

# Multi-Objective Training of Neural Networks

**M. P. Cuéllar**

*Universidad de Granada, Spain*

**M. Delgado**

*Universidad de Granada, Spain*

**M. C. Pegalajar**

*Universidad de Granada, Spain*

## INTRODUCTION

Traditionally, the application of a neural network (Haykin, 1999) to solve a problem has required to follow some steps before to obtain the desired network. Some of these steps are the data preprocessing, model selection, topology optimization and then the training. It is usual to spend a large amount of computational time and human interaction to perform each task of before and, particularly, in the topology optimization and network training. There have been many proposals to reduce the effort necessary to do these tasks and to provide the experts with a robust methodology. For example, Giles et al. (1995) provides a constructive method to optimize iteratively the topology of a recurrent network. Other methods attempt to reduce the complexity of the network structure by mean of removing unnecessary network nodes and connections like in (Morse, 1994). In the last years, evolutionary algorithms have been shown as promising tools to solve this problem, existing many competitive approaches in the literature. For example, Blanco et al. (2001) proposed a master-slave genetic algorithm to train (master algorithm) and to optimize the size of the network (slave algorithm). For a general view of the problem and the use of evolutionary algorithms for neural network training and optimization, we refer the reader to (Yao, 1999).

Although the literature about genetic algorithms and neural networks is very extensive, we would like to remark the recent popularity of multi-objective optimization (Coello et al., 2002, Jin, 2006), specially to solve the problem of simultaneous training and topology optimization of neural networks. These methods have shown to perform suitably for this task in previous works, although most of them are proposed for feedforward models. They attempt to optimize the

structure of the network (number of connections, hidden units or layers), while training the network at the same time. Multi-objective algorithms may provide important advantages in the simultaneous training and optimization of neural networks: They may force the search to return a set of optimal networks instead of a single one; they are capable to speed-up the optimization process; they may be preferred to a weight-aggregation procedure to cover the regularization problem in neural networks; and they are more suitable when the designer would like to combine different error measures for the training. A recent review of these techniques may be found in (Jin, 2006).

## BACKGROUND

Multi-objective algorithms have become popular in the last years to solve the problem of the simultaneous training and topology optimization of neural networks, because of the innovations they can provide to solve it. Certain authors have addressed this problem through the evolution of single ensembles as for example with DIVACE-II (Chandra et al., 2006), which also implements different levels of coevolution. In other works, the networks are fully evolved and the evolutionary operators are designed to deal with both training and structure optimization. Some authors have addressed the problem of the structure optimization attending to reduce either the number of network neurons or either the number of network connections. In the first methods (Abbass et al., 2001; Delgado et al., 2005; González et al., 2003), the optimization is easier since the codification of a network contains a smaller number of freedom degrees than the last methods; however, they have a disadvantage in the sense that the networks obtained are fully connected. On the other hand, the methods in

the second place (Jin et al., 2004; Cuéllar et al, 2007) attempt to reduce the number of connections but it is not ensured that also the number of network nodes is also minimum. Nevertheless, experimental results have shown that the networks obtained with these proposals have a low size (Jin et al, 2004).

The hybridation of multi-objective evolutionary algorithms with traditional gradient-based training algorithms has also provided promising results. While the evolutionary algorithm makes a wide exploration of the solution space, the gradient-based algorithms are capable to address the search to promising areas during the evolution and to exploit the solutions suitably. This hybridation is usually carried out by including the gradient-based training method as a local search operator in the evolutionary process. Then, the local search operator is applied after the mutation and before the evaluation of the solutions. Some examples are the system MPANN developed by H.A. Abbass (2001), and the works by Y. Jin et al. (2006).

In the next section, we make an study of different aspects concerning the multi-objective optimization of neural networks. Concretely, we make an study of the objectives to be achieved in the multi-objective algorithm and the multi-objective algorithms used. We focus our analysis on recurrent neural networks (Haykin, 1999; Mandic and Chambers, 2001), since these models have a high complexity due to the recurrence. The experiments are illustrated in problems of time-series prediction, since this type of problems has multiple applications in many research and enterprise areas and the neural models used are suitable for this application, as suggested by previous works (Aussem, 1999).

