Experimental Analysis of Dependency Factors of Software Product Reliability using SonarQube

Sanjay Joshi
Persistent Systems Limited, India

Bharat Deshpande
BITS Pilani, K K Birla Goa Campus, India

Sasikumar Punekkat
Malardalen University, Sweden
Background

• Current Reliability Models are having limitation due to
  • Assumptions
  • Applicable for certain phases only
  • Not taking holistic view of environment for deriving reliability

• Objective
  • To identify most influential factors making impact on reliability
    • By performing experiments on different software products in diverse domains and developed using diverse technologies.

• These experiments were performed in a large software development organization in India, which has laboratory setup necessary for performing such experiments.
Background

- In this study, we hypothesize reliability to be a function of
  
  - Process Parameters (Schedule Variance, Effort Variance, Productivity)
  - Technology
  - Design parameters (Commercial Off The Shelf (COTS) complexity, Design Complexity)
  - Testing Parameters (Unit Test Defects, Integration Test Defects, System Test Defects, Defect Leakage and Post Delivery Defects)
  - Operational Parameters (Execution Time, Skill of Developer/Tester)

(These factors were identified as the most influential ones as perceived by stakeholders)
Research Context

• Experiments involved studying impact of factors (identified during field survey) on reliability

• For each application performed minimum 30 combinations.

• For identifying most influential input factor contributing to reliability

• Capturing reliability for baseline purpose using “SonarQube”
Research Context

• Expected output: Hypothesis run on data obtained and confirm the result

• Independent parameters
  • Skill
  • Design Complexity
  • Technology
  • COTS Complexity

• Dependent parameters
  • UT Defects
  • IT Defects
  • ST Defects
  • Review Efficiency
  • Post Delivery Defects
  • Execution Time
  • Process Metrics
    • Schedule variance
    • Effort variance
    • Productivity

• Design Experiments on controlled environment
  • One parameters is variable and other are kept constant
Experimental Framework

• Application Selection Criteria for Experiments
  • Domain
  • Technology / Platform
  • Criticality
  • Design Complexity
  • Development Methodology
  • Size (KLOC or FP)

• Application Selected (Few examples)
  • Enterprise Risk Management Portal
  • ECG Monitoring System
  • Photoshop Application
  • e-Finance
Methodology

• The main objective of performing experiments is to find cause and effect relationship=> \( Y = f(x_1, x_2, x_3, x_4, \ldots, x_n) \)

• Experiments were conducted in a multinational software product organization having centers across the globe

• Series of experiments were conducted in controlled environment, where one parameter is considered as variable and other parameters are taken as constant
Methodology

• Minimum of 10K lines of code was the criteria set for developing the application

• The design document was provided to all developers

• SonarQube was run on error free code to give the reliability factor for each skill level
Planning phase:
Set criteria for input parameters and identify runs

Choose Application Domain Ai, e.g., E-finance
Choose independent factor Fj, e.g., skill

Assign task to develop instances of code (Cijk) for different levels of Fj to multiple teams

Perform the task and monitor experiments

Run SonarQube and capture reliability \( R(Cijk) \)

Are Runs Over for \((Ai, Fj)\)?

Prepare graph of \( R \) versus independent factor Fj (e.g., Skill)

Calculate Correlation between Reliability & Fj

Correlation \( \geq 0.8? \)

Yes

High Correlation

No

Correlation \(< 0.5? \)

Yes

Medium Correlation

No

Low Correlation

For every application and every independent factor repeat experiment

Ranking of factors and Validation through MTBF
Experiments (Setup)

Skill Levels 1 to n

Application Development

Constant Parameters

Variable Parameters

Output varying Parameters

Capturing data through tools

Developed Application (Warning and Compilation error free)

Scanning through SonarQube

Reliability

Statistical analysis and Conclusion
Experimental Findings

Chi Square Test

```r
> a <- table(data_skill3$Skill.Level, data_skill3$Reliability)
> chisq.test(a)
```

Pearson's Chi-squared test
data: a

X-squared = 178.25, df = 147, p-value = 0.04044

Chi Square Test

```r
> a4 <- table(data_skill4$Skill.Level, data_skill4$Reliability)
> chisq.test(a4)
```

Pearson's Chi-squared test
data: a4

X-squared = 196, df = 144, p-value = 0.002588

Chi Square Test

```r
> a7 <- table(data_skill7$Skill.Level, data_skill7$Reliability)
> chisq.test(a7)
```

