

Analysis of Backpropagation using Biogeography Based Optimization

Sunali Mehta¹, Sheetal Kundra²,³Harish Kundra
^{1,2}Department of Computer Science & Engineering, PTU, RIEIT, Railmajra, INDIA

Abstract- Biogeography Based Optimisation (BBO) is a population based evolutionary algorithm (EA) motivated by the migration mechanisms of ecosystems. BBO is a nature inspired technique. The idea was given by D.Simon in Dec.2008. Geographical areas that are well suited as residences for biological species are said to have a High Suitability Index (HSI). Features that are correlated with HSI include such factors as rainfall, diversity of vegetation, land area and temperatures. The variables that characterize the habitability are called Suitability Index Variables (SIVs). Habitats with a high HSI tend to have a large number of species, while those with a low HSI have small number of species. It is an example of natural process that can be modelled to solve general optimisation problems. In this paper, this nature inspired behaviour has been considered to optimise the neural network i.e Backpropagation by BBO. The Backpropagation is a Neural Network. In this , the units are connected in a feed- forward fashion with input units fully connected to the units in the hidden layer and hidden units are fully connected to the units in output layer. When a Backpropagation network is cycled , an input pattern is propagated forward to the output units through an intervening input-to-hidden and hidden-to-output weights. The idea behind this is that errors for the units of the hidden layer are determined by backpropagating the errors of the units of the output layer and analysis of the Backpropagation are done by BBO. The BBO searches the optimal or near optimal connection weights and threshold.

Keywords — Biogeography, Evolutionary algorithms, Habitat suitability index, Suitability index variable, optimization

I. INTRODUCTION

The science of biogeography can be traced to the work of nineteenth century naturalists such as Alfred Wallace and Charles Darwin. Until the 1960s, biogeography was mainly descriptive and historical. In the early 1960s, Robert MacArthur and Edward Wilson began working together on mathematical models of biogeography, their work culminating with the classic 1967 publication *The Theory of Island Biogeography*. Their interest was primarily focused on the distribution of species among neighboring islands. They were interested in mathematical models for the extinction and migration of species. Since MacArthur and Wilson's work, biogeography has become a major area of research. A recent search of Biological Abstracts reveals that 25,452 papers were written in the year 2005 that were related to the subject of biogeography. However, a search of INSPEC, an engineering research index, reveals that no biogeography papers have ever been written in the year 2005 that were related to the subject of biogeography. However, a search of INSPEC, an engineering research index, reveals that no biogeography papers have ever been written. In view of this, part of the motivation of this paper is to merge the burgeoning field of biogeography with engineering in order to see how the two disciplines can be of mutual benefit. The application of biogeography to engineering is similar to what has occurred in the past few decades with genetic algorithms (GAs), neural networks, fuzzy logic, particle swarm optimization (PSO), and other areas of computer intelligence.

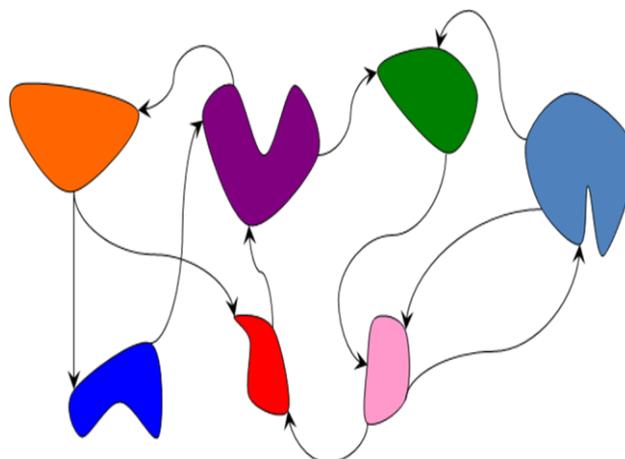


Fig.1 Species migrate between islands.

Mathematical models of biogeography describe how species migrate from one island to another, how new species arise, and how species become extinct. The term “island” here is used descriptively rather than literally. That is, an island is any habitat that is geographically isolated from other habitats. We therefore use the more generic term “habitat” in this paper (rather than “island”). Geographical areas that are well suited as residences for biological species are said to have a high habitat suitability index (HSI). Features that correlate with HSI include such factors as rainfall, diversity of vegetation, diversity of topographic features, land area, and temperature. The variables that characterize habitability are called suitability index variables (SIVs). SIVs can be considered the independent variables of the habitat, and HSI can be considered the dependent variable.

