FREIE UNIVERSITÄT BOZEN
LIBERA UNIVERSITÀ DI BOLZANO
FREE UNIVERSITY OF BOZEN · BOLZANO

European Master Software Engineering

A.A. 2010/2011

Software Metrics : Project Log

Group Members: *Nicolò Paternoster* [9426]
Carmine Giardino [9099]

Table of Content

1. Introduction

1.1 Project description

Collect failure occurrences to model software quality. *Programming part*: Create an R□package that collects defects occurrences from different OS repositories and plots them with their fitting.

1.2 Used Tools

Source Monitor http://www.campwoodsw.com/sourcemonitor.html

Oss: The students WILL ALSO HAVE TO find, and then use, tools
different from the one shown in the previous slide.
See also http://en.wikipedia.org/wiki/List_of_tools_for_static_code_analysis
R http://www.r-project.org/about.html
Eclipse Metrics Plugin http://eclipse-metrics.sourceforge.net/
SVN Repository https://babbage.inf.unibz.it/teaching/2010/11/group3

Present data in XML

1.3 Milestones

Lab	Title	Milestone
26/11/2010	GQM definition	define the GQM of your project
3/12/2010	1) Design of the tool 2) First assessment	design your tool according to your GQM
10/12/2010	Using the tool: case studies	Data collection Apply your tool to real data and get the measures defined in your GQM
17/12/2010	1) Analysis of the results and conclusions 2) Second assessment	Analyze and present the results together with the limitations of your work

2. First Lab

2.1 Log

Title: Collect failure occurrences to model software quality.
Programming part: Create a R package that collects defect occurrences from different OS repositories and plot them with their fitting (online)

We have to produce technical documentation and user manual.

The difficulty is that we need to learn a new language : R
http://www.r-project.org/
Present it in a web-like interface

1) Create the scenario (enough detailed to be able to write the GQM)
Extract the busisness goals (unreversable)
Measurement goals
Functional and non functional requirements

2)Homework

Write a report for non-expert customers (teachers) that wants to know the scenario ,BG,and the measure, and if the customer is interested in functional or non functional*
**affects customer satisfaction,business*
//Time for Questions

Create a package for R as final outcome that could be implemented in the R project.
There are several script .
Online in the sense that the user will not see the code, but just the presented data.
Software quality here is modelling the occurrences of defects over time .
*Non linear regression**

How far real data from the model? Fitting
Compare different fitting models
Our tool might tell the defects slippage
How many defects remains in the a software on a release date? (Question that sw manager wants)

Connect to Bugzilla repositories (the access to different repositories could be different)
and take the information about defect . Focus on DATE
Gather occurrences
The user can select a time window

3. GQM Definition

3.1 Scenario:

A small software house has just released a new product (a.k.a. A) to a big company (the customer) with an one year contract. But 2 months after the release date of A, the big company complains about the high number of defects occurrences of the product. Even if they like product's features and functionality and they would like to keep using it, they are not satisfied in terms of defects occurrences. The salesman of the big company warns that if the growth rate of defect occurrences will not decrease by the 20% respect to the current value (rate outlined after a discussion with an external consultant) within next 10 months they will not renew the contract. For this purpose the project manager wants to understand if they will be able to achieve this result. They have at disposal an open source database that collects information about defects of that project and they want to produce out of this data the prediction of temporal growth of defects occurrences in the subsequent 10 months.
Furthermore the internal project manager wants to receive a report containing information about progress every month.

3.2 Business Goal

BG - Beginning from two months after the product A is released the project manager of the small software house wants to achieve the failure growth rate of requested by the big company (reduce it by 20%) before the contract expires.

3.2 Measurements Goals

Glossary:

Definition of the time frame

t_0 = release date

t_i = date of customer's request (t_0 + 2 months)

t_f = upper bound for estimation , expire date of the contract (t_i + 10 months)

$model(t)$ = the function which represent prediction model of cumulative failures

satisfactory threshold = a decrease of failure occurrences growth rate by 20% respect to the growth rate by the date of customer's request (t_i).

MG 1

Analyze the open database which contains information about product failure during the last two months from the point of view of the tester in the small software house in order to evaluate it in terms of cumulative defects occurrences during the 2 months after the release.

