A Project Report

On

**BUG PREDICTION IN SOFTWARE SYSTEMS USING SOURCE CODE METRICS**

By

Ankita A.Chandak (2010A7PS066H)

S. Monica Lakshmi (2010A7PS117H) Subhajit Ray(2010A7PS452H)

UNDER THE SUPERVISION OF

Dr. N.L.Bhanu Murthy

Asst.Professor,CS dept

BITS­Pilani, Hyderabad Campus

SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS OF BITS C3235: COMPUTER ORIENTED PROJECT



BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE­PILANI

HYDERABAD CAMPUS

NOVEMBER 2013

**ACKNOWLEDGEMENT**

We are deeply indebted to our project guide, Dr.Bhanu Murthy Asst.Professor,CS dept.,

BITS-Pilani, Hyderabad Campus, for sharing his

invaluable time and expertise. He has been a true mentor and motivator. His able guidance has truly helped us in our progress in the project work.

Ankita A.Chandak

S. Monica Lakshmi

Subhajit Ray

**ABSTRACT**

Reliably predicting software defects is one of software engineering’s holy grails. Researchers have devised and implemented a number of bug prediction approaches varying in terms of accuracy, complexity and the input data they require.

In this project a few such approaches have been handpicked and implemented. Data sets have been generated/extracted from publicly available sources and then used to train/test our implemented models. The results thus obtained have been used for extensive comparison of the explanatory and predictive power of well­known bug prediction approaches.

**INTRODUCTION**

Many organizations want to predict the number of defects (faults) in software systems, before they are deployed, to gauge the likely delivered quality and maintenance effort. To help in this, numerous software metrics and statistical models have been developed, with a correspondingly large literature. Most of the wide range of prediction models use size and complexity metrics to predict defects. Others are based on testing data, the "quality" of the development process, or take a multivariate approach. The authors of the models have often made large contributions to a subject otherwise bereft of empirical studies. However, there are a number of serious theoretical and practical problems in many studies. The models are weak because of their inability to cope with the, as yet, unknown relationship between defects and failures. There are fundamental statistical and data quality problems that undermine model validity. More significantly many prediction models tend to model only part of the underlying problem and seriously misclassify it.

In this project, we have implemented and examined two approaches for the prediction of bugs in software systems:

1. CHANGE COUPLING, which examines the implicit relationship between two or more software components that have been observed to frequently change together during the evolution of a software system.
2. CHANGE BURST, which examines the effect of continuous burst of changes on software systems.

I. CHANGE COUPLING

Change coupling is the implicit and evolutionary dependency of two software components that have been observed to frequently change together during the evolution of a software system. The more they change together, the stronger the change coupling dependency is. However, there is no consensus on the formal definition of change coupling, and several alternative measures exist. The authors of the paper have formally defined 4 measures of change coupling emphasizing different aspects. To measure the correlation of change coupling with software defects change coupling measures have been defined for each entity in the system. The entity in our case is a class, as classes are a cornerstone of the object oriented paradigm. The measures thus defined concern the coupling of a class with the entire system.

The following definitions have used the concept of n-coupled classes.

A *transaction* corresponds to a commit in the software repository. It is a set of files which were modified and committed to the repository, together with the commit timestamp, the author, and the comment written by the author at commit time.

Two classes are considered *n-coupled* when there are at least n transactions which include both the classes.

Thus, all our change coupling measures are functions of n. Given two classes c1 and c2, they are n-coupled if the following condition holds:



where T is the set of all the transactions. Given a class c we define the set of coupled classes (SCC) as:





The figure shows an example scenario with 5 classes and 6 transactions. In this case SCC(c1, 3) = {c2, 5g} , SCC(c1, 4) = {c2, c5} and SCC(c1, 5) = {c5}.

The change-coupling measures examined by us are:

*A. Number of Coupled Classes (NOCC)*

The first per-class measure of change coupling is the number of classes n-coupled with a given class c. This measure emphasizes the raw number of classes with which a given class is coupled with. NOCC is defined as:



The NOCC measure is the cardinality of the set of coupled classes. In the example figure NOCC(c1, 3) = 2.

*B. Sum of Coupling (SOC)*

The sum of coupling is the sum of the shared transactions between a given class c and all the classes n-coupled with c. SOC is defined as:



The SOC measure is the sum of the cardinalities of the sets of transactions which include the class c and the classes n-coupled with c. Compared to NOCC, SOC also takes into account the strength of the couplings.

In the given figure, SOC(c1, 3) = 4 + 5 = 9.

*C. Exponentially Weighted Sum of Coupling (EWSOC)*

EWSOC is a variation of SOC, where the shared transactions are exponentially weighted according to their distance in time, emphasizing recent changes over past changes. EWSOC is defined as:



T(c) are all the transactions, sorted by time, which include the class c.



The figure above shows an example of computation of EWSOC for the class c1 for n = 3. In this case T(c1) = {t1, t3, t4, t5, t6}, |T(c1)| = 5 (t2 is not included in the computation since c is absent in it) and therefore



*D. Linearly Weighted Sum of Coupling (LWSOC)*

The last per-class measure of change coupling is another variation of the sum of coupling, in which the shared transactions are linearly weighted according to their distance in time. Like EWSOC, LWSOC emphasizes recent changes, penalizing past changes less. LWSOC is defined as:



In the second figure,



II. CHANGE BURST

A change burst is a sequence of consecutive changes. The history of a system is split into a series of calendar days where each day may or may not have commits.

