A Comprehensive Survey on Swarm Optimization Algorithms for Data Mining

1T. Keerthika, 2Ramya Sankari S, 3Kavinn Krishnan A K, 4Karthick Sekar G

1Assistant Professor
Department of Information Technology
Sir Krishna College of Engineering and Technology,
Coimbatore.

234 I BTech IT
Department of Information Technology
Sir Krishna College of Engineering and Technology,
Coimbatore.

Abstract

In Artificial Intelligence research, swarm intelligence is comparatively a novel domain. Due to the wide variety of engineering problem applications, it has pinched more awareness by the research community. Swarm intelligence is rely upon social behavior that can be down to earth in nature, for example, subterranean insect provinces, herds of flying creatures, fish schools and colonies, where a measure of people with confined abilities for perplexing issues can accompany with the intellectual solutions. This paper displays an exhaustive survey of various Swarm intelligence algorithms. The essential purpose of these algorithms is to utilize many effortless agents applying roughly no rule which in revolve leads to an embryonic global deeds. This paper presents appraise of swarm intelligence algorithms and their application for problem solving techniques.

Keywords: Swarm intelligence, PSO, ACO, ABC, FSO
1. Introduction

The Swarm Intelligence (SI) articulation was presented by Gerardo Beni and Jing Wang in 1989, cellular robotic systems contents [1]. Swarm intelligence can be expressed as common deeds came out from communal insects functioning under very few policies. Single agent entity do not posses much capability by themselves. Therefore, in a colony of agents, task have been accomplished by each individual agents. Through mutual aid, colony has a potential of solving intricate problems. This communal activity which comes out from agents turn as swarm intelligence. Swarms are probable to retort to eminence factors such as food, safety etc. distant from the regular computations. Communication among agents with the partial restrictions plays a major subject in self organization [2]. Effortless rules are followed by their agents and still there is no essential arrangement available about how agent should act as an individual. Performance optimization and robustness plays a major goal in SI systems.

2. Swarm Intelligence

Swarm intelligence is the analysis of computational frameworks roused by the ‘Collective knowledge’. Collective intelligence rises through the participation of expansive quantities of homogeneous agents in the environment. Illustrations incorporate schools of fish, groups of winged creatures, and settlements of ants. Such knowledge is decentralized, self-sorting out and circulated all through a domain. In nature such frameworks are normally used to take care of issues, for example, efficient food foraging, prey dodging, or colony repositioning [2]. The data is normally put away all through the taking an interest homogeneous agents, or is put away or imparted in the environment itself, for example, using pheromones in ants, moving in honey bees, and nearness in fish and feathered creatures. In the vein of evolutionary computation, versatile systems are measured in terms of ‘algorithms' or 'strategies' and are regularly applied to search and optimization areas [7].

A. Artificial Bee Colony Algorithm

The Artificial Bee Colony (ABC) algorithm was projected by Dervis Karaboga in the year 2005[8]. It natures inspire optimisation technique. Food foraging is being express as real honey bees. It finds near optimal solution to the difficult optimisation problems [3].

Main components of ABC
- Food sources
- Employed foragers
- Unemployed foragers

The Food sources
Keeping in mind the end goal to choose a food source, a forager honey bee assesses a few properties related with the food source, for example, its closeness to the hive, wealth of the vitality, taste of its nectar, and the simplicity or trouble of separating this vitality. For the straightforwardness, the nature of the food source can be presented by just a single amount despite the fact that it relies upon different parameters said above.

**Employed foragers**

An employed forager is utilized at a particular food source which it is right now misusing. It conveys particular source data and offers it with different honey bees holding up in the hive. The data incorporates the direction, distance and the gainfulness of the food source. Every food source has just a single utilized honey bee.

**Unemployed foragers**

The honey bee that searches for a food source to utilize is known as Unemployed foragers. When it is searches the environment randomly then it will act as a scout bee or else onlooker bee while utilizing the employed bee information in order to search food source. It can be either a scout who looks through the earth arbitrarily or a spectator who tries to discover a nourishment source by methods for the data given by the utilized honey bee. Around 5-10% is the mean number of scout bee.

**Waggle dance**

The mode of communication among bees to share information about food source. There is a specific area in hive where all bees are gathered to communicate. Onlooker bees observe different dances from dance floor and choose most profitable food source since more information is circulating about the more profitable source. Probability of choosing a food source is directly proportional to profitability. Scout bees convert to employed foragers after requirement by delivering information about food source. Thus, the requirement is directly proportional to profitability.

