The Eighteenth International Symposium on Artificial Life and Robotics 2013 (AROB 18th '13),
Daejeon Convention Center, Daejeon, Korea, January 30-February 1, 2013

Adaptive learning neural networks for system identification of a magneto-rheological damper

Chia-Nan Ko[1], Guan-Yu Liu[2], Yu-Yi Fu[3], and Pi-Yun Chen[4]

[1,2,3]Nan-Kai University of Technology, Taiwan
[4]National Chin-Yi University of Technology, Taiwan

[1]t105@nkut.edu.tw

Abstract: In this article, adaptive learning neural networks (ALNNs) are proposed to identify nonlinear systems. In the proposed NNs, integrating support vector regression (SVR) and adaptive learning algorithm is adopted to optimize the structure of neural networks. In the evolutionary procedure, first, SVR is adopted to determine the number of hidden layer nodes and the initial structure of the NNs. After initialization, adaptive learning algorithm (ALA) with nonlinear time-varying learning rate is then applied to train NNs. In ALA, a computationally efficient optimization method, particle swarm optimization (PSO) method, is adopted to simultaneously find optimal learning rates. Due to the advantages of SVR and adaptive learning algorithm, the proposed NNs (SVR-ALNNs) have good performance for identifying a magneto-rheological (MR) damper system. Simulation results are illustrated the feasibility and superiority of the proposed SVR-ALNNs.

Keywords: adaptive learning neural networks, support vector regression, magneto-rheological damper system.

1 INTRODUCTION

Due to the characteristic of structure simplicity, low power requirement, large force capacity, and high dynamic range, a magneto-rheological (MR) damper attracts attentions as developed semi-active control devices for structural control applications recently. In order to describe the performance of an MR damper, Spencer et al. in 1997 proposed a phenomenological model based on a Bouc-Wen model [1], Chang and Chang and Zhou [2] developed a neural network model, Zhou and Chang [3] proposed an adaptive fuzzy control, and Du et al. designed a model of an MR damper by evolving radial basis function networks [4]. A nonlinear black-box model was proposed to identify an MR damper system to design a force-sensorless control method [5].

In this article, adaptive learning neural networks (ALNNs) are proposed to identify an MR damper system. In the proposed NNs, integrating support vector regression (SVR) and adaptive learning algorithm is adopted to optimize the structure of neural networks. In the evolutionary procedure, first, SVR [6, 7] is adopted to determine the number of hidden layer nodes and the initial structure of NNs. After initialization, ALA with nonlinear time-varying learning rate is then applied to train NNs. In the adaptive learning algorithm, a computationally efficient optimization method, particle swarm optimization (PSO) method, is adopted to simultaneously find optimal learning rates. Due to the advantages of SVR and adaptive learning algorithm, the proposed NNs (SVR-ALNNs) have good performance for identifying the MR damper system. Simulation results are illustrated the feasibility and superiority of the proposed SVR-ALNNs.

2 SVR-BASED INITIAL NNs

Neural networks can be used to estimate the input–output relation of a nonlinear system. In this paper, radial basis function neural networks (RBFNNs) are adopted because they have a simple structure. When the Gaussian function is chosen as the radial basis function, RBFNNs can be expressed in the form

$$\hat{y}_j(t+1) = \sum_{i=1}^{L} w_{ij} \exp\left(-\frac{\left\| \hat{\mathbf{x}} - \mathbf{m}_i \right\|^2}{2\sigma_i^2} \right) \quad \text{for} \quad j = 1, 2, \cdots, p \tag{1}$$

where $\hat{\mathbf{x}}(t) = \left[\hat{x}_1(t) \quad \cdots \quad \hat{x}_m(t) \right]^{\mathrm{T}}$ is the input vector; $\hat{\mathbf{y}}(t) = \left[\hat{y}_1(t) \quad \cdots \quad \hat{y}_p(t) \right]^{\mathrm{T}}$ is the output vector; w_{ij} is the synaptic weight; $\mathbf{m}_i$ and σ_i are the center and width of Gaussian functions respectively; and L is the number of Gaussian functions, which is also equal to the number of hidden layer nodes.

It is very difficult, if not impossible, to solve the above problem directly. Usually, the initial values of L, w_{ij}, $\mathbf{m}_i$, and σ_i are chosen first. Then a training algorithm is applied to the NNs to search for the optimal combination of these values in an iterative manner.

