Improved Software Defect Prediction

SPIN: British Computer Society
23 February 2006

Norman Fenton
Queen Mary University of London and Agena Ltd
Using fault data to predict reliability

Pre-release testing

Post-release operation

delivery
Pre-release vs post-release faults?
Pre-release vs post-release faults: actual
Software metrics....?
Regression models....?
Solution: causal models (risk maps)

- Spec & doc Process quality
- New functionality implemented
- Development Process quality
- Total defects in
- Defects PRE
- Defects fixed
- Defects found
- Testing Process quality
- Rework Process quality
- Defects POST
- Defects found
- Defects fixed
- Total defects in
- Development Process quality
- Spec & doc Process quality
- New functionality implemented
Running the defects risk map

- New functionality implemented: 150
- Spec & doc Process quality
- Process quality
- Development Process quality
- Testing Process quality
- Rework Process quality
- Total defects in
- Defects found
- Defects fixed
- Defects POST

Median value 257

Total defects: 150
Defects found: 0
Defects fixed: 150
High quality development

Defects PRE
0

New functionality implemented
150

Spec & doc Process quality
Very high

Development Process quality
Very high

Total defects in

Defects found

Defects fixed

Defects POST

Testing Process quality

Rework Process quality

Defects

PRE

POST

Median value 40

Median value 25

Total defects 150
Defects found 0
Defects fixed Very high
Very high

Defects

PRE

POST

Median value 40

Median value 25
... but suppose more than expected defects found in testing

Total defects in

Defects PRE

0

Defects found

400

Defects fixed

Defects POST

Development Process quality

Very high

Spec & doc Process quality

Very high

Testing Process quality

New functionality implemented

150

Median value 204

So hard information about testing has ‘overturned’ evidence about the high quality process
How good must your testing and rework be to get 0 residual defects?

Spec & doc Process quality

Development Process quality

Testing Process quality

Rework Process quality

Defects PRE

New functionality implemented

Total defects in

Defects found

Defects fixed

Defects POST

Total defects in

Defects found

Defects fixed

0

150
What’s special about this approach?

• Structured
• Visual
• Robust
The specific problem for Philips

- Defects found
- Predicted
- Actual
- Time
Background to work with Philips

- Reliability models
- Fixed Risk Map (AID)
- MODIST
- Agena Risk

Years: 2000 - 2004
Projects in the trial

MM&C (9) 28%
TV (16) 50%
DVD (7) 22%

(actual number of projects shown in brackets)
Factors used at Philips

• Existing code base...
• Complexity and management of new requirements ...
• Development, testing, rework, process ...
• Overall project management ...
Actual versus predicted defects

Correlation coefficient 95%
Validation Summary (Philips’ words)

“Bayesian Network approach is innovative and one of the most promising techniques for predicting defects in software projects”

“Our evaluation showed excellent results for project sizes between 10 and 90 KLOC”

“Initially, projects outside this range were not within the scope of the default model so predictions were less accurate, but accuracy was significantly improved after standard model tailoring”

“AgenaRisk is a valuable tool as is”
Specific benefits

• Accurate predictions of defects at different phases
• Know when to stop testing
• Identify where to allocate testing
• Minimise the cost of rework
• Highlight areas for improvement
• Use out of the box now if you have no data
• Approach fully customisable
• Models arbitrary project life-cycles
Summary

- Risk maps – the way forward
- Validation results excellent
...And

You can use the technology NOW

www.agenarisk.com