Neural networks letter

Quantum inspired PSO for the optimization of simultaneous recurrent neural networks as MIMO learning systems

Bipul Luitel, Ganesh Kumar Venayagamoorthy *

Real-Time Power and Intelligent Systems Laboratory, Missouri University of Science Technology, Rolla, MO-65409, USA

ABSTRACT

Training a single simultaneous recurrent neural network (SRN) to learn all outputs of a multiple-input–multiple-output (MIMO) system is a difficult problem. A new training algorithm developed from combined concepts of swarm intelligence and quantum principles is presented. The training algorithm is called particle swarm optimization with quantum infusion (PSO-QI). To improve the effectiveness of learning, a two-step learning approach is introduced in the training. The objective of the learning in the first step is to find the optimal set of weights in the SRN considering all output errors. In the second step, the objective is to maximize the learning of each output dynamics by fine tuning the respective SRN output weights. To demonstrate the effectiveness of the PSO-QI training algorithm and the two-step learning approach, two examples of an SRN learning MIMO systems are presented. The first example is learning a benchmark MIMO system and the second one is the design of a wide area monitoring system for a multimachine power system. From the results, it is observed that SRNs can effectively learn MIMO systems when trained using the PSO-QI algorithm and the two-step learning approach.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Simultaneous recurrent neural networks (SRNs) are a class of neural network architectures where the recurrence is instantaneous (Geib & Serpen, 2004). SRNs are appropriate for approximating complex nonlinear systems with fewer neurons because they model the response of a dynamic nonlinear system even with fixed weights. Furthermore, SRNs have the capability of approximating non-smooth functions which cannot be approximated by conventional Multilayer Perceptrons (MLPs). An Elman SRN has its feedback from the hidden layer output to the context layer inputs, and in this study it is represented in vector notation as:

$$H(k, n) = f(A * I(k) + B * H(k, n - 1) + K) \tag{1}$$

$$O(k) = g(C * H(k, R) + K') \quad n = 1, 2, \ldots, R \tag{2}$$

where I is the set of inputs, H is the set of neuron outputs from the hidden layer and O is the set of outputs from the output layer. A is the set of weights from the input layer to the hidden layer, B is the set of weights from the context layer to the hidden layer, C is the set of weights from the hidden layer to the output layer, n is the index of internal recurrence, k is the index of the input sample, R is the number of internal recurrences, K and K' are the biases, and

f and g are neuron activation functions in the hidden and output layers respectively.

Another important feature of recurrent neural networks is their ability to implement associative memory (Michel & Farrell, 1990). Various types of neural networks have been studied for associative memory (Kwan, 2002). Unlike other associative memories that store patterns, this work demonstrates SRNs store dynamics which can be stimulated by excitations of similar dynamics. The feedback connectivities between the hidden/output and input layers in SRNs store information that potentially acts as associative memories.

Although SRNs are powerful neural networks, the training process is intensive and more difficult when there are multiple outputs to learn, i.e. learning of multiple-input–multiple-output (MIMO) system. Traditional training algorithms such as backpropagation through time suffer from local minima and hence it is hard to train SRNs using these techniques because of the recursive calculations involved (Cai, Prokhorov, & Wunsch, 2007). Computational intelligence (CI) based algorithms have gained popularity in the training of neural networks because of their ability to find a global solution in a multi-dimensional search space. Swarm and evolutionary based algorithms such as Particle Swarm Optimization (PSO) (Del Valle, Venayagamoorthy, Mohagheghi, Hernandez, & Harley, 2008) have shown promises in the training of SRNs. In this study, the quantum principle obtained from Quantum PSO (QPSO) (Sun, Feng, & Xu, 2004) has been combined with standard PSO to form a new hybrid algorithm called PSO with Quantum Infusion (PSO-QI). For training, a two step learning approach is introduced to improve the ability of SRNs to learn multiple outputs.

* Corresponding author.
E-mail addresses: iambipul@ieee.org (B. Luitel), gkumar@ieee.org
(G.K. Venayagamoorthy).

2. PSO-QI algorithm and two-step learning

2.1. Particle swarm optimization with quantum infusion

PSO-QI is a hybrid algorithm that uses the quantum principle from QPSO to create a new offspring in PSO. After the positions and velocities of the particles are updated using standard PSO equations, a randomly chosen particle from PSO's *pbest* (the previous particle position giving the best fitness value) population is utilized to carry out the quantum operation; and thus, create an offspring by mutating the *gbest* (the best particle among all the particles in the swarm). The fitness of the offspring is evaluated and the offspring replaces the *gbest* only if it has a better fitness. This ensures that the fitness of the *gbest* is equal to or better than its fitness in the previous iteration. Thus, it is improved and pulled toward the best solution over iterations.

