CLUSTERING USING KFCM TECHNIQUE AND PRIORITIZATION USING GWO IN CLOUD BASED TESTING

Sunitha B[1], Dr. Y.S.S.R Murthy[2]

1Computer Science, Research Scholar, Mewar University(India)

2Computer Science Engineering, Research Supervisor, Mewar University(India)

ABSTRACT

One of the significant activities in any software development is testing. To find any defects that are introduced or uncovered as a result of the changes in software development testing is called Regression testing, it is widely used in software testing. In most of the studies, new algorithms are developed aiming to explore the code and detect the risky areas of the program more effectively than before. Cloud-based testing offers a challenging activity via lower costs, pay-peruse and elimination of upfront capital expenditures. To find an accurate test case optimization and reduce the drawbacks of various algorithms, in our proposed technique use optimal test case prioritization in Cloud based regression testing with aid of KFCM and GWO. The proposed idea consists of two major steps. In the first one, KFCM technique is used for scheduling the test cases based on their similarity. The second step is to apply GWO to find an optimal solution. In KFCM algorithm uses a centroid function to compute the similarity or distance between test cases in the cloud. Based on the similarity the test cases are grouped or clustered, in each cluster the test cases are relatively independent. Finally we have to prioritize the test cases with the help of gray wolf optimization technique. In GWO, minimum sum of feature values are selected as the best test cases. This idea will enhance the performance of prioritizing test cases using regression testing in cloud. Thus we will obtain effective prioritized test cases. Our approach will be executed on JAVA with Cloud Sim platform. The performance will be evaluated with various evaluation metrics.

Keywords: *Cloud testing, Cloud Sim, GWO, Kernel fuzzy means prioritization,*

SECTION I

1.INTRODUCTION

Software testing Process is a set of comprehensive activities that are carried out for finding errors or bugs in the developed software. Given a set of system requirements the Software development process is aimed at evaluating a software item(s) that are studied thoroughly from the initial phase to the retirement phase. Regression testing is a widely used method of testing that focuses on the modified software versions to check whether it behaves as per specifications [1,2]. The methodology used in the regression testing makes sure that the recently modified changes do not exist anymore [3]. The legacy approaches to software testing have become costly in terms of effort, cost and overheads. So many organizations are looking for practices that provide business and profits for long term to organizations.

With cloud-based testing, organizations things become much simpler and don't need to worry about finding servers; procuring testing tools and installing them or using licensed software's. Testers often are provided with various services to access ready to use virtual labs that are online. With cloud testing the test management and execution tools, middleware and storage necessary for creating a test environment closely mirrors the real environment.

Grey Wolf Optimizer (GWO) algorithm is a new optimization method which is employed to solve optimization problems of different varies (S. Mirjalili et al, 2014). As other heuristic algorithms in the area of evolutionary computation, in optimization processes GWO has not required to the gradient of the function. GWO a mathematical model and the computer simulation which mimics the similar resemblances which match with a hunting mechanism of grey wolves in nature.

This paper proposes a new method which utilizes GWO (Grey Wolf Optimizer) algorithm with an aid of KFCM (Kernel Fuzzy C Means) technique in cloud testing. Also a comparative study and analysis of existing K means technique with the proposed KFCM is studied.

2. RELATED WORKS

Krishnapuram and Keller changed the constraint of memberships in FCM and proposed the possibilistic c-means (PCM) algorithm [4]. The advantages of PCM are, it overcomes the need to specify the number of clusters and it is highly robust in a noisy environment. However, still exist some weaknesses in the PCM, i.e., it highly depends on a good initialization and has the poor performance to identify similar clusters [5] [6]. Usually, the FCM can provide a reasonable initialization and an approximate scale that determines the degree to which the objective function is important when compared with the first.

