Solving Component Identification Problem by using Evolutionary Method

Anil Kumar Pandey, Tulika Pandey

Abstract: Today system is day by day increasing and the maintenance of a system is a very difficult task. To explain the component identification problem for architecture which is support to a component based technology. To identify component and optimal , first we describe the Ant clustering approach and then we explain the method how the ABC helps for groups for cluster the Usecase in this paper proposes a .ABC is a population based NP type problem with heuristic that can be used to find approximate solutions to difficult combinatorial optimization problems. This paper presents an effective ABC approach with ant colony in this paper we also find the optimal number of components or clusters. Finally the ABC applies into different exiting system.

Keywords: Modeling, Complexity, Software,UML

I. INTRODUCTION

Swarm Intelligence is an evolutionary, distributed intelligent approach for help us to find optimized type problems. It is managed by the environment of colonies, or swarm of social related. History of social approach of organisms (individuals) in swarm prompted the design of very efficient optimization and grouping methods.[4]

There are different grouping based (cluster) based on Artificial intelligence tools are Particle Swarm Optimization (PSO) and Generate Algorithm (GA). PSO is a new evolutionary computation method developed by Kennedy and Eberhart in 1995.

Both are stochastic optimization method, designed on the social behavior of animals. For example a flock of birds or a swarm of bees or a. PSO is a population-based find method where the alone referred to as particles are cluster into a swarm. Every particle in the swarm show as candidate result to the optimization. In a PSO system, Every particle is flown through the different direction search environment and manage place in search space as per own experience. Also that of neighboring particles.[8] A particle therefore makes use of the best position encountered by it and the best position of its neighbors to position itself toward an best results. The effect is that particles fly toward an best and while still searching a wide area around the current best solution.[12] .The performance of every is measured according to a predefined fitness value and it represent the problem being solved. The advantage of PSO is quick result with others existing methods like Genetic Algorithms (GA),[2,3]. For implementing PSO successfully, one of the important points is that is search how to map problem solution into the PSO, it is towards direct link with feasibility and performance. Applications of PSO include function approximation, clustering, optimization of mechanical structures, and solving systems of equations.

There are many algorithm have been developed for solving numerical and combinatorial optimization problem , clustering with ABC method is an more better than others conventional clustering techniques .[5,6]

II. REVIEW OF EXITING APPROACH

2.1 UML Design

UML is used to design the blueprint of the presentable language for visualizing, specifying, constructing, and documenting software artifacts.UML give a standard type to define a system's blueprints, includingλ conceptual things such as business processes and system functions as well as concrete things such as programming language and database or maintainable components. UML does have an official standard Backed by OMG (Object Management Group). OMG is a not-for-profit industry specifications consortium. OMG members define and maintain the UML.

There are different kind of diagram as per below defined.

- Use Case diagram
- Integration diagram
- Class diagram
- Sequence diagram
- Collaboration diagram '
- Package diagram
- Object diagram
- State chart diagram

Every diagram used as per requirement but the Use case diagram is presented abstract form of a system and more closely to users. A Use-Case used different notations [5] for represent a logical representation of a system. it is give the relation between customer and a machine that fulfill the user goal.

Use-Cases are a kind of diagram for feature need for software machine. The external customer who interacts with software called as an 'Actor'. Use-Cases are fundamentally in text form. A use case defines a activity in a step by step that gives something of calculated value to an actor and is denoted by shaped called horizontal ellipse. .

An actor is a user or external system that apply a role in one or more interactions with your system. Actors are representing as stick. Associations between actors and use cases are indicated by solid lines. An association exists whenever an actor is involved with an interaction described by a use case

Revised Manuscript Received on February 15, 2020.

