International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-1 Issue-6, November 2012

Adaptive Random Testing Through Dynamic Partitioning by Localization with Distance and Enlarged Input Domain

Korosh Koochekian Sabor, Mehran Mohsenzadeh

Abstract*: *Based on the intuition that evenly distributed test cases have more chance for revealing non-point pattern failure regions, various Adaptive Random Testing (ART) methods have been proposed. A large portion of this methods such as ART with random partitioning by localization have edge preference problem. This problem would decrease the performance of these methods. In this article the enlarged input domain approach is used for decreasing edge preference problem in ART with random partitioning by localization. Simulations have shown that failure detection capability of ART by localization with distance and enlarged input domain is comparable and usually better than that of other adaptive random testing approaches.

Index Terms*: *Random testing, adaptive random testing, localization, enlarged input domain.

I. INTRODUCTION

It has been widely recognized that exhaustive testing (testing a program with all possible inputs) is not feasible [1]. With the aim of improving failure detection capability several different methods for selecting test cases have been proposed [2], [3], [4], [5], [6]. Random testing is a standard approach for automatic selection of test cases [7], [8]. In random testing test cases are selected randomly until a stopping condition such as detecting a failure, executing predefined number of test cases or ending of a time limitation is met. On the one hand since random testing is simple to perform and can be used to calculate reliability estimates, it has substantial practical advantages [9] on the other hand random testing may be ineffective on the ground that it does not uses of any information about the program under test [10].

As described by Chan et al in [11] typical failure patterns in programs are block pattern, strip pattern and point pattern. The block pattern (Fig. 1) in which failures are clustered in one or more continuous areas is the most common case. The strip pattern (Fig. 2) involves a contiguous area elongated in one dimension while quite narrow in the other. And lastly the point pattern (Fig. 3) which involves failures that are distributed in small groups in input domain.

Revised Manuscript Received on 30 November 2012
*****Correspondence Author(s)
Mr. Korosh koochekian sabor*, Department of Computer Engineering, Science and Research Branch Islamic Azad University, Tehran, Iran.
Dr. Mehran Mohsenzadeh, Department of Computer Engineering, Science and Research Branch Islamic Azad University, Tehran, Iran.

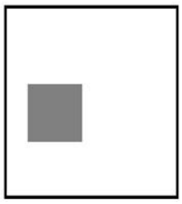


Fig. 1. block pattern

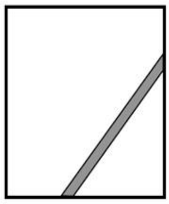


Fig. 2 strip pattern

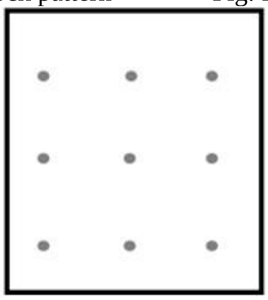


Fig. 3 point pattern

A simple but effective improvement in random testing is presented in [12], [13]. On the ground that failures in programs are shaped according to the represented failure patterns, the new approach uses of the available information about the location of previously executed test cases and tries to improve the performance by decreasing number of executed test cases required to reveal the first failure. It is obvious that when the failure pattern is of block or strip type the chance of revealing a failure would increase by distributing test cases more evenly. On the ground of this simple idea numerous ART methods have been proposed.

Some of ART methods such as ART through dynamic partitioning by localization have edge preference problem. Edge preference problem happens when test cases on the boundaries of input domain have more chance of being selected compared with test cases in the center part of the input domain. This problem would be intensified significantly as the input domain dimension increases. As the result the test cases would not be evenly distributed and the performance of the ART approach would decrease.

ART through dynamic partitioning localized by distance integrates the concept of localization into ART through dynamic partitioning method. Although the localization concept enhances evenly distribution of test cases in ART through dynamic partitioning, it intensifies the edge preference problem.

In this paper we investigate the effect of enlarge input domain for eliminating edge preference in ART through dynamic partitioning localized by distance.

*Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)*

Distance-based ART (D-ART) randomly selects a set of candidate test cases from input domain, But would only selects one test case for execution from the candidate set based on the criterion of maximizing the minimum distances between the candidate test cases and all of the previously executed test cases [13].

