Chapter 56
Particle Swarm Optimization in Bioinformatics, Image Processing, and Computational Linguistics

Badal Soni

https://orcid.org/0000-0002-9617-9468
National Institute of Technology, Silchar, India

Satashree Roy

National Institue of Technology, Silchar, India

Shiv Warsi

National Institue of Technology, Silchar, India

ABSTRACT

Since its inception, particle swarm optimization and its improvement has been an active area of research, and the algorithm has found its application in multifarious domains such as highly constrained engineering problems as well as artificial intelligence. The focal point of this paper is to make the reader aware of the innumerable applications of particle swarm optimization, especially in the field of bioinformatics, digital image processing, and computational linguistics. This review work is designed to serve as a comprehensive look-up guide and to navigate through the algorithm's scope and application in recent times in the aforementioned fields.

INTRODUCTION

The field of optimization aims at deriving a set of optimal solutions for given real valued function(s). Over the years, it has remained one of the major research frontiers in the field of mathematics and computer science.

There are a number of limitations that constrain the performance of an optimization algorithm, with the major ones being (a) nature of the problem, (b) domain and range of the functions being optimized, and (c) time, space and computational bounds. Optimization problems are of two types: test functions, and real-world problems, such as mechanical design problems, transmission network management, vehicle routing optimization problem, power system operation problems, to name a few (Kaveh & Talatahari, 2009; Zahara & Kao, 2009; Galiana et al., 2003; Mahmoudi & Zhou, 2016; Kheshti & Ding, 2017) Some of the well-known test functions[1] are the highly multimodal Ackley function, Schwefel function and the extremely nonlinear Schaffer's f6 function. The reliability and efficiency of such optimization algorithms is usually checked against these set of existing benchmark functions that can also be modified to test the algorithm's robustness.

Optimization algorithms can be broadly categorized into two main categories; (a) *Exact Optimization* and (b) *Heuristic Optimization* algorithms. The former aims to obtain the exact optimum solution(s) of a problem; while the latter outputs a well-approximated non-optimal solution(s) for the exact optimal solution(s). Exact algorithms (Rothlauf, 2011) like numerical methods (Newton-Raphson method, for example), dynamic programming, non-linear programming, branch and bound algorithms etc. produce optimal solutions but are inefficient for solving real world problems which are highly constrained, multimodal, time-varying and multidimensional. Heuristic algorithms are usually preferred over the exact algorithms due to limitations of computational power and time. Therefore, they are the preferred choice of optimization methods to solve problems that have non-polynomial time complexity and hence gradually become intractable with increasing scale of the problem. Particle Swarm Optimization (PSO) (Eberhart & Kennedy, 1995) is one of the notable heuristic search methods without any concrete mathematical proof of its convergence to exact global optima.

The well-known heuristic methods (Silver, 2004) developed take inspiration from the characteristics of natural occurrences or organisms such as annealing in metallurgy, biological evolution, behavior of bird flocks, schools of fish, colonies of ants and swarms of bees. Such nature-inspired heuristics can be divided into (a) *Single-solution based* and (b) *Population based.* The former, comprising of algorithms such as simulated-annealing (Laarhoven & Aarts, 1987), iterated local search(Lourenco &Stutzle, 2003) and variable neighborhood search (Mladenovic & Hansen, 1997), randomly select an initial candidate solution, which is simply a position in the search space of the problem, and then improve upon this single solution to reach the required optimum. On the other hand, any population based algorithm generates multiple such candidate points that are randomly spread across the entiredomain as much as possible. It then proceeds to iteratively improve the solutions by taking into consideration the fitter positions found by the population and combine the candidates to arrive at the final optimal solution. The function to be optimized is taken as a measure of fitness. It is to be noted that both the above mentioned categories do not guarantee that solution returned is the global optimal solution.

Major subcategories of population based heuristics include Evolutionary computation (CoelloCoello, 2005), Genetic algorithm (Man et al., 1996) and Swarm intelligence (Bonabeau et al., 1999). Evolutionary computation is a family of algorithms inspired by biological evolution and Darwin's concept of "survival of the fittest". A population of candidate solutions is generated and in each generation the less desired solutions are removed while introducing small random changes to the remaining ones. Genetic algorithms make use of bio-inspired operators such as (a) *selection*, i.e. fitter solutions have a higher probability of being chosen for the next generation of solutions, and (b) *mutation and crossover*, using which a new solution is generated from a pair of "parent" solutions selected previously. Swarm

Related Content

Bayesian Multilevel Analysis of Postnatal Care Utilization in Ethiopia
Dereje Bekele Dessieand Abay Sahle Nigussie (2021). *International Journal of Applied Research in Bioinformatics (pp. 12-21).*
www.irma-international.org/article/bayesian-multilevel-analysis-of-postnatal-care-utilization-in-ethiopia/267821

Protein Docking and Drug Design
Aditi Gangopadhyay, Hirak Jyoti Chakrabortyand Abhijit Datta (2018). *Applying Big Data Analytics in Bioinformatics and Medicine (pp. 207-241).*
www.irma-international.org/chapter/protein-docking-and-drug-design/182948

Informatics and Data Analytics to Support Exposome-Based Discovery: Part 1 - Assessment of External and Internal Exposure
Dimosthenis A. Sarigiannis, Spyros P. Karakitsios, Evangelos Handakas, Krystalia Papadaki, Dimitris Chapizanisand Alberto Gotti (2018). *Applying Big Data Analytics in Bioinformatics and Medicine (pp. 115-144).*
www.irma-international.org/chapter/informatics-and-data-analytics-to-support-exposome-based-discovery/182945

Chemosensitivity Prediction of Tumours Based on Expression, miRNA, and Proteomics Data
I. Tsamardinos, G. Borboudakis, E. G. Christodoulouand O. D. Røe (2012). *International Journal of Systems Biology and Biomedical Technologies (pp. 1-19).*
www.irma-international.org/article/chemosensitivity-prediction-tumours-based-expression/67103

The Benefits of Big Data Analytics in the Healthcare Sector: What Are They and Who Benefits?
Andrea Darrel, Margee Hume, Timothy Hardieand Jeffery Soar (2015). *Big Data Analytics in Bioinformatics and Healthcare (pp. 406-439).*
www.irma-international.org/chapter/the-benefits-of-big-data-analytics-in-the-healthcare-sector/121469