Anil Kumar Pandey, Sam Higginbottom University of Agriculture, Technology and Sciences (SHUATS), Allahabad, India. E-mail: anipandey@gmail.com

Tulika Pandey, Sam Higginbottom University of Agriculture, Technology and Sciences (SHUATS), Allahabad, India.
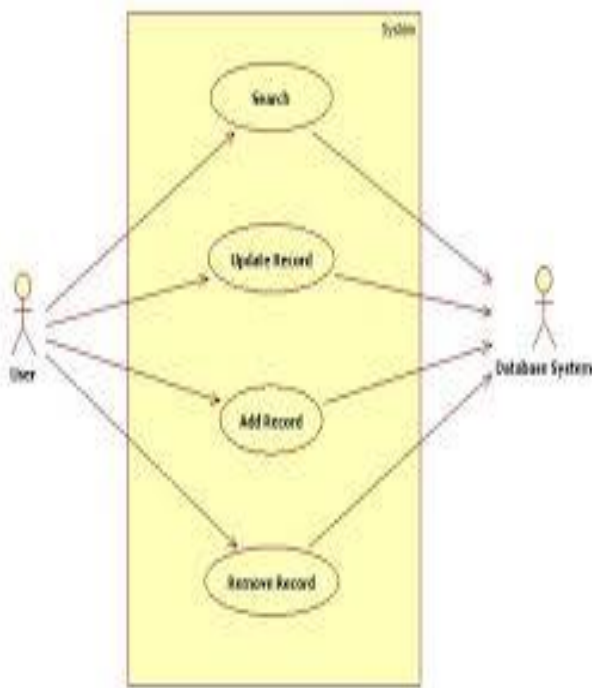
Figure1: Example of UseCase Diagram

The UML is called a unified modeling language. It is a graphical representation provides us with syntax for describing the major elements of software. There are three techniques one is excellent for design and implement and second one had worked extensively with the programming language, third one is object modeling technique was best for analysis.

The featured known as Use Cases of object oriented techniques is a powerful technique for understanding the behaviour of an entire system. A Use Case used basically represents behaviour from a user viewpoint. This diagram is a valuable aid during analysis and design. Use Cases support to understand software requirements**.**

2.2 Clustering

Clustering terms used for the objects separate similar into subgroups. The aim of this separation is two wrap objects into single related group (called component) are needed to be same interaction with nearest similar properties between them..

There are many ways to clustering algorithms available like partitioning methods, hierarchical methods, density-based clustering and grid-based clustering. For the purpose of our identification of components, a representative of the class of partitioning methods, and agglomerative average link clustering, which is a hierarchical approach. With extend we compare against one-dimensional self-organizing maps. According to V. E Castro, cluster or group not to define in one single property and that is the reasons for available different methods of clustering.

 Cluster means " a group of objects in a similar grid" . There are different cluster models purposed by different authors. A model define by cluster is key to understanding the differences between the various approaches like hierarchical, centroid and distributive model. Hierarchical is clustering builds models based on distance connectivity. K-means represents each cluster by a single mean vector and distribution models clusters are modeled using statistical

distributions, such as multivariate normal distributions used by the Expectation-maximization algorithm**.**

There are two type of clustering hard clustering and soft clustering and every object similar to each cluster to a certain degree. For example a animals having different categories but we can divide into a fixed groups or cluster

III.   METHODOLOGY FOR ABC-I APPROACH FOR IDENTIFY COMPONENT

Ant-based clustering is proposed to motivate by the groups observed in real ant colonies. This method is basic rules straightforward. Ants are designed by actually agents that randomly move in their environment, a grid with limitation. UseCase are scattered within this environment. Then Ants can be picked up UseCase according to assign complexity, travel and dropped by the Ants. The picking and dropping method are based on similarity of UseCase within the ants neighbor. Ants are picking up usecase that are surrounded by similar ones. In this approach, a group of the usecase on the grid is obtained. the ant select the path for finding a food is shown in figure 2.
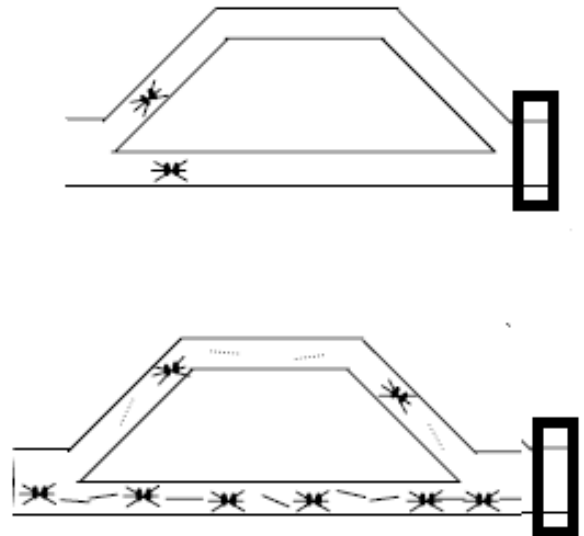

Figure no 2 Ants movement process

3.1 Flow chart for identify components

The process of finding the component of given set of usecases is define with the help of flow chart as per below in figure ().