We thus see a system S as a sequence of builds S = <s1, s2 ….. >, where each build differs from the previous one si =/= si - 1.

We assume that each build is created out of individual components (classes). A component C of S also has a history across builds, and thus comes in a series of individual component versions ci; we thus define

C = <c1, c2 ….>. If ci =/= ci - 1 holds, then the component C has changed in build si.

For each component, we now determine its change bursts as sequences of consecutive changes. These change bursts are determined by two parameters:

* Gap size. We would like to permit short gaps in these sequences, such that a one-day distraction will not break the burst. Therefore the gap size G parameter is introduced, which determines the minimum distance between two changes. If two changes have a distance that is smaller than G, they will be part of the same burst.
* Burst size. On the other hand, we would like to consider only bursts of a certain significance or a certain length. The burst size B determines the minimum number of changes in a burst. If the number of changes in a burst b is smaller than B, it will not be considered.



In the example figure, the arrow at the top shows the sequence of builds; the builds in which a component c has changed are marked with a dot. The rectangles below show the change bursts. If we set the gap size and the burst size to 1, then all directly consecutive changes will be merged to bursts. Increasing the gap size yields longer bursts; increasing the burst size eliminates shorter bursts.

Change burst metrics:

For each component C, we determine change metrics, temporal metrics, people metrics, and churn metrics. In their formal definitions, a component history is

C = <c1, c2 ….> and its bursts are bursts(C) = <B1, B2 ….>

*A. Change Metrics*

Metrics for a component C concerning the size and extent of its change bursts.

1. NumberOfChanges : This is the number of builds in which the component C has changed. The more a component is changed, the more likely it is to have defects.
2. NumberOfConsecutiveChanges : This is the number of consecutive builds for a given gap size G. This measure takes into account all consecutive changes, not just bursts exceeding a given size.
3. NumberOfChangeBursts : This is the number of change bursts for a given gap size G and burst size B. We would assume that change bursts are risky; thus, the number of change bursts may be predictive for defect-prone components.
4. TotalBurstSize : This is the number of changed builds in all change bursts. Assuming that change bursts indicate risky activities, a high number of changes during these bursts could be particularly risky.
5. MaximumChangeBurst : This is the maximum number of changed builds in all change bursts.

*B. Temporal Metrics*

Metrics for a component highlighting when change bursts occurred.

1. TimeFirstBurst : We determine when the first burst occurred, normalized to the total number of builds. We assume that early change bursts may have the longest impact during the project.
2. TimeLastBurst : As TimeFirstBurst, but looking at the last burst instead. We assume that the last activities before release define the final shape of the component and thus may be particularly predictive.
3. TimeMaxBurst. As TimeFirstBurst, but looking at the greatest burst instead. Let Bmax be the burst with the most changes.

*C. People Metrics*

In the following definitions, let people(C) be the set of developers who committed the changes to the component C.

1. PeopleTotal. The number of people who ever committed a change to the component C. The number of developers working on a component may be related to defects.
2. TotalPeopleInBurst. Across all bursts, the number of people involved. We see change bursts as defining moments; hence, the number of people involved may be particularly predictive.
3. MaxPeopleInBurst. Across all bursts, the maximum number of people involved. Again, we are looking for extremes, and check for the greatest change burst in terms of people.

*D. Churn Metrics*

In the following definitions, let churn(c) be the number of lines that were added, deleted, or modified during the changes to the component c.

1. ChurnTotal. The total churn over the lifetime of a component C. We assume that the more has changed, the higher the likelihood defects will be introduced.
2. TotalChurnInBurst. The total churn in all change bursts. Again, we see change bursts as defining moments; hence, the amount of change involved may be particularly predictive.
3. MaxChurnInBurst. Across all bursts, this is the maximum churn. Again, we are looking for extremes across change bursts.

**METHODOLOGY**

Our work relies on the following modules:

Commit Fetching

Module

Commit Filtering

Module

Change Coupling

Module

Change Burst Module

Training and Prediction Module

1. Commit Fetching module: This module contacts the GIT repository for a particular project and fetches the SHAs of all commits made between STARTDATE and ENDDATE. GIT uses SHAs to maintain integrity and security of its repository tree. Once the SHAs have been fetched, the actual commit details are fetched for each SHA in a multi-threaded environment.
2. Commit Filtering Module: The commit details fetched in the previous module are filtered for missing values. Each string in the commit data in entered in a dictionary and replaced with an integer, to make data retrieval/computations faster.
3. Change Coupling Module: This module takes input from the filtering module and then computes the values for NOCC, SOC, EWSOC and LWSOC in a multithreaded environment. It computes these values for different values of n(coupling threshold) and stores them in separate files.
4. Change Burst Module: This module also takes input from the filtering module. It computes the change burst features as explained previously and stores their values in separate files.
5. Training and Prediction Module:
   1. Regression: This module creates and evaluates different regression models for change coupling and change burst data in which the independent variables (for predicting) are respectively, change coupling measures and change burst measures, while the dependent variable (the predicted one) is the number of bugs. Our experiments follow the methodology proposed by Nagappan, which consists of the following steps: building regression models (stepwise regression), evaluating explanative power and evaluating prediction power. The regression was done for 50 folds with 2/3rd training set and 1/3rd testing set. The same was done taking 90% training and 10% testing.
   2. Classification: We applied ID3 algorithm, Random forest and naïve Bayes classifier for decision making and found the precision, recall and accuracy for the all n (till n = 40) values in change coupling and all combinations (from 0 to 6) of burst gap and burst size in change burst.

