2017 IEEE International Conference on Software Quality, Reliability and Security (Companion Volume)

Influence of the Distance Calculation Error on the Performance of Adaptive Random Testing

Yuanchao Qi   Ziyuan Wang*
School of Computer Science and Technology
Nanjing University of Posts and Telecommunications
Nanjing, 210023, China
qiyuanchao.cn@gmail.com, wangziyuan@njupt.edu.cn

Yongming Yao
Tongda College
Nanjing University of Posts and Telecommunications
Yangzhou, 225127, China
yaoym@njupt.edu.cn

***Abstract*—Adaptive random testing, which is an enhanced random testing technique, has been studied for over 10 years. This paper aims to research the influence of the distance problems on the performance of adaptive random testing, where we focus on its efficiency of fault detecting. Experimental results suggest that, the fault detecting ability of adaptive random testing is influenced by the distance calculation error and the modeling mistake with the block failure pattern. It is interesting that, the former does the negative influence to the performance, and the latter does the positive one.**

Keywords*—*Adaptive random testing; F-measure; distance; error.

I. Introduction

Random Testing (RT) is a widely used software testing method in practice [1]. In this method, testers just only select test cases randomly to test the programs without any information of selected test cases. So that RT is very simple to perform and easy to be automated. Based on the RT, Adaptive Random Testing (ART) was developed to enhance the efficiency of fault detecting [2]. During the last decade, researchers have been studying ART to expand its application range [3][4][5] and to improve its performance [6][7].

The key issue of ART is to select test case with higher distance to all existing test cases in each step. How to calculate such a distance? It is necessary to answer this question before performing ART in practice. And there is a conjecture that, the different distance calculation results may lead to the different performance of ART. A series of simulation experiments have been conducted to identify the factors that influence the performance of ART [8]. However, none of those pay attention on the possible errors during calculating distance of test cases, these errors may be inevitable in practice.

This paper aims to research the influence of the distance problems on the performance of adaptive random testing, where we focus on its efficiency of fault detecting. Experimental results suggest that, the fault detecting ability of adaptive random testing is influenced by the distance calculation error and the modeling mistake with the block failure pattern. It is interesting that, the former does the negative influence to the performance, and the latter does the positive one.

The rest of this paper is organized as follows: Section II introduces the background about the details of ART and F-measure. Section III raises the two distance problems that could influence the performance of ART. Experiment and its results are described in section IV. Conclusions are given finally.

II. Preliminaries

In this study, we assume that the random selection of test cases is based on a uniform distribution.

Failure-causing area means the inputs in this domain will produce incorrect outputs. For an input domain D, d and m denote the size of input domain and the size of the failure area respectively. The failure rate θ is defined as $\frac{m}{d}$.

F-measure is the metrics we used to measure the effectiveness of ART. It means the expected number of test cases required to detect the first failure. For random selection of test cases with replacement, the F-measure is equal to $\frac{1}{\theta}$, or equivalently $\frac{d}{m}$. In another word, lower F-measure , more effective the testing strategy is. Because fewer test cases are required to reveal the first failure.

A. Adaptive Random Testing

Adaptive random testing was proposed by Chen et al [2] to enhance the effectiveness and efficiency of random testing. Here the Algorithm 1 describes the first proposed ART method, the *Fixed-Size Candidate Set Adaptive Random Testing* algorithm (FSCS-ART).

Algorithm 1 FSCS-ART

```
T = {}/* T is the set of previously executed test cases */
randomly generate an input t
test the program using t as a test case
while stopping criteria not reached do
  D = 0
  randomly generate next k candidates c_1, c_2, ..., c_k
  for each candidate c_i do
    calculate the minimum distance d_i from T
    if d_i > D then
      D = d_i
      t = c_i
    end if
    add t to T
    test the program using t as a test case
  end for
end while
```

Fig. 1 shows FSCS-ART in operation, on a program with a 2-dimensional input space and the candidate number $k = 3$. Firstly, there are three previous executed test case t_1, t_2, t_3. We want to select the 4th test case, so three candidates, c_1 to c_3 are randomly generated as shown. Then we must calculate d_i for each candidate. Fig. 1(a) describes this process for candidate c_1, the minimum distance $d_{1,1}$ should be picked among $d_{1,1}$ (means the distance between t_1 and c_1), $d_{1,2}$, and $d_{1,3}$. Fig. 1(b) shows the nearest t_i for each candidate. The lines in Fig. 1(b) indicate d_i for the respective candidates. We choose the largest d_i, which is d_3 in Fig. 1(c). Finally, we choose c_3 as test case t_4. Fig. 1(d) shows the four test cases we have chosen by now. We repeat the process until the stopping criterion is satisfied.

