11 October 1994
Electronics Letters Online No: 19950077

M. Arai and J.G.P Binner (*Department of Materials Engineering and Materials Design, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom*)

T.E. Cross (*Department of Electrical and Electronic Engineering, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom*)

References

1 ATHEY, T.W., STUCHLY, M.A., and STUCHLY, S.S.: 'Measurement of radio frequency permittivity of biological tissues with an open-ended coaxial line : Part 1', *IEEE Trans.*, 1982, **MTT-30**, (1), pp. 82–86

2 MOSIG, J.R., BESSON, J.C.E., GEX-FABRY, M., and GARDIOL, F.E.: 'Reflection of an open-ended coaxial line and application to nondestructive measuremet of materials', *IEEE Trans.*, 1981, **IM-30**, (1), pp. 46–51

3 GRANT, J.P., CLARKE, R.N., SYMM, G.T., and SPYROU, N.M.: 'A critical study of the open-ended coaxial line sensor technique for FR and microwave complex permittivity measurements', *J. Phys. E: Sci. Instrum.*, 1989, **22**, pp. 757–770

4 ARAI, M., BINNER, J.G.P., CARR, G.E., and CROSS, T.E.: ' High temperature dielectric property measurements of engineering ceramics'. Microwaves: Theory and Applications in Materials Processing, Am. Ceram. Soc., 1993, Vol. 36, pp. 483–492

Nonlinear time series modelling and prediction using Gaussian RBF networks with enhanced clustering and RLS learning

S. Chen

Indexing terms: Neural networks, Time series

An improved clustering and recursive least squares (RLS) learning algorithm for Gaussian radial basis function (RBF) networks is described for modelling and predicting nonlinear time series. Significant performance gain can be achieved with a much smaller network compared with the usual clustering and RLS method.

Introduction: A powerful learning method for RBF networks is clustering and least squares learning [1,2]. The RBF centres are obtained by means of a κ-means clustering algorithm while the network weights are learnt using the RLS algorithm. The κ-means clustering algorithm is an unsupervised learning method based only on input training samples. It partitions the input data set into n clusters and obtains the cluster centres by attempting to minimise the total squared error incurred in representing the data set by the n cluster centres [3]. The traditional κ-means clustering algorithm can only achieve a local optimal solution, which depends on the initial locations of cluster centres. A consequence of this local optimality is that some initial centres can become stuck in regions of the input domain with few or no input patterns, and never move to where they are needed. This wastes resources and results in an unnecessarily large network.

Recently, an improved κ-means clustering algorithm has been proposed [4], which overcomes the above-mentioned drawback. By using a cluster variation-weighted measure, the enhanced κ-means partitioning process always converges to an optimal or near-optimal configuration, independent of the initial centre locations. This enhanced κ-means clustering algorithm is ideal for learning RBF centres from time series samples for the purpose of modelling.

Method: The RBF network structure considered is the normalised Gaussian RBF network

$$f(\mathbf{x}(k)) = \sum_{i=1}^{n} w_i \phi_i(k) \qquad (1)$$

where

$$\phi_i(k) = \frac{\exp(-\|\mathbf{x}(k) - \mathbf{c}_i\|^2/\sigma_i^2)}{\sum_{j=1}^{n} \exp(-\|\mathbf{x}(k) - \mathbf{c}_j\|^2/\sigma_j^2)} \qquad (2)$$

A normalised Gaussian basis function features either localised behaviour similar to that of a Gaussian function or nonlocalised behaviour similar to that of a sigmoid function, depending on the location of the centre [5]. This is often a desired property.

The RBF centres are learnt using the improved κ-means clustering method [4]

$$\mathbf{c}_i(k+1) = \mathbf{c}_i(k) + M_i(\mathbf{x}(k))[\eta(\mathbf{x}(k) - \mathbf{c}_i(k))] \qquad (3)$$

where the membership function

$$M_i(\mathbf{x}) = \begin{cases} 1 & \text{if } v_i\|\mathbf{x} - \mathbf{c}_i\|^2 \le v_j\|\mathbf{x} - \mathbf{c}_j\|^2 \text{ for all } j \ne i \\ 0 & \text{otherwise} \end{cases} \qquad (4)$$

and v_i is the variation or 'variance' of the ith cluster. To estimate variation v_i, the following updating rule is used:

$$v_i(k+1) = \alpha v_i(k) + (1-\alpha)\left[M_i(\mathbf{x}(k))\|\mathbf{x}(k) - \mathbf{c}_i(k)\|^2\right] \qquad (5)$$

