

# Hybrid Binary Butterfly Optimization Algorithm and Simulated Annealing for Feature Selection Problem

Mohd Faizan, Babasaheb Bhimrao Ambedkar University, India\*

Fawaz Alsolami, King Abdulaziz University, Saudi Arabia

Raees Ahmad Khan, Babasaheb Bhimrao Ambedkar University, India

## ABSTRACT

Feature selection is performed to eliminate irrelevant features to reduce computational overheads. Metaheuristic algorithms have become popular for the task of feature selection due to their effectiveness and flexibility. Hybridization of two or more such metaheuristics has become popular in solving optimization problems. In this paper, the authors propose a hybrid wrapper feature selection technique based on binary butterfly optimization algorithm (bBOA) and simulated annealing (SA). The SA is combined with the bBOA in a pipeline fashion such that the best solution obtained by the bBOA is passed on to the SA for further improvement. The SA solution improves the best solution obtained so far by searching in its neighborhood. Thus, the SA tries to enhance the exploitation property of the bBOA. The proposed method is tested on 20 datasets from the UCI repository, and the results are compared with five popular algorithms for feature selection. The results confirm the effectiveness of the hybrid approach in improving the classification accuracy and selecting the optimal feature subset.

## KEYWORDS

Binary Butterfly Optimization Algorithm (bBOA), Classification, Feature Selection, Hybrid Optimization Algorithm, Simulated Annealing

## 1. INTRODUCTION

Feature Selection is a crucial research area in the development of efficient classification algorithms for high dimensional datasets (Khaire and Dhanalakshmi, 2019). A proper feature selection technique should enhance the performance of the classification model. High dimensional dataset suffers from the famous problem of “curse of dimensionality” (Xue et al., 2016). The original dataset contains irrelevant and redundant features that degrade the performance of the algorithm and also increase computational overheads. Feature selection methods aim to reduce the feature space by removing superfluous and insignificant elements. These methods can be broadly classified into *wrapper* and *filter* methods. In the *wrapper* approach, the classifier is used to evaluate the quality of the selected features. The filter methods use concepts from information theory to obtain a relevant set of features. Wrapper methods are usually slower than filter methods. However, they can generate better classification performance than filter methods (Grande et al., 2007).

The feature selection technique must be capable of searching every possible subset of the original feature set to obtain the optimum feature subset. However, it is impractical to achieve this task owing to the high computational cost that comes along with it which consequently results in the selection of

DOI: 10.4018/IJAMC.2022010104

\*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

a sub-optimum feature subset. Nowadays, evolutionary algorithms and metaheuristic techniques are popular for solving the problem of feature selection. The metaheuristic approaches did not guarantee the best results; however, it will produce better solutions within time bounds (Talbi et al., 2009). Some of the most popular metaheuristics are genetic algorithm (GA) (Holand et al., 1992), differential evolution (DE) (Storn and Price, 1997), tabu search (TS) (Hedar et al., 2006), simulated annealing (SA) (Kirkpatrick et al., 1983), particle swarm optimization (PSO) (Kennedy and Eberhart, 1995), ant colony optimization (ACO) (Dorigo et al., 2006), artificial bee colony (ABC) (Yi and He, 2014) and Cuckoo Search (CS) (Yang and Deb, 2009), ant lion optimization (ALO) (Emary et al., 2016) and sine-cosine algorithm (SCA) (Mirjalili, 2016).

A good metaheuristic should be capable of maintaining a balance between exploration and exploitation strategies during the search process (Talbi et al., 2009). Population-based methods like PSO or ACO are better in exploring the search space while single solution-based methods employ the right exploitation strategies. To achieve a balance between the two, hybrid methods can be used that combine the merits of two or more metaheuristic to achieve better optimization results.

Butterfly Optimization Algorithm (BOA) is a recent nature-inspired swarm algorithm that takes inspiration from the food searching behavior of butterflies in a natural environment (Arora and Singh, 2018). BOA has produced superior results for continuous optimization problems when compared with other recent metaheuristics. The binary version of BOA called bBOA was proposed by Arora and Anand (2019) for feature selection problems in wrapper mode. The binary version bBOA has the same structure as that of native BOA except that it utilizes two transfer functions for generating binary solutions.

