Journal of Theoretical and Applied Information Technology
10th March 2013. Vol. 49 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ESTABLISHING STRUCTURE FOR ARTIFICIAL NEURAL NETWORKS BASED-ON FRACTAL

[1]Yang Zong-chang
[1] School of Information and Electronical Engineering,
Hunan University of Science and Technology, Xiangtan 411201, China
E-mail: [1]yzc233@163.com

ABSTRACT

The artificial neural network (ANN) is a widely used mathematical model composed of interconnected simple artificial neurons, which has been applied in a variety of applications. However, how to determine number of neurons in the hidden layers is an important part of deciding overall neural network architecture. Many rule-of-thumb methods for determining the appropriate number of neurons in the hidden layers are suggested. In this study, to the puzzling problem of establishing structure for the Artificial Neural Networks (ANN), from a microscopical view, two concepts called the fractal dimension of connection complexity (FDCC) and the fractal dimension of the expectation complexity (FDEC) are introduced. Then a criterion reference for establishing ANN structure based on the two proposed concepts is presented that, the FDCC might not be lower than its (FDEC), and when FDCC is equal or approximate to FDEC, the ANN structure might be an optimal one. The proposed criterion is examined with good results.

Keywords: *Artificial Neural Networks, Structure Establishing, Connection Complexity, Fractal Dimension.*

1. INTRODUCTION

An artificial neural network (ANN), also called a neural network, is a widely used mathematical model composed of an interconnected group of simple artificial neurons that also called nodes, neuroses, processing elements or units, are connected together to form a network with mimicking a biological neural network. Artificial neural network (ANN) uses a connectionist approach to computation in processing information, and is used with algorithms designed to change the strength of the connections in the network to yield a desired signal flow. In most cases, an artificial neural network can be seen as an adaptive system that alters its structural weights during a learning step. Artificial neural network is widely used to model complex relationships between its inputs and outputs, and complex global behavior can be determined by the connections between its processing elements and element parameters in the network.

Since its renaissance in early 1980s, artificial neural networks (ANN) research has received a great deal of attention from the science and technology circles over the world [1-7]. Until now, besides so much attention has been given ANN, it has also been reported fairly good performances for its nonlinear learning capability [1-7]. However, to determine number of neurons in hidden layers [8-18] is a very important part of deciding overall neural network architecture for many practical problems employing neural networks [19-20]. However, how to determine its structure especially of the hidden layer is a puzzling problem. Many methods [8-18] for determining the appropriate number of neurons to use in the hidden layers are introduced with varied degrees of success, such as, a method for estimating the number of hidden neurons based on decision-tree algorithm in [9], a network structure equation by error function in [11], and one hidden layer train algorithm method on energy space approaching strategy in [12], and an algorithm using an incremental training procedure in [15], and some guidelines based on a geometrical interpretation of the multilayer perceptron (MLP) for selecting the architecture of the MLP in [16], and employing the singular vector decomposition to estimate the number of hidden neurons in a feed-forward neural network in [17], and some rule-of-thumb methods in [18]. Among the proposed solutions for this problem, some either focus on the special training procedures that needs a large amount of operations and inconvenient for engineering applicability, or rule-of-thumb methods that short of generality.

Journal of Theoretical and Applied Information Technology
10th March 2013. Vol. 49 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A fractal can be view as a mathematical set, which usually has a fractal dimension exceeding its topological dimension and may be one fraction dimension falling between two integers. Fractals are typically self-similar patterns, where self-similar indicates that fractals may be exactly the same measured at various scales.

Addressing the important and puzzling problem that to determine number of neurons in the hidden layers when deciding overall neural network architecture, from a more macroscopical view, a fractal-based approach is investigated in the study.

2. FRACTAL AND FRACTAL DIMENSION

In the well-known problem of the length of British coastline, the author of the paper [21]--Mandelbrot discussed the research published by Richardson. Richardson had observed and found the famous formula as follows:

$$L(r) = Kr^{1-D_f} \qquad (1)$$

Deep meaning for the exponent D_f was not specified by Richardson. In the paper of [21], Mandelbrot discussed self-similar curves, which have fractional dimensions between 1 and 2. This introduced concept provides a new vision for describing many objects around us that have structure on many sizes, whose normal examples include coastlines, plant distributions and rivers, architecture, etc.

By taking logarithm to *Eq.*(1) and making necessary mathematical operations, we get,

$$D_f = \log(L(r)) / \log(1/r) \qquad (2)$$

Simply speaking, fractals are statistically self-similar. Where, self-similar means that fractals may be exactly the same measured at various scales.