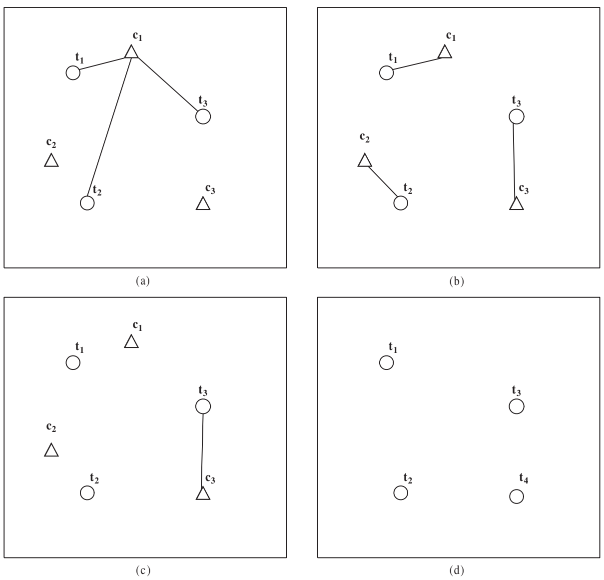

Fig. 1. FSCS-ART in operation

III. Distance Problem

Section II described the FSCS-ART algorithm and operating progress. The core concept of the ART is to select the farthest test case from selected test cases among the candidates. Distance is the key metrics to this selection. So wrong distance's value may influence the performance of ART. Here we consider two typical failure cases that cause distance problems in the ART algorithm.

A. Distance Calculation Error

In this paper, we use Euclidean Distance to describe the distance between two test cases in our examples and experiments. Euclidean Distance is easy to calculate and is possible to get wrong value as well.

Fig. 2 simply describes a distance calculation error. Test cases and candidates in Fig. 2 are the same as Fig. 1. In the common process in ART, candidate test case c_3 will be selected as the 4th test case at this round of selection and the failure area covers c_3, so c_3 will be the failure input to recover the mistake in this program.

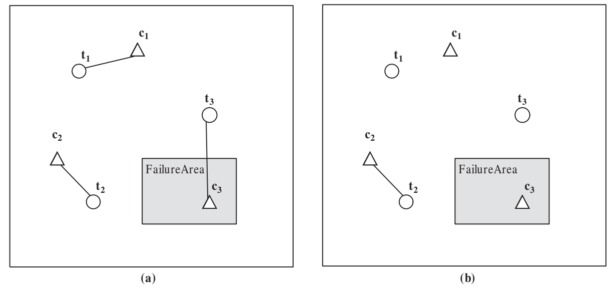

Fig. 2. Example of distance calculation error

However, what if there is a random error during the calculation of distance and the distance value $d_{2,2}$ is bigger than $d_{3,3}$. So ART will select c_2 to be the 4th test case and discard c_1 and c_3, the chance to reveal the failure could be missed.

B. Modeling Mistake

Modeling is an important step to solve problems. Modeling mistake is also a common problem. In this paper, we study a simple modeling mistake: mismatch.

Fig. 3 describes this simple modeling mistake. Fig. 3(a) is the right modeling situation. In this circumstance, candidate c_3 will be the next test case as usual and this test case will trigger the failure. The dash line in Fig. 3(a) is modeling mistake position that Fig. 3(b) is encountered. The vertical line in Fig. 3(b) is the original boundary that this input domain should be. Relative positions between many test cases are changed, in this situation, c_2 is the best candidate test case, so the chance to reveal the failure are missed once again.

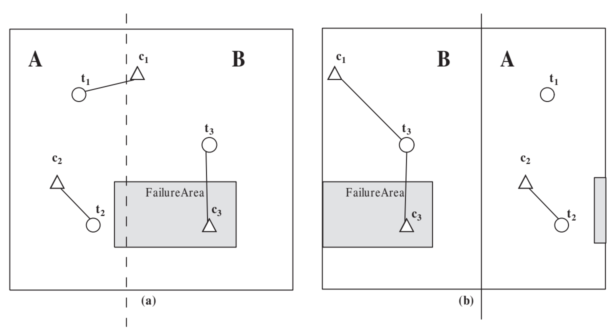

Fig. 3. Example of modeling mistake

These two distance problems may influence the performance of ART, so we decided to do some experiments to study this problem.

IV. Experiment

The goal of our study is to get the ART's F-measures in different failure rate and various level of distance problems. Random Testing is also proceeded in different failure rate as the reference.

Especially, we seek to answer the following research questions:

- *RQ1*: Does distance calculation error and modeling mistake really influence the performance of ART?
- *RQ2*: How does these problems affect the ART?
- *RQ3*: Does failure rate θ influence the experiment result?

A. Experiment Setup

We use simulation experiment to analyse the subject. As shown in Fig. 2, we analyse the performance of ART and RT in 2 dimension plane of square. Failure pattern is set as block pattern, in another word, failure area is a rectangle in the input domain.

