Available online at www.sciencedirect.com

ScienceDirect

Expert Systems with Applications 34 (2008) 1630–1641

Expert Systems with Applications

www.elsevier.com/locate/eswa

Optimizing back-propagation networks via a calibrated heuristic algorithm with an orthogonal array

Tai-Yue Wang *, Chien-Yu Huang

National Cheng Kung University, Department of Industrial and Information Management, 1 University Road, Tainan 70101, Taiwan

Abstract

Improving the performance of neural networks is of considerable importance. Although previous studies have investigated how to design the optimal neural network, the heuristic algorithms developed to support the optimization process contain flaws. These heuristic algorithms do not perform efficiently and they require prior expert knowledge. This study commences by employing an orthogonal array using the Taguchi method to calibrate the factor levels of a heuristic algorithm and to estimate the percent contribution from various individual factors. Subsequently, the calibrated heuristic algorithm is used to optimize a back-propagation network (BPN). Changing the level of each individual factor systematically and then analyzing the main and interactive effects of the design factors by using the analysis of variance (ANOVA) leads to the optimal heuristic algorithm factor levels with regard to experimental cost. The proposed optimization procedure is demonstrated on the classification problems using the University of California's Department of Information and Computer Science (ICS) server. The results indicate that the quality of the solution from the proposed approach is superior to that from a non-calibrated conventional design.

Keywords: Back-propagation network; Heuristic algorithm; Orthogonal array

1. Introduction

The back-propagation network (BPN) is the most popular form of artificial neural network (ANN), and has been successfully applied to solve many problems (Lee, Booth, & Alam, 2005; Liu, Liu, & Zhang, 2001; Mendyk & Jachowicz, 2005; Quah & Srinivasan, 1999; Salchenberger, Cinar, & Lash, 1992). Notwithstanding its popularity, traditional BPNs using the Gradient Steepest Descent Learning Algorithm (GSDLA) may be inconsistent and unpredictable for certain applications (Sexton, Dorsey, & Johnson, 1998a). This kind of BPN suffers two major shortcomings: (1) the GSDLA sometimes converges slowly to the optimal solution, and (2) the GSDLA may yield a poor solution since it can become trapped at a locally optimized solution (Yasumasa et al., 1994). Hence, improving the application performance of BPNs remains an important research issue.

Search techniques such as Genetic Algorithms (GA), Tabu Search (TS), and Simulated Annealing (SA), represent potential means of improving the performance of BPNs. It has been shown that these techniques are particularly effective in optimizing the BPN performance (Sexton, Alidance, & Dorsey, 1998b; Sexton, Dorsey, & Johnson, 1999; Yasumasa et al., 1994). Sexton used GA, TS, and SA for optimizing BPN, and has validated the priority of the BPN together with a heuristic algorithm. Moreover, in 1999, Sexton compared the relative performances of SA and GA in optimizing an ANN for a specific BPN topology (after a lot of trial-and-error). Optimizing the BPN parameters and topology is crucial to enhancing performance, since both aspects significantly influence the quality of the solution.

Many heuristic algorithms require the factor levels to be optimized in order to improve performance. Clearly, different factor levels of a heuristic algorithm will influence the

---

[*] Corresponding author. Tel.: +886 6 2757575; fax: +886 6 2362162.
E-mail address: tywang@mail.ncku.edu.tw (T.-Y. Wang).

0957-4174/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2007.01.013

BPN performance (Kirkpatrick, Gelatt, & Vecchi, 1983; Sexton et al., 1999). Hence, it is necessary to calibrate the influential heuristic algorithm's factors before the algorithm is applied to the BPN optimization task. It has been shown that the Taguchi method (Taguchi, 1986) is an effective tool for factor calibration (Forouraghi, 2000; Khaw, Lim, & Lim, 1995; Ko, Kim, & Kim, 1999), and Taguchi's orthogonal array can be applied to establish the optimal factor levels of a heuristic algorithm. Khoei, Masters, and Gethin (2002) presented an experimental investigation of the aluminum recycling process, in which the Taguchi method was used to determine the optimal configuration of process parameters such that the process performance and quality were both enhanced. Casab, Orsolya, Anna, Eya, and Lstyan (1999) also applied the Taguchi method in the field of ELISA (Enzyme Linked Immunosorbent Assay) optimization, and successfully reduced the interaction effects of the optimized variables such that it was possible to identify the optimum conditions, even when significant interactions existed between the assay variables.