Where N= UseCase

C= objects in one of the clusters

  K= Number of components

By using the example of figure no 5 system   First Usecase of the system is alloted to component number 2 ,Usecase is assign to group number 1 etc. . To build a result, Ant uses the pheromone value to assign each Usecase of software system best group l.

When we start the pheromone table, τ is initialized to some random value,

Variable τ_0.The pheromone value, τ_{ij} at position (i , j) shown in table 2.

For the problem of separating *N* into clusters the pheromone table is of size of multiplication of N and C_i. Thus, each row is associated with C_i. Pheromone concentrations. The pheromone table evolves as per iterate. At any iteration, every one of the ants will develop such result using the steps of pheromone-mediated relation with a view to obtain a near-optimal partition of the given *N* test into C_i groups satisfying the goal. After generating a population of *R* trial solutions, a local search is performed to find the objective function (fitness value) of these solutions.

The pheromone table is then updated depending on the quality of solutions produced by the a Ants and after the modified pheromone table, the Ants

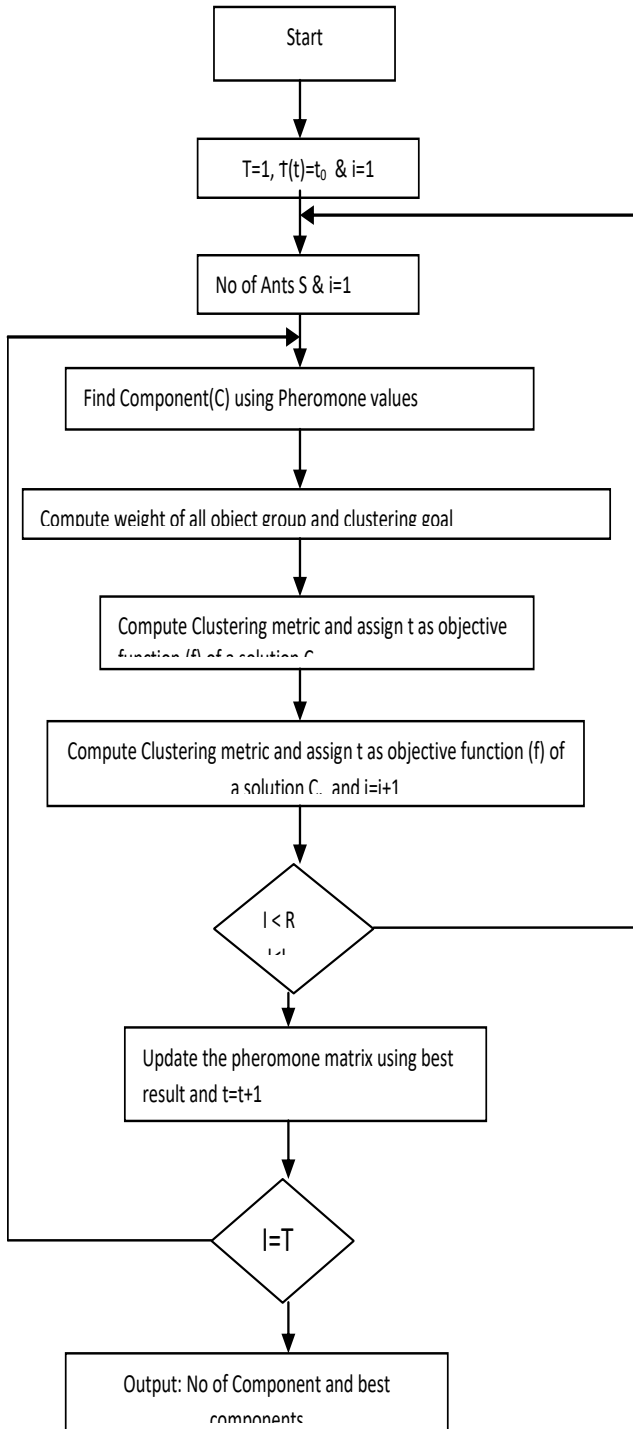The flow chart is shown in below.

