

An Approach to Filter the Test Data for Killing Multiple Mutants in Different Locations

Nagendra Pratap Singh, Rishi Mishra, Sailesh Tiwari, and A. K. Misra

Abstract—Mutation testing is a fault-based testing technique that can be used for testing software at unit level, integration level and specification level. In addition to assessing the test data adequacy, mutation testing has also been used to support other testing activities such as test data generation, regression testing etc. Several works has been done on automatic generation of test data that can be effectively kill mutants. Constraint-based test data generation (CBT) is one of the automatic test data generation techniques using mutation testing, however, existing approaches of test case generation generally generate test data by killing one mutant at one time. Thus, more test cases are needed for achieving a given mutation score. In this paper, an approach is proposed by filtering the test data according to necessity condition and reachability condition by killing multiple mutants, mutated at the different location at one time and filtered test data also achieved same or approximate same mutation score. In proposed approach, some test data is filtered out of large test data that is sufficient to kill multiple mutants, located at different location. So this approach reduces the testing cost and time.

Index Terms—Constraint-based testing, Mutation testing, mutation operator.

I. INTRODUCTION

Software testing is a technique for software quality assurance. It is a time consuming process and accounts for about 50% of the cost of software development [1],[2]. An important problem of software testing is how to generate the effective test data. If the problem of automatic test data generation can be well solved, then the cost of software testing can be significantly reduced. Mutation testing is based on mutation analysis. Mutation testing is a fault-based testing technique which is originally introduced by Hamlet [3] and DeMillo et al. [4] for assessing the effectiveness of test suites in unit testing. Mutation testing is a method of software testing, which involves small syntactic change in programs source code [5]. These syntactic changed programs are called mutant programs which are created by replacing well-defined mutation operators. Mutation testing is a powerful testing technique for achieving correct or closes to correct program. Mutation testing is based on three fundamental assumptions,

One is known as competent programmer hypothesis (CPH) or the competent programmer assumption and second is known as coupling effect [4]. Third assumption is the

presence of an oracle for classifying the output of a test execution as correct or not.

The competent programmer hypothesis was introduced by DeMillo et al. [4]. It says that programmers create a correct version of program. But it may be possible that there may be faults in the program delivered by a competent programmer, we assume that these faults are merely small simple faults which can be corrected by small syntactical changes. Therefore, in Mutation Testing, faults constructed from several simple syntactical changes are applied, which represent the faults that are made by competent programmers. An example of the CPH can be found in Acree et al., s work [6] and a theoretical discussion can be found in Budd et al., s work [7].

The Coupling Effect was also introduced by DeMillo et al., [7]. While the CPH concern with a programmers behavior, the Coupling Effect is observed empirically. DeMillo et al., states that Test data that distinguishes all programs differing from a correct one by only simple errors is so sensitive that it also implicitly distinguishes more complex errors. Offutt [8], [9] extended the Mutation Coupling Effect Hypothesis with a precise definition of simple and complex faults. According to this definition, a simple fault is represented by a mutant that is created by making a single syntactical change, while a complex fault is represented as a mutant that is created by making more than one change. Offutt says the coupling effect hypothesis is complex faults are coupled to simple faults in such a way that a test data set that detects all simple faults in a program will detect a high percentage of the complex faults or says If a test suite kills a mutant, it also kills mutants of mutant [9]. Third assumption is presence of an oracle for classifying the output of a test execution as correct or not the implementation of oracle is itself a complex problem and beyond the scope of our work.

In mutation testing, for an original program (OP) a set of faulty programs M called as mutants (mutant program) are generated by replacing some syntax to the original program OP. In a Table-1 and 2 shows OP and its all mutant programs M_1, M_2, M_3, M_4, M_5 that are generated by replacing the relational operator ($>$) by the operators $<, <=, >=, ==, \neq$ respectively. A rule that generates a mutant from the original program is known as a mutation operator shown in Table land 2, we use ROR operator. There are many other operators that can be used. Typical mutation operators are designed to change the variables and expressions by replacement, insertion or deletion operators. These typical mutation operators were implemented in the Mothra mutation system [10].

Manuscript received September 10, 2012; revised November 21, 2012.

Nagendra Pratap Singh is with the Computer Science and Engineering Department at NVPEMI, Kanpur, India (e-mail: nagendrasngh447@gmail.com).

