# Empirical Software Engineering

# CSE 8340 — Spring 2014

Prof. Jeff Tian, tian@lyle.smu.edu
CSE, SMU, Dallas, TX 75275
(214) 768-2861; Fax: (214) 768-3085
www.lyle.smu.edu/∼tian/class/8340.14s

## Module IIb: TBM in Risk Ident.

- Telecom Case Study
- TBM and Applications
- Results and Discussions

# Overview

- Project background:
  - ▷ Reliability and QA:
  - ▷ SMU research under NSF&THECB grants 1998–: Tian, Nguyen, and others.
  - ▷ Industrial support: Nortel Networks Frame, Allen, Appan, and others.
- Planned activities:
  - ▷ Defect analysis: TBDMs (Tree-based defect models)
  - ▷ Reliability improvement: TBRMs (Tree-based reliability models)
  - ▷ Other: UMMs, testing, HT, etc.
- JSS paper by Tian/Nguyen/Allen/Appan.

# Overview: Context

- Objectives of study:
  Understanding problem prone modules for quality management and improvement.

- Objects of study:
  - ▷ Large telecommunication software
  - ▷ Nortel Networks products: NT-X
  - ▷ Developed/released/used recently
  - ▷ 5 different releases
  - ▷ Latest finished around 2000/2001

- Development environment:
  - ▷ Waterfall-like process
  - ▷ Change/incremental development
  - ▷ Project monitoring through various measurements and tools (e.g., Datrix/EMERALD/COMET)

# Overview: Approach (Design)

- General approach to the study:
  - ▷ Various measurement data.
  - ▷ Identify problem prone modules
  - ▷ Characterize problem prone modules
  - ▷ Conclusions based on above analyses.

- Classification: observational

- 5 different releases
  - ▷ Consistency among releases
    – assessment and understanding
  - ▷ Guidance for new release
    – prediction and control

## Risk Identification: Why?

- Risk and 80:20 rule
  - ▷ Risk: (high) probability of undesirable situations or consequences
  - ▷ 80:20 rule: 80% of problems traceable to 20% of components
  - ▷ Need risk identification

- Problem-prone modules
  - ▷ Likely to contain substantially more internal or development defects.
  - ▷ (Fault-prone: in-field failures)
  - ▷ Identification of the modules
  - ▷ Corrective/remedial actions

- Identify problem prone modules
  - ▷ Data: past defect and other metrics.
  - ▷ Technique: risk identification.
  - ▷ Followup: Characterization.

# Risk Identification: How?

- Techniques used in Nortel Networks:
  - ▷ EMERALD: mainly multiple regression and logistic analysis, with limited use of neural networks.
  - ▷ COMET: principal component analysis (PCA) and discriminant analysis.

- New techniques:
  - ▷ Tree-based modeling with S-PLUS
  - ▷ CART with SAS

- Primary technique here: TBM.
  - ▷ generic comparison: Tian SQP paper.
  - ▷ specifics: later

# Product and Defect Metrics

- Defect metrics:
  - ▷ `DF:` defect fixes
  - ▷ Applied in response to testing failures.
  - ▷ `DF` vs. failure/fault counts:
    - – `DF` captures propagation information.
    - – `DF` is identified with specific modules.
  - ▷ Available in project data depository
  - ▷ Data transformed to percentages.

$$DF = \frac{DF_{raw}}{DF_{max}} \times 100\%$$

- Product metrics:
  - ▷ From EMERALD, a Nortel tool/product
  - ▷ Underlying analyzer for procedure-level metrics
  - ▷ Module level metrics ∼ `DF`

# Product Metrics: Details

- 53 raw product metrics:
  - ▷ volume,
  - ▷ testability, decision complexity, dead code, independent path, structuredness,
  - ▷ readability,
  - ▷ section dependability,
  - ▷ software science.

- 6 synthetic product metrics:
  - ▷ `OurRange`: # metrics $\notin$ acceptable range – a rough indicator of module quality
  - ▷ `Level` (or procedure type)

# Risk Identification: EMERALD

- Techniques used:
  - ▷ Mixture of old and new.
  - ▷ Multiple regression & logistic analysis.
  - ▷ Neural network etc.

- EMERALD output:
  - ▷ `OpRisk`: likelihood of field defect.
  - ▷ Values: “green”, “R”, ..., “RRRRRR”.
  - ▷ Other output also possible.
  - ▷ But not `DF` for this model
  - ▷ Identifying but not characterizing
  - ▷ Try other models

# TBDMs: Why?

- Risk identification:
  - ▷ Assumption (in traditional techniques):
    - linear relation
    - uniformly valid result
  - ▷ Reality of defect distribution:
    - isolated pocket
    - different types of metrics
    - correlation/dependency in metrics
    - qualitative differences
  - ▷ Need new risk id. techniques.

- Risk characterization:
  - ▷ Identified, then what?
  - ▷ Result interpretation.
  - ▷ Remedial/corrective actions.
  - ▷ Extrapolation to new product/release.
  - ▷ TBDMs appropriate.

