

1.

Effects of Branching Strategy on Software Quality

It's been a nice empirical study on branching of software repositories.

It's the **first** empirical study that **evaluates and quantifies** between software quality and different aspects of branching the repository.

2.

Branching plays an important role in the development of large software projects.

Branches provides isolation to the developers

Features are developed in an isolated way without doing any interruption to the main branch or the production version of the system

Regarding Integration failures:

Branching has some disadvantages as well

If a system is using too many branches then it needs to pay additional overhead for merging and integrating the development works from too many branches which may some times lead to integration failures.

3.

This paper represents effects of branching strategies

so first of all I need to tell you what are the main strategies of branching being focused here.

-- Branching strategies are categorized in two main categories

- * Branching according to SOFTWARE ARCHITECTURE

- * Branching according to ORGANIZATIONAL STRUCTURE

4.

The relationship between branching structure and software quality remains an important open question

With more projects employing branches in their development, understanding the impact of branching is increasingly relevant

To address this

The main contribution of this paper becomes:

- * Defining metrics to understand the effects of branching

- * Quantifying the effects of branching on software quality
- * Examining the mismatch between branching structure and organizational structure
- and finally
- * recommendations for proper usage of branching been provided in this research work

5.

To achieve these phenomenons this research work was plotted on two releases of one of the largest projects of Microsoft.

6.

The development data has been collected from the release of windows 2003. The historical development data has been collected at the binary level Because failure record is usually gets reported at this level at Microsoft

To gather the metrics for each component, the **commit histories and software failure data** have been collected

in order to collect the different metrics at the component level **a mapping of source files to binaries** has been used

7.

The high level research question is "How much and in what ways does branching affects software quality"

To find the answer to this research question this work was needed to:

Evaluate software quality at component level

Then it has to come up with some hypotheses with their corresponding rationale along with.

8.

The hypotheses were burned down from the two general aspects:

- * Effects on component quality and
- * Architectural and Organizational congruence

therefore the study has been focused on three factors from the first one and two factors from the second one.

9.

First hypotheses is about the Branch Activity

where it's been proposed that software components with high branching activity have more failures

Here branch activity is considered as a ratio of number of churns for branching divided by number of churns made for development

10.

The second hypotheses says that if development is done in scattered way in multiple branch families then there is more chance of integration failure

11.

Third hypotheses is about the branch depth.....

12.

Here is an example of multiple branch families

13.

Fourth hypotheses has been burned down from the other general aspect that I mentioned before

It's about the architectural and organizational congruence

It says that if branches have higher mismatch with architectural or organizational structure then there are more chance of having integration or merge failure

14.

According to Conway's law ...

Architectural mismatch and organizational mismatch have been calculated based on

15.

The analysis has mainly been done by multiple linear regression

where it's been modeled a relationship between an observed phenomena and an observed outcome

A variance inflation factor (VIF) analysis has been employed to remove the highly correlated factors from the linear model

16.

This figure shows the case study result for the linear model

the base model consists of size, churn, complexity, development changes and number of files

four other models are also created by adding other factors like Branch Activity then Branch Scatter and Branch Scatter Entropy..

...

+ sign denotes positive effect and - sign denotes negative impact

The base model shows the model-fit is 72% and interestingly adding other factors increased the model fit up to 79%

So it can be said that Branch Activity, Branch Scatter, Branch Depth and Branch Families have high relation with integration failures.

17.

This is the metric for windows7

Here also, the model-fit or adjusted R² value has been increased with the addition of the factors

The R² value is less than Vista might be because of windows7 has less post release failures

18.

One question is still lingers

How to best align the branching structure?

Traditionally branch structures are aligned in one of two ways:

to match architecture of software system or to match the organizational structure

The advantage of matching the branching structure with the architectural structure is that changes to a component mostly happen on the same branch, so integrations get minimized.

Aligning the branching structure along the organizational structure means that branches match team boundaries.

In such case each team leader will have his own branch family.

The individual branches within the branch family will be assigned to different sub-teams.

Another linear regression model has been built for this

It examines the relationship of organizational and architectural mismatch of individual branches with branch quality.

19.

Based on the study in this research work we can come up with some recommendations of practical implementation of branching in software development:

- * we need to reduce branching activity
- * we should try to develop a component in as minimum branches as possible
- * branch structure should be aligned to organizational structure as much as possible