Rishi Mishra is with the I. T. Manager in United Bank of India, Lucknow, India (e-mail: rishi.msr@gmail.com).

TABLE I

OP	M1	M2
<pre>int big(a, c) { if(a>c) printf("big=a"); else printf("big=c"); }</pre>	<pre>int big(a, c) { if(a<c) printf("big=a"); else printf("big=c"); }</pre>	<pre>int big(a, c) { if(a<=c) printf("big=a");else printf("big=c"); }</pre>

TABLE II

M3	M4	M5
<pre>int big(a, c) { if(a>=c) printf("big=a"); else printf("big=c"); }</pre>	<pre>int big(a, c) { if(a=c) printf("big=a"); else printf("big=c"); }</pre>	<pre>int big(a, c) { if(a≠c) printf("big=a");else printf("big=c"); }</pre>

After generating the mutants, all generated test data is applied on original program (OP) and as well as on created mutants that distinguish mutants from the original program. Intuitively, a test data that kill more mutants prove its effectiveness on others. Specifically, the ratio of killed mutants to the nonequivalent mutants is used as the mutation adequacy score to indicate test suite quality in mutation testing.

Although, mutation testing is an effective means to find the effectiveness of test suites. Generating the test data that can achieve a satisfactory mutation adequacy score can be very labor-intensive [11]. For this problem, several approaches have been proposed in the literature. The Constraint-based testing (CBT) technique was developed by DeMillo and Offutt [12]. They used symbolic evaluation, control-flow analysis, and information of mutants for generating test data automatically. Test data generated by CBT can kill more than 90% of the mutants for most programs [13]. There was a problem in CBT approach to handling the nested expressions, arrays and loops. Another approach is proposed by Offutt and Jin for test data generation, known as Dynamic Domain Reduction (DDR) [14]. This approach applies some limitations in CBT approach. So we can say that DDR approach is better than CBT approach but both approaches generates test data by solving one mutants necessity condition and reach ability conditions [15], [16]. According to the paper [15], [16] the number of possible mutants is proportional to the product of the number of data references and the number of data objects. Which indicates one test data is generated according to one mutant, that means if there are N number of mutants then N number of test data are required to killing the mutants which become a big burden. So another novel approach is proposed by Ming-Hao Liu et al. [17] that generate test data for killing a multiple mutants that are placed in same location. This approach generates one test data according to multiple mutants that are mutated at the same location at one time. Thus this approach can generate smaller test suite that can achieve the same mutation testing score. The experimental results of this approach show that it is more cost-effective.

Our proposed approach is the extension of [17]. In this approach, generated test data is filtered by killing a multiple mutants that are placed in different location. This approach generates one test data according to multiple mutants that are mutated at the different location at one time. Thus, this

approach can generate smaller test suite that can achieve the same mutation score. Thus, the cost of testing is reduced.

II. BACKGROUND DETAILS

According to the paper [12], [13], [15], [17] the test data to kill a mutant must be satisfied three conditions known as reachability condition, necessity condition, sufficiency condition. Lets OP is original program, M is mutant of OP on a particular statement S and test data for OP is T then three conditions is defined as follows:-

A. Reachability Condition

The test data T must be reach to statement S because S is replaced by mutant. Mutant is a syntactic change to an executable statement, and the other statements in the mutated program are syntactically equal to the statements in the original program. If T cannot be reached to S then it is guaranteed that T will never kill the M.

1- int a, b, c;	Reachability Condition Mutants are:- (a>=c), (a<c), (a<=c), (a==c), & (a≠c)
2- if(a>b)	
{	
3- if(a>c)	
4- printf("a is largest");	
5- else	
6- printf("c is largest");	
}	
7- else	
{	
8- if(c>b)	
9- printf("c is largest");	
10- else	
11- printf("b is largest");	
}	

Fig. 1. For example a program show in figure-1, if we replace.

Statement (a>c) of line number 3 by mutant operator ROR as shown in table 1 and 2 then the reachability condition of this S is (a>b) which is shown in line number 2. That means if T not satisfy the reachability condition then mutant cannot be executed.

B. Necessity Condition

A generated test case that kills mutants needs to have this characteristic: it must differentiate the mutants behavior that is, why the mutant is represented by a single change to the original program, the execution state of the mutant program must differ from that of the original program after some execution of the mutated statement. This characteristic is known as the necessary condition.