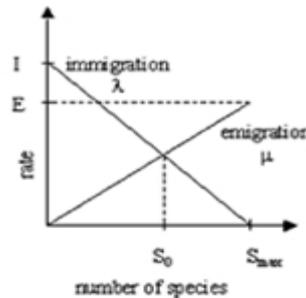


Fig. 2 Species model of single habitat

Habitats with a high HSI tend to have a large number of species, while those with a low HSI have a small number of species. Habitats with a high HSI have many species that emigrate to nearby habitats, simply by virtue of the large number of species that they host. Habitats with a high HSI have a low species immigration rate because they are already nearly saturated with species. Therefore, high HSI habitats are more static in their species distribution than low HSI habitats. By the same token, high HSI habitats have a high emigration rate; the large numbers of species on high HSI islands have many opportunities to emigrate to neighboring habitats. (This does not mean that an emigrating species completely disappears from its home habitat; only a few representatives emigrate, so an emigrating species remains extant in its home habitat, while at the same time migrating to a neighboring habitat.) Habitats with a low HSI have a high species immigration rate because of their sparse populations. This immigration of new species to low HSI habitats may raise the HSI of the habitat, because the suitability of a habitat is proportional to its biological diversity. However if a habitat’s HSI remains low, then the species that reside there will tend to go extinct, which will further open the way for additional immigration. Due to this, low HSI habitats are more dynamic in their species distribution than high HSI habitats. Biogeography is nature’s way of distributing species, and is analogous to general problem solutions. Suppose that we are presented with a problem and some candidate solutions. The problem can be in any area of life (engineering, economics, medicine, business, urban planning, sports, etc.), as long as we have a quantifiable measure of the suitability of a given solution. A good solution is analogous to an island with a high HSI, and a poor solution represents an island with a low HSI. High HSI solutions resist change more than low HSI solutions. By the same token, high HSI solutions tend to share their features with low HSI solutions. (This does not mean that the features disappear from the high HSI solution; the shared features remain in the high HSI solutions, while at the same time appearing as new features in the low HSI solutions. This is similar to representatives of a species migrating to a habitat, while other representatives remain in their original habitat.) Poor solutions accept a lot of new features from good solutions. This addition of new features to low HSI solutions may raise the quality of those solutions. We call this new approach to problem solving biogeography-based optimization (BBO). BBO has certain features in common with other biology-based algorithms. Like GAs and PSO, BBO has a way of sharing information between solutions. GA solutions “die” at the end of each generation, while PSO and BBO solutions survive for-ever (although their characteristics change as the optimization process progresses). PSO solutions are more likely to clump together in similar groups, while GA and BBO solutions do not necessarily have any built-in tendency to cluster. Section II reviews the idea of biogeography, and neural. Section III discusses how neural network can be analyzed with biogeography based optimization.

II. BIOGEOGRAPHY AND NEURAL NETWORK

An A Biogeography Based Optimization (BBO) is a population based evolutionary algorithm (EA) motivated by the migration mechanisms of ecosystems. It is based on the mathematics of biogeography. In BBO, problem solutions are represented as islands, and the sharing of features between solutions is represented as emigration and immigration. The idea of BBO was first presented in December 2008 by D. Simon (Simon, 2008). It is an example of natural process that can be modeled to solve general optimization problems. One characteristic of BBO is that the original population is not discarded after each generation; it is rather modified by migration. Also for each generation, BBO uses the fitness of each solution to determine its emigration and immigration rate. In a way, we can say that BBO is an application of biogeography to EAs. In BBO, each individual is considered as a habitat with a habitat suitability index (HSI), which is similar to the fitness of EAs, to measure the individual. Also, an SIV (suitability index variable) which characterizes the habitability of an island is used. A good solution is analogous to an island with a high HSI, and a poor solution indicates an island with a low HSI. High HSI solutions tend to share their features with low HSI solutions. Low HSI solutions accept a lot of new features from high HSI solutions.