Because of the randomness of ART, one test is not enough. As the position and size of Failure area also affect the ART performance, multiple experiments are required. So we take 100 repeated experiments to eliminate the randomness of ART and 100 random failure areas to eliminate the effects of different position of failure area. For one failure rate, 10000 times experiments are conducted. 15 different failure rates θ = $\{0.75, 0.5, 0.25, 0.1, 0.075, 0.05, 0.025, 0.01, 0.0075, 0.005, 0.0025, 0.001, 0.00075, 0.0005, 0.00025\}$ are considered.

1) Distance Calculation Problem

We use e = [0,1) to define the error of distance calculation. e = 0 means there is no error during the calculation. Take e = 0.2 for example, if the right distance value is d = 10, so the program will return a random number between $d(1-e) = 8$ and $d(1+e) = 12$ which represents an unforeseen distance error.

A special number e=1.0 doesn't mean the program will return a random number between two distance values. It means the distance calculation is completely out of control and the program would return a simple random number.

2) Modeling Mistake

We use t = [0,1] to represent a simple modeling mistake. In our 2 dimension square input domain, the mismatch occurs at only one dimension. t means the position where the modeling mistake takes place. So t = 0 and t = 1 are the same, they mean there are nothing changed in the modeling. And t = 0.5 is the max modeling mistake.

B. Experiment Result

Table I and Table II represent every average F-measure the FSCS-ART and RT got in different θ and situation.

In order to make the F-measure easy to see and compare in Fig. 4 and Fig. 5., F-ratio is used in these figures.

$$F\text{-ratio} = F\text{-measure} \times\theta$$

The RT's F-ratio is close to 1, the lower the F-ratio is, the better the algorithm will be.

1) Distance Calculation Problem

As shown in Fig. 4, RT is very close to 1. FSCS-ART rises in the first stage, then decreases and at last rises again. We can see, with the increasing of e, the performance of FSCS-ART is getting worse. When $e = 1.0$, FSCS-ART degenerates to RT. When $\theta = 0.025$, the distance calculation error influences the most.

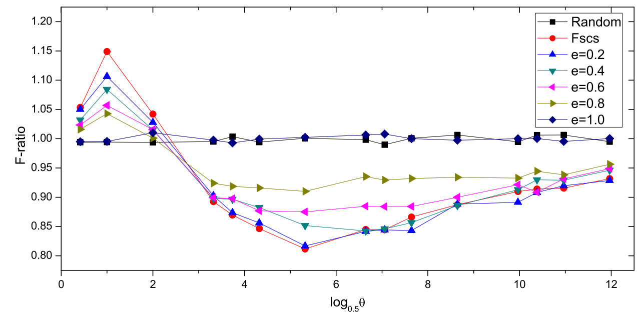

Fig. 4. F-ratio for different distance calculation errors for FSCS-ART in various failure rate θ

2) Modeling Mistake

As shown in Fig. 5, the modeling mistake does influence the performance of FSCS-ART, but to our surprise, the mismatch problem gives the positive influence on the fault-detecting ability of that. The line of $t = 0.1$ and $t = 0.9$ are similar, line of $t = 0.3$ and $t = 0.7$ are similar. $t = 0.5$ makes the performance of FSCS-ART best of all.

With the decrease of θ, the impact of modeling mistake is getting lesser and lesser. When $\theta = 0.00025$, six points in the figure are almost superposition. The influence of modeling mistake is negligible.

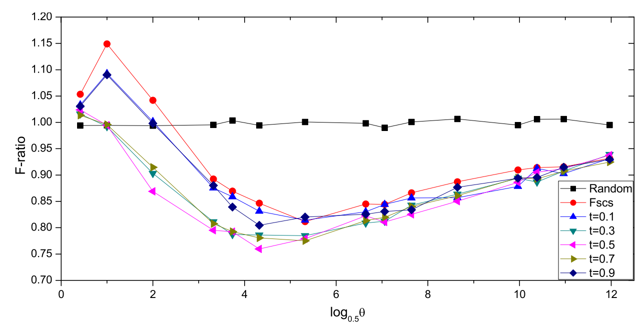

Fig. 5. F-ratio for different modeling mistakes for FSCS-ART in various failure rate θ

V. Conclusions

In this paper, we analyze the influence of two types of mistakes during the ART algorithm. One is the distance calculation error and another is the modeling mistake. The simulation experiments are doing on the square input domain with the block failure pattern.

The results show that two kind of mistakes have quite different influence on the performance of ART with diverse failure rate. Distance calculation error makes ART worse while modeling mistake makes things better. Distance calculation error has less impact on ART when the failure rate is very small and the influence of modeling mistake is almost negligible.

