information
    - – IDRMs (input domain rel. models)
    - – identify high risk areas
  - ▷ Automatic partitioning via TBRMs.
  - ▷ Remedial actions for improvement.

Applications: Overview

- Product coverage:
  - ▷ Commercial products from IBM.
  - ▷ Improvement over original process.
  - ▷ Evolutionary approach:
    1. individual techniques.
    2. integration and refinement.
  - ▷ Recent work with Nortel Networks.

- Scope of Engagement:
  - ▷ Data definition and collection.
  - ▷ Data visualization and analysis.
  - ▷ Test progress tracking.
  - ▷ Reliability analysis with SRGMs.
  - ▷ Reliability improvement with TBRMs.

Applications: Testing & Data

- Data and tracking:
  - ▷ Integration with schedule information.
  - ▷ Normalization effect.
  - ▷ Summary reports and visualization.
  - ▷ Consistency checking automation.

- Customer usage information gathering
  – Operational profile construction.

- Coverage and input-domain analysis:
  - ▷ Functionality/function/static/dynamic.
  - ▷ Different levels of coverage for different testing phases.
  - ▷ Focused coverage through TBRM.

SRGMs: Application Experence

- Time measurement:
  - ▷ Calendar time.
  - ▷ Execution time: Musa models.
  - ▷ Logical time: runs, transactions, etc.
  - ▷ Usage dependent or independent?
  - ▷ Measurement implementation/cost?
- Model applicability and effectiveness:
  - ▷ Calendar time models useless.
  - ▷ Exec. time models costly & sensitive.
  - ▷ Activity-based time measurement (runs, transactions, etc.) suitable.
  - ▷ Context sensitive modeling for sub-groups or sub-phases $\Rightarrow$ TBRMs.

SRGM Conclusions

- Modeling result interpretation:
  - ▷ Accuracy of models:
    - – assessment, model goodness-of-fit.
    - – prediction: training & testing sets
  - ▷ Product purity at exit.
  - ▷ Bound estimations: multiple models.
  - ▷ Convergence of modeling results.

- Evolving to usage-based data/model:
  - ▷ Assurance of homogeneity:
    - – if 'yes', run-based data/model;
    - – if 'no', transaction measurement.
  - ▷ Suitable for input domain analysis.
  - ▷ Serve as cross validation for TBRMs.

Assessing Existing Approaches

- Time domain reliability analysis:
  - ▷ Customer perspective.
  - ▷ Overall assessment and prediction.
  - ▷ Ability to track reliability change.
  - ▷ Issues: assumption validity.
  - ▷ Problem: how to improve reliability?

- Input domain reliability analysis:
  - ▷ Explicit operational profile.
  - ▷ Better input state definition.
  - ▷ Hard to handle change/evolution.
  - ▷ Issues: sampling and practicality.
  - ▷ Problem: realistic reliability assessment?

An Integrated Approach

- Combine strengths of the two.

- Using TBRM for individual modeling:
  - ▷ Input state: categorical information.
  - ▷ Each run as a data point.
  - ▷ Time cutoff for partitions.
  - ▷ Data sensitive partitioning
    $\Rightarrow$ Nelson models for subsets.

- Integrated reliability analyses:
  - ▷ TBRM: partitioned subset reliability.
  - ▷ Use both input and timing information.
  - ▷ Monitoring changes in trees.
  - ▷ Enhanced exit criteria.
  - ▷ SRGM: overall reliability near exit.
  - ▷ Integrate into the testing process.

TBM: Technique for Integration

- Basic ideas:
  - ▷ TBM: tree-based models.
  - ▷ Tree: nodes=data-set, edges=decision.
  - ▷ Data: 1 response variable Y and n predictor variables $X_1, \ldots, X_n$.
  - ▷ Construction: recursive partitioning. (controlled growth vs growing&pruning)

- Usage and applications:
  - ▷ Basic usage: $Y = Tree(X_1, \ldots, X_n)$
  - ▷ Applicability: mixed-types of data.
  - ▷ Past applications: social sciences
  - ▷ In SE: risk identification by Selby & Porter, Tian & Troster, etc.

- Details: Tian/SQE book Ch.21.

