Implementation of Improved IR System through Swarm Intelligence Technique

Ms. Shraddha G. Lohar,Ms. Priyanka J. Howale, Ms. Sanketa S. Pradhan, Mr. Mehul D. Redekar
Asst. Prof. Anagha N. Chaudhari
Department of Information Technology
Pimpri Chinchwad College Of Engineering, Pune,India
shraddhalohar237@gmail.com,howalepriyanka28@gmail.com,spradhan175@gmail.com,mehulredekar@gmail.com,
anagha.chaudhari@gmail.com

Abstract—Information retrieval is concept related to various optimization techniques and clustering algorithms.High dimensional Textual dataset is required for information retrieval .swarm intelligence is one of the most important technique under evolutionary algorithms.Evolutionary algorithm and Swarm Intelligence Technique are used for the optimization.Swarm intelligence technique includesalgorithm for optimization.Comparative study of thos gives best algorithm on the basis of there computational time and efficiency.This best algorithm is using for implementation of information retrieval system.

Keywords — *High Dimensional Textual Datasets, Swarm Intelligance Tehnique, Clustering Optimization, Computational Time, Efficiency.*
_____*****_____

I.   INTRODUCTION

For the complex data sets there is a problem in retrieval the necessary information from particular records. As the original datasets are multidimensional in nature, so for retrieving the particular information, datasets need to be multidimensionality reduced.

Hence,for these there are different optimization techniques or algorithms , and with the help of those algorithms the datasets are first reduced and then that datasets are provided as an input to the algorithms i.e Particle Swarm Optimization, Ant Colony Optimization and Then clusters are obtain for information retrieval system.

➢   *Swam Intelligence(SI)*

Swarm Intelligence algorithms are -

➢   *Particle Swarm Optimization(PSO)*

➢   *Ant Colony Optimization(ACO)*

Using all these algorithms and with the help of comparison between these algorithms there is a retrieval of information from the particular data sets, and as well as development of IR system also takes place.

II.   MODULES

Module 1(Study and Preparation of Dataset):
Input Dataset is required for optimization of IR using Evolutionary Algorithm. Here Instead
of applying already available dataset for System ,we are preparing new dataset in this module.

Module 2(Comparison of Swarm Intelligence Technique algorithms) :

Comparison of Swarm intelligence techniques algorithms based on computation time, reduction of dataset using optimization.
Algorithms:
•   Particle Swarm Optimization
•   Ant Colony optimization

Module 3(Information Retrieval) :

Clusters obtain for the Best algorithm are used for information retrieval from system .There are many applications of IR so any kind of information may retrieve from the system .information retrieval is based on the input given to that System(Input Dataset).
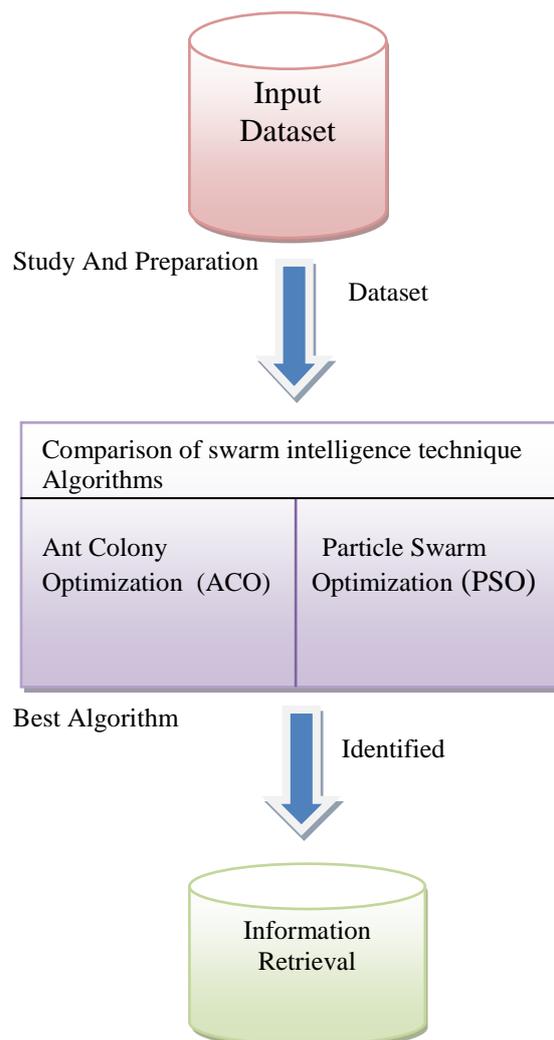
III.   SYSTEM ARCHITECTURE