Given a program OP and a mutant program M by changing a statement S in OP, for a test data T to kill M, it is necessary that the state of M immediately following some execution of S be different from the state of OP at the same location [12]. According to paper [17] necessity condition for killing the mutants (same location) of the program shown in the Fig. 1. is calculated as:-

Necessity condition for (a>c)

$$(a>c) \neq (a>=c)$$

$$(a>c) \neq (a<c)$$

$$(a > c) \neq (a <= c)$$

$$(a > c) \neq (a == c)$$

$$(a > c) \neq (a \neq c)$$

Then combined these five we get

$$\{(a > c) \neq (a >= c)\} \&\& \{(a > c) \neq (a < c)\} \\ \&\& \{(a > c) \neq (a <= c)\} \&\& \{(a > c) \neq (a == c)\} \\ \&\& \{(a > c) \neq (a \neq c)\}$$

which is equivalent to

$$(a < c) \&\& (a == c) \quad (1)$$

C. Sufficiency Condition

The necessity condition never guaranteed that to be sufficient to kill the mutant. For a test case to kill a mutant, it must create different output, in which case the final state of the mutant program differs from the original program. Although filtering the test cases that meet this sufficiency condition is certainly desirable, it is impractical in practice [13].

There are some approaches [14], [17] that utilize the reachability conditions and the necessity condition to generate test data that does not guarantee to kill all mutants, but can kill most of the mutants [13].

III. OUR APPROACH

The proposed approach in paper [17], Ming-Hao Liu et al. says one test data are required to kill multiple mutants that are mutated in same location at one time. Thus this approach generate small test suite that achieve same test adequacy score (mutation score). But in our approach generate small test suite that achieve same test adequacy score and each test data of a test suite is able to killed multiple mutants that are mutated in different location at one time.

In Fig. 1. the all mutants of statement $(a > c)$ is known as same location mutants because they have same reachable condition.

The all mutants of the statement $(a > c)$ and $(c > b)$ is known as different location mutants because they have inside the same conditional statement which is responsible to find out the reachable condition of both statements. Although if two mutated statements are inside different conditional statement that is also known as different location mutants.

There are some assumptions in proposed approach, first the mutants that are located in different location having same reachability conditions and the necessity conditions are similar in structure. Second the combined necessity conditions of different locations mutant in one condition that is also combined with shared reachability condition which is known as Final Filtering Condition (FFC). Third assumption is the conditional statements that lead test case to reach the inner block of code containing mutants are used to provide the reachability condition. The main aim of proposed approach is to replace the mutants on different location by using one mutation operator known as Relation Operator Replacement (ROR). ROR mutation operator can

produce more than one mutant on the particular location (see Fig.1). Our approach follows the following four steps:-

- 1) Find the reachability condition of mutants that are located in different locations.
- 2) Find the necessity condition of each different location mutants.
- 3) Combining necessity condition of each different location mutants with shared reachability condition (if exist) by using conjunction (&&) operator and generate Final Filtering Condition (FFC) for each different locations.
- 4) Generate reduced test data by using FFCs.

Table I and Table II shows original program OP and its all mutant programs $M1, M2, M3, M4, M5$ that is generated by replacing the relational operator ($>$) by the operators $<, <=, >=, ==, \neq$ respectively. Program in Fig. 1 containing three relational operators $(a > b)$ $(a > c)$ and $(c > b)$, each one of these three relational operators is replaced by any one of other relational operators like $<, <=, >=, ==, \neq$, so there are total fifteen mutants are generated.

Proposed approach generates FFC condition that filter the already generated test data. The generated test data before filtering is able to kill the multiple mutants in same location with some adequacy score but the filtered test data is also able to kill multiple mutants in same location as well as different location with the same adequacy score as compared to that before filtering. The step-by-step implementation of proposed approach (by using the program shows in Fig. 1) is as follows:-

A. Step-1

Out of three relational expressions in program (figure-1) the expression $(a > b)$ is a reachability condition for other two relational expressions $(a > c)$ and $(c > b)$ because both expression $(a > c)$ and $(c > b)$ depends on the expression $(a > b)$. That means if condition $(a > b)$ is true then expression $(a > c)$, otherwise expression $(c > b)$ is executed.