A metaheuristic based on hill-climbing called Simulated Annealing was proposed by Kirkpatrick et al (1983). The method tries to improve the current solution at each iteration by generating a trial solution in its vicinity. The improved solution with better fitness value is always accepted while the solution with no improvement is taken with a specific probability value to avoid getting trapped in the local optima. The probability of acceptance of a worse solution is dependent on the temperature parameter. The temperature parameter decreases with every iteration at a fixed rate called the cooling schedule.

Despite producing superior results, the weak exploitation ability of BOA prematurely converges it to local optima. Also, the random transition between the exploration and exploitation phase in BOA is determined by the value of the switching probability which sometimes distracts the BOA from attaining the global optima (Arora et al., 2018). On the other hand, the SA algorithm possesses excellent exploitation properties.

In this paper, we introduce a hybrid wrapper method of bBOA and SA algorithms that can enhance the exploitation capability of the native bBOA for feature selection. The hybrid method combines bBOA with SA to further improve the fitness value of the candidate solutions. The SA algorithm enhances the exploitation property of bBOA by iteratively adjusting the best solution obtained so far at the end of every iteration. In this way, the SA algorithm is embedded into the bBOA and is executed in each iteration of the bBOA. The efficiency of the proposed approach has been evaluated against 20 benchmark datasets from the UCI machine learning repository.

The rest of the paper is organized as follows: Section 2 describes the related work in the area of feature selection using metaheuristic algorithms. Section 3 describes the background and the proposed hybrid approach. Section 4 contains the result and discussion part and Section 5 draws the conclusion and future work.

## 2. RELATED WORK

The objective of feature selection is to select a reduced subset of features by removing irrelevant and redundant features. The task of feature selection requires the creation and evaluation of every possible subset of features which makes it an NP-hard problem (Chandrashekar and Sahin, 2014).

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/article/hybrid-binary-butterfly-optimization-algorithm-and-simulated-annealing-for-feature-selection-problem/284575](http://www.igi-global.com/article/hybrid-binary-butterfly-optimization-algorithm-and-simulated-annealing-for-feature-selection-problem/284575)

## Related Content

---

### Early Warning from Car Warranty Data using a Fuzzy Logic Technique

Mark Last, Yael Mendelson, Sugato Chakrabarty and Karishma Batra (2010). *Scalable Fuzzy Algorithms for Data Management and Analysis: Methods and Design* (pp. 347-364).

[www.irma-international.org/chapter/early-warning-car-warranty-data/38576](http://www.irma-international.org/chapter/early-warning-car-warranty-data/38576)

### RFID Technologies in Healthcare Setting: Applications and Future Perspectives

Alessia D'Andrea, Fernando Ferri and Patrizia Grifoni (2016). *International Journal of Computers in Clinical Practice* (pp. 15-27).

[www.irma-international.org/article/rfid-technologies-in-healthcare-setting/152590](http://www.irma-international.org/article/rfid-technologies-in-healthcare-setting/152590)

### Optimal Power Flow and Optimal Reactive Power Dispatch Incorporating TCSC-TCPS Devices Using Different Evolutionary Optimization Techniques (2019). *Optimal Power Flow Using Evolutionary Algorithms* (pp. 210-243).

[www.irma-international.org/chapter/optimal-power-flow-and-optimal-reactive-power-dispatch-incorporating-tcsc-tcps-devices-using-different-evolutionary-optimization-techniques/212082](http://www.irma-international.org/chapter/optimal-power-flow-and-optimal-reactive-power-dispatch-incorporating-tcsc-tcps-devices-using-different-evolutionary-optimization-techniques/212082)

### Discrete Particle Swarm Optimization for the Multi-Level Lot-Sizing Problem

Laurent Deroussi and David Lemoine (2011). *International Journal of Applied Metaheuristic Computing* (pp. 44-57).

[www.irma-international.org/article/discrete-particle-swarm-optimization-multi/52792](http://www.irma-international.org/article/discrete-particle-swarm-optimization-multi/52792)

### ANN-Based Self-Tuning Frequency Control Design for an Isolated Microgrid

H. Bevrani, F. Habibi and S. Shokooi (2013). *Meta-Heuristics Optimization Algorithms in Engineering, Business, Economics, and Finance* (pp. 357-385).

[www.irma-international.org/chapter/ann-based-self-tuning-frequency/69891](http://www.irma-international.org/chapter/ann-based-self-tuning-frequency/69891)