TABLE I. F-MEASURE OF DIFFERENT DISTANCE CALCULATION ERRORS FOR FSCS-ART IN VARIOUS FAILURE RATE θ

θ	Random	FSCS	e=0.2	e=0.4	e=0.6	e=0.8	e=1.0
0.75	1.3254	1.40467	1.4003	1.37587	1.3647	1.35485	1.32668
0.5	1.98873	2.29798	2.21272	2.1686	2.1137	2.08573	1.991
0.25	3.9755	4.16817	4.11158	4.06098	4.06248	4.00445	4.04075
0.1	9.95225	8.9242	9.0257	8.97818	8.99325	9.23958	9.97635
0.075	13.3801	11.5946	11.6471	11.9533	11.9702	12.248	13.2388
0.05	19.887	16.9309	17.1235	17.647	17.5364	18.3153	19.9869
0.025	40.0357	32.467	32.6701	34.0608	34.9964	36.4044	40.0915
0.01	99.8346	84.4957	84.1729	84.2352	88.4604	93.5297	100.645
0.0075	131.963	112.581	112.567	112.792	117.87	123.913	134.409
0.005	200.136	173.227	168.608	171.298	176.819	186.416	200.017
0.0025	402.581	354.912	355.271	354.267	359.988	373.606	398.961
0.001	994.754	909.738	891.475	912.896	921.55	932.603	1000.04
0.00075	1341.49	1218.87	1210.99	1239.73	1211.2	1259.31	1333.69
0.0005	2012.75	1831.43	1839.57	1858.17	1862.81	1876.64	1990.92
0.00025	3979.77	3727.55	3714.6	3784.4	3795.87	3825.56	4001.1

TABLE II. F-MEASURE OF DIFFERENT MODELING MISTAKES FOR FSCS-ART IN VARIOUS FAILURE RATE θ

θ	Random	FSCS	t=0.1	t=0.3	t=0.5	t=0.7	t=0.9
0.75	1.3254	1.40467	1.37765	1.3547	1.3649	1.3515	1.3742
0.5	1.98873	2.29798	2.18527	1.984	1.9889	1.9917	2.18095
0.25	3.9755	4.16817	4.00577	3.61307	3.4763	3.65865	3.99262
0.1	9.95225	8.9242	8.75372	8.11238	7.95188	8.07503	8.8024
0.075	13.3801	11.5946	11.4497	10.4962	10.5632	10.5671	11.1894
0.05	19.887	16.9309	16.6257	15.7248	15.195	15.6111	16.0901
0.025	40.0357	32.467	32.5739	31.3978	31.1636	31.0052	32.8094
0.01	99.8346	84.4957	83.0343	80.8793	82.2478	81.3632	82.5478
0.0075	131.963	112.581	112.564	108.302	108.11	109.302	110.763
0.005	200.136	173.227	171.279	168.541	165.055	168.048	166.828
0.0025	402.581	354.912	342.792	345.482	340.152	344.347	350.663
0.001	994.754	909.738	878.398	894.003	886.111	894.619	893.997
0.00075	1341.49	1218.87	1216.96	1182.14	1206.84	1190.16	1193.88
0.0005	2012.75	1831.43	1805.39	1815.24	1824.16	1814.72	1829.92
0.00025	3979.77	3727.55	3725.93	3756.15	3750.18	3698.31	3719.34

VI. ACKNOWLEDGMENTS

This work is supported by the Nature Science Foundation of Jiangsu Province (BK20130879).

REFERENCES

[1] Hamlet R. Random testing[J]. Encyclopedia of software Engineering, 1994.

[2] Chen T Y, Leung H, Mak I K. Adaptive random testing[M] //Advances in Computer Science-ASIAN 2004. Higher-Level Decision Making. Springer Berlin Heidelberg, 2004: 320-329.

[3] Chen T Y, Tse T H, Yu Y T. Proportional sampling strategy: a compendium and some insights[J]. Journal of Systems and Software, 2001, 58(1): 65-81.

[4] Chan K P, Chen T Y, Towey D. Restricted random testing[M] //Software QualityECSQ 2002. Springer Berlin Heidelberg, 2002: 321-330.

[5] Mayer J. Lattice-based adaptive random testing[C] //Proceedings of the 20th IEEE/ACM international Conference on Automated software engineering. ACM, 2005: 333-336.

[6] Mayer J. Adaptive random testing by bisection with restriction[C] //International Conference on Formal Engineering Methods. Springer Berlin Heidelberg, 2005: 251-263.

[7] Chen T Y, Kuo F C, Merkel R G, et al. Mirror adaptive random testing[J]. Information and Software Technology, 2004, 46(15): 1001-1010..

[8] Chen T Y, Kuo F C, Zhou Z. On the Relationships between the Distribution of Failure-Causing Inputs and Effectiveness of Adaptive Random Testing[C] //SEKE2005: 306-311.