Review on Automated Test Data Generation Using Computational Intelligence Techniques

Kumari Nivedita[1] ,Swadhin Kumar Barisal[2]

[1] Institute of Technical Education and Research, SOA University Bhubaneswar, Odisha -751030, India

[2] Institute of Technical Education and Research, SOA University , Bhubaneswar, Odisha -751030, India

Abstract

Software testing is the necessity in the field of software development and its needs to work more effectively in order to develop error free software. The major difficulty was obtained in the development process is generating test data, is provided input to the component under test. Previously many researchers have been done in order to reduce time and bug in the manual testing phase. Henceforth researchers were focusing on the evolutionary techniques. Somehow many of the evolutionary technique work properly and the result was up to the makeable label. But the bugs were obtained in that process was not resolved properly. After analysis, the hybrid approach was working somehow good in order to generate test data. We use some population-based meta-heuristics algorithm as an illustration Particle Swarm Optimization, Differential Evolution, and Genetic Algorithms to solve this problem. These techniques are followed by path coverage and maximum fitness.

Keywords: Genetic Algorithms, Particle Swarm Optimization, Differential Evolution and, Maximum fitness.

1.Introduction

Software testing is one of the time taking tasks during software development, in order to deliver error free software. In the previous year, many methodologies were used for generating test data. In the year 1999, Roy Pargas, Mary Jean Harrold, and Robert R. Peck proposed one prototype tool which generates test data by the help of GA, called TGen [5]. Before proposing this tool they have tried the random data [3] for generating test data using path coverage program. But the theory was failed as if the path is infeasible [5]. By the analysis of many research papers, it has been concluded that GA probably not genuine in order to generate test data in the meantime, therefore we move forward for some other evolutionary techniques for

the generation of test data. We are presenting the model for the generation of test data (figure.1).Which comprises of program analyzer, path selector, and test data generator. The initial code is executed by program analyzer, which identifies the best data for next level i.e. for Path selector, Identifies the test data which covers maximum node. Then the reduced data set from the Random data set is now used for the test data generation techniques in order to generate test data.
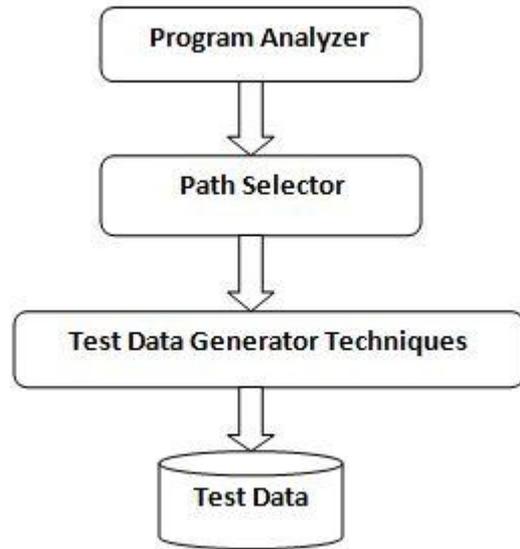

Figure 1 Architecture of test data generation

The GA is applied to CFG in order to obtain the no. of path covers by each node [5], which is to be sorted in the terms of maximization. Out of that, those data covers maximum no of the node that will have the highest fitness value, are selected for next step.
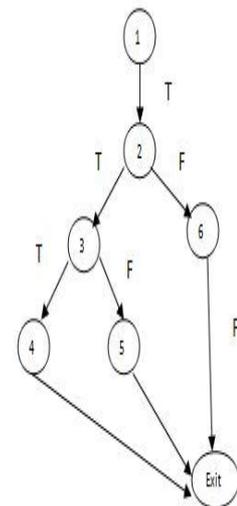
Figure 2 Program on the left and its CFG on the right

International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)
Web Site: www.ijettcs.org Email: editor@ijettcs.org
Volume 6, Issue 3, May- June 2017   ISSN 2278-6856

PSO used for finding the optimal path in the program. The main principal behind this combination (interaction) is called stigmergy [1], is a mechanism of indirect coordination between agent and action.

2. Computing Techniques Used
For generating test data we are using some metaheuristic technique such as:
- Genetic Algorithm
- Particle swarm optimization
- Differential evolution

In this section, the various existing method is used for automated test data generation. After working with them some limitations are pointed out. Further by the use of some existing method one hybrid algorithm [1] [2] is developed, that is to be used for the comparison between other algorithms.

3. Limitations of Existing Method
Earlier many software testing processes use an evolutionary technique for test data generation within the test cases. Some are successfully designed and implemented and some are not run successfully. Therefore test data generation with less time and the lesser cost is still an immense question. Consequently by combining some of the techniques and trying to rectify the better one is our main task

4. Literature Review
Saumya Dixit and Pradeep Tomar presents a Hybrid algorithm i.e. GPSHA (2015). This takes the advantage of two important techniques GA, PSO. These two techniques and hybrid algorithm were tested in the set of program in order to find the effectiveness of GPSHA over GA and PSO.The author achieves his goals as GPSHA gives better result in test data generation with the minimum time [1].