Inspired, we present a fractal-based solution for determining number of neurons in the hidden layers of ANN in the following section.

3. TWO CONCEPTS OF FRACTAL DIMENSION FOR ANN

The Mapping Neural Network Existence Theorem[22] states that, given any continuous function, $\Phi: I^N \to R^M$, $Y = \Phi(X)$ ($I \in [0,1]$), where X and Y are vectors with n and m components respectively. This function can be implemented by a 3-layer neural network with n inputs, one hidden-layer of $2n+1$ neurons and m outputs, its structure, namely (n, $2n+1$,m). The case is under the ideal condition for $I \in [0, 1]$.

From the Mapping Neural Network Existence Theorem [22], it indicates the inherent property of mapping of an ANN [23, 24].

Suppose an ANN with N inputs and M outputs, by ignoring the hidden layers and specific structures, we can get a simplified topological structure of the neural network. In the simplified topological structure, with only 2 layers that the input layer and output layer is considered, its expectation implementation function can be seen as one kind of "mapping" function and illustrated in Fig.1, where A_N denotes the input layer/unit and A_M does the output layer/unit,

$$A_N \xrightarrow{Mapping} A_M$$

Fig. 1 Simplified topological structure of neural networks

The fractal dimension is a mathematical concept, which measures the geometrical complexity of an object. In follows, we try to propose two conceptions based on the fractal dimension for ANN:

Definition1: *we define Ω_E as the called "**Expectation Complexity**" of the neural network in the simplified topological structure* **(***Fig.1*) *by*

$$\Omega_E = (1 + \rho_E{}^k) \times S_E (k > 0) \qquad \textbf{(3)}$$

Where, $S_N = Size(A_N)$ denotes size of input layer and $S_M = Size(A_M)$ for the output layer. The ratio $\rho_E = Max(S_N, S_M) / Min(S_N, S_M)$ indicates its mapping complexity that is expected to be implemented, and $S_E = S_{(N,M)} = Size(A_N, A_M)$ size of both the input and output layers, denotes its structural complexity, and the user-defined parameter k>0.

In the simplified topological structure **(**Fig.1), the *Expectation Complexity* Ω_E indicates a measurement for the ANN with the 2 layers in the simplified topological structure considered. We fetch its scale of measurement (γ_E) by $1/\gamma_E = (1 + 1/\rho_E)$, with consideration of

Journal of Theoretical and Applied Information Technology
10th March 2013. Vol. 49 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

possibility of ρ_E =1 that $\log(\gamma_E)=0$ if taken $\gamma_E=\rho_E$.

Then we define a called ***"Fractal Dimension of Expectation Complexity"*** by,

$$D_E = Log(\Omega_E)/Log(1/\gamma_E) \quad \textbf{(4)}$$

The typical structure of multi-layer completely connected neural networks consists of the input layer, hidden layers and output layer (Fig. 2).

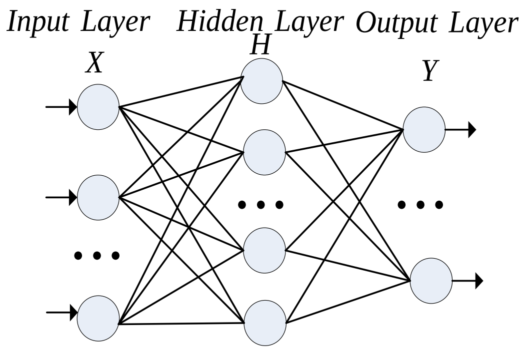

Fig. 2 Structure of multi-layer neural networks

Suppose N inputs in the input layer (unit), M outputs in the out layer (unit), and l hidden layers (units) $H_i(i=1...l)$. There is l+2 layers in the structure of ANN, $(A_N, H_{1,}H_{2,}..._, H_l, A_M)$ (Fig.3). Number of neurons in each layer is, $(L_{1,}L_{2,}..._, L_{l+1}, L_{l+2})$ where L_1=*size* (A_N) =N, L_{l+2}=*size* (A_M) =M, and so on.

