Chapter 6
Ant Colony Optimization Applied to the Training of a High Order Neural Network with Adaptable Exponential Weights

Ashraf M. Abdelbar
Brandon University, Canada

Islam Elnabarawy
Missouri University of Science and Technology, USA

Donald C. Wunsch
Missouri University of Science and Technology, USA

Khalid M. Salama
University of Kent, UK

ABSTRACT

High order neural networks (HONN) are neural networks which employ neurons that combine their inputs non-linearly. The HONEST (High Order Network with Exponential SynapTic links) network is a HONN that uses neurons with product units and adaptable exponents. The output of a trained HONEST network can be expressed in terms of the network inputs by a polynomial-like equation. This makes the structure of the network more transparent and easier to interpret. This study adapts ACO$\mathbb{R}$, an Ant Colony Optimization algorithm, to the training of an HONEST network. Using a collection of 10 widely-used benchmark datasets, we compare ACO$\mathbb{R}$ to the well-known gradient-based Resilient Propagation (R-Prop) algorithm, in the training of HONEST networks. We find that our adaptation of ACO$\mathbb{R}$ has better test set generalization than R-Prop, though not to a statistically significant extent.

DOI: 10.4018/978-1-7998-0414-7.ch006

1. OVERVIEW

In this work, we present an adaptation of the $ACO_{\mathbb{R}}$ (Socha & Dorigo, 2008) Ant Colony Optimization (ACO) (Dorigo & Stützle, 2004) algorithm applied to the training of the HONEST (Abdelbar, 1998; Abdelbar & Tagliarini, 1996; Abdelbar *et al.*, 2002; Elnabarawy & Abdelbar, 2013; Tsai, 2009; Tsai, 2010) (High Order Network with Exponential SynapTic links) neural network. High Order Neural Networks (HONN) are neural networks which include neurons that combine their inputs nonlinearly, and are thus able to capture nonlinear interactions and codependences among their inputs (Rumelhart *et al.*, 1986). HONEST is a HONN that employs connections with associated adaptable exponents and neurons with product units. In previous work, HONEST has been found to have good generalization capabilities, however, the sensitivity of the adaptable exponents made the network difficult to train and prone to local minima traps (Abdelbar, 1998).

$ACO_{\mathbb{R}}$ (Liao *et al.*, 2014; Socha & Dorigo, 2008) is a fairly-recent Ant Colony Optimization (ACO) algorithm for continuous optimization problems. $ACO_{\mathbb{R}}$ has been applied to several continuous optimization problems, including training MLP neural networks (Socha & Blum, 2007). $ACO_{\mathbb{R}}$ algorithm does not make any use of gradient information, which makes the algorithm robust and applicable to problems where the error function is not differentiable. However, it also means that the algorithm misses out on some important signals that can be obtained from gradient information. Although pure gradient descent methods, such as classical Back-Propagation (Werbos, 1994), can suffer from many problems, including local minima traps, there are methods that perform well by making limited use of gradient information. An example is the Resilient Propagation (R-Prop) algorithm (Reidmiller and Braun, 1993), a robust and powerful technique, that uses only the sign of the partial derivative, and ignores the magnitude.

We present an adaptation of $ACO_{\mathbb{R}}$, called Gradient-based $ACO_{\mathbb{R}}$ ($\mathcal{G}$-$ACO_{\mathbb{R}}$), that uses gradient information to help steer the evolutionary direction of the $ACO_{\mathbb{R}}$ algorithm.

We begin by introducing the HONEST network more fully in Section 2, and presenting the standard $ACO_{\mathbb{R}}$ algorithm in Section 3. Then, in Section 4, we present our $\mathcal{G}$-$ACO_{\mathbb{R}}$ applied to the HONEST network. Experimental results are presented in Section 5, and final remarks are offered in Section 6.

2. THE HONEST NEURAL NETWORK

The HONEST network can be considered to be a generalization of the sigma-pi model (Rumelhart *et al.*, 1986), and is also similar in some ways to the ExpoNet (Narayan, 1993) and GMDH (Ivakhnenko, 1971; Puig *et al.*, 2007) networks. An HONEST network is a feedforward network that always contains exactly three layers—although Tsai (2009; 2010) has considered variations of HONEST that use more layers. Let the external inputs to the network be denoted $x_1, x_2, \ldots, x_n$, let the output of the hidden layer neurons be denoted $h_1, \ldots, h_r$, and let the external output of the network be denoted $y_1, \ldots, y_m$. A connection from an input unit x_j to a hidden neuron h_k does not have an associated weight as in MLP networks, but rather has an associated adaptable exponent p_{kj}. Each hidden unit h_k computes the product of its inputs after first raising each input to the power of the exponent associated with its incoming connection:

Related Content