The initial variations $v_i(0)$, $1 \le i \le n$, are set to the same small number, and α is a constant slightly less than 1. The learning rate for centres, η, is self-adjusting based on an 'entropy' formula [4]

$$\eta = 1 - H(\bar{v}_1, \cdots, \bar{v}_n)/\ln(n) \qquad (6)$$

where

$$H(\bar{v}_1, \cdots, \bar{v}_n) = \sum_{i=1}^{n} -\bar{v}_i \ln(\bar{v}_i) \text{ with } \bar{v}_i = v_i / \sum_{j=1}^{n} v_j \qquad (7)$$

The widths σ_i^2, $1 \le i \le n$ can be calculated, after the clustering process has converged, from the variances of the clusters. Because the optimal κ-means clustering distributes the total variation equally among the clusters, a universal width can be used for all the nodes. The network weights w_i, $1 \le i \le n$ are learnt using the usual RLS algorithm.

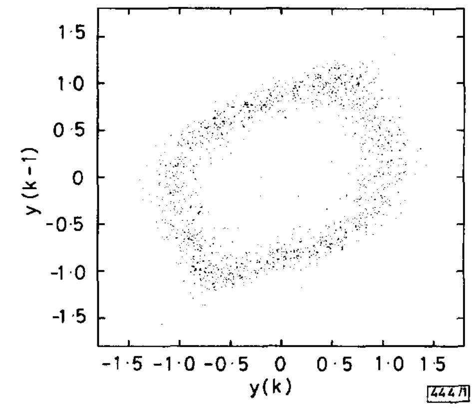

Fig. 1 *Noisy observations of two-dimensional time series example*

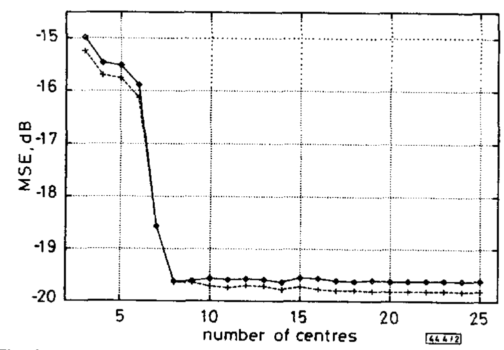

Fig. 2 *Mean square error against centre number for two-dimensional time series example*

Results: The method was applied to nonlinear time series modelling and prediction. The first example was a simulated two-dimensional system

$$\begin{aligned} y(k) = \; & [0.8 - 0.5\exp(-y^2(k-1))]y(k-1) \\ & - [0.3 + 0.9\exp(-y^2(k-1))]y(k-2) \\ & + 0.1\sin(\pi y(k-1)) + e(k) \end{aligned} \quad (8)$$

The noise $e(k)$ had a zero mean and variance 0.01. Two thousand samples of the time series are depicted in Fig. 1. The first 1000 points were used as the training set and the last 1000 as the test set. Fig. 2 shows the mean square error (MSE) as a function of the centre number.

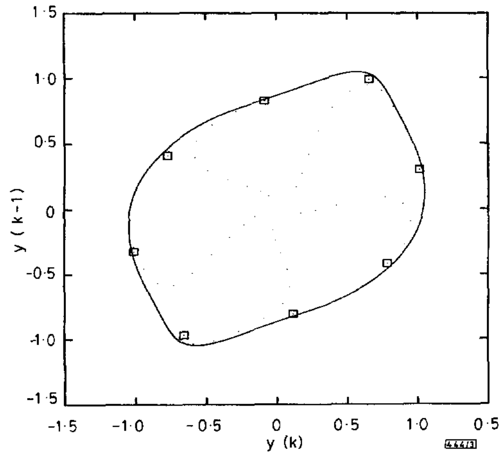

Fig. 3 *Noise-free two-dimensional system and RBF centre locations*

—— limit cycle
□ centre locations

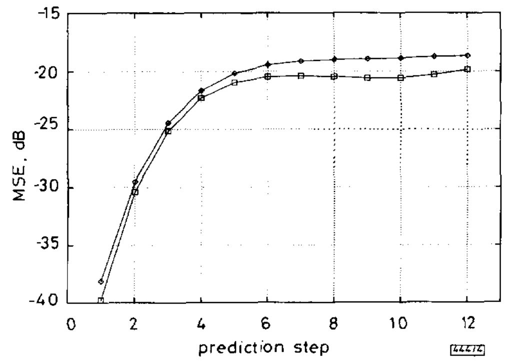

Fig. 4 *Multistep prediction performance for noisy Mackey-Glass time series*

Eight centres are sufficient for modelling this time series. The noise-free system is a limit cycle shown in Fig. 3, where the eight centre locations obtained by the enhanced κ-means clustering algorithm from noisy data are also depicted. From Fig. 3 it can be seen that an optimal centre configuration was obtained. When the network model output was fed back to the input, the iterative network output generated a limit cycle which was indistinguishable from the system limit cycle. The results shown here are better than some previous results [2, 6]. Furthermore, the present method requires a smaller network size.