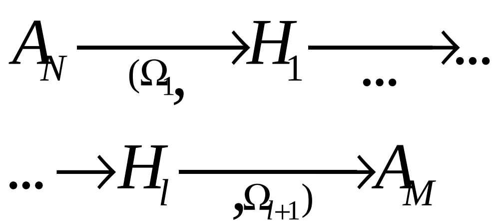

Fig. 3 Topological structure of Multi-layer neural network

In Fig.3, it shows that the topological structure has l+1 connected sub fractal structures from (A_N, H_1) to ($H_{1,}$ A_M). Then we define its called Connection Complexity as follows,

Definition2: *Define* $\boldsymbol{\Omega_C}$ *called* ***"Connection Complexity"*** *for the neural network with multi-layer connected topological structure (Fig.3), and the connection complexity of each sub-structure as* $\boldsymbol{\Omega_i}$ *(1≤ i ≤ l+1),*

$$\Omega_C = \{\Omega_1, \Omega_2, ..., \Omega_{l+1}\} \quad \textbf{(5)}$$

In the case, the connection complexity Ω_C is measured on l+2 layers that *ANN*= $(A_N, H_{1,}H_{2,}..._, H_l, A_M)$, contains l+1 connected sub-structures (Fig.3).

According to *Eq*.(2), we have,

$$\Omega_i = (1+\rho_i^{\,k})\times S_{(H_i,H_{i+1})} \quad \textbf{(6)}$$

Where, $(1\le i\le l+1)$

$\rho_i = \sum size(H_j)(\forall H_j \in \Omega_i)/(N+M)$, ,
$S_{(H_i,H_{i+1})} = size(H_i, H_{i+1})$.

Finally, we define D_C as the called ***"Fractal Dimension of Connection Complexity",*** which is sum of the fractal dimensions on all its sub-structures:

$$D_c = \sum_{i=1}^{l+1} \log(\Omega_i)/\log(1/\gamma_i) \quad \textbf{(7)}$$

Where, in *Eq*.(7), fetch γ_i by $1/\gamma_i = (1+1/\rho_i)$ with consideration of possibility of ρ_i =1 that $\log(\gamma_i)=0$ if taken $\gamma_i=\rho_i$.

4. CRITERION REFERENCE FOR ESTABLISHING ANN STRUCTURE TRAINING OF ANN PARAMETERS

Based on the presented two concepts of fractal dimension for ANN in Section2, namely the fractal dimension of expectation complexity (D_E)((*Eq*.(4)) and fractal dimension of connection complexity (D_C)(*Eq*.(7)) , we propose a criterion reference for establishing ANN structure as follows,

To establish ANN structures, the fractal dimension of connection complexity (D_c) might not be smaller than its fractal dimension of expectation complexity (D_E),

$$D_c >= D_E \quad \textbf{(8)}$$

When the fractal dimension of connection complexity (D_c) is equal or approximate to its fractal dimension of expectation complexity (D_E), i.e., $D_c=D_E$ or $D_c \approx D_E$, the established structure might be an optimal one.

Journal of Theoretical and Applied Information Technology
10th March 2013. Vol. 49 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. EXPERIMENTAL RESULTS

The presented criterion reference for establishing ANN structure is applied to the following tests.

In the references of [9-13], with specific cases, the authors reported their optional choices for establishing the structures of artificial neural networks based on their presented methods.

Comparing with their choices of the references [9-13], our solutions based on our proposed fractal-based criterion for establishing the structures of the ANNs given in the references of [9-13], are listed in Table1-Table5, respectively, where fetch k=2 in *Eq*.(3) and *Eq*.(6), and N, M denotes the ANN with N nodes in the input-layer and M nodes in the output-layer, and H denotes the hidden-layer size to be determined.

TABLE 1: Experimental Result (3-Layer ANN) Comparing With That Of The Reference Of [9]

ANN structure: (***N*=8, *H*, *M*=7**)					
Size(H)	1	2	3	4	5
ΔD	-4.0916	-3.5181	-2.9701	-2.4153	-1.8425
Size(H)	6	*7*	**8**	9	10
ΔD	-1.2463	-0.6237	**0.0266**	0.7051	1.4119
Size(H*) (H*, our optimal)			**8**		
Our suboptimal, size(H*)			**8 ,7 or 9**		
The choice(result) in the reference of **[9]**			**8 or 7**		

TABLE 2: Experimental Result (3-Layer ANN) Comparing With That Of The Reference Of [10]

ANN structure: (***N*=4, *H*, *M*=2**)					
Size(H)	1	2	3	4	5
ΔD	-6.9686	-5.9439	-4.7459	-3.3609	-1.796
Size(H)	*6*	**7**	*8*	9	10
ΔD	-0.0665	**1.8103**	3.8178	5.9417	8.1694
Size(H*) (H*, our optimal)				**7**	
Our suboptimal, size(H*)				**7 , 6 or 8**	
The choice(result) in the reference of **[10]**				**9**	

TABLE3: Experimental Result (3-Layer ANN) Comparing With That Of The Reference Of [11]