Omur Sahin, Bahriye Akay (2016) introduces some metaheuristics algorithms to solve the optimization problem. Here author wants to compare the population-based metaheuristic algorithm with the Random search algorithm (RS). By the experiment, the author concluded that metaheuristic techniques are more useful than RS for large size problem [3].

Harsh Bhasin, Naresh Chauhan, Sandhya Pathak (2015) presents the GA based technique for the generation of test data. Previously for manual testing PUT was used, takes more time for processing, so testing requires automation of data. Therefore, the author decides to use GA for automatic test data generation under the PUT and this was successfully achieved with the less time [4].

Sanjay Singla, Dharminder Kumar, H M Rai and Priti Singla (2011) presents one hybrid algorithm that is based on the combination of GA and PSO, used to generate test data by data flow coverage through the dominance concept of two distinguished nodes. The behavior of a hybrid algorithm is analyzed by the no. of the small program and compared with GA and PSO individually with the same set of the program in order to find the best method for test cases [2].

Kulvinder Singh and Rakesh Kumar (2010) present the GAs (Genetic Algorithm and specification) technique in order to reduce faults and bugs occurred in the genetic algorithm. Some new test code is applied for this specification technique [6].

Sapna Varshney, Monica Mehrotra, presents a differential evolution based approach for generating Test Data. The author tries to compare the proposed approach DE with the other highly adaptive techniques Random Search, GA, PSO, and it was found that DE gives better performance than another one. DE works with one data flow dependency for generating test data at a time [7].

Roy P Pargas, Mary Jean Harrold, Robert R.Peck (1999) designed one prototype model tool i.e. TGen, for generating test data by the use of GA. And this tool is going to compare with GA algorithm. Input is random, for both. The author concluded that TGen works more efficiently in order generate test data [5].

Shujuan Jiang, Jiaojiao Shi, Yanmei Zhang, Han Han trying to reduce the adaptative characteristics of PSO. It improves the quality of PSO by reducing the evolution equation and after that; it makes an adaptive adjustment based on the inertia scheme [8].

Ahmed A. A. Esmin and Stan Matwin (2013) present one another hybrid algorithm HPSOM, with the help of PSO and GA mutation method. This algorithm is designed to maintain the automatic balance between local and global optima searching abilities [9].

P. Maragathavalli (2011) presents the test data generation using evolutionary computation. The evolutionary testing is used to optimize search based technique in order to generate test data by the use of an evolutionary algorithm. Here author compares the GA based testing technique with random testing and finds that for complex problem GA is defeating random testing [21].

Anastasis A. Sofokleous, Andreas S. Andreou (2008) proposes a frame for finding test data generation using Genetic Algorithm. This frame consists of program analyzer (used for path analyzing)and test data generator(uses two algorithms Batch-Optimistic (BO) and the Close-Up (CU)) for finding the best solution, and that result is compared with the previously defined dynamic algorithm in the case of efficiency, performance etc[19].

Moataz A. Ahmed, Irman Hermadib (2007) presents an approach which reduces cost and time during testing. For covering single path and generating test data GA gave the best result for some time but whenever it is trying to use for multiple path coverage's simultaneously then GA fails.

Therefore author designed GA-based test data generator for multi-path coverage at the same time [23].

Akshat Sharma, Rishon Patani, and Ashish Aggarwal (2016) present some method by using a genetic algorithm for generating test data. The author wants to compare the random method for test data generation with a genetic algorithm. Therefore author uses some genetic algorithm techniques which are using population as input and evaluates multiple test cases for that population. Simultaneously we also drive the same population for the random method and compare them in order to increase the efficiency for with reduce time in software testing process [15].

Jaya Srivastaval and Twinkle Dwivedi approaches that software testing is a very important feature in the field of software development life cycle. For detecting the fault with minimum no. of effort, automation for test data generation is the best choice. Hence for optimization of multiple sets of test cases author uses Genetic algorithm approach, as it automatically generates the best test cases. Randomly generated test case data are optimized by Genetic Algorithm, and it detects more no. of fault possible from source code which helps to improve the software quality as well as software testing process [14].

5. Findings

By this research analysis, we have obtained that GA gives better result in obtaining the test data by the no. of test cases. But GA may fail when we take large size population, as it takes more time to evaluate the problem. So researchers are trying to make hybrid algorithm by the help of PSO and GA. PSO is used in GA and it may reduce the total evaluation time for producing test data generation. It is observed by testing some program in the hybrid algorithm. Henceforth this hybrid algorithm is trying to compare with the DE Algorithm. By research analysis among all evolutionary technique some time DE gives best test data. As it works with one function in its algorithm, which identifies the better one among parents and child in the selection steps only. Therefore DE is used for comparison.

6. Conclusion

For generating test data use of computational intelligence technique is the best choice. Integration of some of the evolutionary approaches likewise GA, PSO and make one hybrid approach is the main idea. Software testing is the process of eliminating the common data and delivers the error free software. Genetic algorithm uses CFG for identifying path, within that PSO is applied for finding the best path to another one. Rest of GA process is used for generation of test data. The algorithm continues till the maximum no. of coverage has not been reached.