B. Step-2

The Necessity condition for location-1(line number-3) is shows in equation (1) similarly the Necessity condition for location-2 (line number-8) is

$$(c < b) \&\& (c == b) \quad (2)$$

C. Step-3

In this step the necessity condition of location-1 which shows in equation (1) is combined with their reachability condition $(a > b)$ by using conjunction operator.

$$(a > b) \&\& [(a < c) \&\& (a == c)] \\ [(a > b) \&\& (a < c)] \&\& [(a > b) \&\& (a == c)] \\ [(b < a < c)] \&\& [(a == c) > b] \quad (3)$$

Similarly the necessity condition of second location which shows in equation (2) is combined with their reachability condition $(a <= b)$ by using conjunction operator.

$$(a <= b) \&\& (c < b) \&\& (c == b) \\ [(a <= b) \&\& (c < b)] \&\& [(a <= b) \&\& (a == c)]$$

$$\begin{aligned}
 & [((a,c)<b) \parallel ((a==b)>c)] \ \&\& \ [((a<(b==c)) \parallel (a==b==c))] \\
 & [((a; c) < b) \parallel ((a == b) > c)] \\
 & \&\& [((a < (b == c)) \parallel (a == b == c))] \quad (4)
 \end{aligned}$$

According to figure-1 line no. 2 (reachability condition of line no.3 and line no.8) is also replaced by five mutants using ROR mutation operator. In line no. 2 ($a>b$) is not depends on any branch predicate so there are no any reachability condition to kill all five mutants. Hence the necessity condition for ($a>b$) is show in equation 5.

$$(a < b) \ \&\& \ (a == b) \quad (5)$$

D. Step-4

Thus the combination of FFCs (Eqn-3, 4, 5) is used to Generate a filtered test data that is used to kill mutants.

IV. EXPERIMENTAL RESULT

Let us take an example of program shown in Fig. 1 that can find the largest integer from the given three integers as inputs. The following test case are generated to test the Program shown in Fig. 1.

T1:-All three integers are different.

T2:-Any two integers are same.

T3:-All integers are same.

Assume that the value of a,b,c are 5,4,6 respectively then all possible test data in each test case is shown in TABLE-III.

TABLE III

Test Case	Test Data
T1	$t_{11}=(5,4,6), t_{12}=(5,6,4), t_{13}=(6,5,4),$ $t_{14}=(6,4,5), t_{15}=(4,5,6), t_{16}=(4,6,5)$
T2	$t_{21}=(5,5,4), t_{22}=(5,4,5), t_{23}=(4,5,5),$ $t_{24}=(5,5,6), t_{25}=(5,6,5), t_{26}=(6,5,5),$ $t_{27}=(4,4,6), t_{28}=(4,6,4), t_{29}=(6,4,4)$ $t_{210}=(4,4,5), t_{211}=(4,5,4), t_{212}=(5,4,4)$ $t_{213}=(6,6,4), t_{214}=(6,4,6), t_{215}=(4,6,6)$ $t_{216}=(6,6,5), t_{217}=(6,5,6), t_{218}=(5,6,6)$
T3	$t_{31}=(5,5,5), t_{32}=(6,6,6), t_{33}=(4,4,4)$

Let us select 15 test data out of 27 test data randomly which is $t_{11}, t_{12}, t_{13}, t_{14}, t_{15}, t_{16}, t_{24}, t_{25}, t_{26}, t_{31}, t_{213}, t_{214}, t_{27}, t_{28}, t_{29}$ and using these test data find the out put of the actual program and their mutant programs which is shown in TABLE IV.

Again by using our approach, filter the previesly selected 15 Test data by using Eqn-3,4,5. Test data t_{11} and t_{214} satisfy the Eqn-3. Combination of two test data, one test data from $t_{12}, t_{16}, t_{25}, t_{28}, t_{213}$ and second test data t_{31} satisfied the Eqn-4. Our approach select any one combination of two test data out of six possible combination because all six possible combination is able to kill the same number of mutant, TABLE-V shows the results of two combinations (t_{12}, t_{31}) and (t_{28}, t_{31}) out of six possible combinations.