ANN structure: (***N*=2, *H*, *M*=1**)					
Size(H)	1	2	3	4	5
ΔD	-5.2344	-3.1643	-0.3569	**3.0501**	6.9294
Size(H)	6	7	8	9	10
ΔD	11.1876	15.7592	20.5976	25.6681	30.9446
Size(H*) (H*, our optimal)				**4**	
Our suboptimal, size(H*)				**4 , 3 or 5**	
The choice(result) in the reference of **[11]**				**5**	

TABLE 4: Experimental Result (3-Layer ANN) Comparing With That Of The Reference Of [12]

ANN structure: (***N*=2, *H*, *M*=3**)					
Size(H)	1	2	3	**4**	5
ΔD	-4.0274	-2.8298	-1.3633	**0.3712**	2.3493
Size(H)	6	7	8	9	10
ΔD	4.5408	6.9172	9.4542	12.1322	14.9352
Size(H*) (H*, our optimal)				**4**	
Our suboptimal, size(H*)				**4 , 3**	
The choice(result) in the reference of **[12]**				**3**	

TABLE 5: Example Of Establishing Structure For Multi-Layer ANN [11]

***The input-layer size N*=3, *the output layer size M*=1 D_E= 12.8228, *l* (**number of hidden layers)				
l	**1**	**2**	**3**	
Size(H*)	**7**	**(3,4)**	**(2,2,4)**	**(2,2,3)**
D_c	15.8979	13.3769	13.8049	11.1700
ΔD	3.0752	0.5541	0.9822	-1.6528
Our optimal structures	**(3,7,1)**	**(1,4,3,1)**	**(3,2,2,4,1) or (3,2,2,3,1)**	
Structure in the Reference of **[11]**			**(3,2,2,3,1)** [3]	

From results in Table.1-Table 5, it is shown that the presented criterion based the fractal for establishing the structure of ANN yields satisfying results, which are agree well with their solutions reported in the references [9-13].

6. CONCLUSION AND FUTURE STUDY

Artificial neural network is a well-known computational model that composed of an interconnected group of simple artificial neurons, tries to simulate some properties of biological neural networks with the aim of solving particular tasks.

In the artificial intelligence field, artificial neural networks have been applied successfully to a wide variety of fields. However, how to determine the number of neurons in hidden layers is a very important part of deciding overall neural network architecture for many practical problems employing neural networks.

The fractal is one classical mathematical concept that fractals are typically self-similar patterns, where self-similar indicates that fractals may be exactly the same at varied scales. Fractal patterns with various degrees of self-similarity have been rendered or found in nature, science and technology fields.

Addressing the important and puzzling problem that how to determine number of neurons in the hidden layers of ANN, in this study, we introduce a

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

fractal-based approach from a more macroscopical view. Regarding the ANN structure, two concepts called fractal dimension of connection complexity (FDCC) and fractal dimension of the expectation complexity (FDEC) is proposed. Then a criterion reference for establishing ANN structure is presented that, the FDCC might not be lower than its (FDEC), and when FDCC is equal or approximate to FDEC, the ANN structure might be an optimal one.

The proposed approach is examined by experiments. Experimental results indicate practicality of the proposed that the presented criterion reference based the fractal for establishing the structures of ANNs yields satisfying results, which are agree well with the optimal solutions by employing other different methods.

To further extend and improve the proposed fractal-based approach for establishing the structures of ANNs is still included in our further study.

ACKNOWLEDGMENT

The project was supported by Scientific Research Fund of Hunan Provincial Education Department (09C399) and research fund of Hunan University of Science and Technology (E50811).

REFRENCES:

[1] LILIANA, and T. A. NAPITUPULU, "Artificial Neural Network Application in Gross Domestic Product Forecasting an Indonesia Case", *Journal of Theoretical and Applied Information Technology*, Vol. 45. No. 2, 2012, pp.410-415.

[2] X.F. Li, and J. P. Xu, "The Improvement of BP Artificial Neural Network Algorithm and Its Application", *International Conference on E-Business and E-Government* (*ICEE*), May 07-09, 2010, pp.2568-2571.

[3] W. J. Shi , X. Z. Wang , D.Q. Zhang, F. Wang, and M. Y. Ma, "A Novel FOCAL Technique based on BP-ANN ", *International Journal for Light and Electron Optics,*Vol. 117, No. 4, 2006, pp.145-150.

[4] X. MA, "Interest Degree Analysis Based on Browsing Behaviours", *Journal of Theoretical and Applied Information Technology*, Vol. 45. No. 2, 2012, pp. 587-592.

