

Software Metrics and Quality Engineering

CSE 8314 — Fall 2017

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/8314.17f

Module V: Recent Development and Advanced Topics

- New Metrics and Applications
- New Models and ESE Guidelines
- Data Collection/Extraction/Mining
- Hypothesis Testing

New Metrics and Applications

- New metrics might be needed:
 - ▷ new language/technology
 - ▷ new application domain
 - ▷ new vs. adapted/adopted metrics

- New language/technology
 - ▷ CK metrics for OO
 - ▷ Other new languages/technologies

- New application domains (+technologies):
 - e.g., Web, net-centric, SOA, Cloud, IoT, etc.

- NCSS complexity metrics (SDPS slides)

New Models and ESE

- *Empirical Research in S/w Eng. (ESE):*
2002 paper by Kitchenham, Pfleeger, Pickard, Jones, Hoaglin, Emam, Rosenberg (TSE 28(8):721-734).

- Why a guideline?
 - ▷ More ESE research activities
 - ▷ Maturing of SE and ESE
 - ▷ Practical concerns:
 - reader/students
 - researchers/meta-analyst
 - reviewers/editors
 - journals/conferences etc

- More details in CSE 8340

ESE Guideline: Basis

- Internal basis for the guideline:
 - ▷ Research experience in ESE
 - ▷ Researcher experience in ESE
 - ▷ from both author/reviewer perspectives
 - ▷ Other CS/SE work

- External basis for the guideline:
 - ▷ Scientific method
 - ▷ Implicit guidelines used for emp. studies in other mature disciplines (most natural science)
 - ▷ (Explicit) guidelines for emp. studies in other disciplines (e.g., medical)
 - ▷ External experts as co-authors

- Result: *Preliminary* guidelines.

ESE Guideline: Sources

- Authors as information sources:
 - ▷ Diverse background
 - ▷ Experience in SE/ESE/statistics/others

- Other important sources:
 - ▷ Similar guidelines for medical journals
 - ▷ Meta-analysis studies
(studies of empirical studies and results)
 - ▷ Papers about statistical applications:
 - positive (guide, "what should be done")
 - negative ("what was wrong/to avoid")
 - ▷ Other "soft" sciences
 - ▷ List of specific references in paper

ESE Guideline: Topic Areas

1. Experimental context
2. Experimental design
3. Conduct experiment and data collection
4. Analysis
5. Presentation of result
6. Interpretation of result

Data Collection/Extraction/Mining

- Data collection
 - ▷ Data source identification
 - ▷ Data collection procedures
 - ▷ Tools: computing vs extracting
 - ▷ IBM data: complexity/defect/activity/etc.

- Data extraction:
 - ▷ Tapping into pre-existing data sources
 - ▷ Web measurement example (paper in Blackboard)

- Data mining:
 - ▷ (unstructured/big) data source
 - ▷ mining/extensive processing necessary
 - ▷ AutoODC work at SMU (paper in Blackboard)

Validation and Hypothesis Testing

- Hypothesis: An assumption or concession made for the sake of argument.
 - ▷ Simple hypothesis: One value of the population parameter ($\mu = 115$).
 - ▷ Composite hypothesis: A range of values that the population parameter may assume ($\mu \neq 115$).
 - ▷ Null Hypothesis (H_0): Status quo.
 - ▷ Alternative Hypothesis (H_a): Believed to be true.
 - ▷ Both H_0 and H_a can be simple or composite.

- Hypothesis Testing: Choose between two competing hypotheses about the value of a population parameter using the knowledge obtained from a sample.
 - Example HT: slides online (pp.45-48)

HT and Taylored Metrics

- Another HT example
 - ▷ Part of OS/SRE work
 - ▷ SEDE'2017 paper by Y. Tian, J. Tian, and N. Li
 - ▷ experience factore captured and HT applied
 - ▷ result comparison with HT applied too
 - ▷ see paper and slides on Canvas

- An example of taylored metrics
 - ▷ Part of accurate Markov OP work at SMU
 - ▷ SERA'2017/Springer book chapter by G. Karami and J. Tian
 - ▷ impact of accurate Markov OP on
 - test coverage and efficiency (waste)
 - reliability
 - ▷ metrics on the above
 - ▷ see paper and slides on Canvas