# OBSERVATIONS

**A. Regression model**

The table below gives the results of both the above mentioned approaches on the data of Eclipse, PDE and Equinox in terms of explanative power (adjusted R2 ), and predictive power (Spearman’s correlation).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| n |  | n |  | n |  | n |  |
| 0 |  | 10 | 0.1507 | 20 | 0.3443 | 30 | 0.0651 |
| 1 | 0.5697 | 11 | 0.1192 | 21 | 0.3571 | 31 | 0.0348 |
| 2 | 0.5617 | 12 | 0.2035 | 22 | 0.3036 | 32 | 0.0736 |
| 3 | 0.4365 | 13 | 0.2505 | 23 | 0.2890 | 33 | ­0.0196 |
| 4 | 0.5630 | 14 | 0.2573 | 24 | 0.2607 | 34 | 0.0315 |
| 5 | 0.5595 | 15 | 0.2824 | 25 | 0.2457 | 35 | ­0.0160 |
| 6 | 0.4812 | 16 | 0.2830 | 26 | 0.2165 | 36 | 0.1061 |
| 7 | 0.3992 | 17 | 0.3021 | 27 | 0.1517 | 37 | 0.0929 |
| 8 | 0.3882 | 18 | 0.3688 | 28 | 0.1951 | 38 | 0.0930 |
| 9 | 0.3499 | 19 | 0.3377 | 29 | 0.1853 | 39 | 0.0933 |

Table1.1: Change Coupling Spearman’s coefficient for Eclipse

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| n |  | n |  | n |  | n |  |
| 0 | 0.3815 | 10 | 0.422 | 20 | 0.3812 | 30 | 0.1917 |
| 1 | 0.3868 | 11 | 0.4171 | 21 | 0.3593 | 31 | 0.1719 |
| 2 | 0.4215 | 12 | 0.4252 | 22 | 0.3551 | 32 | 0.1648 |
| 3 | 0.4401 | 13 | 0.419 | 23 | 0.3168 | 33 | 0.1503 |
| 4 | 0.4458 | 14 | 0.4093 | 24 | 0.2972 | 34 | 0.1503 |
| 5 | 0.4375 | 15 | 0.4079 | 25 | 0.28 | 35 | 0.1439 |
| 6 | 0.4344 | 16 | 0.3874 | 26 | 0.2856 | 36 | 0.0788 |
| 7 | 0.4327 | 17 | 0.3703 | 27 | 0.2465 | 37 | 0.01056 |
| 8 | 0.4237 | 18 | 0.3602 | 28 | 0.2188 | 38 | 0.01056 |
| 9 | 0.417 | 19 | 0.3651 | 29 | 0.2435 | 39 | 0.01056 |

Table1.2: Change Coupling Adjusted R square for Eclipse

The best value was found for n = 1, and n = 4 for Spearman’s coefficient and Adj.R2 Thus overall we take n = 4 as the best n to be considered. Later we find the precision, accuracy and recall for all the n values and compare the results for optimum ‘n’ value.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| n |  | n |  | n |  | n |  |  |
| 1 | 0.3500 | 9 | 0.2778 | 17 | 0.1937 |  | 25 | 0.0297 |
| 2 | 0.3465 | 10 | 0.2242 | 18 | 0.1293 |  | 26 | 0.0756 |
| 3 | 0.3825 | 11 | 0.3226 | 19 | 0.1304 |  | 27 | 0.0751 |
| 4 | 0.3158 | 12 | 0.2889 | 20 | 0.0947 |  | 28 | 0.0755 |
| 5 | 0.2316 | 13 | 0.2953 | 21 | 0.0844 |  | 29 | 0.0755 |
| 6 | 0.2298 | 14 | 0.2199 | 22 | 0.0569 |  | 30 | 0.0753 |
| 7 | 0.2127 | 15 | 0.2145 | 23 | 0.0677 |  | 31 | 0.0755 |
| 8 | 0.1986 | 16 | 0.1859 | 24 | 0.0726 |  |  |  |

Table 2.1: Change Coupling Spearman’s coefficient PDE

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| n |  | n |  | n |  |
| 0 | 0.1214 | 8 | 0.4215 | 16 | 0.4908 |
| 1 | 0.1336 | 9 | 0.4637 | 17 | 0.5041 |
| 2 | 0.1203 | 10 | 0.4662 | 18 | 0.4487 |
| 3 | 0.1086 | 11 | 0.4944 | 19 | 0.5377 |
| 4 | 0.1257 | 12 | 0.5217 | 20 | 0.5154 |
| 5 | 0.1675 | 13 | 0.5051 | 21 | 0.4663 |
| 6 | 0.2308 | 14 | 0.4741 | 22 | 0.4997 |
| 7 | 0.3358 | 15 | 0.4692 | 23 |  |

Table 2.2: Change Coupling Adjusted R Square for PDE

Here Spearman’s coefficient and adj.R2  did not give same N as the optimum

The precision for n=1, n=12 and n=19 are tabulated below in table 2.3

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ID3 | |  | |  | | Random | Forest |  | Naive | | Bayes | |  | |
| n | Precision | | Recall | | Accur | | Precision | Recall | Accuracy | Precision | | Recall | | Accur. | |
| 1 | 0.858 |  | 0.854 |  | 0.229 |  | 0.476 | 0.356 | 0.356 | 0.267 |  | 0.356 |  | 0.3556 |  |
| 12 | 0.067 |  | 0.258 |  | 0.239 |  | 0.058 | 0.242 | 0.242 | 0.067 |  | 0.246 |  | 24.558 |  |
| 19 | 0.061 | | 0.246 | | 0.242 | | 0.058 | 0.242 | 0.24165 | 0.059 | | 0.242 | | 0.2416 | |