References

[1] Saumya Dixit, & Pradeep Tomar, "Automated Test Data Generation Using Computational Intelligence", 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) 2-4 Sept 2015

[2] Sanjay Singhla, Dharminder Kumar, H M Rai & Priti Singla," A Hybrid PSO approach to Automate test Data Generation For Data Flow Coverage with Dominace Concepts" International Journal of Advanced Science & Technology 2011

[3] Omur Sahin, & Bahriye Akay, "Comparisons of Metaheuristic Algorithms and Fitness Functions on Software Test Data Generation", Applied Soft Computing vol.49, pp: 1202-1214, 2016

[4] Harsh Bhasin, Naresh Chauhan, Sandhya Pathak, "Test Data Generation Using Computational Intelligence Technique", Journal of Computer Sciences and Applications, vol. 3, no. 2 pp: 56-60, 2015

[5] Roy P. Pargas, Mary Jean Harrold, Robert R P, "Test-data Generation Using Genetic Algorithms. "Software Testing Verification and Reliability vol. 9, no. 4 pp: 263-282, 2012.

[6] Kulvinder Singh and Rakesh Kumar, "Optimization of Functional Testing using Genetic Algorithms." , International Journal of Innovation, Management, and Technology, vol. 1, no. 1, ISSN: 2010-0248, April 2010

[7] Sapna Varshney & Monica Mehrotra, "A Differential Evolution Based Approach to Generate Test Data For Data-flow Coverage", Computing, Communication and Automation (ICCCA), International Conference on IEEE, 2016.

[8] Shujuan Jiang, Jiaojiao Shi, YanmeiZhang, Han Han, "Automatic Test Data Generation Based On Reduced Adaptive Particle Swarm Optimization Algorithm", Neurocomputing vol. 158, pp: 109–116, 2015

[9] Ahmed A. A. Esmin & Stan Matwin, " HPSOM: A Hybrid Particle Swarm Optimization Algorithm With Genetic Mutation", International Journal Of Innovative Computing, Information And Control ICIC International vol. 9, no. 5, pp. 1919-1934, May 2013

[10] Ruchika Malhotra & Mohit Garg, "An Adequacy of test data generation techniques using genetic algorithms", Journal of Information processing Systems, 2011

[11] Maneela Tuteja & Gaurav Dubey, "A Research Study on importance of testing & Quality Assurance in Software Development Life Cycle", International Journal of Soft Computing & Engineering (IJSCE), ISSN: 2231-2307, volume-2, Issues – 3, July 2012.

[12] Yeresime Suresh & Snatanu Ku Rath,"A genetic algorithm based approach for the test data generations basis path testing", in pro. International conference on soft computing & software engineering, 2013

[13] D. P. Mohapatra," Automated Test Case Generation and Its Optimization for Path Testing Using Genetic Algorithm and Sampling ", 2009 IEEE.

[14] Srivastaval, Jaya, and Twinkle Dwivedi. "Software Testing Strategy Approach On Source Code Applying Conditional Coverage Method." International Journal of Software Engineering & Applications (IJSEA), Vol.6, No.3, May 2015

[15] Sharma, Akshat, P. Rishon, And Ashish Aggarwal. "Software Testing Using Genetic Algorithms." Int. J. Comput. Sci. Eng. Surv(IJCSES) 7.2 (2016): 21-33.

[16] P. R.Srivastava1 And Tai-Hoon Kim2, "Application of Genetic Algorithm In Software Testing", International journal of Software Engineering And Its applications Vol. 3,No.4, October 2009

[17] Kukkonen, Saku, and Carlos A. Coello Coello. "Generalized Differential Evolution for Numerical and Evolutionary Optimization." *NEO 2015*. Springer International Publishing, 2017. 253-279.

[18] Mishra, Deepti Bala, et al. "A Systematic Review of Software Testing Using Evolutionary Techniques." Proceedings of Sixth International Conference on Soft Computing for Problem Solving, 2017.

[19] Sofokleous, Anastasis A., and Andreas S. Andreou. "Automatic, evolutionary test data generation for dynamic software testing." Journal of Systems and Software 81.11 (2008): 1883-1898.

[20] Mishra, Deepti Bala, Rajashree Mishra, Kedar Nath Das, and Arup Abhinna Acharya. "A Systematic Review of Software Testing Using Evolutionary Techniques." In Proceedings of Sixth International Conference on Soft Computing for Problem Solving, pp. 174-184, 2017.

[21] Windisch, Andreas, Stefan Wappler, and Joachim Wegener. "Applying particle swarm optimization to software testing." In Proceedings of the 9th annual conference on Genetic and evolutionary computation, pp. 1121-1128. ACM, 2007.

[22] Maragathavalli, P. "Search-based software test data generation using evolutionary computation." arXiv preprint arXiv:1103.0125 (2011).

[23] Ahmed, Moataz A., and Irman Hermadi. "GA-based multiple paths test data generator." Computers & Operations Research 35.10 (2008): 3107-3124.