A. Neural Network

Neural Network is developed by McCulloch and Pitts in 1943. Neural Network has certain processing capabilities of the human brain. The term neural network was traditionally used to refer to a network or circuit of biological neurons. The modern usage of the term often refers to artificial neural networks, which are composed of artificial neurons or nodes. Thus the term has two distinct usages:

Biological Neural Networks is an information processing paradigm, inspired by biological system, composed of a large number of highly interconnected processing elements (neurons) to solve specific problems.

Artificial Neural Networks is an artificial representation of the human brain that tries to simulate its learning process. An artificial Neural network is often called a Neural network or simply Neural net. It is a mathematical model or computation model. In this paper Backpropagation Neural Network is used to analyze the BBO. It was presented by Rumelhart, Hinton and Williams in 1986. The idea behind this is that errors for the units of the hidden layer are determined by back-propagating the errors of the units of the output layer. This method is often called the backpropagation learning rule. It is a supervised learning process. Backpropagation is a systematic method of training multi-layer artificial neural networks. For better understanding, the backpropagation learning algorithm can be divided into two phases: propagation and weight update.

Phase 1: Propagation

Each propagation involves the following steps:

1. Forward propagation of a training pattern's input through the neural network in order to generate the propagation's output activations.
2. Backward propagation of the propagation's output activations through the neural network using the training pattern's target in order to generate the deltas of all output and hidden neurons.

Phase 2: Weight update

For each weight-synapse:

1. Multiply its output delta and input activation to get the gradient of the weight.
2. Bring the weight in the opposite direction of the gradient by subtracting a ratio of it from the weight.

This ratio influences the speed and quality of learning; it is called the *learning rate*. The sign of the gradient of a weight indicates where the error is increasing; this is why the weight must be updated in the opposite direction. Repeat the phase 1 and 2 until the performance of the network is good enough.

III. NEURAL NETWORK WITH BIOGEOGRAPHY BASED OPTIMIZATION

In this section we describe how Backpropagation is analysed with biogeography based optimization.

A. Algorithm

1. Initialize the BBO parameters- Initialize, Cost, and Feasibility.
2. Initialize a random set of habitats, each habitat corresponding to a potential solution to the given problem.
3. For each habitat, map the Habitat Suitability Index to Initialize, Cost, Feasibility.
4. Probabilistically use Cost function and Feasibility to modify each non-elite habitat, and then recompute each HSI.
5. For each habitat, minimize the cost function of Neural Network (Backpropagation) using Backpropagation
6. Is acceptable solution found? If yes then go to Step 8.
7. Number of iterations over? If no then go to Step 3 for the next iteration.
8. Stop

B. Flowchart of work

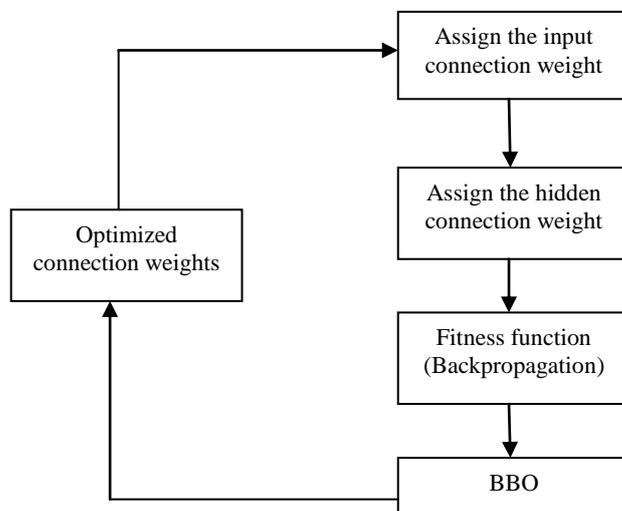


Fig3. Flowchart of the present work

In this figure the BBO algorithm optimizes the connection weights between layers in Backpropagation. A Backpropagation is able to work parallel with input variables and consequently handle large sets of data swiftly. The connection weights in Backpropagation are optimized by using BBO. It has three phases in the first phase; the BBO searches the optimal or near optimal connections weights and threshold. The connection weights and threshold are initialized in two random values before search process. It needs three sets of parameters.

- (a) The first set is the set of connection weights of input layer and the hidden layer.
- (b) The second is the set of connection weight between the hidden layer and the output layer.
- (c) The third step represents the threshold. a fully connected feed forward network with one hidden layer and one output is shown above.