Table 2.2

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| n |  | N |  | n |  |
| 1 | 0.6313 | 9 | 0.3503 | 17 | 0.1096 |
| 2 | 0.3562 | 10 | 0.3763 | 18 | 0.0787 |
| 3 | 0.4363 | 11 | 0.2785 | 19 | 0.0497 |
| 4 | 0.2369 | 12 | 0.2101 | 20 | 0.0448 |
| 5 | 0.2935 | 13 | 0.2178 | 21 | 0.1022 |
| 6 | 0.1881 | 14 | 0.0708 | 22 | 0.1594 |
| 7 | 0.3326 | 15 | 0.1216 | 23 |  |
| 8 | 0.4426 | 16 | 0.1328 | 24 |  |

Table 3.1 Change Coupling Spearman’s coefficient for Equinox

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| n |  | n |  | n |  |
| 0 | 0.3068 | 8 | 0.2808 | 16 | 0.239 |
| 1 | 0.3771 | 9 | 0.3199 | 17 | 0.4532 |
| 2 | 0.2585 | 10 | 0.3412 | 18 | 0.3719 |
| 3 | 0.3029 | 11 | 0.2458 | 19 | 0.3719 |
| 4 | 0.2444 | 12 | 0.238 | 20 | 0.06758 |
| 5 | 0.2425 | 13 | 0.2435 | 21 | ­0.002543 |
| 6 | 0.2405 | 14 | 0.2742 | 22 | ­0.003851 |
| 7 | 0.2741 | 15 | 0.2419 | 23 |  |

Table 3.2 Change Coupling Adjusted R square for Equinox

Both Rho and adj.R2 result into n=1 as optimum coupling. Precision, Recall and Accuracy and incorrect values for n=1 from ID3 : 0.863, 0.896, 39.0909 %, 4.5455 % from Random forest: 0.863, 0.896, 39.0909 %. 4.5455 % from Naive Bayes Classifier : 0.334, 0.436, 43.6364 %, 56.3636 %

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 0 | 0.5821 | 0.5575 | 0.7571 | 0.7859 | 0.7264 | 0.7339 | 0.7559 |
| 1 | 0.3247 | 0.4880 | 0.6865 | 0.7954 | 0.8060 | 0.7072 | 0.7596 |
| 2 | 0.3504 | 0.5510 | 0.6990 | 0.7743 | 0.8114 | 0.7949 | 0.7892 |
| 3 | 0.4035 | 0.4714 | 0.7549 | 0.7014 | 0.7534 | 0.8144 | 0.7682 |
| 4 | 0.3332 | 0.5384 | 0.7590 | 0.7224 | 0.7914 | 0.7869 | 0.7927 |
| 5 | 0.4113 | 0.6309 | 0.6885 | 0.7189 | 0.7943 | 0.7463 | 0.8150 |
| 6 | 0.5810 | 0.5714 | 0.7145 | 0.7688 | 0.7984 | 0.8152 | 0.7185 |

Table 4.1: Change Burst Spearman’s Coefficient for Eclipse

Rows corresponds to gap size and column to burst size

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 0 | 0.5151 | 0.5202 | 0.5442 | 0.5708 | 0.5805 | 0.5211 | 0.5062 |
| 1 | 0.5197 | 0.5237 | 0.5328 | 0.5798 | 0.5783 | 0.5679 | 0.5773 |
| 2 | 0.5276 | 0.5349 | 0.5332 | 0.552 | 0.5502 | 0.5774 | 0.5634 |
| 3 | 0.5329 | 0.5337 | 0.5351 | 0.5614 | 0.555 | 0.5634 | 0.5965 |
| 4 | 0.5389 | 0.5388 | 0.5307 | 0.5461 | 0.5554 | 0.5563 | 0.57 |
| 5 | 0.5407 | 0.5403 | 0.5188 | 0.5365 | 0.5335 | 0.5487 | 0.5489 |
| 6 | 0.5346 | 0.5344 | 0.5189 | 0.523 | 0.523 | 0.5485 | 0.5423 |

Table 4.2: Change Burst Adjusted R square for Eclipse

Rows corresponds to gap size and column to burst size

Precision, Recall and Accuracy and incorrect values for burst size = 6 and gap size = 5 from ID3 : 0.856 , 0.86, 59.59 % from Random forest: 0.667, 0.791, 79.06% from Naive Bayes Classifier :0.782, 0.726, 72.57 %

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 |  | 1 |  | 2 | 3 | 4 | 5 | 6 |
| 0 |  | ­0.1746 |  | ­0.1470 | 0.7472 | 0.6977 | 0.7411 | 0.7437 | 0.7478 |
| 1 |  | 0.7530 |  | 0.7246 | 0.7449 | 0.7240 | 0.7269 | 0.7404 | 0.7471 |
| 2 |  | 0.8102 |  | 0.7619 | 0.7521 | 0.7307 | 0.7261 | 0.7357 | 0.7417 |
| 3 |  | 0.8054 |  | 0.7785 | 0.7588 | 0.7424 | 0.7232 | 0.7210 | 0.7372 |
| 4 |  | 0.8068 |  | 0.7860 | 0.7509 | 0.7401 | 0.7255 | 0.7190 | 0.7254 |
| 5 |  | 0.8088 |  | 0.8056 | 0.7648 | 0.7340 | 0.7403 | 0.7454 | 0.7121 |
| 6 |  | 0.8074 |  | 0.8024 | 0.7668 | 0.7424 | 0.7427 | 0.7291 | 0.7346 |