Two methods named ART by random partitioning and ART by bisection with no need to distance calculation have been proposed in [14]. The fundamental difference between these two methods and other ART methods is that rather than specifying the next test case which should be executed these two methods specify the region where the next test cases should be selected of. The main difference between these two methods is the partitioning scheme used. In random partitioning the input domain is partitioned according to the last previously executed test case and the biggest partition is selected as the next test case generation region. In ART by bisection the input domain is iteratively divided into equal sized partitions and a partition would be selected as the next test case generation region only if there are no executed test cases within that region.

In this article Elements of the input domain are known as failure-causing input, if they produce incorrect outputs. The input domain is denoted by D. The failure rate is calculated as division of number of failure causing inputs by total number of inputs in the input domain. F-measure which is defined as number of test cases required for detecting the first failure is used as effectiveness metric. F-Measure is more prominent for testers and is closer to the reality since usually when a failure is detected, testing would normally stop and debugging would start. In addition F-measure reflects distribution of test cases in a more consistent way.

This article is organized as follows. In section 2 ART Through Dynamic Partitioning by Localization with Distance is discussed. In section 3 the algorithm of the proposed method named ART through dynamic partitioning localized by distance and enlarge input domain is presented. In Section 4, we present our simulation results. Finally, conclusion is presented in Section 5.

II. ART THROUGH DYNAMIC PARTITIONING BY LOCALIZATION WITH DISTANCE

ART through dynamic partitioning choose the next test case based on the idea that if the test case is selected from the biggest region then it would have more chance of being far from previously executed test cases. But as it is shown in Fig. 4 this approach would not always lead to the expected result. To overcome this problem in [15] ART through dynamic partitioning localized by distance have been proposed.

ART through dynamic partitioning localized by distance would be executed in two steps, in the first step the test case Generation region is selected and some test cases are selected randomly from this test case generation region as the candidate set, In the second step the distance calculation with the goal of maximizing the minimum distance between the candidate test cases and nearby previously executed test cases is done. It is noteworthy that executed test cases are divided into two categories: nearby executed test cases and far executed test cases. In this method distance calculation would only be done with the nearby executed test cases. Nearby executed test cases are those which are located in the vertices of the test case generation regions, while the other executed test cases are classified as far executed test cases.

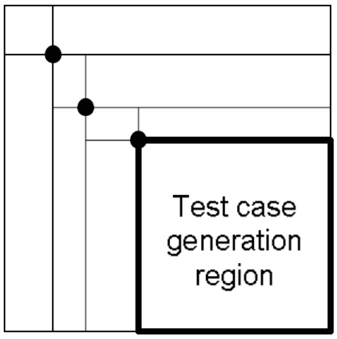


Fig. 4 ART through dynamic partitioning [15].

For example in Fig. 5 there exists one nearby executed test case which has been shown by a black filled circle and 4 candidate test cases shown by black filled triangle. In this case test case number 4 has the farthest distance to the nearby executed test case so it would be selected as the next test case for execution.

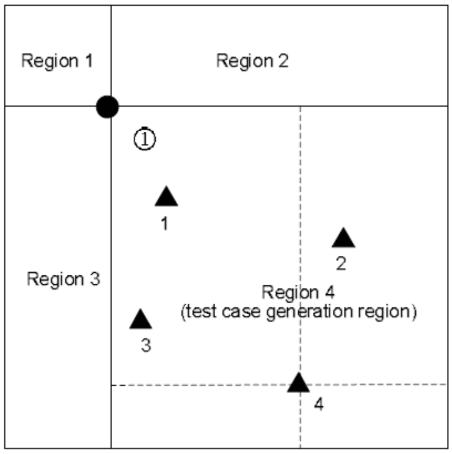


Fig. 5 ART through dynamic partitioning localized by distance [15].

This method resolves the problem in Fig. 4 but causes edge preference to occur. As the result the performance will decrease.

The main reason for edge preference is that less nearby previously executed test cases near the boundaries exist, So candidate test cases near the boundaries would have more chance of being selected.

For eliminating edge preference an approach named enlarged input domain has been proposed in [16]. In this method the input domain would be enlarged by factor f in each dimension and test cases would be selected from the enlarged input domain but they would be executed only if they are within the original input domain. Fig. 6 shows the original input domain D and the enlarged input domain D'.