According to the uncertainty principle, the position and velocity of a particle in the quantum world cannot be determined simultaneously. Thus QPSO differs from standard PSO mainly in the fact that exact values of x and v cannot be determined. Hence the probability of finding a particle at a particular position in the quantum search space is mapped into its actual position in the solution space by a technique called "collapsing". In Quantum Delta-Potential-Well based PSO (QDPSO) (Sun et al., 2004), a delta potential well based probability density function is used to avoid explosion and help the particles converge. By using Monte Carlo Simulation (Sun et al., 2004), the position equation in QDPSO is given by (3):
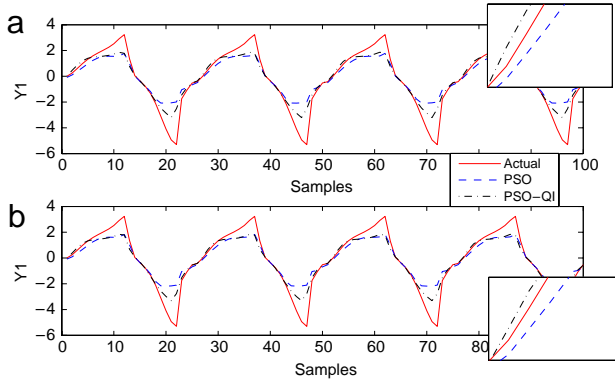
$$x(k) = J(k) \pm \frac{L(k)}{2} \ln(1/$$

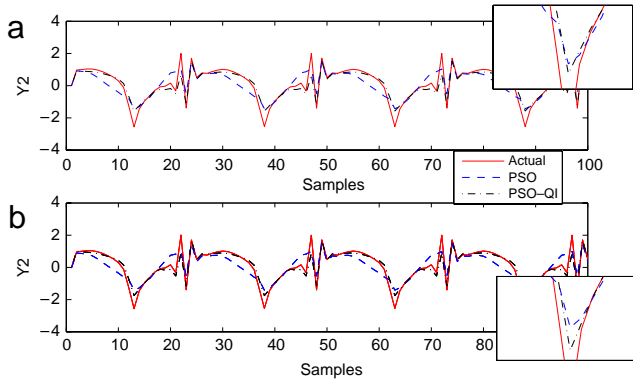**Fig. 1.** Output of SRN for Y1 in (a) Step 1 and (b) Step 2.

Fig. 2. Output of SRN for Y2 in (a) Step 1 and (b) Step 2.

Table 1
MSE results for the case studies.

		PSO		PSO-QI	
		Step 1	Step 2	Step 1	Step 2
Case I	Y1	1.2494	1.1487	0.7132	0.6407
	Y2	0.3150	0.2996	0.1427	0.1012
Case II	G1	0.0147	0.0075	0.0147	0.0057
	G2	0.0281	0.0111	0.0218	0.0068
	G3	0.0225	0.0118	0.0181	0.0067
	G4	0.0344	0.0167	0.0319	0.0105

where $u_i(k)$ and $y_i(k)$ are the ith input and output at instant k. The inputs to the SRN are the present values of system inputs and outputs. The SRN outputs are the one-step ahead prediction of the system outputs. An SRN of 4 inputs, 10 hidden nodes and 2 output nodes is then trained on 100 samples of uniform random numbers between [0, 1] and tested on 100 data samples of input vector $[\sin(2\pi k/25), \cos(2\pi k/25)]$. The dimension of the SRN training problem is 160 and 10 in Steps 1 and 2 respectively. The testing results obtained are shown in Figs. 1 and 2. The MSEs are shown in Table 1.

3.2. Case II: Design of a wide area monitoring system

In the second case study, a wide area monitor (WAM) used for the two area four machine power system described in Venayagamoorthy (2007), and shown in Fig. 4, is considered. The WAM is modeled by an SRN with 8 input nodes, 15 hidden nodes and 4 output nodes (405 weights). The inputs to the SRN are the current deviations in reference voltage V_{ref}, caused by the pseudorandom binary signal excitations and speed deviations of the four machines. The outputs are the one-step ahead predictions of their speed deviations. The operating conditions are similar to
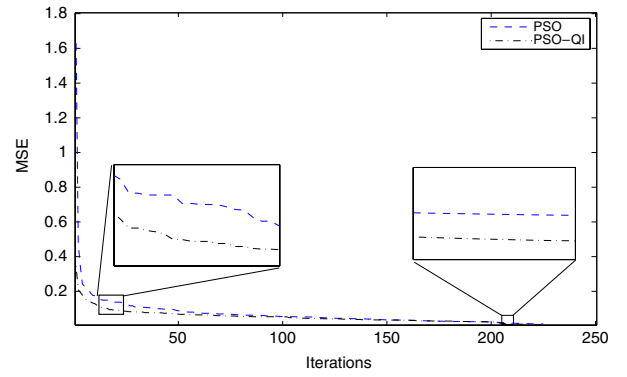