The second example was a Mackey-Glass time series prediction. To make the task more realistic, a small amount of noise was added to the time series samples, giving rise to a signal/noise ratio of 50dB. The network model obtained was used to compute multistep prediction over a noisy test set not used in training. The results are illustrated in Fig. 4, where it can be seen that good predictive accuracy was achieved using a network of only 20 centres.

Conclusions: An enhanced clustering and RLS learning algorithm has been applied to nonlinear time series modelling and prediction using Gaussian RBF networks. The improved κ-means clustering algorithm ensures that an optimal centre configuration can be achieved, resulting in a smaller network size with better performance.

IEE 1995 *23 November 1994*
Electronics Letters Online No: 19950085

S. Chen (*Department of Electrical and Electronics Engineering, University of Portsmouth, Anglesea Building, Portsmouth PO 3DJ, United Kingdom*)

References

1 MOODY, J., and DARKEN, C.J.: 'Fast learning in networks of locally-tuned processing units', *Neural Comput.*, 1989, **1**, pp. 281–294
2 CHEN, S., BILLINGS, S.A., and GRANT, P.M.: 'Recursive hybrid algorithm for non-linear system identification using radial basis function networks', *Int. J. Control,* 1992, **55**, (5), pp. 1051–1070
3 DUDA, R.O., and HART, P.E.: 'Pattern classification and scene analysis' (John Wiley, New York, 1973)
4 CHINRUNGRUENG, C., and SÉQUIN, C.H.: 'Enhanced κ-means techniques for partitioning the input domain in function approximation', *J. Appl. Sci. Comput.*, 1994, **1**, (1)
5 CHA, I., and KASSAM, S.A.: 'Interference cancellation using radial basis function networks', submitted to *Signal Process.*, 1994
6 AN, P.E., BROWN, M., HARRIS, C.J., and CHEN, S.: 'Comparative aspects of neural network algorithms for on-line modelling of dynamic processes', *Proc. Inst. Mech. Eng. I, J. Syst. Control. Eng.*, 1993, **207**, pp. 223–241

All-optical multiplexing of femtosecond signals using an AlGaAs nonlinear directional coupler

J.U. Kang, G.I. Stegeman and J.S. Aitchison

Indexing terms: Multiplexing, Aluminium gallium arsenide, Directional couplers, Nonlinear optics

The authors demonstrate multiplexing of femtosecond signals without the usual 3dB loss of conventional passive multiplexers using an AlGaAs nonlinear directional coupler controlled by strong pump pulses.

One of the most powerful applications of integrated all optical switching devices is the multi/demultiplexing of high repetition rate signals. All-optical demultiplexing using nonlinear directional couplers (NLDCs) has been successfully demonstrated in recent years [1]. All-optical multiplexing has also been demonstrated using fibres [2, 3]. However, because of the weak nonlinearity in the fibre, lengths of the order of kilometres are required. Thus, such devices suffer from long transit times which can be unacceptable, and from environmental effects of temperature and pressure. The passive means of multiplexing, using Y-junctions, results in a 3dB excess loss. Using strong pump pulses to control a weak signal pulse train through crossphase modulation, we demonstrate an efficient and compact method of multiplexing femtosecond signals using an NLDC, operated with photon energy below half the bandgap.

The experiment was performed with an NaCl colour centre laser operating at a wavelength which corresponds to the low loss telecomuunication window of 1.55µm. Additive pulse modelocking (APM) of the laser produces 600fs pulses at 76MHz. A 2.2cm long, strip-loaded AlGaAs, half beat length waveguide coupler was used for this experiment where the guiding region is made of $Al_{0.18}Ga_{0.82}As$ and the cladding regions of $Al_{0.24}Ga_{0.76}As$. Waveguides were formed using standard optical lithography and reactive ion etching.

A schematic diagram of the experimental arrangement is shown in Fig. 1. Using a polarising beamsplitter, the strong pump (transverse magnetic field) and the weak signals (transverse electric field)