AN ENHANCED CLUSTERING ALGORITHM FOR GAUSSIAN RBF NETWORKS

F. Albu, N. Dumitriu, A. Popescu

Abstract A powerful learning method for RBF networks is clustering and least squares learning [1,2]. The RBF centers are obtained by means of a k-means clustering algorithm while the network weights are learnt using the LMS algorithm. The k-means algorithm is an unsupervised learning method based only on input training samples. It partitions the input data set into *n* cluster centers. The traditional k-means clustering algorithm can only achieve a local optimal solution, which depends on the initial locations of cluster centers. A consequence of this local optimality is that some initial centers can become stuck in regions of the input domain with few or no input patterns, and never move to where they are needed. An improved k-means clustering algorithm has been proposed [2], which overcomes the above-mentioned drawback. By using a cluster variation-weighted measure, the enhanced k-means partitioning process always converges to an optimal or near-optimal configuration, independent of the initial center locations [3].

Introduction

The algorithm for training a RBF network is a combination of the following three methods: 1) using a k-means clustering algorithm to find the proper function centers; 2) using a p-nearest neighbor heuristic rule to determine the width of basis functions; and 3) using the least-mean-square error to learn the weights between the hidden layer and output layer [4]. In order to obtain a neural network solution to a given problem, it is essential to define the proper network architecture. In the case of RBF network models, the network architecture is determined by the number of nodes in each layer and the location of the function centers. However, the determination of the proper number of hidden layer nodes and the location of the function centers require various techniques for different problems. Especially, when using large training data sets, the selection of the number of hidden layer nodes and the location of centers will have profound effects on the performance of the RBF network solution. The most natural choice is to let the number of nodes in the hidden layer equals to the number of training samples; and locate a hidden node for each training data point. This choice of function centers leads to the most basic form of RBF networks. The advantages of this solution are: 1) it is simple; 2) the choice of function centers is uniquely determined; 3) it fully utilizes each training sample; and 4) it is highly desirable in the case of high input dimension and sparse training data. However, there are also many problems associated to this simple solution. For example, computation costs, ill-condition problem, over-fitting problem.

In order to overcome the various problems associated with the basic form of RBF networks, a number of methods have been used select the location of function centers. A number of methods to deal with the problems are discussed in the following sections.

I. The k-means clustering algorithm

Instead of placing the function centers on each of the training data points or on a randomly selected subset of the data points, a more commonly used method is to use the k-means clustering algorithm to find the function centers.

In general, clustering algorithms are used to group some given objects defined by a set of numerical properties in such a way that the objects within a group are more similar than the

objects in different groups. Therefore, a particular clustering algorithm needs to be given a criterion to measure the similarity of objects, how to cluster the objects into groups. The k-means clustering algorithm uses the Euclidean distance to measure the similarities between objects. K-means clustering algorithms need to assume that the number of groups (clusters) is known a priori knowledge.

The standard k-means clustering is a general clustering algorithm to cluster *N* objects into *m* groups of given number *m*. The method minimizes the total squared Euclidean distance E of the form:

$$E = \sum_{i=1}^{N} \sum_{j=1}^{m} M_{ij} \left\| x_i - c_j \right\|$$

where x_i ; i=1, 2, … *N* are the *N* objects and c_j ; j=1, 2, … *m* are the *m* centers. M_{ij} is the cluster membership function which is defined by a N×m matrix of 0's and 1's with exactly one 1 per row which identifies the groups to which a given object belongs. In this algorithm, the similarity between objects is defined by the Euclidean distance: the smaller the distance between two objects represents that the two objects are more similar. A version of k-means algorithm which can be used to select the function centers for RBF networks is presented in the following lines:

Step 1. Initialise the function centers

Set the initial function centers to the first m training data or to the m randomly chosen training data.

Step 2. Group all patterns with the closet function center

For each pattern x_i, assign x_i to group j*, where $\left\| x_i - c_{j^*} \right\| = \min_j \left\| x_i - c_j \right\|$

Step 3. Compute the sample mean for the function center

For each group c_j, $c_j = \frac{1}{m_j} \sum_{x_i \in group\ j} x_i$

where m_j is the number of patterns in group j.

Step 4. Repeat by going to step 2, until no change in cluster assignments

The traditional k-means clustering algorithm can only achieve a local optimal solution, which depends on the initial locations of cluster centres (Fig. 1). 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. An improved k-means clustering algorithm has been proposed [2], which overcomes the above-mentioned drawback. The RBF centres are learnt using the improved k-means clustering method [2]

$$c_i(k+1) = c_i(k) + M_i(\boldsymbol{x}(k))[\eta \boldsymbol{x}(k) - c_i(k)]$$

where the membership function:

$$M_i(x) = \begin{cases} 1 \text{ if } v_i \|\boldsymbol{x} - c_i\|^2 \le v_j \|\boldsymbol{x} - c_j\|^2 \text{ for all } j \ne i \\ 0 \text{ otherwise} \end{cases}$$

and v_i is the variation of the *i*th cluster. To estimate variation v_i, the following updating rule is used:

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

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 centers is:

$$\eta = 1 - H(\overline{v_1}, ..., \overline{v_n}) / \ln(n)$$

where

$$H(\overline{v_1}, ..., \overline{v_n}) = \sum_{i=1}^{n} -\overline{v_i} \ln(\overline{v_i}) \text{ with } \overline{v_i} = v_i / \sum_{j=1}^{n} v_j$$

The widths $\sigma_i^2, 1 \le i \le n$ can be calculated, after the clustering process has converged, from the variances of the clusters. Because the optimal k-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 LMS algorithm.