Table 5.1: Change Burst Spearman’s Coefficient for PDE

Rows corresponds to gap size and column to burst size

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 0 | 0.5635 | 0.6042 | 0.6253 | 0.6351 | 0.6316 | 0.63 | 0.6237 |
| 1 | 0.6135 | 0.6177 | 0.6287 | 0.637 | 0.6347 | 0.6316 | 0.6273 |
| 2 | 0.6187 | 0.6149 | 0.6301 | 0.6407 | 0.6334 | 0.6327 | 0.6296 |
| 3 | 0.6166 | 0.6122 | 0.6349 | 0.639 | 0.6293 | 0.6345 | 0.628 |
| 4 | 0.6122 | 0.6132 | 0.6346 | 0.6447 | 0.6304 | 0.6347 | 0.6348 |
| 5 | 0.6101 | 0.6079 | 0.6336 | 0.6443 | 0.6346 | 0.6404 | 0.6413 |
| 6 | 0.6219 | 0.611 | 0.6324 | 0.6404 | 0.6348 | 0.6397 | 0.6398 |

Table 5.2: Change Burst Adjusted R square for PDE

Rows corresponds to gap size and column to burst size

Precision, Recall and Accuracy and incorrect values for burst size = 2 and gap size = 0 from ID3 :0.3 , 0.299, 11.7878 %, 27.7014 % from Random forest: 0.249, 0.306, 30.6483 %, 69.3517 % from Naive Bayes Classifier :0.306, 0.342, 34.1847 %, 65.8153 %

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 0 | 0.8135 | 0.7878 | 0.7530 | 0.8182 | 0.8186 | 0.8345 | 0.8221 |
| 1 | 0.5588 | 0.4432 | 0.7604 | 0.7934 | 0.8263 | 0.8230 | 0.8325 |
| 2 | 0.6301 | 0.6010 | 0.7353 | 0.7736 | 0.8117 | 0.8202 | 0.8221 |
| 3 | 0.7175 | 0.6638 | 0.7830 | 0.7916 | 0.8180 | 0.8035 | 0.8153 |
| 4 | 0.6946 | 0.7120 | 0.7609 | 0.7916 | 0.8047 | 0.7821 | 0.7896 |
| 5 | 0.7887 | 0.7030 | 0.7815 | 0.7858 | 0.7966 | 0.8003 | 0.8117 |
| 6 | 0.7122 | 0.6510 | 0.7607 | 0.7852 | ­0.0287 | 0.8103 | 0.8081 |

Table 6.1: Change Burst Spearman’s coefficient for Equinox

Rows corresponds to gap size and column to burst size

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| 0 | 0.532 | 0.5441 | 0.6007 | 0.5499 | 0.5033 | 0.5033 | 0.5033 |
| 1 | 0.5544 | 0.5529 | 0.6325 | 0.634 | 0.519 | 0.519 | 0.519 |
| 2 | 0.5526 | 0.5809 | 0.6162 | 0.6349 | 0.5984 | 0.5268 | 0.5238 |
| 3 | 0.553 | 0.584 | 0.6391 | 0.6207 | 0.6088 | 0.5544 | 0.5216 |
| 4 | 0.5441 | 0.5976 | 0.5927 | 0.5796 | 0.6244 | 0.6018 | 0.5928 |
| 5 | 0.536 | 0.6033 | 0.577 | 0.6129 | 0.642 | 0.651 | 0.6452 |
| 6 | 0.5271 | 0.5446 | 0.5502 | 0.5856 | 0.5568 | 0.6742 | 0.6737 |

Table 6.2: Change Burst Adjusted R square for Equinox

Rows corresponds to gap size and column to burst size

Precision, Recall and Accuracy and incorrect values for burst size = 0 and gap size = 5 from ID3 : 0.538, 0.667, 29.0909 %, 14.5455 % from Random forest:0.427, 0.464, 46.3636 %, 53.6364 % from Naive Bayes Classifier : 0.506, 0.464, 46.3636 %, 53.6364 %

**B. Classification**

It is observed that the best values for precision and recall are obtained in decision tree using ID3 algorithm while the accuracy is best for random forest and least for decision tree.

The table below shows the summarized results.

