Static Analysis Tools Predicting Pre-Release Defect Density

Presented by Peter Rigby

MSR Course

November 16, 2006

Introduction and Goal

- Goal use static analysis tools to predict the pre-release defect density.
- Pre-release defect density is measured as the number of defects per KLOC found by other defect detection techniques (e.g., testing)

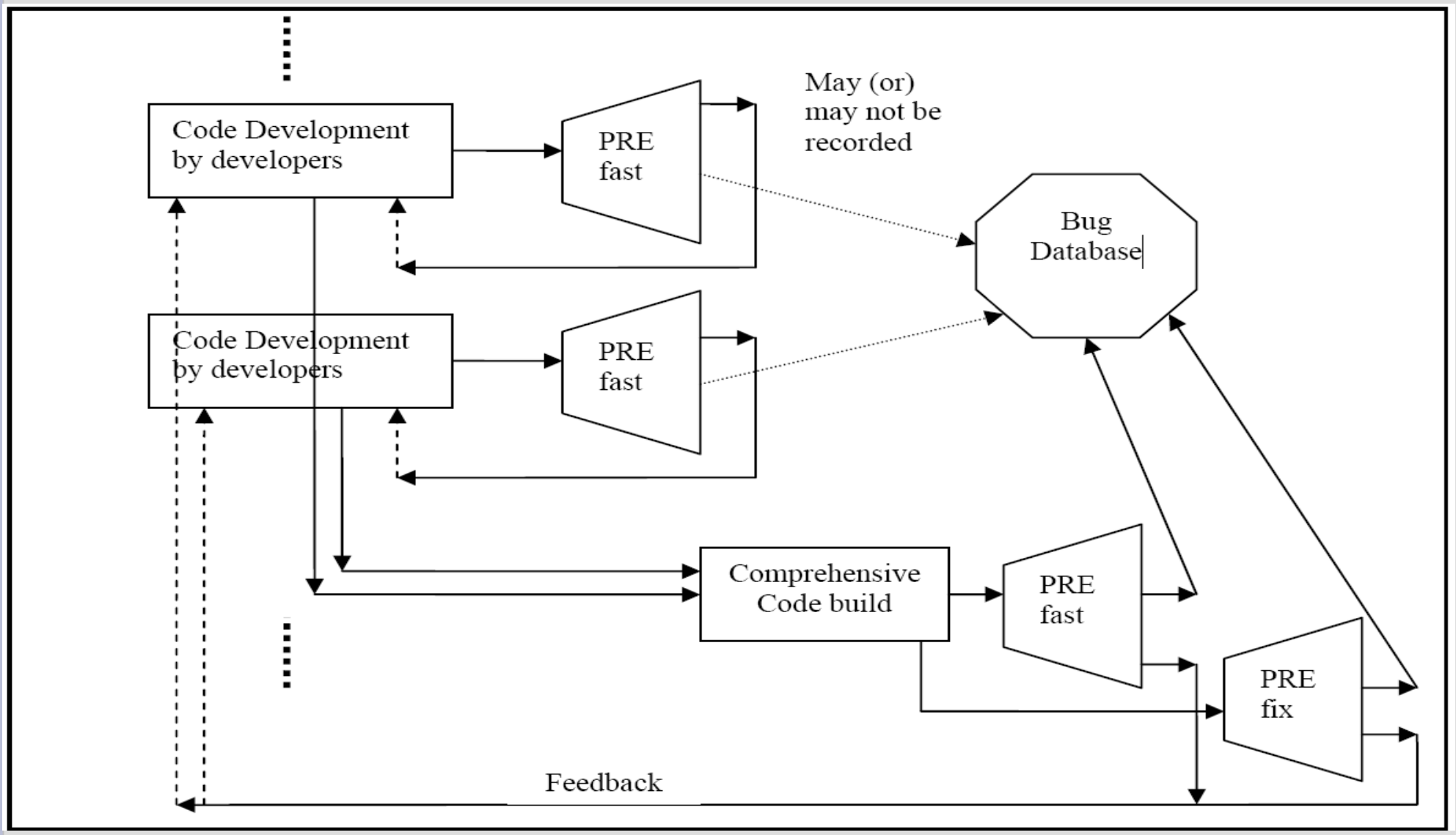
Hypotheses

- Question: Are static analysis tools leading indicators of faulty code?
- H1: Static analysis defect density can be used as an early indicator of pre-release defect density;
- H2: Static analysis defect density can be used to predict pre-release defect density at statistically significant levels;
- H3: Static analysis defect density can be used to discriminate between components of high and low quality (fault and not fault-prone components)