**ABC algorithm**

Steps involved in ABC optimization algorithm are

- A global probabilistic selection of food source (possible solution) by scout bees.
- It tries to improve food source quality by searching out new food source of better quality.
- Onlooker bees perform a local selection known as greedy selection.
- A random selection procedure carried out by scouts for replacing abandon resource.
- Optimization problem can be identified with the possible food source.
- The number of the employed bees or the onlooker bees equal to the number of solutions in the population.
- D dimensional vector represents a number of optimization parameters as solutions.
A scout bee tries to improve solution by taking in consideration a large number of food sources along with its quality.

\[ \sum = \frac{\text{fit}}{\sum_{n=1}^{\text{SN}} \text{fit}_n} \]

**B. Particle Swarm Optimization (PSO) algorithm**

The Particle Swarm Optimization (PSO) algorithm was proposed by James Kennedy and Eberhart in the year 1995[4]. The technique used to find approximate solutions to difficult or impossible maximization and minimization numeric problems based on Artificial Intelligence is called Particle Swarm Optimization. Optimization problem deals with some substantial quantities such as velocity and position in flocking a bird, non natural particles are build to fly within the seek space. From the intended skill and from the swarm, the knowledge is being spread to the particle. The significance of the particle purposefully optimizes the goal of swarm. As a result, swarm can unite to build up high-quality motion in confined regions of the dilemma. Implementation of PSO algorithms is simple and also it achieves universal finest solution among privileged probability [11].

**How it works**

Firstly, with a group of random particles (solutions) PSO is initialized and then generations are updated. Every particle tries to amend its existing pose and speed according to its distance between its current positions.

**C. Ant colony optimization (ACO) algorithm**

To guide nest mates to food ager ants lay attractive trail pheromones [5]. Pheromones are used to repel foragers from unrewarding routes. This improves the effectiveness of foraging networks. Here we present the empirical evidence. Special ants act as a no entry ants to signal the aging ants. This finding consumes some sophisticated control mechanisms used in self organized ant colonies. Each ant built its own tour from a starting city. Ant chooses a town to go to with a probability. When a tour is completed lay pheromone on each edge is visited. By laying behind pheromones all along their trails, ants perform communication. Others use this to persuade as their path choice. Sturdy pheromone concentrations be likely to be follow. In search mechanism, the pheromone to be used. It also makes the system amenable to parallel implementations. Ant system was the first member of ACO class of algorithms. It consists of some metaheuristic intelligence methods. A similarity with estimation of distribution algorithms has been shared by the ACO. It loses its attractive strength as the time passes which makes the pheromone trail evaporate. It takes more time for pheromones to evaporate. Rapid Discovery is one of the merit and Random decisions sequence, automatic changing of probability distribution, Composition of products are considered as a demerits of ACO. The idea of this ant colony is to mimic the behavior of the simulated ants [6].

**Some General Considerations**
• The best condition encourages ants to search in the nearby places.
• Local updating of this condition makes less evaporation and makes it more comfortable.
• Heuristic improvements can modify the living of ants.

Applications
• Employed in bus routes and garbage collection.
• Machine scheduling: The transportation time is minimized.
• Lacquering machines has been fed.
• Online optimization of telecommunication networks
• Personnel placement in airline companies
• Product Composition

D. Fish Swarm Optimization (FSO) algorithm
In sight of population and stochastic search, Fish swarm optimization algorithm plays a major role in swarm awareness progression algorithms [9]. The Objective function of FSOA is to find out additional quantity of food quantity in sight of swarm, prey, progressive, and chase practices. The key initiative of the fish is always depend on social behaviors and the communal movements. The fish colonies are maintain based on a series of instinctive behaviors, and accordingly demonstrate intelligent behaviors. The review of FSO algorithm mainly deals with the continuous progression, various grouping and also its applications. Many optimization methods will be available they have a similarity through this technique and therefore improvement during the performance. Its disadvantages are more time interval, global and local search imbalancement [10]. FSO Algorithm has some notable parameters such as visual and step. These parameters in turn help to achieve majority intense fitness value and also provide balance among the local and global searches. The visual scope of each fish is the main logical idea behind this algorithm.

The amount of points in its visual scope is denoted as npi.

While npi visual = 0, it denotes empty. While the visual scope is crowded, then the fishes swarm move to the subsequent point. If it is not crowded, the fishes swarm to the central region of the visual scope.

while the visual scope of si is not crowded,
\[
\frac{np^i\text{visual}}{x} < \theta,
\]

Where x be the population size number θ is crowded parameter.

Defining the central point by,
\[
c^i = \sum_{jC = \text{visual}^i s^j} np^i\text{visual}
\]
Where "min" indicate the key point by means of least task value.

Searching behavior activated when the condition dissatisfied.

Some of the advantages of FSO algorithm are global search ability, tolerance of parameter setting, and better robustness.

**Applications**
- Image processing
- Data mining
- Feed-forward neural networks
- Signal processing

**3. Conclusion**

This survey paper encompasses of four significant nature stimulated swarm intelligence algorithms of Data mining. Thus, major features of every optimization techniques have been identified as a part of this survey. In Problem solving technique, these optimization algorithms have affluent applications, merits and demerits. In future prospect, the researcher will have adequate opportunities to carry out with different optimization techniques and fabricate newer heights of fineness in performance. Auxiliary synergies may possibly be exploring with such swarm intelligence algorithms by means of bio-inspired algorithms. Thus, the application of meta-heuristics and hybrid algorithms gets improvements in performance.

**References**

2. Yichen Hu, “Swarm Intelligence”.