Figure 4: flow chart for identify component

3.2 Brief procedure for identifying components

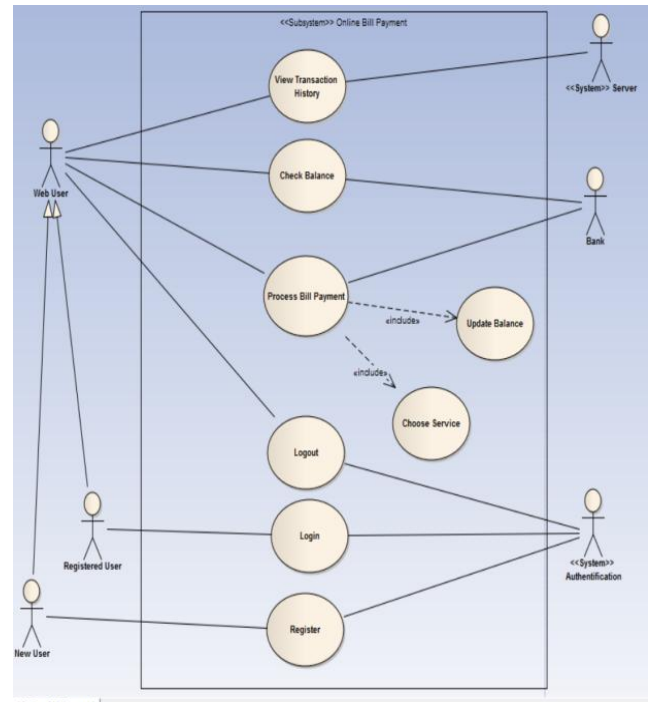Let us assume a one Usecase diagram for given system as given below

Figure 5: Usecase Diagram

In the figure 5 The Total UseCases are in table is N = 8 test object defined by n = 4 attributes (Complexity ,coupling, actor relation and cohesion) as shown in Below. The usecase are to be clustered into three (K = 3) and number of ants A = 10.

Now first we first we calculate different attributes as per given below into table 1

Table 1: Usecase and its different complexity

Usecase and its complexity				
	1	2	3	4
UC1	3.4344	3.559893	4.846065	2.740776
UC2	1.576434	4.171606	3.093204	4.166356
UC3	3.556108	3.189764	3.987396	1.176885
UC4	4.827876	4.316019	4.230079	2.247841
UC5	2.225135	4.269753	2.001373	1.825343
UC6	3.450514	4.821528	3.358715	1.630024
UC7	1.664327	1.522111	2.54973	2.217689
UC8	3.303293	4.75219	1.364269	1.422864

Now iteration, *t* with a view the steps Ants create solutions by applying data given by the pheromone table updated at the end of iteration, *t* − 1.

We created the pheromone table

Table 2: Random generation of Pheromones value.

N	1	2	3
	Pheromone table created by ABC approach		
1	0.521165	0.834193	0.690603
2	0.990478	0.217414	0.455092
3	0.591327	0.320109	0.166082
4	0.696646	0.902829	0.380505
5	0.689657	0.943419	0.915403
6	0.469222	0.889981	0.405896
7	0.948943	0.271523	0.276254
8	0.361614	0.255501	0.277841

The pheromone values for the Usecase as shown in Table 2 are: $\tau_{11} = 0.014756$, $\tau_{12} = 0.015274$ and $\tau_{13} = 0.08800$.

It represent that at the present loop, Usecase number 1 has greater probability than others belonging to cluster no. 2 because τ_{12} value is maximum.

For creating a result for component C, the Ant opt the cluster for each Usecases of component with the help of probability, component having the maximum pheromone values is opted. Repeated steps we can find the all possible groups by using each ants .the result after apply process as per given below table.