Nowadays, the Taguchi method is the method most commonly chosen when analyzing interaction effects and calculating the percent contribution of separate factors in order to screen and rank them (Roy, 1990). Furthermore, it is also an appropriate approach for solving problems characterized by continuous, discrete, and qualitative design variables (Lin & Tseng, 2000). Consequently, in this paper, we employ the signal to noise ratio (S/N), the analysis of variance (ANOVA), and the analysis of means (ANOM) from the Taguchi method (Nelson & Dudewicz, 2002; Roy, 1990) to investigate the main effects and interactions of a heuristic algorithm's factor levels. We anticipate that the optimal combination of the adopted heuristic algorithm factors and the appropriate levels of each of these factors will yield a superior result for the BPN's performance.

The present study performs a three-step investigation into the BPN optimization procedure; namely (1) an orthogonal array of the Taguchi method is employed to identify the factor levels for a heuristic algorithm, (2) the calibrated heuristic algorithm is used to design the BPN's parameters and topology in order to enhance its performance, and (3) the ANOVA of the Taguchi method is used to estimate individual percent contributions in order to rank and screen these controllable factors of the adopted heuristic algorithm.

Fig. 1 provides an outline of the current research, which can be broadly summarized as follows:

Step 1. *Identify the BPN's decision parameters and topology*: The decision parameters and topology of a BPN include the initial weight values factor, learning rate, momentum factor, number of hidden layers, number of neurons in the first layer, and number of neurons in the second layer. After that, the calibrated heuristic algorithm is used to specify appropriate network values for the BPN.

Step 2. *Calibrate the heuristic algorithm factor levels*: A suitable orthogonal array is employed to examine the levels of the heuristic algorithm factors in order to ensure the robustness of the experimental design.

Step 3. *Rank and screen factors of the heuristic algorithm*: To estimate the percent contributed by individual factors, information about the relative merits of individual factors is obtained using the analysis of variance of the experimental results.

Step 4. *Pool factors of the heuristic algorithm*: In order to reduce experimental cost and difficulty, it may be necessary to pool one or more of the controllable factors of the heuristic algorithm during the ranking and screening processes.

Step 5. *Optimize the BPN parameters and topology*: The calibrated heuristic algorithm is used to optimize the BPN in order to enhance the BPN's application performance.

Step 6. *Compare the performance of different factors of the heuristic algorithm*: The solution quality of the BPN is compared when optimized by a heuristic algorithm with various combinations of calibrated factor levels.

2. Methodology

This study employs an orthogonal array to screen and rank the heuristic algorithm factors by exploring their main and interactive effects. The appropriate quantitative factors and their permissible levels are first determined. Then an orthogonal array is used to establish relevant design treatment combinations. The ranking and screening of these factors are then achieved by analyzing the response table, which compiles the relevant data generated from each trial run of the design array. This approach enables the appropriate heuristic algorithm factor levels to be determined while simultaneously minimizing experimental costs.

2.1. Identifying the BPN's decision parameters and topology

Before using the calibrated heuristic algorithm to optimize the BPN parameters and topology, it is first necessary to identify the BPN's decision parameters and topology for working performance. In most cases, one hidden layer is sufficient to compute arbitrary decision boundaries for the outputs. However, some researchers have used a BPN with two hidden layers for more complicated applications. The number of neurons in hidden layers in the BPN is clearly related to the number of hidden layers used in the BPN (Khaw et al., 1995; Sexton et al., 1998a; Zurada, 1995). As for the number of neurons and number of hidden layers required, the suggestions of previous studies (Khaw et al., 1995; Zurada, 1995) were followed. In the current BPN used, the maximum number of hidden layers is set to 2 and the maximum number of neurons in each hidden layer is specified to be 63 for the search space of the BPN.