In support vector machine (SVM), an SVR method can approximate an unknown function. An output of the RBFNNs and its corresponding training pairs will be used for simulations. Meanwhile, assume that a set of basis functions, $g_l(\mathbf{x})$, $l = 1, 2, \cdots, M$, is given. Then the problem of funct-ion approximation is transformed into one of finding the parameters of the basis linear expansion.

$$f(\mathbf{x}, \boldsymbol{\theta}) = \sum_{l=1}^{M} \theta_l g_l(\mathbf{x}) + b \tag{2}$$

where $\boldsymbol{\theta}$ is a parameter vector to be identified and b is a constant to be determined.

From [8], one sees that the solution is to find $f(\mathbf{x}, \boldsymbol{\theta})$ that minimizes

$$R(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N} L_{\varepsilon}(y_i - f(\mathbf{x}_i, \boldsymbol{\theta})) \tag{3}$$

subject to the constraint

$$\|\boldsymbol{\theta}\|^2 < C \tag{4}$$

where $L_{\varepsilon}(\cdot)$ is the ε-insensitive loss function defined as

$$L_{\varepsilon}(e) = \begin{cases} 0 & \text{for } |e| \le \varepsilon \\ |e| - \varepsilon & \text{otherwise} \end{cases} \tag{5}$$

for some non-negative numbers ε.

By using the Lagrange multiplier method, it has been shown in [6] that the minimization of (3) leads to a dual optimization problem. A set of basis functions $g_l(\mathbf{x}_r)$ is replaced via the kernel function.

$$K(\mathbf{x}_r, \mathbf{x}_s) = \sum_{l=1}^{M} g_l(\mathbf{x}_r) g_l(\mathbf{x}_s) \tag{6}$$

Then the dual optimization problem can be expressed as follows:
Minimize

$$\begin{aligned} Q(\alpha, \alpha^*) = \varepsilon\sum_{r=1}^{N}(\alpha_r + \alpha_r^*) - \sum_{r=1}^{N} y_r(\alpha_r - \alpha_r^*) \\ + \frac{1}{2}\sum_{r,s=1}^{N}(\alpha_r^* - \alpha_r)(\alpha_s^* - \alpha_s)K(\mathbf{x}_r, \mathbf{x}_s) \end{aligned} \tag{7}$$

subject to the constraint

$$\sum_{r=1}^{N}\alpha_r^* = \sum_{r=1}^{N}\alpha_r \quad 0 < \alpha_r^*, \alpha_r < C \tag{8}$$

and the solution is in the form of the following linear expansion of kernel functions [6]

$$f(\mathbf{x}, \alpha, \alpha^*) = \sum_{l=1}^{M}(\alpha_l^* - \alpha_l)K(\mathbf{x}, \mathbf{x}_l) + b \tag{9}$$

Note that some of the $(\alpha_l^* - \alpha_l)$ values are not zeros and the corresponding vectors $\mathbf{x}_l$ are called support vectors (SVs).

In this study, since the Gaussian function is used as the kernel function, (9) can be rewritten as

$$f(\mathbf{x}, \lambda) = \sum_{l=1}^{\#SV} \lambda_l \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_l\|^2}{2\sigma_l^2}\right) + b \tag{10}$$

where l is the number of SVs, $\lambda_l = (\alpha_l^* - \alpha_l) \neq 0$, and $\mathbf{x}_l$ denotes SVs. Comparing (10) with (1), $\#SV$ l, λ_l, and $\mathbf{x}_l$ in (10) can be regarded as L, i, w_{ij}, and $\mathbf{m}_i$ in (1), respectively.

3 ADAPTIVE LEARNING NNs

To overcome the existing problems in robust backpropagation learning algorithms, the annealing robust learning algorithm (ARLA) is adopted in the training procedure of the presented NNs by Chung et al. [7]. A cost function for the annealing learning is defined here as

$$J_j(t) = \frac{1}{N}\sum_{k=1}^{N}\rho\left[e_j^{(k)}(t); \beta(t)\right] \text{ for } \quad j = 1, 2, \cdots, p \tag{11}$$

where

$$e_j^{(k)}(t) = y_j^{(k)} - \sum_{i=1}^{L} w_{ij} \exp\left(-\frac{\left\|x^{(k)} - m_i\right\|^2}{2\sigma_i^2}\right) \tag{12}$$

t is the epoch number, $e_j^{(k)}(t)$ is the error between the kth desired output $y_j^{(k)}$ of the annealing learning NNs at epoch t for the jth input–output training data in an identification system, $\beta(t)$ is a deterministic annealing schedule acting as the cut-off point, and $\rho(\cdot)$ is a logistic loss function defined as

$$\rho[e_j^{(k)}; \beta] = \frac{\beta}{2}\ln\left[1 + \frac{(e_j^{(k)})^2}{\beta}\right] \text{ for } \quad j = 1, 2, \cdots, p \tag{13}$$