Table 3: Random generation of Pheromones value

Use Case	UC1	UC2	UC3	UC4	UC5	UC6	UC7	UC8
A1	2	1	3	2	2	3	2	1
A2	1	1	2	2	2	3	3	1
A3	1	1	1	2	2	3	2	1
A4	2	1	1	2	3	3	2	1
A5	2	2	1	2	2	3	2	1
A6	2	1	1	2	2	3	3	1
A7	2	1	1	2	2	3	2	1
A8	2	3	1	2	2	3	2	3
A9	2	1	1	2	1	3	2	1
A10	1	1	2	2	2	3	1	1

To define how the model works, let us developing first result phase, C_1 as below in Table 4. By using random function generation, we generate different numbers from uniform between 0 and 1. The numbers created are equal to the length of solution component. Now we created random numbers like
(0.593241, 0.691452, 0.886142, 0.888432, 0.343672, 0.867721, 0.0814324, 0.248767). And Usecase 1, 2, 3,4,5,6, 7 and 8 are allotted its appropriate clusters.

measured in terms of fitness function. This fitness value is explain as the sum of squared Euclidean distances between each Usecases and the center of belonging groups or components.

Let be a given Usecase of U_i where i=8, and complexity $\{x1, x2, . . . , xN\}$ in partitioned into a number, say K, of clusters or groups. The mathematical equation can be described as

Table 4: Random generation of Pheromones value

N	C1	C2	C3
	For first A1		
	Clustring		
1	0	1	0
2	1	0	0
3	0	0	1
4	0	1	0
5	0	1	0
6	0	0	1
7	0	1	0
8	1	0	0

$C_m = y*A_{iv}$

Table 5: Random generation of Pheromones value

K	n			
	1	2	3	4
1	4.95	3.2	1.45	0.2
2	4.825	3.4	1.42	0.22
3	5.05	3.55	1.5	0.3

Now we find the fitness values of each assign cluster by each ant by below procedure steps

Step1: calculate the values of w_{ij}. The value of w_{ji} is either 0 or 1. it is assigned one if it depend on Usecase (i) belong to the to the cluster (j) otherwise is assigned to zero. After apply the procedure we develop the below table show in table number X. the value if i=1 to N and j=1 to K.

Table 6: Random generation of Pheromones value

N	C1	C2	C3
	For first A1		
	Clustring		
1	0	1	0
2	1	0	0
3	0	0	1
4	0	1	0
5	0	1	0
6	0	0	1
7	0	1	0
8	1	0	0

Step2: After calculate the values if w_{ij}. We find the center of each cluster (c_{ij}). The value of center of cluster is equal to sum of $c_{ij} = \sum(w_{ij} \times A_{ij})/\sum w_{ij}$ where i=1 to N. After that we find the values by using table W and A.

Table 7: Random generation of Pheromones value

K	n			
	1	2	3	4
1	4.95	3.2	1.45	0.2
2	4.825	3.4	1.42	0.22
3	5.05	3.55	1.5	0.3

Step 3 : Now by using step 2 and step 3 we substitute the weight and cluster center table in below equation

$$ff = \sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} w_{ji}(A_{jl} - CM_{il})^2$$

and find the fitness function of each ants for assigned the allotted the Usecases. Many of the ACO algorithms work on as a method of local find with a improve optimal solution.

Table 8: fitness values of each Ant

Ant	Objective function(ff)
A1	2.44
A2	2.23
A3	1.51
A4	2.22
A5	1.41
A6	2.78
A7	2.98
A8	1.97
A9	2.03
A10	2.11

Show the best optimal groups for creating software is and its fitness is .which is maximum is A7 in table 8 . show the components of are C1= (UC2, UC3,UC8) , C2=(UC1,UC4,UC5,UC7) and C3= (U6)

IV. COMPARISON ALGORITHMS SOLUTIONS

We applied the Ant based clustering approach for identified the group .Here eight UseCase are used for a particular application. Every steps are implemented in PyCharm and used i5 intel system for experiments.Usecase are described and created a group as per best fitness and UseCase complexity is created using a random number.