For the purpose of graphical display, the equalizer order is chosen 2. Let the channel transfer function be

$$H_1(z) = 0.5 + z^{-1}$$

In Fig. 1 are plotted the desired states, 1000 samples which form the data clusters and the RBF centers founded after using the enhanced k-means clustering algorithm. The SNR was 10 dB. This algorithm ensures a near-optimal center configuration.

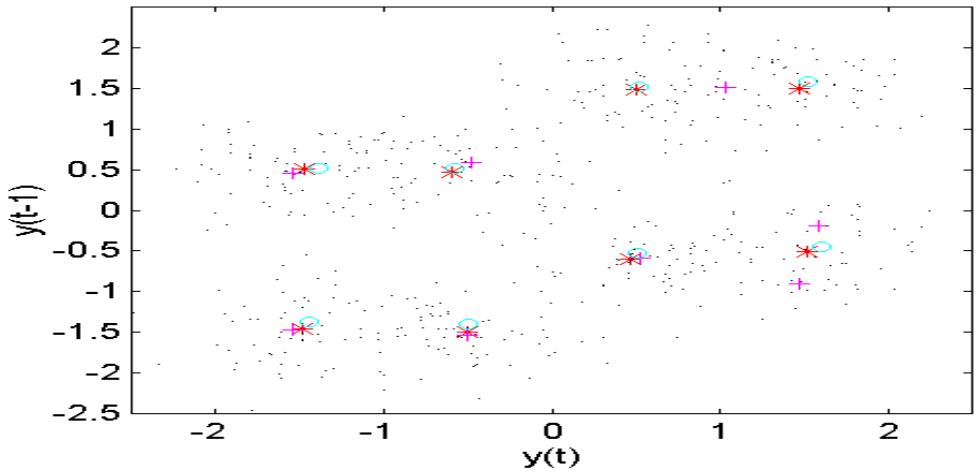

Fig.1 Data clusters, *-desired states, o-RBF centers using the enhanced k-means clustering method, + -RBF centers using k-means algorithm

Simulations were conducted to illustrate the difference in performance between the clustering algorithms proposed for channel equalization [4-6]. We used 100 000 test samples for measuring BER performance.

The channel used for various SNR's was: $H_2(z) = 0.3482 + 0.8704 \cdot z^{-1} + 0.3482 \cdot z^{-2}$. We used a 4-64-1 RBF network, and the delay was 1. Because of the superiority of enhanced k-means algorithm in finding the location of desired centers (Fig. 1) a RBF network using this clustering method obtains better performances (Fig. 2).

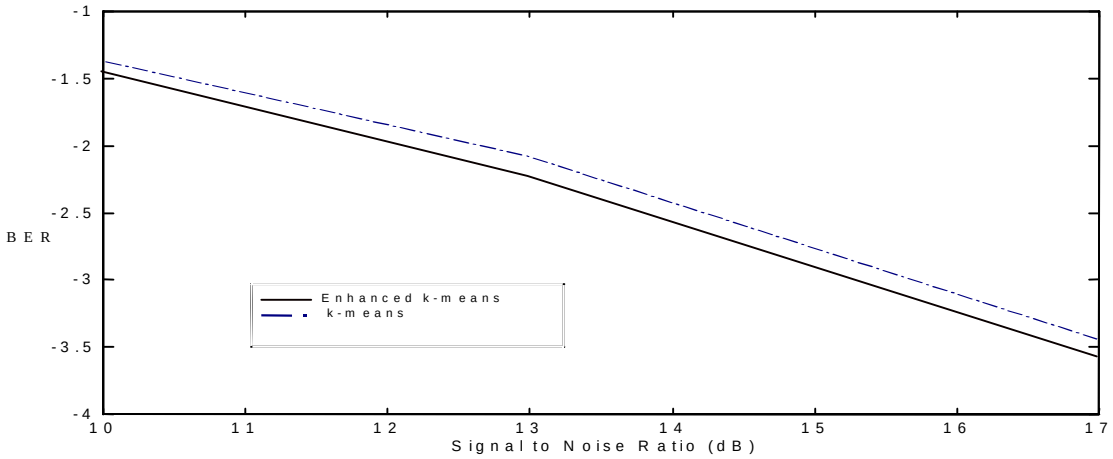

Fig.2 Comparison of performance for H2

Conclusions

An enhanced clustering algorithm has been applied to channel equalization using Gaussian RBF networks. The improved k-means clustering algorithm ensures that a near-optimal center configuration can be achieved.

References

[1] Moody, J., and Darken, C.J.: "Fast learning in networks of locally-tuned processing units", Neural Computation, 1989, 1, pp. 281-294

[2] Chinrungrueng, C., and Sequin, C.H.,: "Enhanced k-means techniques for partitioning the input domain in function approximation", J. Appl. Sci. Comput., 1994

[3] S. Chen, "Nonlinear time series modelling and prediction using Gaussian RBF networks with enhanced clustering and RLS learning", Electronics Letters, Vol. 31, No2, 1995

[4] S. Chen, B. Mulgrew, P. M. Grant, "A clustering technique for digital communications channel equalisation using Radial Basis Function networks", IEEE Transactions on Neural Networks, Vol. 4, July 1993.

[5] G. J. Gibson, S. Siu, C.F.N. Cowan, “The application of Nonlinear Structures to the Reconstruction of Binary Signals”, IEEE Transactions on Signal Processing, Vol. 39, No 8, August 1991
[6] B. Dorizzi, J. C. M. Mota, F. Albu, “A step towards equalisation for radiomobile channel: Neural Networks and Variable Selection”, Workshop on “Emerging Techniques for Communication Terminals”, Toulouse, France,July/1997.
[7] S. Haykin, Neural network: A Comprehensive Foundation, Prentice Hall, 1994.