Jyoti, Kamna Solanki Suggested "A Comparative Study of Regression Testing Techniques" in their paper, the results of a comparative study on five different regression test optimization techniques is represented. The comparison is based on different criteria such as number of test cases chosen, their test execution time, accuracy, user parameters, global variables handling, type of testing etc. The above specified algorithms are found to be suitable for different requirements/environments of regression testing [6].

Cloud computing emerged as a new paradigm in software industry and to successfully provide cloud services and resource sharing, the cloud must be tested before came into service. Cloud testing is one amongst the emerging areas of testing in which web applications make use of cloud computing environment and infrastructure as a service (IaaS) to simulate real world scenarios on how to control the user traffic by using cloud applications across the globe. The work on this paper by Dr. Rahul Malhotra & Prince Jain [7] is explanation on various cloud computing testing techniques and challenges.

Chitra Jain, Gaurav Srivastava proposed "Designing a Classifier with KFCM technique to Achieve Optimization of Clustering". In their paper they suggested that KFCM Algorithm is based on the idea of minimizing the objective function by and reducing the number of iterations required to gather a feasible cluster center. Initially, the cluster centers are scattered and then based on the concept of empirical formula the distance in which distance of test cases (samples) from the center of the clusters is calculated and the centroid cluster thus obtained, its cluster membership degree is calculated. The Cluster center is updated iteratively and the difference between the objective function and cluster center get minimized [8].

Er. Tamanna Narula, Geetika Sharma proposed a paper " Framework for Analyzing and Testing Cloud based applications". This paper explained that Cloud based testing is new paradigm in software testing method where in this system, the use Cloud computing use web applications environments ("cloud") and gather to simulate Real-world user traffic as a means of evaluating stress testing and load testing web sites. Unit testing activities provide an importance where on-demand software testing services can be utilized. Paper emphasizes on how automatic test case generation for unit tests [9] and concept of Symbolic execution are migrated to cloud environment. The problem of regression testing case selection is solved by prioritizing test cases [10].

The work proposes a new meta-evolution called Grey Wolf Optimizer (GWO) inspired by grey wolves (Canis lupus). The GWO depicts a leadership hierarchy and mechanism of hunting grey wolves in nature. For simulating the leadership hierarchy four types of grey wolves parameters such as alpha (α), beta (β), delta (δ), and omega (ω) are considered. In addition, there are three main steps of hunting, searching for prey, encircling prey, and attacking prey, are implemented. This well-known Algorithm is then benchmarked on several well-known test functions, and the results are verified by a comparative study with PSO ,GSA, DE EP algorithms etc. The results show that the GWO algorithm is able to provide very competitive results compared to these well-known meta-heuristics [11].

Fuzzy C-Means can determine, and updating the membership values iteratively using the test cases with already existing number of clusters i.e. K. Thus, every test case present in the test suite carries a membership value for entire cluster. FCM has been prominently used in various fields which is discussed in [12]-[15]. Many numbers of variants of the FCM algorithm had been discussed. Sikka et al. [16] pointed out some of these algorithms. In this paper, they implemented FCM using three basic options and compared number of iterations, accuracy with K-Means algorithm.

In this paper, a kernel-based fuzzy c-means algorithm (KFCM+GWO) is proposed. KFCM uses a new kernel-induced metric in the cloud space to replace the original Euclidean norm metric in FCM. By replacing the inner product by an appropriate 'kernel' function, performing a nonlinear mapping to a high dimensional feature space without increasing the number of parameters. Many learning systems have used this 'kernel method' and this motivated to find a minimal objective function GWO (Grey Wolf Optimizer).

The rest of this paper is organized as following: In Section 2, we introduce the KFCM technique, and Section 3 presents the GWO algorithm. Proposed work along with some conclusions and future scope are given in Section 4.