To evaluate the performance of training NNs, the root mean square error (*RMSE*) of the training data is adopted and defined as

$$RMSE = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(e_j^{(k)}\right)^2} \quad \text{for} \quad j = 1, 2, \cdots, p \tag{14}$$

Based on a gradient descent type of learning algorithm, the synaptic weights w_{ij}, the centers $\mathbf{m}_i$, and the widths σ_i of Gaussian functions are updated, respectively, as

$$\Delta w_{ij} = -\eta_w \frac{\partial J_j}{\partial w_{ij}} = -\frac{\eta_w}{N}\sum_{k=1}^{N}\varphi_j(e_j^{(k)}; \beta)\frac{\partial e_j^{(k)}}{\partial w_{ij}} \tag{15}$$

$$\Delta \mathbf{m}_i = -\eta_m \frac{\partial J_j}{\partial \mathbf{m}_i} = -\frac{\eta_m}{N}\sum_{j=1}^{p}\sum_{k=1}^{N}\varphi_j(e_j^{(k)}; \beta)\frac{\partial e_j^{(k)}}{\partial \mathbf{m}_i} \tag{16}$$

$$\Delta \sigma_i = -\eta_\sigma \frac{\partial J_j}{\partial \sigma_i} = -\frac{\eta_\sigma}{N}\sum_{j=1}^{p}\sum_{k=1}^{N}\varphi_j(e_j^{(k)}; \beta)\frac{\partial e_j^{(k)}}{\partial \sigma_i} \tag{17}$$

$$\varphi_j(e_j^{(k)}; \beta) = \frac{\partial \rho(e_j^{(k)}; \beta)}{\partial e_j^{(k)}} = \frac{e_j^{(k)}}{1 + (e_j^{(k)})^2 \big/ \beta(t)} \tag{18}$$

where η_w, η_m, and η_σ are the learning rates for the synaptic weights w_{ij}, the centers $\mathbf{m}_i$, and the widths σ_i, respectively, and $\varphi(\cdot)$ is usually called the influence function. The adaptive annealing schedule has the following properties [7]:

(A) β_{initial}, $\beta(t)$ for the first epoch, has large values.

(B) $\beta(t) \to 0^+$ for $t \to \infty$.

In the ALA, a nonlinear time-varying evolution concept is adopted over iterations, in which the learning rates η_w, η_m, and η_σ with a high value $\eta_{\max}$ and nonlinearly decreases to

$\eta_{\min}$ at the maximal number of epochs, respectively. This means that the mathematical expressions are given as shown as

$$\eta_w = \eta_{\min} + \left(1 - \frac{h}{epoch_{\max}}\right)^{pw} (\eta_{\max} - \eta_{\min}), \quad (19)$$

$$\eta_m = \eta_{\min} + \left(1 - \frac{h}{epoch_{\max}}\right)^{pm} (\eta_{\max} - \eta_{\min}), \quad (20)$$

$$\eta_\sigma = \eta_{\min} + \left(1 - \frac{h}{epoch_{\max}}\right)^{p\sigma} (\eta_{\max} - \eta_{\min}), \quad (21)$$

where $epoch_{\max}$ is the maximal number of epochs and h is the current number of epochs. In the updated procedure, appropriate functions for the learning rate and η_w, η_m, and η_σ can promote the performance of ALNNs. However, simultaneously determining the optimal combination of pw, pm, and $p\sigma$ is a time-consuming work. An efficient evolutionary algorithm, PSO, will be adopted to determine the optimal solution $(pw, pm, p\sigma)$.