For change burst, rows corresponds to gap size and column to burst size. Also the each gap­burst size combination represents three values, first corresponding to decision tree using ID3, second to random forest and third to Naive Bayes classifier.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| **0** |  | 0.874 | 0.865 | 0.864 | 0.872 | 0.866 | 0.867 |
|  |  | 0.728 | 0.651 | 0.653 | 0.672 | 0.684 | 0.65 |
|  |  | 0.787 | 0.783 | 0.789 | 0.78 | 0.781 | 0.794 |
|  |  |  |  |  |  |  |  |
| **1** | 0.874 | 0.874 | 0.87 | 0.857 | 0.878 | 0.854 | 0.866 |
|  | 0.653 | 0.656 | 0.654 | 0.654 | 0.653 | 0.688 | 0.649 |
|  | 0.778 | 0.79 | 0.779 | 0.785 | 0.778 | 0.777 | 0.793 |
|  |  |  |  |  |  |  |  |
| **2** | 0.874 | 0.874 | 0.87 | 0.85 | 0.883 | 0.834 | 0.86 |
|  | 0.705 | 0.699 | 0.693 | 0.697 | 0.666 | 0.67 | 0.651 |
|  | 0.78 | 0.793 | 0.788 | 0.787 | 0.782 | 0.777 | 0.79 |
|  |  |  |  |  |  |  |  |
| **3** | 0.881 | 0.876 | 0.871 | 0.861 | 0.883 | 0.848 | 0.857 |
|  | 0.688 | 0.668 | 0.654 | 0.707 | 0.654 | 0.667 | 0.662 |
|  | 0.787 | 0.798 | 0.777 | 0.783 | 0.785 | 0.776 | 0.794 |
|  |  |  |  |  |  |  |  |
| **4** | 0.881 | 0.876 | 0.87 | 0.879 | 0.879 | 0.857 | 0.86 |
|  | 0.709 | 0.68 | 0.726 | 0.667 | 0.667 | 0.86 | 0.649 |
|  | 0.784 | 0.793 | 0.777 | 0.786 | 0.786 | 0.785 | 0.792 |
|  |  |  |  |  |  |  |  |
| **5** | 0.882 | 0.876 | 0.862 | 0.861 | 0.879 | 0.863 | 0.862 |
|  | 0.652 | 0.698 | 0.681 | 0.714 | 0.655 | 0.661 | 0.65 |
|  | 0.787 | 0.782 | 0.776 | 0.783 | 0.787 | 0.795 | 0.79 |
|  |  |  |  |  |  |  |  |
| **6** | 0.889 | 0.876 | 0.86 | 0.841 | 0.877 | 0.856 | 0.867 |
|  | 0.655 | 0.694 | 0.693 | 0.703 | 0.652 | 0.667 | 0.653 |
|  | 0.786 | 0.788 | 0.777 | 0.79 | 0.778 | 0.782 | 0.793 |

Table 7.1. Change burst precision

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** |
| 0 |  | 0.896 | 0.877 | 0.881 | 0.889 | 0.871 | 0.89 |
|  |  | 0.811 | 0.796 | 0.796 | 0.796 | 0.799 | 0.785 |
|  |  | 0.732 | 0.746 | 0.74 | 0.714 | 0.699 | 0.729 |
|  |  |  |  |  |  |  |  |
| 1 | 0.899 | **0.896** | 0.882 | 0.882 | 0.9 | 0.855 | 0.886 |
|  | 0.808 | 0.799 | 0.799 | 0.799 | 0.796 | 0.802 | 0.782 |
|  | 0.732 | 0.737 | 0.74 | 0.74 | 0.714 | 0.699 | 0.726 |
|  |  |  |  |  |  |  |  |
| 2 | 0.899 | 0.896 | 0.882 | 0.875 | 0.894 | 0.845 | 0.892 |
|  | 0.808 | 0.896 | 0.802 | 0.785 | 0.788 | 0.794 | 0.782 |
|  | 0.729 | 0.896 | 0.74 | 0.737 | 0.726 | 0.714 | 0.726 |
|  |  |  |  |  |  |  |  |
| 3 | 0.904 | 0.9 | 0.883 | 0.877 | 0.898 | 0.848 | 0.888 |
|  | 0.796 | 0.794 | 0.802 | 0.805 | 0.802 | 0.788 | 0.785 |
|  | 0.732 | 0.752 | 0.737 | 0.732 | 0.723 | 0.702 | 0.732 |
|  |  |  |  |  |  |  |  |
| 4 | 0.904 | 0.9 | 0.888 | 0.898 | 0.898 | 0.853 | 0.892 |
|  | 0.808 | 0.796 | 0.814 | 0.791 | 0.791 | 0.892 | 0.782 |
|  | 0.714 | 0.735 | 0.737 | 0.723 | 0.723 | 0.717 | 0.732 |
|  |  |  |  |  |  |  |  |
| 5 | 0.904 | 0.901 | 0.879 | 0.877 | 0.895 | 0.863 | 0.891 |
|  | 0.799 | 0.805 | 0.796 | 0.799 | 0.805 | 0.779 | 0.788 |
|  | 0.729 | 0.743 | 0.74 | 0.737 | 0.726 | 0.735 | 0.723 |
|  |  |  |  |  |  |  |  |
| 6 | 0.908 | 0.901 | 0.877 | 0.875 | 0.894 | 0.86 | 0.887 |
|  | 0.805 | 0.802 | 0.805 | 0.788 | 0.785 | 0.791 | 0.785 |
|  | 0.732 | 0.746 | 0.732 | 0.746 | 0.717 | 0.726 | 0.729 |

