A Review on Numerical function optimization Algorithm and its Applications to Data Clustering & Classification

Neeraj Dahiya¹, Surjeet Dalal², Savita Khatri³

¹, ²Dept. Computer Science and Engineering, SRM University, India
³Dept. of Computer Science and Engineering, BMIET, India

Abstract—Natural phenomenon's and swarms behaviour are the warm area of research among the researchers. A large number of algorithms have been developed on the account of natural phenomenon's and swarms behaviour. These algorithms have been implemented on the various computational problems for the sake of solutions and provided significant results than conventional methods but there is no such algorithm which can be applied for all of the computational problems. In 2009, a new algorithm was developed on the behalf of numerical function for continuous optimization problems. In short span of time, this algorithm gain popularity among researchers and has been applied to large number of problems such as clustering, classification, parameter identification etc. This paper presents the compendious survey on the numerical function optimization algorithm and its applications as well as enlightens the applicability in data clustering & classification.

Index Terms—Classification, Clustering, Numerical function, Optimization, Nature Inspired Algorithm

I. INTRODUCTION

Nature has always been a continuous source of inspiration for researchers and scientists. A large number of algorithms have been developed based on the natural process of evolution, laws, swarms behavior etc. Nature inspired algorithms are the latest state of art algorithms & works well with optimization problems as well as other problems than the classical methods because classical methods are inflexible in nature. It has been proved by many researchers that nature inspired algorithms are convenient to solve complex computational problems such as to optimize objective functions [1, 2], pattern recognition [3, 4], control functions [5, 6], image processing [7, 8], filter modeling [9, 10], clustering [3], classification [11] etc. In last one and half decade several nature inspired algorithms have been developed such as Particle swarm optimization (PSO), Genetic Algorithm (GA), Simulated Annealing (SA), Ant colony optimization (ACO), Artificial Bee colony (ABC) optimization, Big Bang Big Crunch (BB-BC) etc. These algorithms show better results than classical methods in terms of accuracy, convergence computational time etc. Kirkpatrick et al. [12] proposed simulated annealing algorithm based on the annealing process of metals and applied SA to solve many combinatorial optimization problems but SA suffered from convergence problem and trapped in local minima [13, 14]. In 1995, Kennedy et al. [15] developed an algorithm based on the swarm behavior of birds & fish schooling and named it Particle Swarm Optimization (PSO). This algorithm was applied to solve many optimizations such as function optimization [16] but this algorithm suffered with partial optimism and not works with scattering problems. In case of complex and complicated problems PSO algorithm suffered with optimal solution [17]. In 1975, John Holland [18] developed an algorithm based on the natural process of evolution and called it genetic algorithm. Many optimization problems solved by with the help of genetic algorithm but it's suffered from parameter tuning problem due to complicated crossover & mutation operator and convergence speed [19]. ACO was developed by the Dorigo [20] in 1992 based on the behaviors of real world ants. This algorithm applied to solve many optimization problems such as TSP [21], Graph Coloring [22], Network Routing [23] etc. ACO algorithm is sometimes suffered from selection bias [24] and convergence [25, 26] problem. Devis Karbooga et al. [27] in 2006 proposed ABC algorithm based on the behavior of honey Bees i.e. capabilities of Bees to find the food source for numerical optimization problems. Later on ABC algorithm has widely used to solve other problem such as multi-dimensional numeric problems [28], real-parameter optimization [29, 30], constrained numerical optimization problems [31] etc. Dervis Karaboga & Beyza Gorkemli et al. [8] provided comprehensive description about the ABC algorithm and its application in various fields as well as hybridization of ABC & its variants. There are some drawbacks of ABC algorithm such as it requires large number of objective function evaluations [32], handling complex multimodal functions, functions with narrow curve valley, convergence slow down due to stochastic nature [30] and weak exploitation feature [33]. Osman K. Erol et al. [34] in 2006 proposed a new algorithm Big Bang Big Crunch (BB-BC) based on the big bang theory (Theory of Evolution of Universe) for optimization problems. BB-BC algorithm is used to solve large numbers of optimization problems such as multi modal optimization problem [35], multi-objective optimization problem [36],
clustering [37] etc. A brief description about the nature inspired algorithms has been given above that are used to solve the optimization problems but till date there does not exist any algorithm that can solve the entire optimization problems exactly.

II. THE CLUSTERING PROBLEM

Clustering is the process of recognizing natural groupings or clusters in multidimensional data based on some similarity measures[6]. Distance measurement is generally used for evaluating similarities between patterns. In particular the problem is stated as follows: given N objects, allocate each object to one of K clusters and minimize the sum of squared Euclidean distances between each object and the center of the cluster belonging to every such allocated object. The clustering process, separating the objects into the groups (classes), is realized by unsupervised or supervised learning. In unsupervised clustering which can also be named automatic clustering, the training data does not need to specify the number of classes. However, in supervised clustering the training data does have to specify what to be learned; the number of classes. The dataset that we tackled contains the information of classes. Therefore, the optimization goal is to find the centers of the clusters by minimizing the objective function, the sum of distances of the patterns to their centers.