[5] L. LI, Z.-W. LIU, X.-Y. WANG, and J.-P. XU, "Remote Monitoring and Intelligent Analysis Platform for Water Quality in Lake Reservoir", *Journal of Theoretical and Applied Information Technology*, Vol. 47. No. 2, 2013, pp 594-597.

[6] L. GAO, Y.H. ZHANG, M. ZHANG, L.M. SHAO, and J.X. XIE, "A Multi-step Prediction Model Based on Interpolation and Adaptive Time Delay Neural Network for Time Series", *Journal of Theoretical and Applied Information Technology*, Vol. 47. No. 2, 2013, pp. 870 - 874.

[7] L. ZHAO, and Y. X. MAO, "Flaw Identification of Metal Material in Eddy Current Testing using Neural Network Optimized by Particle Swarm Optimization", *Journal of Theoretical and Applied Information Technology*, Vol. 47. No. 1, 2013, pp. 261-265.

[8] M. ETTAOUIL, M. LAZAAR, and Y. GHANOU, "Architecture Optimization Model for the Multilayer Perceptron and Clustering", *Journal of Theoretical and Applied Information Technology*, Vol. 47. No. 1, 2013, pp. 064 – 072.

[9] H.-C.YUAN, "A Novel Method for Estimating the Number of Hidden Neurons of the Feedforward Neural Networks", *Journal of Chinese Computer Systems*, Vol.24, No.4, 2003, pp.658-660.

[10] S.-L. PANG, "Study on Miltilayer Perceptron Credit Scoring Model", *Acta scientiarum naturalium, Universitatis Sunyatseni*, Vol.42, No.4, 2003, pp.119-122.

[11] O.-G. Liu, "Research on A Structure of Multi-Layer Forward Artificial Neural Network", *Journal of Natural Science of Hunan Normal University of China*, Vol.27, No.1, 2004, pp.27-30.

[12] R.-Y. CUI, "A Hidden Layer Training Algorithm for Three-Layered Feedforward Neural Networks Based on Energy Space Approaching Strategy", *Jisuanji Yanjiu yu Fazhan/Computer Research and Development*, Vol.40, No.7, 2003, pp.908-912.

[13] X.-M. LI, "A New Method to Determine Hidden Note Number in Neural Network", *Journal of Jishou University of China (Natural Science Edition)*, Vol.23, No.1, 2002, pp.90-91.

[14] H-X. ZHOU, "An improved algorithm on hidden nodes in multi-layer feed-forward neural networks", *Journal of Zhejiang Normal University*, Vol.25, No.3, 2003, pp.269-271.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[15] D. R. Liu, “A new learning algorithm for feedforward neural networks”, *Intelligent Control*, Sept, 2001, pp.39-44.

[16] C. Xiang, S. Q. Ding and T. H. Lee, “Geometrical interpretation and architecture selection of MLP”, *IEEE Trans. on Neural Networks*, Vol.16, No.1, 2005, pp.84-96.

[17] E. J. Teoh, K. C. Tan and C. Xiang, “Estimating the number of hidden neurons in a feedforward network using the singular vector decomposition”, *IEEE Transactions on Neural Networks*, Vol.17, No.6, 2006, pp.1623-1629.

[18] J. Heaton, Introduction to Neural Networks with Java, Chesterfield: Heaton Research Inc., 2005, pp.125-154.

[19] R. Sikora, T. Chady, P. Baniukiewicz, M. Caryk, and B. Piekarczyk, “The Choice of Optimal Structure of Artificial Neural Network Classifier Intended for Classification of Welding Flaws”, *AIP Conference Proceedings*, February 22, 2010, pp. 631-638.

[20] J. Dobes, L. Posisil, and V. Panko, “Selecting an optimal structure of artificial neural networks for characterizing RF semiconductor devices”, *IEEE International Midwest Symposium on Circuits and Systems*, 1-4 Aug, 2010, pp.1206-1209.

[21] B. Mandelbrot, “How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension”, *Science*, Vol.156, No.3775, 1967, pp. 636-638.

[22] R. Hecht-Nielsen, “Kolmogorov's mapping neural network existence theorem”, *In Proceedings of IEEE First Annual International Conference on Neural Networks*, Vol. 3, 1987, pp. III-11-III-14.

[23] R. Neruda, “Kolmogorov learning for feedforward networks”, *International Joint Conference on Neural Networks* (*IJCNN '01*), 15-19 July, 2001, pp.77-81.

[24] S. Huang, and Y. Huang, “Bounds on the number of hidden neurons in multilayer perceptrons”, *IEEE Transactions on Neural Networks*, Vol.2, No.1, 1991, pp.47-55.