Table 7.2. Change Burst Recall

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **0** | **1** | **2** | **3** | **4** | **5** | **6** |
| 0 |  | 53.39% | 54.87% | 56.93% | 56.93% | 59.59% | 62.24% |
|  |  | 18.88% | 79.65% | 79.65% | 79.65% | 79.94% | 78.47% |
|  |  | 73.16% | 74.63% | 74.04% | 71.39% | 69.91% | 72.86% |
|  |  |  |  |  |  |  |  |
| 1 | 52.80% | 53.39% | 55.16% | 57.52% | 58.11% | 59.00% | 61.65% |
|  | 80.83% | 79.94% | 79.94% | 79.94% | 79.65% | 80.24% | 78.17% |
|  | 73.16% | 73.75% | 74.04% | 74.04% | 71.39% | 69.91% | 72.57% |
|  |  |  |  |  |  |  |  |
| 2 | 52.8024 | 53.3923 | 55.16% | 57.8171 | 57.2271 | 58.11% | 61.06% |
|  | 0.8083 | 53.3923 | 80.24% | 78.4661 | 78.7611 | 79.35% | 78.17% |
|  | 72.8614 | 53.3923 | 74.04% | 73.7463 | 72.5664 | 71.39% | 72.57% |
|  |  |  |  |  |  |  |  |
| 3 | 52.51% | 53.39% | 55.46% | 56.64% | 56.93% | 57.82% | 61.06% |
|  | 79.65% | 79.35% | 80.24% | 80.53% | 80.24% | 78.76% | 78.47% |
|  | 73.16% | 75.22% | 73.75% | 73.16% | 72.27% | 70.21% | 73.16% |
|  |  |  |  |  |  |  |  |
| 4 | 52.51% | 53.39% | 56.34% | 57.23% | 57.23% | 58.11% | 61.06% |
|  | 80.83% | 79.65% | 81.42% | 79.06% | 79.06% | 61.06% | 78.17% |
|  | 71.39% | 73.45% | 73.75% | 72.27% | 72.27% | 71.68% | 73.16% |
|  |  |  |  |  |  |  |  |
| 5 | 52.80% | 53.69% | 55.75% | 56.64% | 57.82% | 59.29% | 60.18% |
|  | 79.94% | 80.53% | 79.65% | 79.94% | 80.53% | 77.88% | 78.76% |
|  | 72.86% | 74.34% | 74.04% | 73.75% | 72.57% | 73.45% | 72.27% |
|  |  |  |  |  |  |  |  |
| 6 | 52.51% | 53.69% | 54.87% | 57.82% | 57.23% | 59.59% | 60.47% |
|  | 80.53% | 80.24% | 80.53% | 78.76% | 78.47% | 79.06% | 78.47% |
|  | 73.16% | 74.63% | 73.16% | 74.63% | 71.68% | 72.57% | 72.86% |