Object study: set of failure
Purpose: evaluate
Focus: failure occurrence
Stakeholder: tester
Context factor: small software house, open failure database, perform the measure every month

Q1 - How many failures were reported in the database each day in the period from t_0 to t_i ?
Qlink1 - Failures reported in the database each day.
M1 - $\frac{\#\,of\,failures}{day},\ \forall\ day \subset \{t_0, \dots, t_i\}$

Q2 - What is the cumulative number of failures occurrences reported after two months?
Qlink2 - total cumulative number of failures after two months

M2 - $Total\ \#\ of\ failures\,,at\,t_{i}$

Q3 - How are the failure occurrences distributed over the two months period?
Qlink3 - Failure occurrences distribution over the two months period
M3 - $cumulative\ \#\ of\ failures\,(\,day\,)\,,^{\forall\ day\subset\{\,t_{0},\ldots,t_{i}\}}$

MG 2

Analyze failure occurrences of the product A during the first two months in order to predict their growth rate at the end of the total 12-months period (t_{f}) from the point of view of the tester in the small software house

Object study: set of failure occurrences every month
Purpose: predict
Focus: growth rate
Stakeholder: tester
Context factor: small software house,perform the measure every month

Q4 - Which model can represent properly the failure occurrences growth?
Qlink4 - the model that fits best the failure occurrences distribution
M4 : $min[d(mdl_{1},data)\,,\ldots,d(mdl_{i},data)]$,
$where\ d(a,b)\ is\ a\ distance\ function$

$and\ 'data'\ is\ obtained\ by\ interpolating\ cumulative\ \#\ of\ defect\ occur$

Q5 - According to the selected model, what will be the cumulative number of failure to the end of the next 10 months?
Qlink5 - cumulative number of failure to the end of the next 10 months
M5 : $model(t),evaluated\ in\ t\ =\ t_{f}$

Q6 -What is the growth rate after the first two months, according to real data?
Qlink6 - growth rate after the first two months
M6 : $\frac{d}{dt}[data(t)]\,,evaluated\ in\ t=t_{i}$

Q7 - According to the model,what will be the growth rate of failure occurrences at the end of the 10 months?
Qlink7 - growth rate of failure occurrences at the end of the 10 months
M7 : $\frac{d}{dt}[model(t)], evaluated\ in\ t = t_f$

Q8 - According to the model,what will be the time necessary to decrease the growth rate of failure occurrences by 20% respect to the growth rate at t_i?
Qlink8 - time necessary to decrease the growth rate of failure occurrences by 20% respect to the growth rate at t_i
M8 :
$f^{-1}(M5 - \frac{M5}{5}), where\ f^{-1}(growthRate)\ is\ the\ inverse\ function\ o$

Q9 - According to the model, how much extra time is needed to reach the *satisfactory threshold* ?
Qlink9 - extra time is needed to reach the *satisfactory threshold*.
M9 : $M8 - t_f$

Q10 - According to the model, what will be the cumulative total # of failure occurrences at the satisfactory threshold?
Qlink10 - Cumulative total # of failure occurrences at the *satisfactory threshold*.
M10 : $model(M8)$

Q11 - In the case in which,according to the prediction model, the product A will not be able to reach the satisfactory threshold, how many failures should the company expect to experience before reaching that level?

Qlink11 - Defect slippage from the time t_f to the needed time to reduce the growth rate of 20% of t_i
M11 : $M10 - model(t_f)$

MG 3

Object study: history of failure occurrence prediction (t_f)
Purpose: compare
Focus: growth rate
Stakeholder: project manager
Context factor: small software house,perform the measure every month

Q12 - Is the current predicted growth rate of failure occurrences better respect to the past month prediction?

Qlink12 - current predicted growth rate (t_f) of failure occurrences respect to the past month prediction

M12 : $M7_i - M7_{i+1}, \forall\, i \subset \{2, \dots, 11\}$

4. Second Lab

4.1 Log

Design of the tool (Class diagram, component diagram …)
If it connects to database, which kind of database, which kind of file,
simple tool,

1. Collect the data
2. Aggregated data view (WEB based)

//HOME Work for next week : Finish your tool, write the user manual, upload everything on SVN

The tool has to work. Work WELL (good coding , understandable)
Presentation of the work very focused on the GOALs

5. Design and specifications

6. User manual