SECTION II

KERNEL FUZZY C-MEANS

The concept of KFCM technique is based on minimizing the objective function by significantly reducing number of iterations required to obtain the cluster center using a centroid function. Initially, the cluster centers are randomly initialized and then the degree of cluster membership is calculated by a Euclidean distance in which distance of each sample(test case) from the centroid cluster thus obtained and then cluster centers are updated iteratively, until objective function (OF) get minimized.

$$OF = \sum_{i=1}^{n} \sum_{c=1}^{C} M_{ic}^{m} \| T_i - C_c \|^2 \qquad \textbf{(1)}$$

Step by step procedure of KFCM technique is given below:

Step by Step procedure of KFCM:

Step 1: Kernel version of the FCM algorithm and its objective function are given bellow:

$$OF = \sum_{i=1}^{n} \sum_{c=1}^{C} M_{ic}^{m} (1 - K(T_i - C_c)) \qquad (2)$$

Step 2: Compute the fuzzy centers C_i

$$C_c = \frac{\sum_{i=1}^{n} M_{ic}^{m} K(T_i, C_c) T_c}{\sum_{i=1}^{n} M_{ic}^{m} K(T_i, C_c)} \qquad (3)$$

Step 3: Calculate the fuzzy membership function M_{ic} using

$$M_{ic} = \frac{(1 - K(T_i, C_c))^{-1/(m-1)}}{\sum_{c=1}^{C} (1 - K(T_i, C_c))^{-1/(m-1)}} \qquad (4)$$

Step 4: Repeat step 2 and 3 until the maximum of OB_m value is achieved

Identify the essential conditions for minimizing OF is revise equation (3) and (4) only when the kernel function K is selected to be the linear function with $K(T_i, C_c) = T_i^T C_c + S$. Where, S is the constant value.

SECTION III

GREY WOLF OPTIMIZER (GWO)

This section summarizes the main steps in grey wolf optimizer (GWO) to optimally run the number of test cases in the cloud. Grey wolf algorithm (Canis lupus) is a new population based algorithm which is introduced in 2014 by Mirjalili et al (Mirjalili,et al, 2014). Grey Wolf Optimizer algorithm inspired by grey wolves. Method mimics grey wolves the social hierarchy and hunting behavior. For simulating the leadership hierarchy in Grey Wolf Optimizer algorithm, four groups are defined: alpha, beta, delta, and omega. Furthermore, the three main steps of hunting, searching for prey, encircling prey, and attacking prey, are simulated. This GWO algorithm requires few parameters need to be set, namely initializing alpha, beta, and delta, number of iterations, and a stopping criterion.

The important main steps of grey wolf hunting are as follows:

a) The Tracking, chasing, and approaching the victim

b) Pursuing, Encircling and Harassing until the victim stops moving.

c) Initiating the attack towards the victim.

For modeling the social hierarchy of wolves until designing Grey Wolf Optimizer, the fittest solution is considered as the alpha (α) ,accordingly, the second and third best solutions are named as beta (β) and delta (δ) respectively. The rest of the remaining candidate's solutions are considered to be omega (ω) .The final position would be in a random position within a circle which is defined by the position of α, β, and δ in the search space. In other words alpha (α), beta (β), and delta (δ) estimate the exact victim's current position and other wolves update their positions randomly around the victim. Pseudo code of the algorithm is shown in its simplest form in Fig. 1.

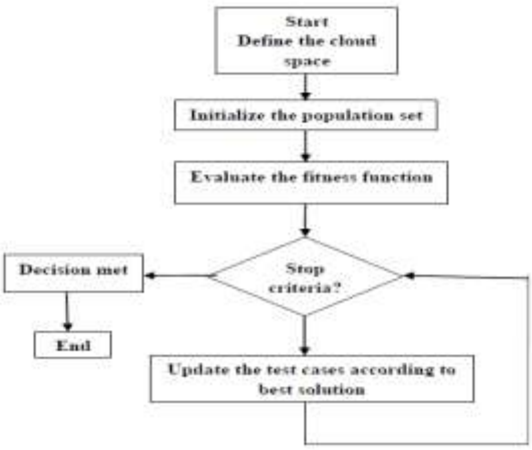

Fig.1 Flow chart of Grey Wolf optimization

SECTION IV

PROPOSED METHOD

In the proposed method, we consider the following main steps:-

1) Initially test case generation based on feature values is carried out.
2) KFCM is used to separate the test cases in cloud environment.
3) GWO is used to prioritize the test cases in cloud .