Table 7.3. Change Burst Accuracy

For change coupling, n corresponds to ‘n­coupled’ class

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| n | Precision | Recall | Accuracy | Incorrectly  Classified  Instance |
|  |  |  |  |  |
| 1 | 0.761 | 0.853 | 23.89% | 4.13% |
|  | 0.651 | 0.796 | 79.65% | 20.35% |
|  | 0.66 | 0.782 | 78.17% | 21.83% |
|  |  |  |  |  |
| 2 | 0.766 | 0.856 | 24.48% | 4.13% |
|  | 0.652 | 0.802 | 80.24% | 19.76% |
|  | 0.663 | 0.69 | 69.03% | 30.97% |
|  |  |  |  |  |
| 3 | 0.756 | 0.86 | 25.37% | 4.13% |
|  | 0.651 | 0.796 | 79.65% | 20.35% |
|  | 0.686 | 0.643 | 64.31% | 35.69% |
|  |  |  |  |  |
| 4 | 0.744 | 0.855 | 29.50% | 5.01% |
|  | 0.654 | 0.802 | 80.24% | 19.76% |
|  | 0.67 | 0.658 | 65.78% | 34.22% |
|  |  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 5 | 0.763 | 0.867 | 36.58% | 5.60% |
|  | 0.652 | 0.799 | 79.94% | 20.06% |
|  | 0.699 | 0.652 | 65.19% | 34.81% |
|  |  |  |  |  |
| 6 | 0.768 | 0.849 |  | 7.96% |
|  | 0.68 | 0.799 | 79.94% | 20.06% |
|  | 0.691 | 0.646 | 64.60% | 35.40% |
|  |  |  |  |  |
| 7 | 0.727 | 0.848 | 51.03% | 9.14% |
|  | 0.653 | 0.805 | 80.53% | 19.47% |
|  | 0.676 | 0.658 | 65.78% | 34.22% |
|  |  |  |  |  |
| 8 | 0.7 | 0.833 | 57.23% | 11.50% |
|  | 0.653 | 0.805 | 80.53% | 19.47% |
|  | 0.671 | 0.667 | 66.67% | 0.333333 |
|  |  |  |  |  |
| 9 | 0.673 | 0.819 | 59.88% | 13.27% |
|  | 0.655 | 0.808 | 80.83% | 19.17% |
|  | 0.687 | 0.678 | 67.85% | 32.15% |
|  |  |  |  |  |
| 10 | 0.645 | 0.803 | 62.54% | 15.34% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.674 | 0.681 | 68.14% | 31.86% |
|  |  |  |  |  |
| 11 | 0.659 | 0.812 | 66.08% | 15.34% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.673 | 0.693 | 69.32% | 30.68% |
|  |  |  |  |  |
| 12 | 0.65 | 0.806 | 68.73% | 16.52% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.661 | 0.699 | 69.91% | 30.09% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
| 13 | 0.648 | 0.805 | 69.32% | 16.81% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.659 | 0.659 | 69.91% | 30.09% |
|  |  |  |  |  |
| 14 | 0.649 | 0.805 | 69.62% | 69.62% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.652 | 0.702 | **70.21%** | 29.79% |
|  |  |  |  |  |
| 15 | 0.648 | 0.805 | 71.98% | 17.40% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.652 | 0.723 | 72.27% | 27.73% |
|  |  |  |  |  |
| 16 | 0.648 | 0.805 | 73.16% | 17.70% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.652 | 0.735 | 73.45% | 26.55% |
|  |  |  |  |  |
| 17 | 0.652 | 0.808 | 74.34% | 17.70% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.662 | 0.752 | 75.22% | 24.78% |
|  |  |  |  |  |
| 18 | 0.655 | 0.81 | 75.22% | 17.70% |
|  | 0.653 | 0.808 | 80.83% | 19.174 |
|  | 0.662 | 0.752 | 75.22% | 24.78% |
|  |  |  |  |  |
| 19 | 0.657 | 0.811 | 75.81% | 17.70% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.657 | 0.761 | 76.11% | 23.89% |
|  |  |  |  |  |
| 20 | 0.657 | 0.811 | 75.81% | 17.70% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0.657 | 0.761 | 76.11% | 23.89% |
|  |  |  |  |  |
| 21 | 0.654 | 0.809 | 76.11% | 17.99% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.663 | 0.767 | 76.70% | 23.30% |
|  |  |  |  |  |
| 22 | 0.665 | 0.776 | 77.58% | 22.42% |
|  | 0.653 | 0.808 |  | 19.17% |
|  | 0.657 | 0.811 | 76.99% | 17.99% |
|  |  |  |  |  |
| 23 | 0.655 | 0.809 | 77.58% | 18.29% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.656 | 0.779 | 77.88% | 22.12% |
|  |  |  |  |  |
| 24 | 0.653 | 0.808 | 79.65% | 18.88% |
|  | 0.653 | 0.808 | 79.65% | 18.88% |
|  | 0.655 | 0.785 | 78.47% | 21.53% |
|  |  |  |  |  |
| 25 | 0.652 | 0.806 | 79.65% | 19.17% |
|  | 0.655 | 0.808 | 80.83% | 19.17% |
|  | 0.656 | 0.788 | 78.76% | 21.24% |
|  |  |  |  |  |
| 26 | 0.652 | 0.807 | 79.94% | 19.17% |
|  | 0.655 | 0.808 | 80.83% | 19.17% |
|  | 0.656 | 0.791 | 79.06% | 20.94% |
|  |  |  |  |  |
| 27 | 0.652 | 0.807 | 79.94% | 19.17% |
|  | 0.655 | 0.808 | 80.83% | 19.17% |
|  | 0.657 | 0.794 | 79.35% | 20.65% |
|  |  |  |  |  |
| 28 | 0.652 | 0.807 | 79.94% | 19.17% |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 0.655 | 0.808 | 80.83% | 19.17% |
|  | 0.657 | 0.794 | 79.35% | 20.65% |
|  |  |  |  |  |
| 29 | 0.652 | 0.807 | 79.94% | 19.17% |
|  | 0.655 | 0.808 | 80.83% | 19.17% |
|  | 0.655 | 0.794 | 79.35% | 20.65% |
|  |  |  |  |  |
| 30 | 0.653 | 0.807 | 80.24% | 19.17% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.654 | 0.802 | 80.24% | 19.76% |
|  |  |  |  |  |
| 31 | 0.653 | 0.807 | 80.24% | 19.17% |
|  | 0.655 | 0.808 | 80.83% | 19.17% |
|  | 0.654 | 0.802 | 80.24% | 19.76% |
|  |  |  |  |  |
| 32 | 0.656 | 0.81 | 80.53% | 18.88% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.655 | 0.805 | 80.53% | 19.47% |
|  |  |  |  |  |
| 33 | 0.656 | 0.81 | 80.53% | 18.88% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.655 | 0.805 | 80.53% | 19.47% |
|  |  |  |  |  |
| 34 | 0.656 | 0.81 | 80.53% | 18.88% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.655 | 0.805 | 80.53% | 19.47% |
|  |  |  |  |  |
| 35 | 0.656 | 0.81 | 80.53% | 18.88% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.655 | 0.805 | 80.53% | 19.47% |
|  |  |  |  |  |
| 36 | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  |  |  |  |  |
| 37 | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  |  |  |  |  |
| 38 | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  |  |  |  |  |
| 39 | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  |  |  |  |  |
| 40 | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.653 | 0.808 | 80.83% | 19.17% |
|  | 0.653 |  |  |  |

Table 8.1. Precision, Recall and Accuracy for change coupling

Previously as seen from Regression, Table1.1 and Table 1.2, n=1 and n=4 were found out to be optimum from all n coupled classes. The result thus is concurrent with the one found by precision recall and accuracy.

**CONCLUSION**

Change Coupling and Change Burst models as described in the paper by Marco D’Ambros, Michele Lanza, Romain Robbes, were applied for bug prediction for Eclipse, PDE and Equinox. A comparison was drawn between the two methodologies. Different regression and classification models were used to evaluate the data and validate predictions made by each methodology. The results were found to be in line with values as described in the paper. This can be further studied to evaluate new metrics for bug prediction. These can be the combination of ones described here or entirely new ones.

**REFERENCES**

[1] D’Ambros, Michele Lanza, Romain Robbes **“**An Extensive Comparison of Bug Prediction Approaches”

[2] D’Ambros, Michele Lanza, Romain Robbes “On the Relationship Between Change Coupling and Software Defects”

[3] Nachiappan Nagappan, Andreas Zeller, Thomas Zimmermann, Kim Herzig, Brendan Murphy “Change Bursts as Defect Predictors”