Tools

- PreFix
  - Uses symbolic execution, applied bottom-up
  - Selects execution path, starts with leafs, creates a symbolic summary for future use
  - Example errors: uninitialized memory
  - Processor intensive
- PreFast
  - Pattern matching in AST
  - Local dataflow analysis
  - Example error: NULL pointer
  - Negligible processor usage

Process


Code Development by developers
PRE fast
May (or) may not be recorded
Bug Database
Code Development by developers
PRE fast
Comprehensive Code build
PRE fast
PRE fix
Feedback

Statistical Techniques

- Principle component analysis
  - Linear transformation onto new coordinate system
  - Component with greatest variance projected onto first coordinate ...
  - Able to eliminate variables that are highly correlated to other metrics
- Discriminant Analysis
  - Similar to PCA, but
  - models the difference between the classes of data.
  - Can be used as a classifier

Correlation and Regression

- Regression analysis identifies which variable(s) produce the best prediction
- Combined Prefast and Prefix is best.

Table 1. Correlation results of Pre-release defects/KLOC
(All correlations are significant at the 0.01 level (2-tailed))

	Prefast defects/ KLOC	Prefix defects / KLOC	Pre-release defects/ KLOC
Prefast defects /KLOC (ρ)	1.000		
Prefix defects /KLOC (ρ)	.380	1.000	
Pre-release defects/KLOC (ρ)	**.368**	**.577**	1.000

Table 2. Regression Fits

Predictors	Linear (R^2)	Better fits ? (R^2)
PREfast alone	0.566	Yes. Cubic (0.604)
PREfix alone	0.495	Yes. Cubic (0.514)
Both PREfast and PREfix	0.627	N/A

Data Splitting

- 199 Components, 2/3 train, 1/3 test
- Positive correlation
- Discriminant Analysis -> 83% correct classification

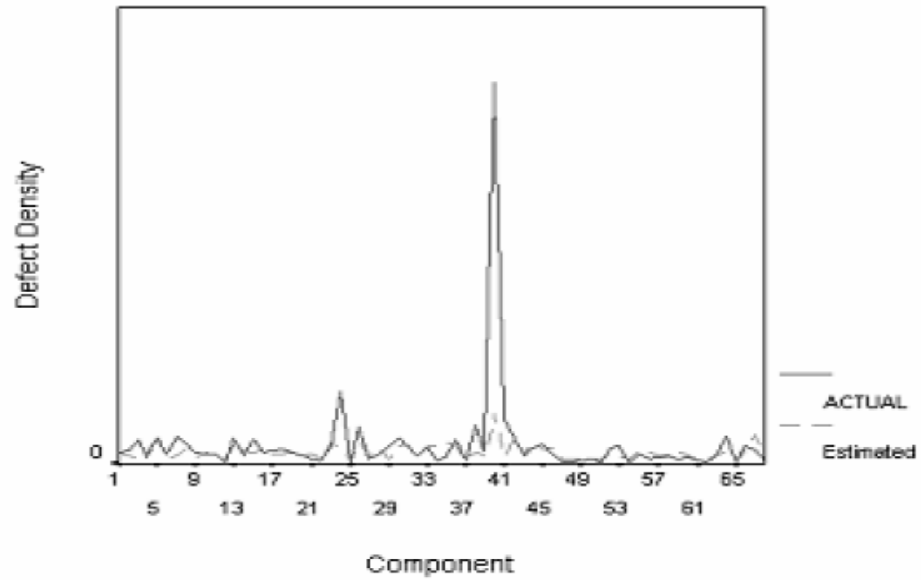


Figure 2. Actual vs. estimated pre-release defect density

Table 3. Fit and Correlation results of random model splitting

S.No	R^2	F-Test (sig)	Correlation Results (Spearman)
1.	0.870	429.79,p<0.0005	0.496, p<0.0005
2.	0.656	339.95,p<0.0005	0.536, p<0.0005
3.	0.841	122.83,p<0.0005	0.526, p<0.0005

Validation and Limitations

- Validation
  - Association -> Correlations
  - Consistency -> Regression tests
  - Discriminative Power -> Discriminant Analysis
  - Tracking -> Split data
  - Predictability -> Correlation
- Limitations
  - Single project
  - Tied to Prefast and Prefix tools

Conclusion

- Strengths
  - Tools used in process so likely not the same errors found in testing
  - Showed predictive power of static analysis
  - 83% of components correctly classified
  - Actual vs estimated graphs
- Weaknesses
  - The correlations are moderate
    - Telling us something we didn't already know?
  - Splitting the data 3 times.
    - What about 10 fold?
  - Validation and hypotheses are redundant
  - Not all Prefast are recorded