A New Approach to Train Multilayer Perceptron ANN Using Error Back-propagation and Genetic Algorithms Hybrid: A Case Study of PVTx Estimation of CH$_4$+CF$_4$ Gas Mixture

Abdolreza Moghadassi$^{a,*}$, Mahmood Reza Nikkholgh$^a$, Sayed Mohsen Hosseini$^a$, Fahime Parvizian$^a$, Seyyed Jelaladdin Hashemi$^b$

$^a$Department of Chemical Engineering, Faculty of Engineering, Arak University, Arak, Iran.
$^b$Petroleum University of Technology, Ahvaz, Iran.
*Email: a_moghadassi@yahoo.com, a-moghadassi@Araku.ac.ir

Received: 14 February 2011; Accepted: 3 May 2011

Abstract

A new algorithm to train Multilayer Perceptron Artificial Neural Network using the genetic and Error Back-propagation algorithms Hybrid has been devised. The new algorithm solves the local minimum trap as a natural result of the standard numerical optimization based methods and by following the global minimums the ANN training accuracy has been highly improved. There are many algorithms for training a Multilayer Perceptron ANN to estimate the PVTx of CH$_4$+CF$_4$ gas mixture. The new devised algorithm is compared and evaluated against these algorithms and indicates a better accuracy.

Keywords: Artificial neural network; Gas mixture; Genetic algorithm; Hybrid; PVTs estimation.

1. Introduction

The ANN is a tool inspired from the biological human being neural network. It is known as a powerful tool in estimating data analysis and provides a fast compact general solution for the modeled complex dynamic systems. Although it is not too long past the mathematical modeling of ANN, many researchers have investigated their applications in various fields [1-7]. The specific powerfullness of ANN is due to their learning and training ability. This specialty causes ANN to be trained by a set of input data with their solutions. The network free parameters which are adjusted to train the network to respond correctly to the input data (Weight function matrices and Bias vectors) during the training period are stored by a specific algorithm.

The important idea that was the key development to ANN during the eighties of twentieth century is the Error Back-propagation algorithm which was offered by David Rummlhdhur and James Mcland in 1986 [8]. Many approaches are offered to train the ANN based on this algorithm [9]. But most of them which are based on standard optimization are generally slow processes and trap in the local extremums and finally diverge [1].

For the first time, John Holland used the genetic algorithm in 1975. Since then, the genetic algorithm has been known as a powerful tool in solving many optimization
problems [10]. It may be used to solve complicated optimization problems that have high oscillation and scattering of data. The dominant aspect of the genetic algorithm is its use of scattered data over the data search domain to find the best solution for the considered problem [10].

There are many works that benefit from the powerfulness of both genetic and ANN in a hybrid form, each of which tries to get its best benefits from the combination [11-16].

In this paper, a new algorithm has been offered based on genetic algorithm powerfulness that increases the training process accuracy by following the global extremums. The new algorithm's evaluation is made in comparison to other algorithms to estimate PVTx of the CH4+CF4 gas mixture.

2. Theory
2.1. The genetic algorithm

The input data to the genetic algorithm are usually binary coded numbers. The solution of the considered problem may be obtained by having a series of input data with their solutions called chromosomes. The chromosomes are normally indicated by simple strings of data which themselves are a combination of Genes. A set of generated chromosomes that are used to begin optimization process by the genetic algorithm is called population which the problem's answer is extracted from it. A schematic view of simple population of chromosomes and genes are shown in Fig. 1. Each chromosome of the population is evaluated by a fitness function. Depending on the fitness function, those chromosomes which pass the required criterion have more chance to enter the next step in the solution search process. This process is undertaken by Selection operator. There are various Selection Operators with Roulette Wheel as their most common one.

The Crossover operator which is the dominant aspect of genetic algorithm from the other optimization methods one or more parts of the chromosomes determined by Selection Operator are picked up and combined to get a new chromosome with a better match [1]. This process is shown in Fig. 2.

Finally, the Mutation Operator changes one or more parts of a chromosome to add a new data to the population. Fig. 3 illustrate the process. The purpose of this operator is to preserve the population diversity and to prevent false convergence [1].
An optimized solution to the problem is obtained through using the repeated method for a number of chromosomes in the considered population. The genetic algorithm calculations approach is shown in Fig. 4.

2.2. ANN training using genetic algorithm

The Feed Forward Multilayer ANN training by Error Back-propagation algorithm has several obvious weaknesses some of which are slow training speed, local extremum trapping possibility and need for training speed adjustment in each training process repetition [1].

The genetic algorithm can be used to overcome the above mentioned difficulties. In fact, the ANN can be considered like other common mathematical models and so try to determine their adjustable parameters to get the best match between the model output and the object.

Therefore, the mean square error (MSE) which is normally the criterion used in the evaluation of the network accuracy can be taken as the object function of the genetic algorithm. The weights and biases of the ANN model are adjusted according to the chosen object function. The network structure is determined firstly and then the whole network is coded by chromosomes which the process is called indirect coding [1]. A two layer ANN with two inputs and one output and its corresponding chromosomes that are indirectly coded are shown in Fig. 5.
The above network has nine adjustable parameters that are coded as a string. The string length depends on the number of bits assigned to each parameter. The bits number has considerable effect on the convergence speed. Assigning low number of bits to each weight makes the evolution process insensitive to the weight parameter. Assigning high number of bits to each weight reduces the speed of convergence due to increase in the chromosome's length and therefore increasing the calculation time [1].