4 SIMULATION RESULTS

The identification scheme of a nonlinear system is depicted in Fig. 1, training input-output data are obtained by feeding a signal $\mathbf{x}(k)$ to the system and measure its corresponding output $\mathbf{y}(k+1)$ Then subject to the same input signal, the objective of identification is to construct an SVR-ALNNs using PSO method, which produces an output $\hat{\mathbf{y}}(k+1)$ to approximate $\mathbf{y}(k+1)$ as closely as possible.

In this section, a magneto-rheological (MR) damper system adopted to verify the feasibility of the proposed SVR-ALNNs. When applying the proposed algorithm, the population size, the maximal iteration number, and the maximal epoch number are chosen to be 30, 100, and 1000, respectively. The variables pw, pm, and $p\sigma$ in learning rate functions (19) to (21) are all chosen as real numbers in the range $[0.1, 0.5]$. Meanwhile, the values of $\eta_{\max}$ and $\eta_{\min}$ are set as 4.0 and 1.0, respectively.

***Example*:**

A phenomenological model has been proposed by Spencer [1] to portray the behaviour of a prototype MR damper. This phenomenological model is based on a Bouc–Wen model, the model as shown in Fig. 2, and is governed by the following seven simultaneous equations:

$$F = c_1\dot{y} + k_1(x - x_0), \quad (22)$$

$$\dot{y} = \frac{1}{(c_0 + c_1)}[\alpha z + c_0\dot{x} + k_0(x - y)], \quad (23)$$

$$\dot{z} = -\gamma|\dot{x} - \dot{y}|z|z|^{n-1} - \zeta(\dot{x} - \dot{y})|z|^n + A(\dot{x} - \dot{y}), \quad (24)$$

$$\alpha = \alpha_a + \alpha_b u, \quad (25)$$

$$c_1 = c_1^a + c_1^b u, \quad (26)$$

$$c_0 = c_0^a + c_0^b u, \quad (27)$$

$$\dot{u} = -\lambda(u - v), \quad (28)$$

where F is the force generated by the MR damper; x is the displacement of the damper; y is an internal pseudo displacement of the MR damper; u is the output of a first order filter; v is the command voltage sent to the current driver; k_1 is the accumulator stiffness; c_0 and c_1 are the viscous damping coefficients observed at large and low velocities, respectively; k_0 is the gain to control the stiffness at large velocities; x_0 is the initial displacement of spring k_1 associated with the nominal damper force due to the accumulator; γ, ζ, A are hysteresis parameters for the yield element, and α is the evolutionary coefficient. A set of parameters which is obtained by Spencer to characterize one MR damper using experimental data and a constrained nonlinear optimization algorithm is listed in Table 1.

The data is generated using a sinusoidal displacement function with an amplitude of ±1 cm and a frequency of 3Hz and a sinusoidal voltage function with mean value of 1.6 V, an amplitude of 0.5V and a frequency of 0.5 Hz. The time duration for this validation data is 4s and the time increment is 0.002s which amounts to a total of 2000 training data sets.

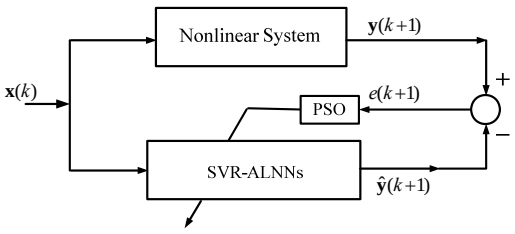

Fig. 1. The proposed PSO-based SVR-ALNNs scheme for identification of a nonlinear system.

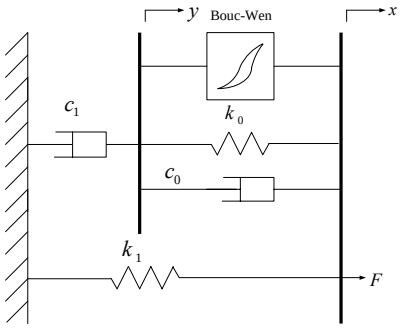

Fig. 2. Phenomenological model of MR damper.