Pearson's Chi-squared test
data: a7

X-squared = 157.89, df = 120, p-value = 0.01162

### Pearson's Chi2 Test

<table>
<thead>
<tr>
<th>Technology (major)</th>
<th>Chi2</th>
<th>Df</th>
<th>p value</th>
<th>Less than 0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>C#</td>
<td>393.75</td>
<td>336</td>
<td>0.0163</td>
<td>Yes</td>
</tr>
<tr>
<td>Sharepoint</td>
<td>441.15</td>
<td>344</td>
<td>0.0000304</td>
<td>Yes</td>
</tr>
<tr>
<td>ASP.NET</td>
<td>196</td>
<td>144</td>
<td>0.002588</td>
<td>Yes</td>
</tr>
<tr>
<td>Java</td>
<td>226.65</td>
<td>180</td>
<td>0.0105</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Chi Square Test

```r
> a7 <- table(data_skill7$Skill.Level, data_skill7$Reliability)
> chisq.test(a7)
```
## Experimental Findings

<table>
<thead>
<tr>
<th>Technology V/s Parameters</th>
<th>C#</th>
<th>Sharepoint</th>
<th>ASP.NET</th>
<th>Java</th>
<th>Inference</th>
<th>Average</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skill</td>
<td>0.890</td>
<td>0.9191</td>
<td>0.981</td>
<td>0.950</td>
<td>Strong</td>
<td>0.935025</td>
<td>2</td>
</tr>
<tr>
<td>UT defects</td>
<td>0.231</td>
<td>0.040</td>
<td>0.224</td>
<td>0.233</td>
<td>No</td>
<td>0.182</td>
<td>9</td>
</tr>
<tr>
<td>IT defects</td>
<td>0.295</td>
<td>0.230</td>
<td>0.262</td>
<td>0.260</td>
<td>No</td>
<td>0.26175</td>
<td>8</td>
</tr>
<tr>
<td>ST defects</td>
<td>0.6840</td>
<td>0.820</td>
<td>0.910</td>
<td>0.985</td>
<td>Good</td>
<td>0.84975</td>
<td>7</td>
</tr>
<tr>
<td>On time</td>
<td>0.040</td>
<td>0.201</td>
<td>0.040</td>
<td>0.105</td>
<td>No</td>
<td>0.0965</td>
<td>13</td>
</tr>
<tr>
<td>Load</td>
<td>0.833</td>
<td>0.789</td>
<td>0.980</td>
<td>0.820</td>
<td>Good</td>
<td>0.8555</td>
<td>6</td>
</tr>
<tr>
<td>Design Complexity</td>
<td>0.990</td>
<td>0.771</td>
<td>0.913</td>
<td>0.911</td>
<td>Good</td>
<td>0.89625</td>
<td>4</td>
</tr>
<tr>
<td>COTS</td>
<td>0.846</td>
<td>0.843</td>
<td>0.921</td>
<td>0.900</td>
<td>Good</td>
<td>0.8775</td>
<td>5</td>
</tr>
<tr>
<td>Review Efficiency</td>
<td>0.997</td>
<td>0.910</td>
<td>0.870</td>
<td>0.960</td>
<td>Strong</td>
<td>0.93425</td>
<td>3</td>
</tr>
<tr>
<td>Post Delivery Defects</td>
<td>0.936</td>
<td>0.990</td>
<td>0.960</td>
<td>0.966</td>
<td>Strong</td>
<td>0.963</td>
<td>1</td>
</tr>
<tr>
<td>SV</td>
<td>0.170</td>
<td>0.170</td>
<td>0.010</td>
<td>0.215</td>
<td>No</td>
<td>0.14125</td>
<td>11</td>
</tr>
<tr>
<td>EV</td>
<td>0.190</td>
<td>0.230</td>
<td>0.065</td>
<td>0.224</td>
<td>No</td>
<td>0.17745</td>
<td>10</td>
</tr>
<tr>
<td>Productivity</td>
<td>Productivity</td>
<td>0.160</td>
<td>0.190</td>
<td>0.051</td>
<td>No</td>
<td>0.124275</td>
<td>12</td>
</tr>
</tbody>
</table>
Sample Data – Skill (EXCEL FILE)

MWSM_Data
Conclusion

One of the noteworthy findings from these experiments are factors contribute significantly towards software product reliability:

- Post-Delivery Defects,
- Skill
- Review Efficiency

With the help of this exercise, we could also eliminate some parameters such as:

- Process metrics (Schedule Variance, Effort Variance and Productivity)
- Unit Test Defects, Integration Test Defect, System Test Defects

These experiments also indicate that following input parameters could make significant impact:

- Load Condition
- Design Complexity
- COTS
Acknowledgement

• We would like to thank Dr. Yogesh Badhe, data scientists from Persistent Systems Ltd. for his valuable support in data analysis

• Also, we would like to thank Dr. Ramprasad Joshi, Birla Institute of Technology & Science, Pilani, Goa Campus for his valuable inputs for documentation while performing experiments

• Punnekkat acknowledges support from FiC (SSF) Project
Thank You