The aim of using proposed KFCM technique is to improve the software efficiency and its competency in cloud. First application is uploaded on to the cloud and test case generation being one of the major steps in the study is applied to identify the test cases. As part of test case generation we considered basic feature values like statement coverage, line coverage, loop coverage etc. The underlying test generation step uses KFCM to cleanly separate the relevant and irrelevant test cases. Since our aim is to improve software efficiency viz performance only the relevant test cases are sent to cloud for prioritization. And GWO algorithm which is illustrated above is used here which will obtain the prioritized test cases.

Analysis Process:

Analysis I: While applying GWO algorithm every cluster is given for test case prioritization and for each cluster thus created is iterated using regression testing. Testing is conducted and at regular intervals at the end of every 5th iteration , 10th iteration and subsequent iteration the performance evaluation is carried out on each cluster i.e. cluster 1,2,3.

It is observed that at the end of 20th iteration and 25th iteration the objective function OF is attained with various values of cluster1 (0.0099) and cluster 2 (0.02) and cluster 3 (0.011).

Analysis II:

The other factor based on performance is the memory used and the CPU time (in ms)will be computed. The study shows that the amount of time i.e. taken by the test cases in cloud is faster than when computed normally in traditional systems. Study also reveals that at the end of 20th iteration and 25th iteration is 619 ms as against 5th iteration 326ms.

Analysis III:

A comparative analysis of the proposed work shows that KFCM + GWO is better than (K means + GWO) in cloud based system. GWO algorithm takes minimum execution time to comprehensive the prioritization procedure related to the available K means and GWO IS shown in fig.2 and less memory usage (calculated in bits) in cloud when compared with K-means.

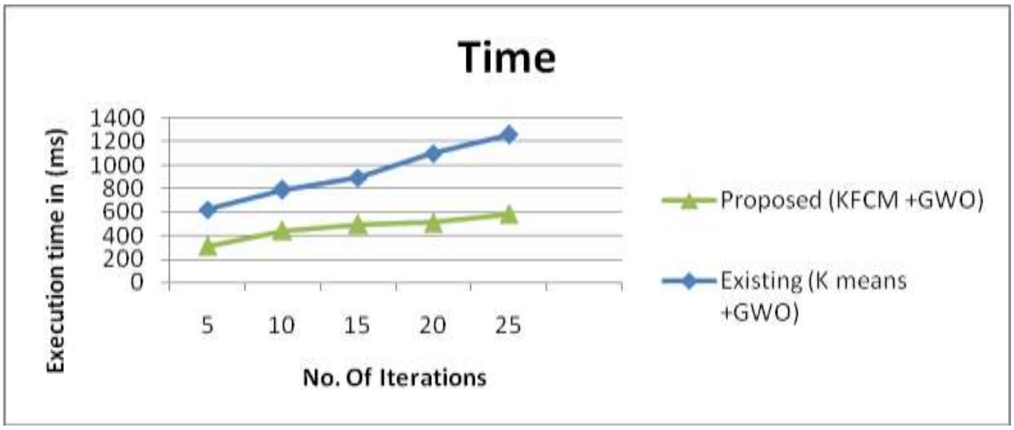

Fig.2 Execution time comparison for proposed Vs existing method

CONCLUSION

In comparative analysis, the work is compared bearing in mind existing K means algorithm with GWO. KFCM+GWO algorithm technique is better than K Means +GWO in cloud base scenario. To find a best solution the distance between any two test cases is considered as a metric in K means technique. Clustering technique using KFCM results are evaluated using the centroid function rather than the distance between the test cases. And also since the applications are put on cloud there is user privacy in diverse web clients. Both virtual and real-time test cases can be applied by the users. Online validation and measurements is quite possible rather than fixed test environment. More performance evaluation metrics can be applied looking into the growing needs and security factors on the cloud.