C. Result

The accuracy of spatial data has been defined as “ The closeness of results of observations, computations , or estimates to the true values or the values accepted as being true”.

Accuracy assessment is an important step in the process of analysing Backpropagation. Accuracy defines “correctness”. Backpropagation are a special class of function . Calculate the error between the actual output and desired output

$$E_p = \sum k(d_k - x_k)^2$$

Where,

$d_k =$ is the desired output for node k

$x_k =$ is the actual output for node k

$E_p =$ is mean square error when path pattern input is given as output

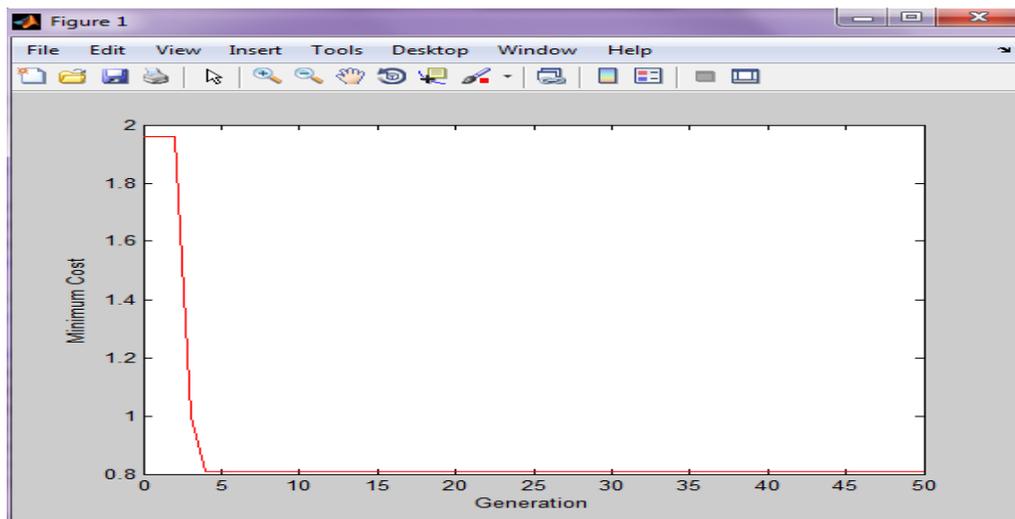


Fig. 3 Describes the minimization of cost function of BBO

IV. CONCLUSION AND FUTURE SCOPE

In this work, popular swarm intelligence technique is used to classify the optimization results. The study of geographical distribution of biological species can be used to derive algorithms for optimization. BBO can be successfully applied to practical problems. In BBO it is not possible to check for feasibility while a new solution is being completed. The feasibility check has to wait until after the new solution is already complete. Analysis of Neural network using BBO algorithm has solved this problem. Moreover it provides the better accuracy then other optimization algorithms. We cannot conclude that BBO is universally better than other methods, or vice versa. However it may be possible in future work to quantify the performance of BBO relative to other algorithms for problems with specific features. This paper is preliminary in nature and, therefore, opens up a wide range of possibilities for future results. It might be fruitful to explore the ideas of species sharing only between similar solutions. Species are more likely to migrate to habitats better close to their place of origin. At present BBO is still greedy in the sense that it builds each rule with the aim of optimizing that rule's quality individually, without directly taking into account the interaction with other rules. A less greedy, but possibly more computationally expensive way to approach the problem would be to associate a particle with an entire rule set and then to consider the quality of the entire rule set when evaluating a particle. Also the nominal part of the rule is always discovered first and separately from the continuous part, it could be advantageous to use a more “co-evolved” approach.. Future research will focus on using this algorithm with other soft computing techniques. So this swarm technique can also be combined with fuzzy-rough approach also in future.

