氏名（本籍）	バサビ チャクラボルティ BASABI CHAKRABORTY　（インド）
学位の種類	博士（情報科学）
学位記番号	情博第30号
学位授与年月日	平成8年3月26日
学位授与の要件	学位規則第4条第1項該当
研究科，専攻	東北大学大学院情報科学研究科（博士課程）システム情報科学専攻
学位論文題目	Study on Information Processing Ability of Hierarchical Fractal Neural Network （階層型フラクタル結合神経回路網の情報処理能力の研究）
論文審査委員	（主査） 東北大学教授　澤田　康次　　東北大学教授　山本　光璋 東北大学教授　星宮　望 （工学研究科）

論文内容要旨

Chapter 1. Introduction

The goal of research on artificial neural network is to develop machines having brain-like information processing capability. Incorporation of human brain's capability of executiong myriad of functions, maintaining a compact size, demands the need of an effecient design of the network in terms of the number of nodes and the number of connections. Till now no definite guideline for designing the most effective organization of the neurons to from the network has been ebolved.

In this thesis a multilayer hirarchical feedforward architecture with fractal connection structure within layers, in which connectivity of any two neurons depends on their destance from each other, has been proposed and its classification capability and generalization behaviour has been examined compared to the conbentional fully connected model and a randomly connected sparse model by simulation with different data sets.

The thesis is presented with five chapters. The first chapter explains the purpose of this study and briefly states the contents of it. The second chapter introduces some of the relevant models and their drawbacks that leads to the devalopment of the proposed new model. The third chapter presents the proposed new model and its operation. The study on discrimination ability of this model by sumulation with defferent data sets has been discussed. In the fourth chapter the generalization behaviour of the model has been studied and the simulation results are presented. The thesis ends with the concliding remark and future direction of research in chapter five.

Chapter 2. Earlier Hierarchical and Fractal Models

Some of the existing hierarchical and fractal models are described below :

- Neocognitron : A hierarchical model with fully localized connections among layers with self organized unsupervised learning used for recognition of handwritten numerals.

- LeCun's net : A hierarchical model with fully localized constrained connections among layers with back-propagation learning applied to han-written digit recognition problem.
- Linsker's model : Multilayered self-organized unsupervised model resembling human visual system.
- Merrill & Port's Model : Fractally configured constrained neural network appled to a simple generalization problem.
- Baram'm Model : Multilayered model with localized connections among layers used for pattern association problem.
- Kim's model : Recurrent fractal model used for pattern association problem, effecient for correlated patterns.

The problems with those models are basically :

- For the models used in pattern recognition problems, both neocognitron and Lecun's net have a complex and problem dependant architecture. For large world problems the software implementation is computationally complex as well as the hardware implementation due to lack of modularity in design.
- Merrill's net is time consuming as it requires searching of best possible configuration from a number of alternatives.
- Linsker's model resembles biological model but its engineering appeal has not been established.

The other existing fractal models have not been tested for pattern recognition problem. A hierarchical fractal model is thus proposed in the hope of achieving better effeciency in pattern recognition problems motivated by the facts that

1. Fractal connection structure ensures sparseness (less number of connections) and modularity for compactness and easy hardware implementation.
2. Hierarchical architecture allows to deal with partial informations (breaking up the input in parts), thereby leading to a smaller net.

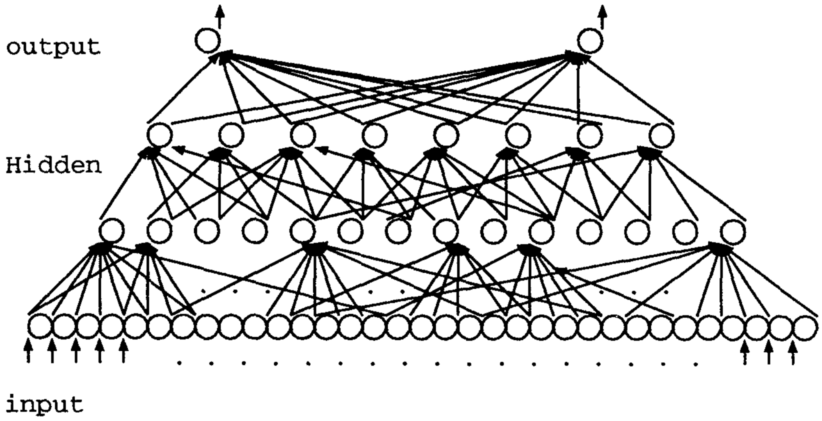

Figure 1 : The Architecture of the proposed Hierarchical Fractal net

Chapter 3. Proposed Hierarchical Fractal model : Performance as a classifier

Proposed Model

Figure 1 represents our proposed model. The neurons are arranged in a layered hierarchy where each layer is an array of neurons in one or two dimension depending on the type of inputs to be processed. Each neuron in the array can be defined by a spatial position coordinate, Q_{ik}, Q_{ijk} representing *i*th neuron or (*i*, *j*) th neuron (in one dimension or in two dimension as the case may be) in the *k*th layer. The upper layer neurons other than the input layer are positioned relative to the position of the immediately previous layer depending on number of neurons available in both the layers. Thus Q_{ik} is calculated as

$$Q_{ik} = [\lceil n_{k-1}(2i-1)/2n_k \rceil, k] \text{ for } i=1,\ 2,\dots,n_k \tag{1}$$

where n_{k-1} and n_k represents the number of neurons in $(k-1)$th and kth layers respectively. The input layer neurons have Q_{ik} value defined as $Q_{ik}=[i, 0]$