Fig.1: System architecture

IV. ALGORITHMS

A. Particle Swarm Optimization(PSO):

Particle Swarm Optimization (PSO) is a global optimization algorithm for dealing with problems in which a best solution can be represented as a point or surface in an n-dimensional space. Hypotheses are plotted in this space and seeded with an initial velocity, as well as a communication channel between the particles. Particles then move through the solution space, and are evaluated according to some fitness criterion after each timestep. Over time, particles are accelerated towards those particles within their communication grouping which have better fitness values. The main advantage of such an approach over other global minimization strategies such as simulated annealing is that the large number of members that make up the particle swarm make the technique impressively resilient to the problem of local minima[1].
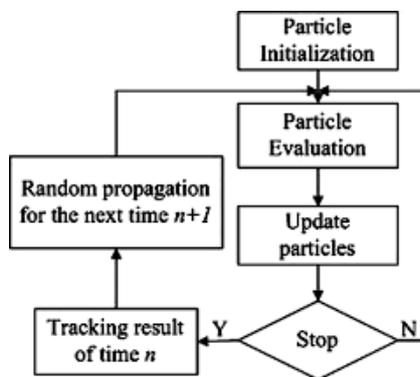

Fig.2: Particle Swarm Optimization

Algorithm: Pseudo code for PSO:
1. Initialize all particles.
2. Calculate fitness value of each particle
3. Find best fitness value out of that all particles
4. If the fitness value is better than the best fitnessvalue (pBest),then Set pBest = currentfitness value
5. If pBest is better than gBest,
Set gBest = pBest
6. Calculate particle VelocityUse gBest and Velocity
7. Update particle Data

B Ant Colony Optimization(ACO):

Ant colony optimization (ACO), introduced by Dorigo in his doctoral dissertation, is a class of optimizationalgorithms modeled on the actions of an ant colony. ACO is a probabilistic technique useful in problems that deal with finding better paths through graphs. Artificial 'ants' simulation agents locate optimal solutions by moving through a parameter space representing all possible solutions. Natural ants lay down pheromones directing each other to resources while exploring their environment. The simulated 'ants' similarly record their positions and the quality of their solutions, so that in later simulation iterations more ants locate better solutions[1].
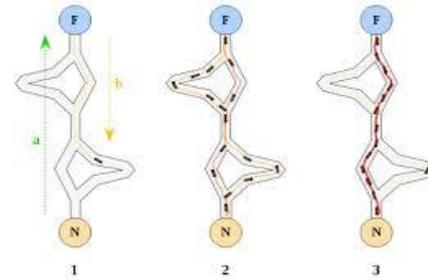

Fig.3: Ant Colony System

Scheme:
1. Construct ant solutions
2. Define attractiveness $\tau 0$, based on experience from solution
3. Define specific visibility function, η, for a give problem (e.g. distance)[9]

Ant walk
1. Initialize ants and nodes (states)
2. Choose next edge probabilistically according to attractiveness and visibility
3. Each ant maintains a taboo list of infeasible transitions for that iteration
4. Update attractiveness of an edge according to the number of ants that pass through
5. Pheromone update Parameter is called evaporation rate
6. Pheromones = long-term memory of an ant colony
ρ small low evaporation slow adaptation
ρ large → high evaporation fast adaptation
"new pheromone" or $\Delta\tau$ usually contains the baseattractiveness constant Q and a factor that you wantto optimize [9]

Algorithm: General Ant Colony Pseudo Code
Initialize all edges to (small) initial pheromone level $\tau 0$;
Place each ant on a randomly chosen city;
For each iteration do:
Do while each ant has not completed its tour:
For each ant do:
Move ant to next city by the probability function
End;
End;
For each ant with a complete our do:
Evaporate pheromones;
Apply pheromone update;
If(ant k's tour is shorter than the global solution)
  Update global solution to ant k's tour
End;
End;

V.   EXPERIMENTAL RESULTS

1.Study and Preparation of Dataset

Textual dataset is necessary for the comparison of Algorithms.following dataset is sound track dataset and it contains list of 200 songtracks**.**

```
SOUNDTRACKS
===========
# "♯JustDating" (2014) {Pilot (#1.1)}
- "Wait Up (Boots of Danger)"
  Performed by 'Tokyo Police Club' (qv)

# "♯LawstinWoods" (2013) {The Incident (#1.6)}
- "This Is the End (I Don't Want You Back)"
  By Bradley Voorhees

# "♯LawstinWoods" (2013) {The Loop & Rocks (#1.9)} {{SUSPENDED}}
- "Unable to Let Go"
  by Bradley Voorhees Featuring MJ Rhythm

# "♯LawstinWoods" (2013) {The Neverending Day (#1.13)}
- "You Forgot My Birthday"
  By Bradley Voorhees featuring Joshua P. Kennedy
```

Fig.4: Dataset

2. Ant Colony Optimization(ACO)

 Ant Colony Optimization algorithm is executed successfully .Soundtrack dataset is given as input to this algorithm.output of this algorithm is as follows (fig.5):

Output shows time required to travel dataset and path of the dataset,which follow by each ant. suppose we initialize 5 ants then for each ant, algorithm calculate best length of path and required time. out of that 5 paths ,best trail is identified.