Test data that satisfy the Eqn-5, is also satisfy the Eqn-4. So our approach select only 4 test data $t_{11}, t_{31}, t_{214}, t_{28}$, and out of 15 test data after filtering. The output of the actual program and their mutant program is shown in TABLE VI.

TABLE IV

T.D.	OP	M1	M2	M3	M4	M5	M6	M7
t_{11}	c	c	a	a	c	a	c	c
t_{12}	b	b	b	b	b	b	b	c
t_{13}								
t_{14}	a	a	c	c	c	a	a	a
t_{15}								
t_{16}	a	a	c	c	c	a	a	a
t_{24}								
t_{25}	c	c	c	c	c	c	c	b
t_{26}								
t_{31}	b	b	b	b	b	b	b	c
t_{213}	c	c	c	c	c	c	c	b
t_{421}								
t_{27}	b	b	b	b	b	b	b	c
t_{28}	a	a	c	c	c	a	a	a
t_{29}								
	b	b	b	b	b	b	c	b
	b	b	b	b	b	b	b	c
	c	a	c	a	a	c	c	c
	c	c	c	c	c	c	c	b
	b	b	b	b	b	b	b	c
	a	a	c	c	c	a	a	a
M8	M9	M10	M11	M12	M13	M14	M15	
c	c	c	c	c	c	c	c	c
c	b	a	b	a	a	b	a	a
a	a	a	a	b	b	b	a	a
a	a	c	a	c	c	c	c	a
b	b	c	c	c	c	c	c	c
c	b	c	b	c	c	b	c	c
b	b	c	c	c	c	c	c	c
c	b	a	b	b	b	b	a	a
c	c	c	c	b	c	c	b	b
c	b	c	a	b	a	a	a	b
c	c	c	c	c	c	c	c	c
b	b	c	c	c	c	c	c	c
c	b	a	b	c	c	b	c	c
a	a	a	a	b	b	b	a	a

TABLE V

T.D.	OP	M1	M2	M3	M4	M5	M6	M7
t_{12}	b	b	b	b	b	b	b	c
t_{31}	b	b	b	b	b	b	b	b
M8	M9	M10	M11	M12	M13	M14	M15	
c	b	c	b	a	a	a	a	a
c	c	b	c	b	c	c	c	b
T.D.	OP	M1	M2	M3	M4	M5	M6	M7
t_{28}	b	b	b	b	b	b	b	c
t_{31}	b	b	b	b	b	b	c	b
M8	M9	M10	M11	M12	M13	M14	M15	
c	b	c	b	c	c	b	c	c
c	c	b	c	b	c	c	c	b

TABLE VI

T.D.	OP	M1	M2	M3	M4	M5	M6	M7
t_{11}	c	c	a	a	c	a	c	c
t_{31}	b	b	b	b	b	b	c	b
t_{214}	c	a	c	a	a	c	c	c
t_{28}	b	b	b	b	b	b	b	c
M8	M9	M10	M11	M12	M13	M14	M15	
c	c	c	c	c	c	c	c	c
c	c	b	c	b	c	c	c	b
c	c	c	c	c	c	c	c	c
c	b	c	b	c	c	b	c	c

In traditional mutation testing approach one test data are require to kill one mutant. So in this example there are 15 test data are required to kill all 15 mutants but by in our approach only four test data are able to kill all 15 mutants, so the testing cost and time is reduced.

V. CONCLUSION AND FUTURE WORK

This paper presents a new approach for filtering the generated test data and filtered test data can be used for killing multiple mutants at different locations effectively. The filtering of test data is based on the combination of necessity condition with reachability condition (if exist). The conjunction operator (&&) is used in combination of these condition which generates Final Filtering Condition (FFC) for each different location. Experimental results show that proposed approach reduces the test data by filtering them.

Both necessity and reachability conditions are satisfied by combination of Final Filtering Condition (FFC). The generated test data before filtering is able to kill the multiple mutants in same location with some adequacy score but the filtered test data is also capable to kill multiple mutants in same location as well as different location with the same adequacy score as compared to that before filtering. Proposed approach reduces the test data effectively by filtering them so that less test data is used for killing the multiple mutants. Therefore, less execution time is required to kill the mutants which reduce the cost of mutation testing activity and prove its efficacy over other approaches given in literature. In future, proposed approach can be implemented as tool to filter the generated test data or it can also be integrated with the existing mutation testing tools such as Mu, JUNIT etc.