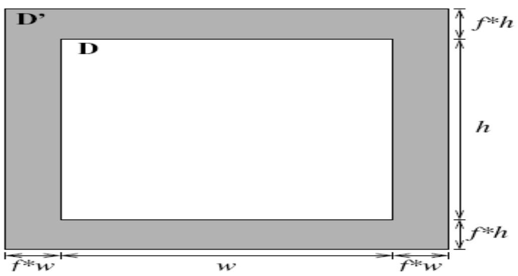


Fig. 6. The original input domain enlarged by factor F in each dimension [16].

III. PROPOSED METHOD

In this paper we will use enlarge input domain method to eliminate edge preference problem in ART through dynamic partitioning localized by distance. Test cases would be selected from the enlarged input domain and would be used for partitioning input domain, but they will be executed only if they are within the original input domain.

In this Paper domain is used to indicate the whole input space and region is used to indicate a part of the whole input domain.

Adaptive random testing through dynamic partitioning localized by distance and enlarge input domain algorithm

1. Enlarge the original input domain by factor f and assume vertexes as i'_{min} , j'_{min} , i'_{max} , Denotes j' Denotes value in X axis and i' (j'_{max} value in Y axis in cartesian coordinates)
2. Initiate the subdomain linked list L with the entire input domain $\{(i'_{min}, j'_{min}, F), (i'_{min}, j'_{max}, F), (i'_{max}, j'_{max}, F), (i'_{max}, j'_{min}, F)\}$ as the only element in it.
3. Remove the last element in the link list L as the test case generation region and set executed test cases set E to be empty.
4. Check the "Flag" of each vertex of the test case generation region. If "Flag" is "T", add it to the nearby executed test case set "E". Denote the number of elements in E as *l*.
5. Select K test cases randomly as the candidate test cases set c={ $c_1, c_2, \dots, c_k$ } and calculate the Euclidian distance between each element in candidate set and nearby executed test cases. The distance between a candidate test case and a nearby executed test case is denoted by dist(c_j, e_i). Choose a candidate test case c_q as the next test case according to the following criterion.

$$\forall j \in \{1,2,\dots,k\} \left(\min_{i=1}^{l} dist(c_q, e_i) \geq \min_{i=1}^{l} dist(c_j, E_i) \right)$$

(1)

6. If c_q is within the original input domain then Execute it. If the execution caused a failure then report the failure and finish the algorithm. Else partition the input domain into four regions by the executed test case and add them to the list L. Sort elements in L According to the region size in ascending order. Go to step3. If c_q is not inside the original input domain then do not execute the test case and only partition the input domain by the selected test case into four region. For each region if it is wholly located outside of the original input domain then decline it, else add it to the list L. Sort elements in L According to the region size in ascending order and go to step3.

The operation of this algorithm is illustrated in Fig. 7 and 8. In these figures the original input domain is shown by the inner dashed line and the enlarged input domain is shown by the outer dashed line. The test cases are selected from the enlarged input domain. In the first step a test case is randomly selected and executed then enlarged input domain is partitioned by the executed test case into four regions. These four regions are shown in Fig. 7.

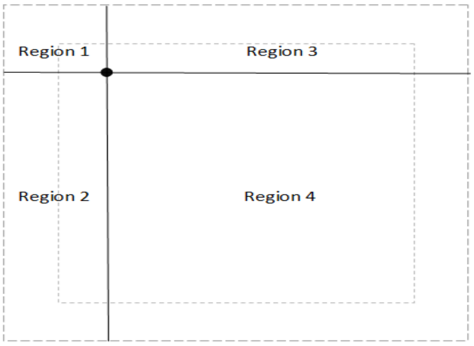


Fig. 7. Four region created by the first non-failure causing executed test case in ART through dynamic partitioning by localization and enlarged input domain

In the second pass as it is shown in Fig. 8 4 candidate test cases are selected from the biggest region. Since candidate test case 4 has the maximum distance to all of the previously nearby executed test cases, it has been selected as the next test case.