Then weight matrix for dataset is shown in (fig.5).after that length of best trail is found and time required for best trail.

```
Output
Initializing pheromones on trails
Entering UpdateAnts - UpdatePheromones loop
New best length of 1104.0 found at time 0
New best length of 1086.0 found at time 1
New best length of 1077.0 found at time 13
New best length of 1063.0 found at time 15
Time complete
Best trail found:136 105 165 198 101 191 116 183 20 9 86 193 235 226 148 238 67 35 164
190 119
253 199 239 185 120 167 66 175 200 217 33 215 133 114 3 27 187 24 7 94
48 237 32 229 151 195 146 157 100 47 37 127 222 68 139 142 181 8 242 19
49 4 81 54 56 234 93 103 144 210 134 12 156 90 73 17 169 11 39 188
22 168 76 150 184 163 255 72 243 96 104 182 46 249 57 45 154 241 177 179
102 52 18 153 171 28 232 126 257 123 0 176 233 59 1 135 201 197 122 142
244 29 118 58 106 218 224 170 43 31 115 194 88 110 82 228 145 174 246 152
202 117 74 107 87 196 212 252 248 166 98 221 97 26 206 71 180 225 50 213
60 62 80 227 6 70 79 121 230 30 251 15 208 223 211 109 205 13 209 159
10 236 160 113 149 137 65 55 155 5 44 173 112 34 91 128 40 95 99 23
92 83 25 219 189 89 16 141 75 84 132 178 21 138 216 85 143 245 61 125
258 38 220 240 250 64 147 186 130 77 69 111 172 204 63 203 214 247 192 51
129 161 256 53 124 108 140 254 41 2 231 36 42 207 14 78 158 131

Length of best trail found: 1063.0
End Ant Colony Optimization demo
```

Fig.5: Ant Colony Optimization Output

3.Particle Swarm Optimization(PSO)

Particle Swarm Optimization is executed Successfully. Input to the algorithm is Dataset and output is various functions of particle swarm optimization. out of this five functions ,function which give best result that only used for further implementation of algorithm.

  1.Hole Function
  2.Peak function
  3.Sphere Function
  4.Step Function

1.Hole Function

In hole function, PSO considered or arrange all the particle in X and Y coordinates axis according to their target (gbest) values. each particle which arrange in x and y coordinate will moved towards the target value. it covered in 31 iteration according to their no of particle. In above yellow circle denotes the no of particle which moved towards the red target value according to their x and y coordinates.
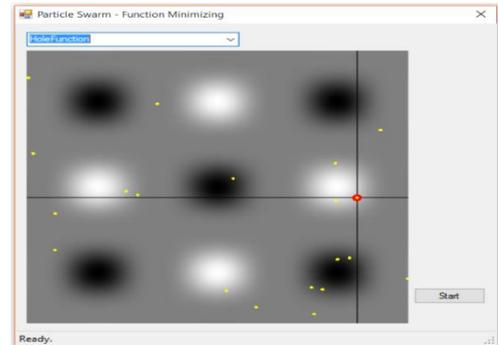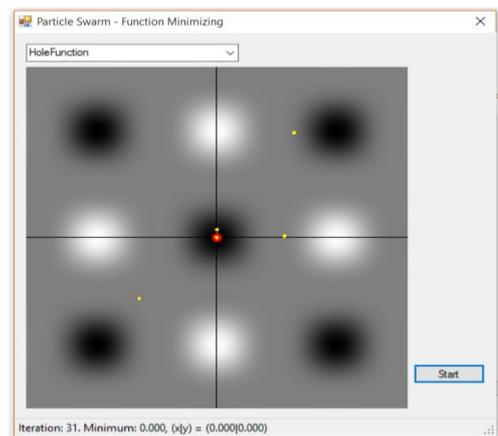

Fig.6.1 : Hole Function

Fig.6.2 : Hole Function

2.Peak function

Peak Function Centre will stand out as a leader in function venue design.