REFERENCES

- [1]. Simon, D. (2008), "Biogeography-based optimization", IEEE Transactions on Evolutionary Computation, vol. 12, no. 6, pp. 702-713.
- [2]. R. MacArthur and E. Wilson, "The Theory of Biogeography". Princeton, NJ: Princeton Univ. Press, 1967.
- [3]. A. Wallace, "The Geographical Distribution of Animals" (Two Volumes). Boston, MA: Adamant Media Corporation, 2005.
- [4]. D. Simon, "Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches", John Wiley & Sons, 2006
- [5]. D. Simon, "Training Radial Basis Neural Networks with the Extended Kalman Filter," *Neurocomputing*, vol. 48, no. 1-4, pp. 455-475, October 2002
- [6]. M. Ergezer and D. Simon, "Oppositional biogeography-based optimization for combinatorial problems," *IEEE Congress on Evolutionary Computation*, New Orleans, LA, pp. 1496-1503, June 2011 .
- [7]. C. Yang and D. Simon, "A New Particle Swarm Optimization Technique," *International Conference on Systems Engineering*, Las Vegas, NV, pp. 164-169, August 2005.
- [8]. M. Ovreiu and D. Simon, "Cardiomyopathy Detection from Electrocardiogram Features," in *Cardiomyopathies – From Basic Research to Clinical Management (J. Veselka, editor) InTech*, pp. 117-134, 2012 .
- [9]. D. Simon, M. Ergezer, and D. Du, "Population Distributions in Biogeography-Based Optimization Algorithms with Elitism," *IEEE Conference on Systems, Man, and Cybernetics*, San Antonio, TX, pp. 1017-1022, October 2009
- [10]. Bratton, D. and Kennedy, J. (2007), "Defining a Standard for Particle Swarm Optimization", In Proceedings of IEEE Swarm Intelligence Symposium, Honolulu, Hawaii, USA, pp. 120-127.
- [11]. Dorigo, M., Maniezzo, V. and Colomi, A. (1996), "The Ant System: Optimization by a colony of cooperating agents", IEEE Transactions on Systems, Man, and Cybernetics-Part B, vol.26, no.1, pp.1-13.
- [12]. Kennedy, J. and Eberhart, R.C.(1995), "Particle Swarm Optimization", In Proceedings of IEEE International Conference on Neural Networks, New Jersey, USA, pp. 1942-1948.
- [13]. Aleksander, I. and Morton "An introduction to neural computing", 2nd edition Neural Networks, Eric Davalo and Patrick Naim.
- [14]. Y. Liu, J. A. Starzyk, Z. Zhu, "Optimized Approximation Algorithm in Neural Network without Overfitting", IEEE Trans. on Neural Networks, vol. 19, no. 4, June, 2008, pp. 983-995.
- [15]. Esprit, I.F.Croall, J.P.Mason, "Industrial Applications of Neural Networks".
- [16]. Eric Davalo and Patrick Naim Klimasauskas, CC., 1989 "DARPA Neural Network Study" (October, 1987-February, 1989). MIT Lincoln Lab. Neural Networks,
- [17]. Hammerstrom, D. (1986), "The 1989 Neuro Computing Bibliography". A Connectionist/Neural Network Bibliography. Y. Liu, J. A. Starzyk, Z. Zhu, "Optimized Approximation Algorithm in Neural Network without Overfitting", IEEE Trans. on Neural Networks, vol. 19, no. 4, June, 2008, pp. 983-995.
- [18]. Arthur Earl Bryson, Yu-Chi Ho (1969). *Applied optimal control: optimization, estimation, and control*. Blaisdell Publishing Company or Xerox College Publishing. pp. 481.
- [19]. Rumelhart, David E.; Hinton, Geoffrey E., Williams, Ronald J. (8 October 1986). "Learning representations by back-propagating errors". *Nature* **323** (6088): 533-536.
- [20]. Industrial Applications of Neural Networks (research reports Esprit, I.F.Croall, J.P.Mason)
- [21]. DARPA Neural Network Study (October, 1987-February, 1989). MIT Lincoln Lab. Neural Networks, Eric Davalo and Patrick Naim
- [22]. Klimasauskas, CC. (1989). The 1989 Neuro Computing Bibliography. Hammerstrom, D. (1986). A Connectionist/Neural Network Bibliography.

Author's Profile



Sunali Mehta is a student of Rayat Institute of Engineering and Technology, Railmajra . She is a pursuing student of Mtech CSE. She did her Btech(IT) from MIET, Jammu.



Mrs Sheetal Kundra is an Assistant Professor of Deptt (CSE & IT) at Rayat Institute of Engineering and Technology, Railmajra.



Prof Harish Kundra is an Associate Professor and Head of Deptt (CSE and IT) at Rayat Institute of Engineering and Technology, Railmajra.