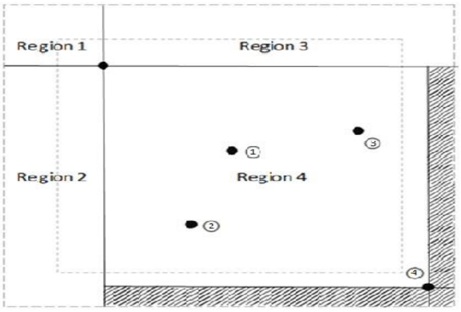


Fig. 8. Four regions created by the second test case which is not in the input domain

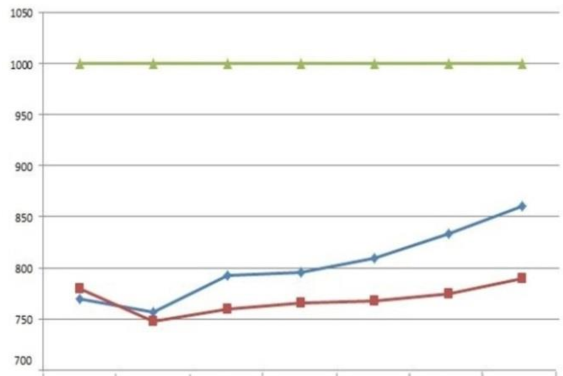
For the reason that the selected test case is outside of the original input domain it is not executed, but has been used for partitioning the enlarged input domain into four regions. Since 3 of these partitions are wholly in the enlarged input domain they are discarded and only one partition which has been left is added to the partitions list.

It is obvious that there are two nearby executed test cases for the third pass in this algorithm, as the result edge preference problem would be eliminated.

IV. SIMULATIONS

In this article a series of simulations in 2-dimesional and 3-dimensional input space domains with the aim of measuring failure detection capability of the proposed approach has been conducted. The execution for each section was run on a lightly loaded Intel Core i5 2300 cpu. In each test run a failure-causing region of the specified size and pattern has been randomly assigned within the input domain. For block pattern a square has been used as failure-causing region. For strip pattern two points on the adjacent borders of the input domain has been randomly selected and have been connected to each other to form a strip with predefined size. For the point pattern 10 circular regions were randomly located in the input domain without overlapping each other. For each simulation, 5000 test runs were executed with failure rate 0.001 and 1,5,10,15,20,25 and 30 as the candidate test cases set sizes and the average of 5000 run with failure rate 0.001 was calculated for each candidate test cases set sizes.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)



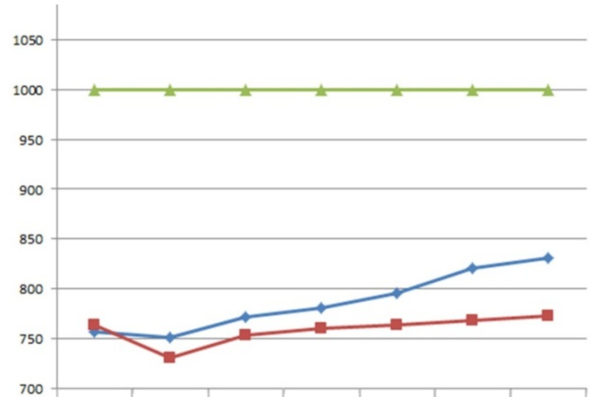
	1	5	10	15	20	25	30
RT	1000	1000	1000	1000	1000	1000	1000
Art By Localization	770	757	793	796	710	832	860
Proposed Method	780	748	760	766	768	775	790

In the first part of simulations the performance of the proposed method has been compared with that of ART through dynamic partitioning localized by distance in a 2-dimensional input domain. Line with triangle shows random testing performance, line with diamond shows ART by localization performance and line with square represents proposed method performance. When the candidate test cases set size is equal to 1 the proposed method and ART through dynamic partitioning localized by distance behave the same as ART through dynamic partitioning. As it is obvious in Fig. 9 when number of candidate test cases set size is equal to 1 the performance of the proposed method is less than that of ART through dynamic partitioning localized by distance. The reason for this lower performance is that although when candidate test cases set size is equal to one no edge preference problem exist, in the proposed method the input domain would be enlarged and this enlargement would lead to larger partitions, as the result the number of test cases required to detect the first failure would increase.

Fig. 9. F-Measure of RT and ART through dynamic partitioning by localization and the proposed method in 2-D input domain with block failure pattern with diverse candidate test case set sizes

As it is obvious in Fig. 9 the maximum performance in both methods is when the candidate test cases set size is equal to 5. By increasing candidate test case set size the performance of both methods decreases but the speed of performance loss in ART through dynamic partitioning localized by distance is faster than that of newly proposed method. The reason for faster performance loss in ART by localization is that increment in the candidate test cases set size intensifies edge preference but in the proposed approach the edge preference effect is much lower than that of ART through dynamic partitioning localized by distance so the speed of performance loss is much lower. The newly proposed method outperforms random testing by 27% in the best case when the size of candidate test cases set is equal to 5.