TBRM in Integrated Analysis

- Tree-based reliability models (TBRMs) using all information:
  - ▷ Input domain partitioning information.
  - ▷ Testing results.
  - ▷ Timing information.
  - ▷ Each run as a data point.

- Model construction:
  - ▷ Response: Result indicator.
    - 1 for success, 0 for failure.
    - ⇒ Nelson model for subsets.
    - Mapping to failure rate or MTBF.
  - ▷ Predictor: Timing and input states.
    - Data sensitive partitioning.
    - Key factors affecting reliability.
    - Homogeneity of product reliability.

Using Integrated Analysis

- Interpretation of trees:
  - ▷ Predicted response: success rate. (Nelson reliability estimate.)
  - ▷ Time predictor: reliability change.
  - ▷ State predictor: risk identification.
- Monitoring reliability change:
  - ▷ Change in predicted response.
  - ▷ Through tree structural change.
- Risk identification and remedies:
  - ▷ Identify high risk input state.
  - ▷ Additional analysis.
  - ▷ Enhanced test cases.
  - ▷ Remedies for components.

TBRMs in Integrated Analysis

- Treatment of product bundles:
  - ▷ TBRM for individual products.
  - ▷ Dynamic change w.r.t. process needs.
  - ▷ SRGM (& TBRM) for bundle near exit.
- Risk identification:
  - ▷ High risk input sub-domains.
  - ▷ Additional analysis for the identified.
  - ▷ Guide for remedial actions.
- Results interpretation:
  - ▷ Progression of trees & tree types.
  - ▷ Usage as exit criteria.

Cross Validation

- Consistency with macro models:
  $\Rightarrow$ Effects on cost, schedule, quality.

- Validate with reliability growth models:
  - ▷ Trend of reliability growth.
  - ▷ Stability of failure arrivals.
  - ▷ Estimated reliability.
  - ▷ Product purity level at exit.

- Process changes & improvements:
  - ▷ Failure detection and fault removal.
  - ▷ Long term effect on development.

- Ultimate test: in-field problems.

Integrated Approach: Implementation

- Modified testing process:
  - ▷ Additional link for data analysis.
  - ▷ Process change and remedial actions.
- Activities and Responsibilities:
  - ▷ Evolutionary, stepwise refinement.
  - ▷ Collaboration: project & quality orgs.
  - ▷ Experience factory prototype (Basili).
- Implementation:
  - ▷ Passive tracking and active guidance.
  - ▷ Periodic and event-triggered.
  - ▷ S/W tool support

Implementation Support

- Types of tool support:
  - ▷ Data capturing
    - – mostly existing logging tools
    - – modified to capture new data
  - ▷ Analysis and modeling
    - – SMERFS modeling tool
    - – S-PLUS and related programs
  - ▷ Presentation/visualization and feedback
    - – S-PLUS and Tree-Browser

- Implementation of tool support:
  - ▷ Existing tools: minimize cost
    - – internal as well as external tools
  - ▷ New tools and utility programs
  - ▷ Tool integration
    - – loosely coupled suite of tools
    - – connectors/utility programs
    - – common depository: S-PLUS

Application Summary

- Tracking and input-domain analysis:
  - ▷ Effectiveness of visualization.
  - ▷ Problems with input-domain assessment.
- Time-domain analysis refinement:
  - ▷ Data normalization by runs/trans best.
  - ▷ Context sensitive modeling promising.
- Integrated approach using TBRM:
  - ▷ Guidance as well as assessment.
  - ▷ Risk focusing $\Rightarrow$ reliability improvement.
  - ▷ Progression of trees.
  - ▷ Usage as exit criteria.
  - ▷ Cross validation.

Future Directions

- Implementation and deployment:
  - ▷ Data: automated data capturing.
  - ▷ OP: evolutionary approach.
  - ▷ Integration: analysis and improvement.
  - ▷ Use in different industrial environments.
- Exploration and improvement:
  - ▷ Customize time/transaction measurement.
  - ▷ Early indicators/predictive modeling.
  - ▷ Customer environment/OP refinement.
  - ▷ Integrate to life-cycle quality models.
  - ▷ Management and cost modeling.
  - ▷ Refinement of modeling techniques.
- Continued research at SMU and collaboration with our industrial partners.