## MULTI-OBJECTIVE EVOLUTIONARY ALGORITHMS FOR NEURAL NETWORKS TRAINING AND OPTIMIZATION

The most recent multi-objective evolutionary algorithms are based in the concept of Pareto dominance as a criterion to determine whether a solution is optimal or not. Let  $F(s) = (f_1(s), f_2(s), \dots, f_k(s))$  be a set of  $k$  objectives to be achieved, and let  $s_1$  and  $s_2$  be two solutions. In a minimization problem, it is said that  $s_2$  is dominated by  $s_1$  if, and only if:

$$(f_i(s_1) \leq f_i(s_2), \forall i : 1 \leq i \leq k) \wedge (\exists j : f_j(s_1) < f_j(s_2), 1 \leq j \leq k) \quad (1)$$

The solutions that are non-nominated by any other solution are called the *non-dominated set* or *Pareto frontier*. The goal of any multi-objective algorithm is to find the solutions in the Pareto frontier. Thus, the selection of the objectives to be achieved in a multi-objective algorithm is a key aspect, since they will be used to guide the search across the search space to obtain the optimal solutions. However, the higher the number of objectives is, the higher the complexity of the search space is. In this work, we attempt to train and optimize the size of an Elman Network (Mandic and Chambers, 2001), for time series prediction problems. This network type has an input layer, an output layer and a hidden layer. The data of the time series is provided in time to the network inputs, and the objective is the network output to provide the future values of the time series at the output. The recurrent connections are in the hidden layer, so that the output of a hidden neuron at time  $t$  is also input for all the hidden neurons at time  $t+1$ . The reader may found a wider information about dynamical recurrent neural networks applied for time series prediction in (Aussem, 1999; Mandic and Chambers, 2001).

$$f_1(s^*) = \min \{f_1(s)\} = \min \left\{ \frac{1}{T} \sum_{t=1}^T (Y(t) - O(t))^2 \right\} \quad (2)$$

$$f_2(s^*) = \min \{f_2(s)\} = \min \{h(s)\} \quad (3)$$

$$f_3(s^*) = \min \{f_3(s)\} = \min \{n(s)\} \quad (4)$$

For the problem of neural network optimization and training, we consider three objectives to be achieved (see equations (2)-(4)). The objective  $f_1(s)$  attempts to minimize the network error, while  $f_2(s)$  is used to optimize the number of hidden neurons and  $f_3(s)$  the number of network connections. In equation (2),  $T$  is the number of training patterns,  $Y(t)$  is the desired output for pattern  $t$  and  $O(t)$  is the network output. In equation (3),  $h(s)$  is the number of hidden neurons for the network  $s$ ; and  $n(s)$  is the number of network connections in equation (4). Another issue related to the objectives is the network codification. For example, in

5 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/chapter/multi-objective-training-neural-networks/10385](http://www.igi-global.com/chapter/multi-objective-training-neural-networks/10385)

## Related Content

---

### Content-Based Image Classification and Retrieval: A Rule-based System Using Rough Sets Framework

Jafar M. Ali (2007). *International Journal of Intelligent Information Technologies* (pp. 41-58).

[www.irma-international.org/article/content-based-image-classification-retrieval/2422](http://www.irma-international.org/article/content-based-image-classification-retrieval/2422)

### Design, Measurements, and Analysis of Enhanced Bandwidth UWB: On-Body Antenna for Ambient Intelligence Environment

Raghvendra Singh, Kanad Ray, Preecha Yupapin and Jalil Ali (2021). *International Journal of Ambient Computing and Intelligence* (pp. 140-158).

[www.irma-international.org/article/design-measurements-and-analysis-of-enhanced-bandwidth-ubw/272042](http://www.irma-international.org/article/design-measurements-and-analysis-of-enhanced-bandwidth-ubw/272042)

### Digital Transformation of Mining Supply Chains: Strategies and Technologies

Kaushal Kishore Mishra, Pawan Pant, Azmee Zaheer and Swati Priya (2026). *Multimedia and AI Applications for Sustainable Smart Mining and Energy* (pp. 211-246).

[www.irma-international.org/chapter/digital-transformation-of-mining-supply-chains/402944](http://www.irma-international.org/chapter/digital-transformation-of-mining-supply-chains/402944)

### Revolutionizing Special Education: The Impact of AI and Data Science on Personalized Learning

M. Diviya, M. Subramanian, M. Prabu and Thangavel Murugan (2025). *Driving Quality Education Through AI and Data Science* (pp. 73-90).

[www.irma-international.org/chapter/revolutionizing-special-education/370075](http://www.irma-international.org/chapter/revolutionizing-special-education/370075)

### Generating Fully Bounded Chaotic Attractors

Zeraoulia Elhadj (2013). *Investigations into Living Systems, Artificial Life, and Real-World Solutions* (pp. 148-153).

[www.irma-international.org/chapter/generating-fully-bounded-chaotic-attractors/75926](http://www.irma-international.org/chapter/generating-fully-bounded-chaotic-attractors/75926)