Since when number of the dimension increases to three or four edge preference would increase, in the second part of simulations we have investigated the performance of the proposed method in a 3-dimensional input space domain.



	1	5	10	15	20	25	30
RT	1000	1000	1000	1000	1000	1000	1000
Art By Localization	757	751	772	781	795	821	832
Proposed Method	763	730	753	760	764	767	773

Fig. 10. F-Measure of RT and ART through dynamic partitioning by localization and the proposed method in 3-D input domain with block failure pattern with diverse candidate test case set sizes

Fig. 10 shows the result of simulation in 3-dimensional input domain. Line with triangle shows random testing performance, line with diamond shows ART by localization performance and line with square represent proposed method performance. By increment in the candidate test case set size the edge preference causes a significant performance loss in ART through dynamic partitioning localized by distance but as it can be seen the performance loss in the proposed method is much lower. The reason for lower performance loss is the same as the first part of simulations.

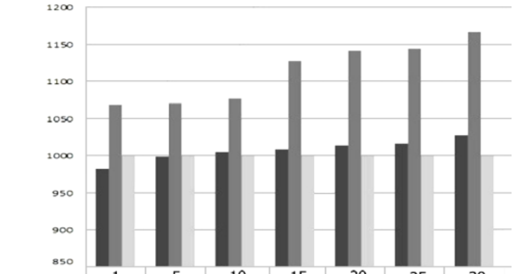
In the third part of simulations the performance of the proposed method has been compared with that of ART through dynamic partitioning localized by distance and pure random testing when failure pattern is of strip or point type.

As it is obvious in Fig. 11 when Failure pattern is strip the performance of both ART through dynamic partitioning localized by distance and proposed method decreases. The reason is that in the case of strip failure pattern distributing test cases more evenly has less performance in improving failure detection capability. But as it is obvious even for strip pattern the performance of ART is better than that of random testing. As the size of candidate test cases set size increases the edge preference effect would increase and when candidate test cases set size is 30, edge preference causes both methods to have Less performance compare to pure random testing. As it is obvious since the proposed method eliminates edge preference, performance loss in the proposed method is less than that of ART through dynamic partitioning localized by distance.

As it can be seen from Fig. 12 when the failure pattern is of point type the performance of the ART through dynamic partitioning localized by distance is similar and in a lot of cases less than that of pure random testing. This Fig shows us that Favorable failure patterns for ART methods are block and strip pattern. The point that should be considered is that in all cases the performance of the proposed method is less than that of ART through dynamic partitioning localized by distance and pure random testing.

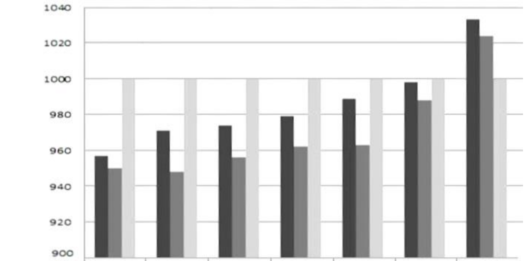
Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The reason is that when the failure pattern is of point type ART through dynamic partitioning localized by distance would not lead to any increase in performance and no edge preference problem would exists. In this case enlarging the input



	1	5	10	15	20	25	30
RT	1000	1000	1000	1000	1000	1000	1000
Art By localization	983	999	1005	1009	1014	1017	1028
Peoposed Method	1069	1071	1077	1128	1142	1145	1167

Fig. 11. Comparison of the F-Measure of RT and ART through dynamic partitioning by localization and the proposed method in 2-D input domain with strip failure pattern with diverse candidate test case set sizes



	1	5	10	15	20	25	30
RT	1000	1000	1000	1000	1000	1000	1000
Art By localization	957	971	974	979	989	998	1033
proposed method	950	948	956	962	963	988	1024

Fig. 12. Comparison of the F-Measure of RT and ART through dynamic partitioning by localization and the proposed method in 2-D input domain with point failure with diverse candidate test case set sizes.

Domain would only result in a larger input domain and larger partitions. As the result numbers of test cases require to reveal first failure would increase.

In a nutshell the simulations show that by enlarging input domain the edge preference problem is eliminated in ART through dynamic partitioning by localization. Eliminating edge preference improved failure detection capability of ART through dynamic partitioning by localization both in 2-dimensional and 3-dimensional input domain when failure pattern was of point or strip type.