The Original ABC Algorithm

The artificial bee colony algorithm is a new population-based metaheuristic approach, initially proposed by Karaboga _7_ and Karaboga and Basturk _8_ and further developed by Karaboga and Basturk _13_ and Karaboga and Akay _14_. It has been used in various complex problems. The algorithm simulates the intelligent foraging behavior of honey bee swarms. The algorithm is very simple and robust. In the ABC algorithm, the colony of artificial bees is classified into three categories: employed bees, onlookers, and scouts. Employed bees are associated with a particular food source that they are currently exploiting or are “employed” at. They carry with them information about this particular source and share the information onto onlookers. Onlooker bees are those bees that are waiting on the dance area in the hive for the information to be shared by the employed bees about their food sources and then make a decision to choose a food source. A bee carrying out random search is called a scout. In the ABC algorithm, the first half of the colony consists of the employed artificial bees, and the second half includes the onlookers. For every food source, there is only one employed bee. In other words, the number of employed bees is equal to the number of food sources around the hive. The employed bee whose food source has been exhausted by the bees becomes a scout. The position of a food source represents a possible solution to the optimization problem, and the nectar amount of a food source corresponds to the quality _fitness_ of the associated solution represented by that food source. Onlookers are placed on the food sources by using probability-based selection process. As the nectar amount of a food source increases, the probability value with which the food source is preferred by onlookers increases, too _7, 8_. The main steps of the algorithm are given in Algorithm 1. In the initialization phase, the ABC algorithm generates randomly distributed initial food source positions of SN solutions, where SN denotes the size of employed bees onlooker bees. Each solution xi _i _ 1, 2, ..., SN_ is a n-dimensional vector. Here, n is the number of optimization parameters. And then evaluate each nectar amount fiti. In the ABC algorithm, nectar amount is the value of benchmark function.

Standard PSO

The PSO algorithm searches the space of the objective functions by adjusting the trajectories of individual agents, called particles, as the piecewise path formed by positional vectors in a quasi-stochastic manner [5, 6]. There are now as many as about 20 different variants of PSO. Here we only describe the simplest and yet popular standard PSO. The particle movement has two major components: a stochastic component and a deterministic component. A particle is attracted toward the position of the current global best g_ and its own best location x_i in history, while at the same time it has a tendency to move randomly. When a particle finds allocation that is better than any previously found locations, then it updates its the new current best for particle i. There is a current global best for all the particles. The aim is to find the global best among all the current best solutions until the objective no longer improves or after a certain number of iterations. For the particle movement, we use x_i to denote the current best for particle i, and g_ _ min or max{f(xi)}(i = 1, 2, ..., n) to denote the current global best. Let xi and vi be the position and velocity for particle i, respectively.

Firefly Algorithm

Now we can idealize some of the flashing characteristics of fireflies so as to develop firefly-inspired algorithms. For simplicity in describing our new Firefly Algorithm (FA), we now use the following three idealized rules: 1) all fireflies are unisex so that one firefly will be attracted to other fireflies regardless of their sex; 2) Attractiveness is proportional to their brightness, thus for any two flashing fireflies, the less brighter one will move towards the brighter one. The attractiveness is proportional to the brightness and they both decrease as their distance increases. If there is no brighter one than a particular firefly, it will move randomly; 3) The brightness of a firefly is affected or determined by the landscape of the objective function. For a maximization problem, the brightness can simply be proportional to the value of the objective function. Other forms of brightness can be defined in a similar way to the fitness function in genetic algorithms.

The Multiobjective ABC Algorithm

As opposed to single-objective optimization, MOEAs usually maintain a nondominated solutions set. Immultiobjective optimization, for the absence of
preference information, none of the solutions can be said to be better than the others. Therefore, in our algorithm, we use an external archive to keep a historical record of the nondominated vectors found along the search process. This technique is used in many MOEAs.5, 16.

Simulation Results for Benchmark Functions

This section deals with the experimental setup of our study. It includes the performance measures, parameters settings, datasets to be used, experiment results and statistical analysis. To prove the effectiveness of the MCSS algorithm, ten datasets are applied in which two datasets are artificial ones and rest of are taken from UCI repository. The proposed algorithm is implemented in Matlab 2010a using a computer with window operating, corei3 processor, 3.4 GHz and 4 GB RAM. Experimental outcomes of MCSS algorithm are compared with other clustering algorithms like K-means, GA [30], PSO [49], ACO [35] and CSS [38].

III. PERFORMANCE MEASURES

The performance of MCSS algorithm is examined over the sum of intra cluster distance and J-measure parameters. The sum of intra cluster distance can be measured in terms of best case, average case and worst case solutions including standard deviation parameter which shows the dispersion of the data. J-measure parameter is used to measure the accuracy of proposed method. Performance measures are described as follows.