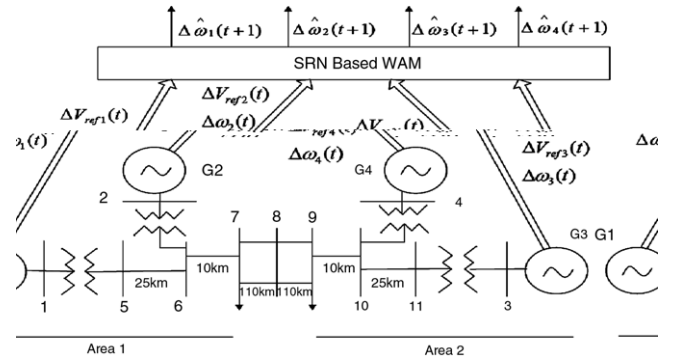**Fig. 3.** Comparison of fitnesses from PSO and PSO-QI algorithms.

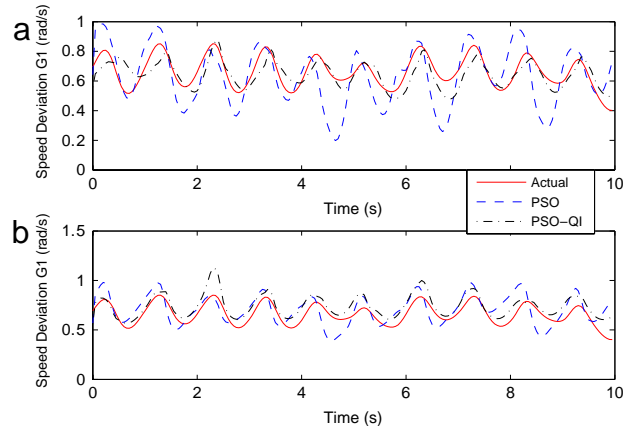**Fig. 4.** Wide area monitor in a two area four machine system.

Fig. 5. SRN output for G1 (Area 1) in (a) Step 1 and (b) Step 2.

(Venayagamoorthy, 2007). The loads are modified to have a power transfer of 510 MW during training and 458 MW during testing from Area 1 to 2. The voltage of all the generators is 1.03 pu in both operating conditions. The dimension of the SRN training problem is 405 and 15 in Step 1 and 2 respectively. The outputs of the SRNs for two generators (G1 and G4) are shown in Figs. 5 and 6. The fitness curve obtained for the two step process is shown in Fig. 3. The results show the improvement from PSO to PSO-QI and from Step 1 (first 200 iterations) to Step 2 (last 25 iterations). The MSEs are shown in Table 1.

4. Conclusion

SRNs learning MIMO systems using a PSO-QI training and a two-step learning approach have been presented. The results of this study show that hard-to-train MIMO SRNs can be successfully
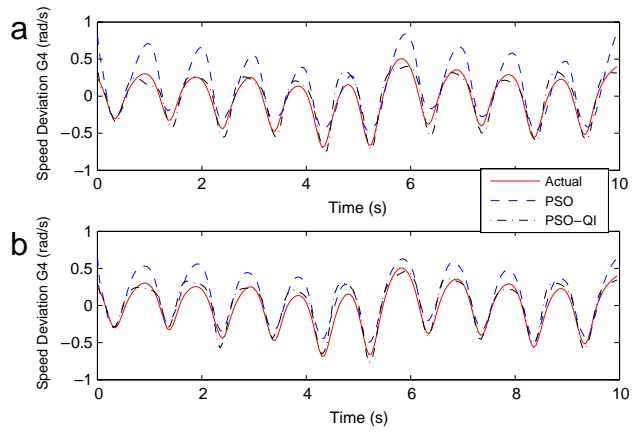
Fig. 6. SRN output for G4 (Area 2) in (a) Step 1 and (b) Step 2.

trained with better accuracy. Due to dimension and training time reductions from Step 1 to 2 in the two-step learning approach, SRN training can successfully cope with multiple inputs and outputs; in other words are scalable. Explicit studies on the associative memory property of ing   55saveertybe,tive