The point which should be considered is that with point failure pattern ART methods usually do not increase performance. The reason is that when the failure pattern is of point type evenly distribution of test cases will not improve failure detection capability.

V. CONCLUSION

In this article a new method for eliminating edge preference problem in ART through dynamic partitioning localized by distance has been proposed. According to this combinational method before starting ART through dynamic partitioning localized by distance input domain would be enlarged by factor f in each dimension and then ART through dynamic partitioning localized by distance would be executed on the enlarged input domain.

The performance of this combinational method has been compared with ART through dynamic partitioning localized by distance and pure random testing and it has been observed that as the size of candidate test cases set grows the performance of the both ART through dynamic partitioning localized by distance and proposed method decrease but this decrease is less in the proposed method. According to the empirical results since this combinational method not only eliminates edge preference but also gains benefit from adaptive random testing approach advantages. We recommend testers that if Adaptive random testing through dynamic partitioning localized by distance is suitable for their goal, to use Adaptive random testing through dynamic partitioning localized by distance and enlarged input domain. In our future works we would investigate the effect of increase or decrease of factor F in the performance of the method. Also we would investigate the performance of enlarge input domain for eliminating edge preference when it is combined with other adaptive random testing methods.

REFERENCES

1. B.Beizer, Software testing techniques (Van Nostrand Reinhold, New York, 1990)
2. J. W. Laski , B. Korel, A data flow oriented program testing strategy, IEEE Transactions on Software Engineering, Vol. 9, No. 3, 1983, 347-354,.
3. J. Offutt and S. Liu, Generating test data from SOFL specifications, The Journal of Systems and Software, Vol. 49, No. 1, 1999, 49-62.
4. J. Offutt, S. Liu, A. Abdurazik and P. Ammann, Generating test data from state-based specifications, Software Testing, Verification & Reliability, Vol. 13,No.1, 2003, 25-53.
5. P. Stocks and D. Carrington, “A framework for specification-based testing”, IEEE Transactions on Software Engineering, Vol. 22, No. 11, 1996, 777-793.
6. L. J. White and E. I. Cohen, “A domain strategy for computer program testing”, IEEE Transactions on Software Engineering, Vol. 6, No. 3, 1980,247-257.
7. R. Hamlet. Random testing. Encyclopedia of Software Engineering, Wiley, New York,1994, 970–978.
8. P. S. Loo and W. K. Tsai. Random testing revisited. Information and Software Technology, Vol.30. No 7, 1988, 402-417.
9. E. J.Weyuker and B. Jeng. Analysing partition testing strategies. IEEE Transactions on Software Engineering, 17:703–711, 1991.
10. G. Myers, The Art of Software Testing, (NewYork: John Wiley & Sons, 1979).
11. F. T. Chan, T. Y. Chen, I. K. Mak, and Y. T. Yu. Proportional sampling strategy: guidelines for software testing practitioners. Information and Software Technology, 38:775– 782, 1996.
12. T. Y. Chen, T. H. Tse and Y. T. Yu, Proportional sampling strategy: a compendium and some insights. The Journal of Systems and Software, 58, 2001, 65-81.
13. I. K. Mak, On the effectiveness of random testing. Master’s thesis, (Department of Computer Science, The University of Melbourne, 1997).
14. T. Y. Chen, G. Eddy, R. Merkel and P. K. Wong, Adaptive random testing through dynamic partitioning , accepted to appear in Proceedings of the 4th International Conference on Quality Software (QSIC2004)

AUTHOR PROFILE

Mr. Korosh koochekian sabor received his bachelor of science in software engineering in 2009 and his master of science in software engineering in 2012. From 2009-2012 he worked as lecturer in numerous universities in Iran. He has developed the biggest web based bus ticket system in Iran in 2010.his interests include analysis of algorithms, optimization, and software engineering (especially software testing and reliability).

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Dr. Mehran Mohsenzadeh has received his Ph.D in software engineering from Department of Computer engineering, Science and Research Branch, Islamic Azad University of Iran, in 2004. His major interests are Data Integration, Data Quality, Data/Web Mining, Software Architecture, Methodologies, Ontological Engineering and has published more than 50 conference papers and 5 journal papers. He is Assistant Professor of Computer Engineering Department, Science and Research Branch, Islamic Azad University of Iran.