Table 1. Parameters for an MR damper

Parameter	Value	Parameter	Value
c_0^a	$21.0 Ns/cm$	α_a	$140 N/cm$
c_0^b	$3.50 Ns/cm\ V$	α_b	$695\ N/cm\ V$
k_0	$46.9 N/cm$	γ	$363\ cm^{-2}$
c_1^a	$283 Ns/cm$	ζ	$363\ cm^{-2}$
c_1^b	$2.95 Ns/cm\ V$	n	2
k_1	$5.00 N/cm$	λ	$190\ s^{-1}$
x_0	$14.3\ cm$	A	301

***Problem 1*:**
In this investigation, when applying SVR, various parameters of $\varepsilon = 0.2$ and $\varepsilon = 0.1$ in (7) for $C = 100$ in (8) are adopted to determine the initial structure of NNs, respectively. In the ARLA, various learning rates, $1.0 \le \eta \le 4.0$, are used to train the NNs. After 1000 training epochs, the *RMSE* values for various learning rates are obtained, respectively. The details of the simulation results are shown in Table 2.

***Problem 2*:**
With the nonlinear learning rates, the ALA is adopted to train NNs, in which the optimal learning rates are determined by linear time-varying evolution PSO method [9]. The optimal sets are obtained as (pw, pm, $p\sigma$)=(4.3658, 0.01, 1.3637) and (pw, pm, $p\sigma$) =(2.4125, 0.1035, 3.0082) for $\varepsilon = 0.2$ and $\varepsilon = 0.1$ respectively. Meanwhile, the final values of RMSE with ALNNs are found to be 0.0070, and 0.0038 shown in Table 2. To show the superiority of the proposed SVR-ALNNs, the comparison of the proposed NNs with SVR-ARLA-NNs has been shown in Table 2. From Table 2, the proposed SVR-ALNNs have obtained promising results after only 1000 training epochs.

5 CONCLUSIONS

This paper presents the integrating SVR and ALA to train NNs for identification of an MR damper system. With the initial structure of the NNs using SVR method, the nonlinear time varying learning rates are simultaneously determined by PSO method to perform the ALA. Then the optimal NNs are obtained to identify the MR damper system. From the simulation results, one can conclude that the proposed SVR-ALNNs using PSO method have good performance for identification of the MR damper system using only few training epochs. Meanwhile, the superiority of the proposed SVR-ALNNs with nonlinear learning rates over SVR-ARLA-NNs with fixed learning rates for identification has been verified. The proposed SVR-ALNNs can be further extended to identify more complex systems.

Table 2. The values of *RMSE* (14) for an MR damper system after 1000 training epochs, in which ARLA with various learning rates and ALA with time-varying learning rates are applied to train NNs.

ε	ALA	ARLA (learning rate η)						
		4.0	3.5	3.0	2.5	2.0	1.5	1.0
0.2	**0.0070**	0.0092	0.0090	0.0114	0.0083	**0.0081**	0.0101	0.0152
0.1	**0.0038**	0.0056	0.0058	**0.0049**	0.0062	0.0080	0.0083	0.0111

ACKNOWLEDGMENT

This work was supported in part by the National Science Cou ncil, Taiwan, R.O.C., under grants NSC 99-2221-E-252-017.

REFERENCES

[1] Spencer Jr BF, Dyke SJ, Sain MK, Carlson JD (1997), Phenomenological model for magnetorheological dampers. Journal of Engineering Mechanics 123(3):230-238

[2] Chang CC, Zhou L (2002), Neural network emulation of inverse dynamics for a magnetorheological damper. Journal of Structural Engineering 128(2):231-239

[3] Zhou L, Chang CC (2003), Adaptive fuzzy control for nonlinear building-magnetorheological dramper system. Journal of Structural Engineering 129(7):905-913

[4] Du H, Lam J, Zhang N (2006), Modelling of a magneto-rheological damper by evolving radial basis function networks. Engineering Applications of Artificial Intelligence 19:869-881

[5] Truong DQ, Ahn KK (2011), Nonlinear black-box models and force-sensorless damping control for damping systems using magneto-rheological fluid dampers. Sensors & Actuators: A. Physical 167(2):556-573

[6] Vapnik VN (1995), The Nature of Statistical Learning Theory. Springer, New York

[7] Chuang CC, Jeng JT, Lin PT (2004), Annealing robust radial basis function networks for function approximation with outliers. Neurocomputing 56:123-139

[8] Pei L, Liu JB, Guinness R, Chen YW, Kuusniemi H, Chen RZ (2012), Using LS-SVM based motion recognition for smartphone indoor wireless positioning. Sensors 12:. 6155-6175

[9] Ratnaweera A, Halgamuge SK, Watson C (2004), Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. IEEE Transactions on Evolutionary Computation 8(3):240-255