S.NO	Software System	No of UseCase	No of Component	Method	SC	SCh	Scom	Fitness Value
1	OBS	30	6	GA	0.969	0.159	0.0779	0.7291
				ABC	0.827	0.059	0.0709	0.8291
2	RAS	81	8	GA	0.907	0.138	0.153	0.7537
				ABC	0.897	0.121	0.103	0.8037
3	AS	87	11	GA	0.867	0.111	0.0599	0.6961
				ABC	0.911	0.211	0.0621	0.6991
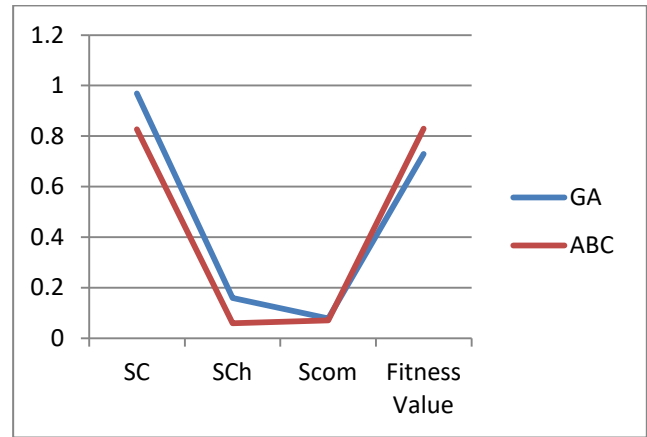
V. CONCLUSIONS

In conclusion, an ant ABC algorithm is design to find the component identification .The ants use pheromone table a identify the component, which help different ants for the best component identification finding result. The pheromone (weight) assigned at place (i.e. placed Usecase i to the group j for a build result) as per its fitness function which is less result among others, and the evaporation rate. The evaporation value is helps to see other groups position of Usecase.

REFERENCES

1. R. J. G. B. Campello, E. R. Hruschka, "A Fuzzy Extension of the Silhouette Width Criterion for Cluster Analysis" Fuzzy Sets and Systems, vol. 157, n. 21, pp. 2858-2875, 2006
2. J.F. JIMENEZa, F.J. CUEVASb, J.M. CARPIO "Genetic Algorithms applied to Clustering Problem and Data Mining " Proceedings of the 7th WSEAS International Conference on Simulation, Modelling and Optimization, Beijing, China, September 15-17, 2007
3. N Md Jubair Basha1 and Dr Chandra Mohan2 "A strategy to identify component using clustering approach for component reusability"Natarajan Meghanathan, et al. (Eds): SIPM, FCST, ITCA, WSE, ACSIT, CS & IT 06, pp. 397–406, 2012. © CS & IT-CSCP 2012 ,
4. Alex A. , "Computing Laboratory, A Review of Evolutionary Algorithms for Data Mining" Freitas University of Kent, UK
5. Mohammad Babrdel Bonab "Modified K--Means Algorithm for Genetic Clustering "IJCSNS International Journal of Computer Science and Network Security, VOL.11 No.9, September 2011
6. Mark Harman and Yuanyuan Zhang" Search Based Software Engineering: Trends, Techniques and Applications"
7. ACM, Inc., 2 Penn Plaza, Suite 701, 2007 New York,
8. Gholam Reza Shahmohammadia Saeed Jalilia Seyed Mohammad Hossein Hasheminejada "Identification of System Software Components Using Clustering Approach" JOURNAL OF OBJECT TECHNOLOGY Published by ETH Zurich, Chair of Software Engineering © Jot, 2010
9. Wiem Mkaouer, Marouane Kessentini" High Dimensional Search-based Software Engineering: Finding Tradeoffs among 15 Objectives for Automating Software Refactoring using NSGA-III "Conference'10, Month 1–2, 2010, City, State, Country. Copyright 2010 ACM 1-58113-000-0/00/0010
10. Mahdie khadame, Sima Emadi "Measurement of Software Reliability Based on Coupling and Cohesion Rate by External and Internal View of Classes "IJCSNS International Journal of Computer Science and Network Security, VOL.16 No.11, November 2016
11. L E Agustin Blas " A New Gropuing Genetic Algorithm for Clustering problem " Expert System with application 38(2012) 9685-5703
12. Eleni Constantinou "Extracting resuable components : A Semi Automated Approach for complex structure " information processing letter 115 (2015)