REFERENCES

[1] Siripong Roongruangsuwan and Jirapun Daengdej,"Test Case Prioritization Techniques", In proceeding of the Journal of Theoretical and Applied Information Technology, pp. 45-60, 2010.

[2] Deepak Garg and Amitava Datta, "Test Case Prioritization due to Database Changes in Web Applications", In proceeding of *IEEE International Conference on Software Testing, Verification and Validation, pp. 726-730, April 2012.*

[3] Albert Pravin and Subramaniam Srinivasan, "Effective Test Case Selection and Prioritization in Regression Testing", In proceeding of the *Journal of Computer Science, Vol. 9, No. 5, pp. 654-659, 2013.*

[4] S. Nachiyappan and S. Justus, "Cloud Testing Tools and Its Challenges: A Comparative Study", *International symposium on big data and cloud computing, Vol.50, pp. 482-489, 2015.*

[5] Shraddha D. Mungekar and Dhanashree K. Toradmalle, "Cloud Testing- An Advancement in Software Testing", *International Journal of Application or Innovation in Engineering and Management, Vol.4, Issue 3, pp. 61-63, March 2015.*

[6] Jyoti, KAMNA "**A** Comparative Study of Five Regression Testing Techniques":A Survey *International Journal of Scientific and Research Vol.3, Issue 8, August 2014*

ijates
ISSN 2348 - 7550

[7] Dr. Rahul Malhotra and Prince Jain "Testing techniques and its challenges in cloud" *The SIJ Transactions on Computer Science Engineering & its Applications (CSEA), Vol. 1, No. 3*

[8] Chitra Jain, Gaurav Srivastava **"**Designing a Classifier with KFCM Algorithm to Achieve Optimization of Clustering and Classification Simultaneously" *International Journal of Emerging Technology and Advanced Engineering Volume 3, Issue 9*

[9] Er. Tamanna Narula, Er. Geetika Sharma "Framework for Analyzing and Testing Cloud based Applications" *International Journal of Advanced Research in Computer Science and Software Engineering Volume 4, Issue 6, June 2014*

[10] Khushwant Virdi , Radhika Kalyan, Navjot Kaur "Software Testing as a Service (STaaS) using Cloud Computing" *Volume 4, Issue 1, January 2015*

[11] sayedalili mirjalili "Grey wolf Optimization" *Advances in Engineering Software Volume 69, March 2014, Pages 46–61*

[12] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy C-Means model," *IEE EFS, vol. 3, no. 3, p. 370, 1995.*

[13] D. Q. Zhang and S. C. Chen, "A novel kernelized fuzzy C-Means algorithm with application in medical image segmentation," *Artif. Intel. Med, vol. 32, pp. 37–50, 2004.*

[14] S. Migaly, J. Abonyi, and F. Szeifert, "Fuzzy Self-Organizing Map Based on Regularized Fuzzy C-Means Clustering," in Advances in Soft Computing, Engineering Design and Manufacturing, J. M. Benitez, O. Cordon, F. Hoffmann et al. (Eds.), *Springer Engineering Series, (Revised papers of the 7th On-line World Conference on Soft Computing in Industrial Applications (WSC7)), 2002, pp. 99–108.*

[15] S.Albayrak and F. Armasyali, "Fuzzy C-Means Clustering on Medical Diagnostic System," in Proc. *Int. XII Turkish Symp, Artif. Intel. NN, 2003.*

[16] K. Sikka, N. Sinha, P. K. Singh, and A. K. Mishra, "A Fully Automated Algorithm under Modified FCM Framework for Improved Brain MR Image Segmentation," *Magnetic Resonance Imaging, vol. 27, no. 7, pp. 994–1004, 2009*