REFERENCES

[1] G. Myers, *The Art of Software Testing*, John Wiley and Sons, New York NY, 1979.
 [2] I. Sommerville, *Software Engineering*, Addison-Wesley Publishing Company Inc., 4th edition, 1992.
 [3] R. G. Hamlet, "Testing programs with the aid of a compiler," *IEEE Transactions on Software Engineering*, vol. 3, no. 4, pp. 279-290, July 1977.
 [4] R. A. DeMillo, R. J. Lipton, and F. G. Sayward, "Hints on test data selection: Help for the practicing programmer," *IEEE Computer*, vol. 11, no. 4, pp. 34-41, April 1978.
 [5] A. Jefferson Offutt, *A Practical System for Mutation Testing: Help for the Common Programmer*

[6] A. T. Acree, T. A. Budd, R. A. DeMillo, R. J. Lipton, and F. G. S. Ward, "Mutation analysis, Georgia institute of technology, Atlanta, Georgia," *Technique Report GIT-ICS*, vol. 79, no. 8, 1979.
 [7] T. A. Budd, R. A. DeMillo, R. J. Lipton, and F. G. Sayward, "Theoretical and empirical studies on using program mutation to test the functional correctness of programs," in *Proceedings of the 7th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages(POPL80)*, Las Vegas, Nevada, pp. 220-233, 28-30 January 1980.
 [8] A. J. Offutt, "The coupling effect: Fact or Fiction," *ACM SIGSOFT Software Engineering Notes*, vol. 14, no. 8, pp. 131140, December 1989.
 [9] A. J. Offutt, "Investigations of the software testing coupling effect," *ACM Transactions on Software Engineering and Methodology*, vol. 1, no. 1, pp. 520, January 1992.
 [10] K. N. King and A. J. Offutt, "A fortran language system for mutation-based software testing," *Software: Practice and Experience*, vol. 21, no. 7, pp. 685-718, October 1991.
 [11] A. J. Offutt and R. H. Untch, "Mutation 2000: Uninting the Orthogonal," *Mutation 2000: Mutation Testing in the Twentieth and the Twenty First Centuries*, San Jose, CA, pp. 45-55, October 2000.
 [12] R. A. DeMillo and A. J. Offutt, "Constraint-based automatic test data generation," *IEEE Transactions on Software Engineering*, vol. 17, pp. 900-910, September 1991.
 [13] R. A. DeMillo and A. J. Offutt, "Experimental results from an automatic test case generator," *ACM Transactions on Software Engineering Methodology*, vol. 2, pp. 109-127, April 1993.
 [14] A. J. Offutt, Z. Jin, and J. Pan, "The dynamic domain reduction approach to test data generation," *Software Practice and Experience*, vol. 29, pp. 167-193, January 1999.
 [15] A. J. Offutt and J. Pan, "Detecting equivalent mutants and the feasible path problem," in *Proceedings of the 1996 Annual Conference on Computer Assurance (COMPASS 96)*, (Gaithersburg MD), IEEE Computer Society Press, pp. 224-236, June 1996.
 [16] W. E. Wong and A. P. Mathur, "Reducing the cost of mutation testing: an empirical study," *Journal of Systems and Software*, vol. 31, no. 3, pp. 185-196, December 1995
 [17] M. H. Liu, Y. F. Gao, J. H. Shan, J. H. Liu, L. Zhang, and J. S. Sun, "An approach to test data generation for killing multiple mutants," *IEEE International Conference on Software Maintenance (ICSM'06)*, 0-7695-2354-4/06, 2006



Nagendra Pratap Singh is Assistant Professor and Head of Computer Science and Engineering Department at Naraina Vidya Peeth Engineering and Management Institute, Kanpur, India. He is widely known about a work on Mutation Analysis and Testing Technique. He has more than 10 years of Experience (Teaching and Research) and awarded M.Tech. Degree from Moti Lal Nehru National Institute of Technology, Allahabad, India. His M.Tech. thesis based on Mutation Testing under the supervision of Prof. A. K. Misra from Moti Lal Nehru National Institute of Technology, Allahabad, India. He is enrolled in Advance Ph.D. Program from Shri G. S. Institute of Technology and Science, Indore, India.