Presentation: Predicting Fault Incidence Using Software Change History

Tao Xia

November 16, 2006

Introduction

- Study an old telecommunication system
- Use info available on 31 March 1994 to predict number of faults that appear between 1 April 1994 and 31 March 1996
- Identify aspects which related to faults potential of the modules

Fault Potential Factors

- Product Measures
  - Lines of code
  - Complexity metrics
- Process Measures
  - Number of past faults
  - Number of deltas
  - Age
  - Organization
  - Number of developers
  - Connections between modules
  - Weighted time damp model

Change Management Data

- IMR database
  - Initialize and complete time
  - Deltas associated with
  - Bug/New
  - Modules involved
- Delta database
  - Change details to each file

Statistical Tools

- Generalized Linear Models
- Simulations to Assess Significance

Results

TABLE 1
Models to Fit Fault Data

Model	Intcp	Common	Intl	US	Error
(A) Stable	-	-	-	-	757.4
(B) Null model	-	-	-	-	3108.8
(C) Organization only	3.46	0	-0.13	-1.39	2587.7
(D) 0.84 log(lines/1000)	0.92	0	0.17	-0.92	1271.4
(E) −0.14 log(lines/1000) + 1.19 log(deltas/1000)	3.31	0	0.46	-0.70	980.0
(F) 1.05 log(deltas/1000)	2.95	0	0.43	-0.72	985.1
(G) 0.07 log(lines/1000) + 0.95 log(deltas/1000) - 0.44age	2.63	0	0.73	-0.65	696.3
(H) 1.02 log(deltas/1000) - 0.44age	2.87	0	0.74	-0.63	697.4

Results

- Complexity matrix
  - Most complexity measures were virtually perfectly predictable form lines of code.
  - They are not good predictors

	1	2	3	4	5	6	7	8	9	10	11	12
1 Lines Of Code	1	.97	.88	.88	.91	.99	.98	.92	.97	.85	.72	.35
2 McCabe V(G)1	.97	1	.88	.90	.88	.95	.95	.89	.93	.86	.76	.29
3 Functions	.88	.88	1	.82	.89	.85	.84	.91	.84	.76	.65	.29
4 Breaks	.88	.90	.82	1	.83	.86	.85	.85	.85	.78	.67	.27
5 Unique Operators	.91	.88	.89	.83	1	.89	.87	1.00	.94	.65	.47	.48
6 Total Operands	.99	.95	.85	.86	.89	1	1.00	.90	.98	.85	.72	.31
7 Program Volume	.98	.95	.84	.85	.87	1.00	1	.88	.97	.87	.74	.28
8 Expected Length	.92	.89	.91	.85	1.00	.90	.88	1	.94	.69	.53	.42
9 Variable Count	.97	.93	.84	.85	.94	.98	.97	.94	1	.77	.60	.38
10 MaxSpan	.85	.86	.76	.78	.65	.85	.87	.69	.77	1	.92	-0.10
11 MeanSpan	.72	.76	.65	.67	.47	.72	.74	.53	.60	.92	1	-0.25
12 Prog Level	.35	.29	.29	.27	.48	.31	.28	.42	.38	-0.10	-0.25	1

Results

- Best Linear Model
  - Number of changes and the age
- Weighted Time Damp Model
  - Best models
  - Larger and more recent MR contributes more to the fault potential

Feedback

- Positive
  - Weighted time damp model
  - Correlation between complexity metrics (not sure where the number come from)
  - Number of developers and frequency change are not good predictors too
- Negative
  - Undefined statistic words (Poisson distribution, gamma distribution)
  - Results validation on other projects