Intra cluster distances

Intra cluster distance can be used to measure the quality of clustering.35-36. It indicates the distance between the data objects within a cluster and its cluster center. This parameter also highlights the quality of clustering that is, minimum is the intra cluster distance, better will be the quality of the solution. The results are measured in terms of best, average and worst solutions.

Standard Deviation (Std.)

Standard deviation gives the information about the scattering of data within a cluster.47, 49. Lower value of standard deviation indicates that the data objects are scattered near its center while high value indicates that the data is dispersed away from its center point.

F-Measure

This parameter is measured in terms of recall and precision of an information retrieval system.50-51. It is also described in terms of weighted harmonic mean of recall and precision. Recall and precision of an information retrieval system is computed using equation 8 which can be described.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Classes</th>
<th>Attributes</th>
<th>Total instances</th>
<th>Instance in each classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>ART 1</td>
<td>3</td>
<td>2</td>
<td>300</td>
<td>(100, 100, 100)</td>
</tr>
<tr>
<td>ART 2</td>
<td>3</td>
<td>3</td>
<td>300</td>
<td>(100, 100, 100)</td>
</tr>
<tr>
<td>Iris</td>
<td>3</td>
<td>4</td>
<td>150</td>
<td>(50, 50, 50)</td>
</tr>
<tr>
<td>Glass</td>
<td>6</td>
<td>9</td>
<td>214</td>
<td>(70, 17, 76, 13, 9, 29)</td>
</tr>
<tr>
<td>LD</td>
<td>2</td>
<td>6</td>
<td>345</td>
<td>(145, 200)</td>
</tr>
<tr>
<td>Thyroid</td>
<td>3</td>
<td>3</td>
<td>215</td>
<td>(150, 30, 35)</td>
</tr>
<tr>
<td>Cancer</td>
<td>2</td>
<td>9</td>
<td>683</td>
<td>(444, 239)</td>
</tr>
<tr>
<td>CMC</td>
<td>3</td>
<td>9</td>
<td>1473</td>
<td>(629, 334, 510)</td>
</tr>
<tr>
<td>Vowel</td>
<td>6</td>
<td>3</td>
<td>871</td>
<td>(72, 89, 172, 151, 207, 100)</td>
</tr>
<tr>
<td>Wine</td>
<td>3</td>
<td>13</td>
<td>178</td>
<td>(59, 71, 48)</td>
</tr>
</tbody>
</table>

Clustering Algorithms using ART1 Dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Parameters</th>
<th>K-means</th>
<th>GA</th>
<th>PSO</th>
<th>ACO</th>
<th>CSS</th>
<th>MCSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best</td>
<td>157.12</td>
<td>154.46</td>
<td>154.06</td>
<td>153.37</td>
<td>153.01</td>
<td>153.18</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>161.12</td>
<td>158.87</td>
<td>158.24</td>
<td>158.52</td>
<td>158.29</td>
<td>158.02</td>
<td></td>
</tr>
<tr>
<td>Worst</td>
<td>166.08</td>
<td>164.08</td>
<td>161.83</td>
<td>162.52</td>
<td>161.32</td>
<td>159.26</td>
<td></td>
</tr>
<tr>
<td>Std</td>
<td>0.34</td>
<td>0.381</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>F-Measure</td>
<td>99.14</td>
<td>99.78</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

IV. CONCLUSION

In this paper, magnetic charged system search algorithm is applied to solve the clustering problems. The idea of proposed algorithm came from the electromagnetic theory and based on the behavior of moving charged particles. A moving charged particle exerts both the forces (electric force and magnetic force) on other charged particles and in turn altered the positions of charged particles. Therefore, in MCSS algorithm, initial population is presented in the form of charged particles. It utilizes the concept of electric and magnetic forces along with newton second law of motion to obtain the updated positions of charged particles. In MCSS, both the electric force (Ek) and magnetic force (Mk) correspond to the local search for the solution while the global solution is exploited using newton second law of motion. The aim of this research is to investigate the applicability of MCSS algorithm for clustering problems. To achieve the same, performance of the MCSS algorithm is evaluated on variety of datasets and compared with K-Means, GA, PSO, ACO and CSS using intra cluster distance, standard deviation and J-measure parameters. Experiment results support the applicability of proposed algorithm in clustering field as well as the proposed method provides good results with most of datasets in comparison to the other methods. For deep insight into the proposed method, a statistical analysis is also performed along with the experimental analysis. Statistical analysis also supports the validation of the proposed MCSS method for solving the clustering problem. Finally, it can be
concluded that proposed method not only gives good results but also improves the quality of solutions.

REFERENCES


Silva A, Neves A, Costa E, J. a s in g the swarm: a p redator p ray app roach to function op timization, In: Proceed in g of the international conference on soft computing MENDEL, 2002.

© 2016 IJRRAS All Rights Reserved