2.3. ANN training using Hybrid algorithms

The inherent capability of the genetic algorithm makes it possible to find the closest point optimized global points. The repetition evaluation of the function is necessary. A common method to increase convergence speed is reduction of generations number as a criterion for convergence matching. Approaching the optimized points, another algorithm will be followed which is faster in convergence and is more capable than genetic algorithm to find the optimized local points. In fact, when the genetic algorithm approaches the optimized global points, the information is transmitted to another algorithm as a first guess for the final solution. The fminunc, patternsearch and fminsearch which optimize MATLAB Functions algorithms may be used next to the genetic algorithm in the MATLAB software.

The new idea in this paper is to combine the genetic algorithm and the fast Error Back-propagation algorithm for the ANN training. At first, a chromosome containing the network adjustable parameters from the genetic algorithm is defined and such a close solution to the optimized global points is obtained. Then at the end of the first algorithm the best obtained solution is used the initial guess at the beginning of network training by Error Back-propagation algorithm. This hybrid algorithm is shown in Fig. 6.

The Levenberg-Marquardt back-propagation algorithm's high speed property might be the best choice for combining with the genetic algorithm. The reason is that such a combination increases both accuracy and convergence speed of the genetic algorithm.

![Fig. 6. The Hybrid method used in ANN training.](image-url)
3. Results and Discussions

The results of the new devised algorithm are compared to the results of the common Back-propagation algorithms in training ANN used in the estimation of PVTx of methane and tetra flour methane gas mixture. These algorithms are used to train a network with ten neurons. The network has three input ports as \( T, P, x_1 \), one output port as \( \rho \) and hidden output transfer functions layers, tansig and purelin. The experimental data obtaining by Douslin et al. are used in training the network [6, 17]. The Back-propagation algorithms used to compare the results with genetic and Levenberg-Marquardt hybrid are shown in Table 1.

The algorithms of Table 1 are trained to estimate the methane and tetra flour methane gas mixture density. The results are tabulated in Table 2. Each algorithm is trained ten times and the average training time, the maximum, the minimum, and the maximum number of iterations to match the convergence criterion are also shown in Table 2. The MSE convergence criterion is taken to be 0.001. As it is expected, the hybrid algorithm shows good results. The ANN's training convergence speeds of the above algorithms are also compared in Fig. 7. The results verify the powerfulness of the devised algorithm.

4. Conclusions

A new algorithm to train Artificial Neural Network using the genetic and Error Back-propagation algorithms Hybrid has been devised. The new algorithm solves the local minimum trap which is the natural result of the standard numerical optimization based methods and by following the global minimums the ANN training accuracy has been highly improved. This is because the genetic algorithm has an extended domain over the search space. It is worth mentioning that due to genetic algorithm's slow process compared to standard optimization based process; there is a need to use a faster meta-heuristic method such as Ant Colony.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>MATLAB Function</th>
<th>Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM</td>
<td>trainlm</td>
<td>Levenberg-Marquardt</td>
</tr>
<tr>
<td>BFG</td>
<td>trainbfg</td>
<td>BFGS Quasi-Newton</td>
</tr>
<tr>
<td>RP</td>
<td>trainrp</td>
<td>Resilient Backpropagation</td>
</tr>
<tr>
<td>SCG</td>
<td>trainscsg</td>
<td>Scaled Conjugate Gradient</td>
</tr>
<tr>
<td>CGB</td>
<td>traincgb</td>
<td>Conjugate Gradient with Powell-Beale Restarts</td>
</tr>
<tr>
<td>CGF</td>
<td>traincgf</td>
<td>Fletcher-Powell Conjugate Gradient</td>
</tr>
<tr>
<td>CGP</td>
<td>traincgp</td>
<td>Polak-Ribiére Conjugate Gradient</td>
</tr>
<tr>
<td>OSS</td>
<td>trainoss</td>
<td>One-Step Secant</td>
</tr>
</tbody>
</table>

Table 1. The different methods of ANN training used in the evaluation of GA-LM Hybrids [20]

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Mean Time, Sec</th>
<th>Min. Time, Sec</th>
<th>Max. Time, Sec</th>
<th>Min. Iteration</th>
<th>Max. Iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA-LM</td>
<td>0.820</td>
<td>0.581</td>
<td>1.052</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>LM</td>
<td>0.863</td>
<td>0.471</td>
<td>1.502</td>
<td>8</td>
<td>43</td>
</tr>
<tr>
<td>BFG</td>
<td>4.464</td>
<td>2.654</td>
<td>7.491</td>
<td>94</td>
<td>308</td>
</tr>
<tr>
<td>RP</td>
<td>5.038</td>
<td>2.323</td>
<td>14.240</td>
<td>193</td>
<td>1273</td>
</tr>
<tr>
<td>SCG</td>
<td>5.486</td>
<td>2.003</td>
<td>9.303</td>
<td>81</td>
<td>408</td>
</tr>
<tr>
<td>CGB</td>
<td>7.497</td>
<td>3.274</td>
<td>13.639</td>
<td>105</td>
<td>471</td>
</tr>
<tr>
<td>CGF</td>
<td>7.278</td>
<td>1.983</td>
<td>9.674</td>
<td>67</td>
<td>414</td>
</tr>
<tr>
<td>CGP</td>
<td>8.616</td>
<td>2.123</td>
<td>16.564</td>
<td>76</td>
<td>630</td>
</tr>
<tr>
<td>OSS</td>
<td>16.670</td>
<td>8.162</td>
<td>25.256</td>
<td>372</td>
<td>1125</td>
</tr>
</tbody>
</table>

Table 2. The comparison of different methods for ANN training
Fig. 7. The different methods for ANN training used in the evaluation of CH$_4$+CF$_4$ gas mixture density.

References