# TBM & TBDMs: Ideas

- TBDMs: tree-based defect models using tree-based modeling (TBM) technique

- Decision trees:
  - ▷ multiple/multi-stage decisions
  - ▷ may be context-sensitive
  - ▷ natural to the decision process
  - ▷ applications in many problems
    - – decision making & problem solving
    - – decision analysis/optimization

- Tree-based models:
  - ▷ reverse process of decision trees
  - ▷ data ⇒ tree
  - ▷ idea of decision extraction
  - ▷ generalization of ”decision”

# TBM: Types and Applications

- Key ”selling” points:
  - ▷ intuitiveness and interpretation
    - – compare to PCA, NN
  - ▷ quantitative & qualitative info.
  - ▷ hierarchy/importance/organization

- Past applications:
  - ▷ social sciences
  - ▷ Selby&Porter: Amadeus project
  - ▷ Tian et al:
    - – NASA/SEL work (area IV)
    - – IBM product defects: with Troster
    - – IBM TBRM: 8317 coverage
    - – SMU: UMM/testing, Nortel work

# TBM: Technique

- Technique: tree-based modeling
  - ▷ Tree: nodes=data-set, edges=decision.
  - ▷ Data attributes:
    - – 1 response & $n$ predictor variables.
  - ▷ Construction: recursive partitioning.
  - ▷ Usage: relating response to predictors
    - – $Y = Tree(X_1, \ldots, X_n)$
    - – understanding vs. predicting
    - – identification and characterization
  - ▷ Works for mixed-types of data.
  - ▷ Tree growing and pruning.

- Algorithm: Fig.1
  - ▷ regression tree and example
  - ▷ classification tree: modify Step 3

# TBDMs: Result for NT-X

- Overall result: Fig. 2
  - ▷ Similar results for other releases
  - ▷ General understanding: simplicity/pruning

- How to read each node?
  – `DF` and node size summary

- Split conditions
  - ▷ Distinguishing characteristics
  - ▷ Root to leaf: order of importance
  - ▷ Metrics selected out of 59 by algorithm:
    - – Halstead program length (`HalLen`)
    - – # basic utility routines (`Level1`)
    - – # include files (`FilIncNbr`)
    - – comments volume average (`ComStrAvg`)
    - – Halstead level (`HalLvl`)

# TBDMs: Result for NT-X

- Identifying problem prone modules
  - ▷ Identified leaf nodes (Table 1)
  - ▷ Comparison to other nodes (Fig. 2)
  - ▷ Isolated pockets: llrl, rr (rrl + rrr)
  - ▷ Groups vs. individual identification

- Characterizing problem prone modules
  - ▷ Split conditions as characterization
  - ▷ Symptoms of problems
  - ▷ Further analysis ⇒ systematic problems
  - ▷ Corrective/remedial actions
  - ▷ Future process/product improvement

# Other TBDM Results

- TBDM performed:
  - ▷ DF $\sim$ metrics (previous)
  - ▷ DF $\sim$ *OpRisk*
  - ▷ DF $\sim$ all

- EMERALD result validation
  - ▷ TBDM set1: tot.fix $\sim$ *OpRisk*
  - ▷ TBDM set2: tot.fix $\sim$ all
  - ▷ Consistent pattern
  - ▷ Reasonable predictions, but...
    - – not much constructive info.

# TBDMs: NT-X vs. IBM LS and NS

- IBM products for comparison:
  - ▷ LS and NS: A legacy and a new system
  - ▷ Large s/w systems: 995, 1302 modules
  - ▷ Metrics: `DF` + (11, 15) other
    – design (6), size (2), complexity (5, 3)
    – change (2 for LS)

- Results for IBM LS and NS:
  - ▷ LS: change, size, data complexity
  - ▷ NS: design and control complexity
  - ▷ Problem-prone modules: Table 2

- Comparison: NT-X similar to IBM LS
  - ▷ Common traits of legacy systems
  - ▷ Implications: similar initiatives

## Recommendation: Integrated Strategy

- Main considerations:
  - ▷ Existing tools and infrastructure
  - ▷ Past experience and domain knowledge
  - ▷ Applicability and effectiveness of new risk identification techniques
  - ▷ Tailoring for your environment

- Specific for Nortel Networks:
  - ▷ Measurement: existing tools/databases
  - ▷ Identification: EMERALD
  - ▷ Characterization: TBDMs (this paper)
  - ▷ Cross validation: both
  - ▷ Follow-up: causal analysis needed – but TBDMs can help/guide

# Recommendation: Lifecycle Integration

- Main considerations:
  - ▷ Process and data availability
  - ▷ Experience/infrastructure/tools/etc.
  - ▷ Different focus, but similar techniques?
  - ▷ Tailoring for your process/product
- Lifecycle integration:
  - ▷ Analysis of inspection/other data
  - ▷ Analysis and feedback loop
  - ▷ Our current/future research projects
  - ▷ QA and improvement focus:
    - – defect prevention
    - – defect detection and removal
    - – defect containment

# Conclusions and Perspectives

- Problems addressed:
  - ▷ Large telecommunication systems.
  - ▷ Multiple releases, diverse components
  - ▷ Uneven `DF` distribution (80:20 rule)
  - ▷ Need risk identification and characterization for corrective/remedial actions
- Conclusions: an effective strategy
  - ▷ Existing measurement tools/infrastructure
  - ▷ EMERALD for risk identification
  - ▷ TBDMs for risk characterization
  - ▷ TBDMs guided follow-up actions
- Future work:
  - ▷ Lifecycle approach to quality
  - ▷ Progression: qualitative ⇒ quantitative