Peak function of particle will collect the other particles behaviour. Target value set center of x and y coordinates. All the particles moved toward their according to their behaviour and velocity of particle
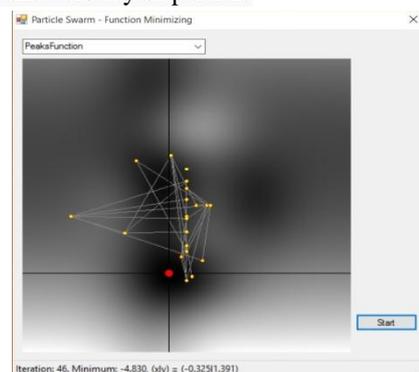

Fig.7 : Peak Function

3.Sphere Function

Sphere function of particle located all the function are in centered to target value.

Sphere function covered the global optimum. It take to covered global optimum in 11 iteration. This function is simple and strongly convex function.Sphere function is well known benchmark for testing local search ability of optimization algorithm
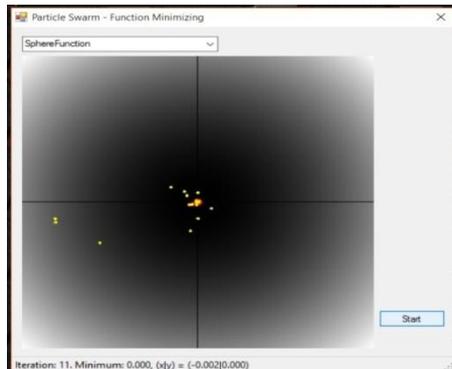

Fig.8: Step Function

4.Step Function

In step function particle, particle are in gbest value depends on experience of each other.

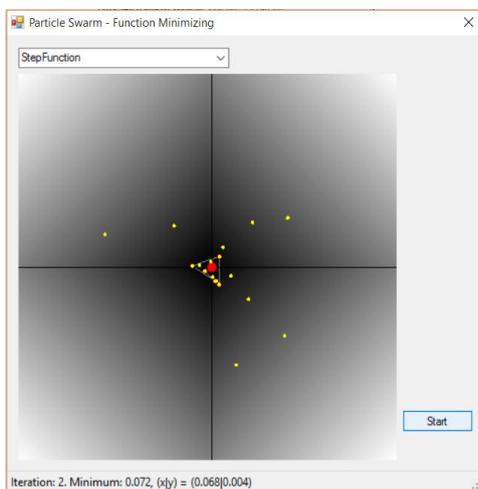It covered to complete in iteration 2 according to no of particle.

Fig.9: Step Function

V.   CONCLUSION

This paper helps us to retrieve relevant data with the help of optimization using evolutionary   algorithm techniques. i.e.Swarm Intelligence Technique. Algorithms applied on input datasets to obtain output in the form of clusters. Clusters are formed from multidimensional datasets where complexities of datasets are reduced using algorithms present under Swarm Intelligence. Comparison between output of these two algorithms(PSO,ACO) in the form of clusters gives best and efficient algorithm for optimization

of Information Retrieval as per user query.paper includes implementation of PSO and ACO algorithms.The future scope for this research would be optimization of information retrieval system, hence forming an improved IR system on high dimensional textual datasets.

VI.   REFERENCES

[1]   Sandeep U. Mane Assistance Professor, Dept. of CSE, Pankaj G. Gaikwad.M. Tech Student, Dept. of CSE, "Nature Inspired Techniques for Data Clustering"2014

[2]   C. Ahn, J. An, J. Yoo,"Estimation of particle swarm distribution algorithms: combining the benefits of PSO and EDAs",2012

[3]   D. Karaboga, B. Akay," A modified artificial bee colony (ABC) algorithm for constrained optimization problems",2011.

[4]   Ant Colony Optimization by Marco Dorigo and Thomas Stützle, MIT Press, 2004. ISBN 0-262-04219-3Particle Swarm Optimization by Maurice Clerc, ISTE, ISBN 1-905209-04-5, 2006

[5]   Karaboga, Dervis . "Artificial bee colony algorithm". Scholarpedia ,2010.

[6]   A.E. Rizzoli, R. Montemanni, E. Lucibello, L.M. Gambardella,"Ant colony optimization for real-world vehicle routing problems",2007

[7]   X. Zhang, L. Tang,"A new hybrid ant colony optimization algorithm for the vehicle routing problem",2009

[8]   S. Mazzeo, I. Loiseau,"An ant colony Algorithmfor the capacitated vehicle routing problem",2004

[9]   "http://www.math.ucla.edu/~wittman/10c.1.s /Lectures/Raids/ACO.pdf"