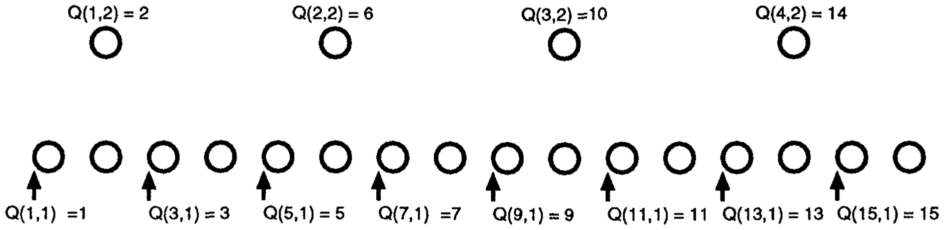

Figure 2 : Spatial Positions of the Upper and Lower layer Neurons

Figure 2 explains the positioning of neurons according to Equation 1 in case of 16 and 4 neurons in the lower and upper layer respectively.

The probability of an upper layer processing element other than the output layer to receive connection from the lower layer elements follow a inverse power law generating a fractal structure which results in a sparsely connected network. The connectivity (or sparseness) of the net can be controkked in a regular manner.

The probability that ith. processing element in the kth. layer receives connection from the jth processing element of the previous layer, defined by CP_{ijk} follows the law

$$\begin{aligned} CP_{ijk} &= A r_{ij}^{D_n - d} \\ i &= 1, 2\dots\ n_k \\ j &= 1, 2\dots\ n_{k-1} \\ 0 &\leq D_n \leq d \end{aligned} \tag{2}$$

where r_{ijk} is the Euclidean distance between ith. processing element in the kth. layer and jth. processing element of the previous layer defined as

$$r_{ijk} = \|Q_{ik} - Q_{j(k-1)}\|,\quad r_{ijk} \geq 1 \tag{3}$$

d denotes dimension of the array of neurons in k th layer. A represents a constant, D_n represents the fractal dimension of the synaptic connection distribution which controlls the sparseness in connectibity of the net.

Simulation Results

Simulation has been done with a generated fractal pattern (Cntor set) and a real life date namely sonar data used by Gorman and Sejnowski in underwater target classification problem. Networks are trained by simple backpropagation algorithm with a set of training samples and the trained net has been used to find out the average classification rate for a test set.

Figure 3 and Figure 4 represent the result for Cantor set data and sonar data respectively. In both the cases the fractal net with fractal dimension around 0.9 performs equally well as the full connection net though in the average connectivity is 68% in the first case and 80% in the later one. The net used is bigger in size in Cntor set data classification than the net sonar data classification. The above result also indicates the fact that the advantage of using fractal connection for connection reduction is much effective for bigger net.

The graph also shows that the performance of the fractal net is better than randomly connected net of same average connectivity in terms of classification rate as well as deviation from run to run.

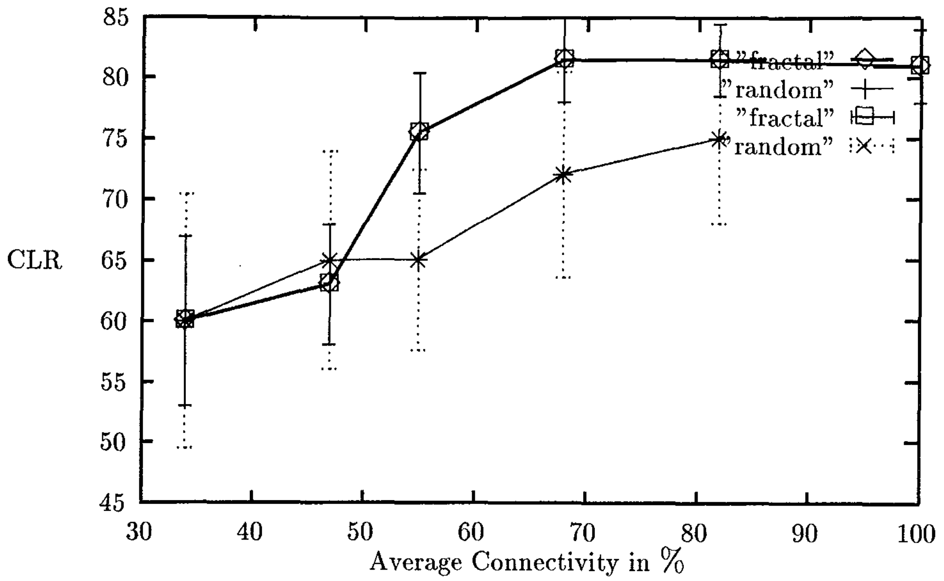

Figure 3 : Classification rates (CLR) for Full net Fractal net and Random net with increasing average connectivity for Cantor set

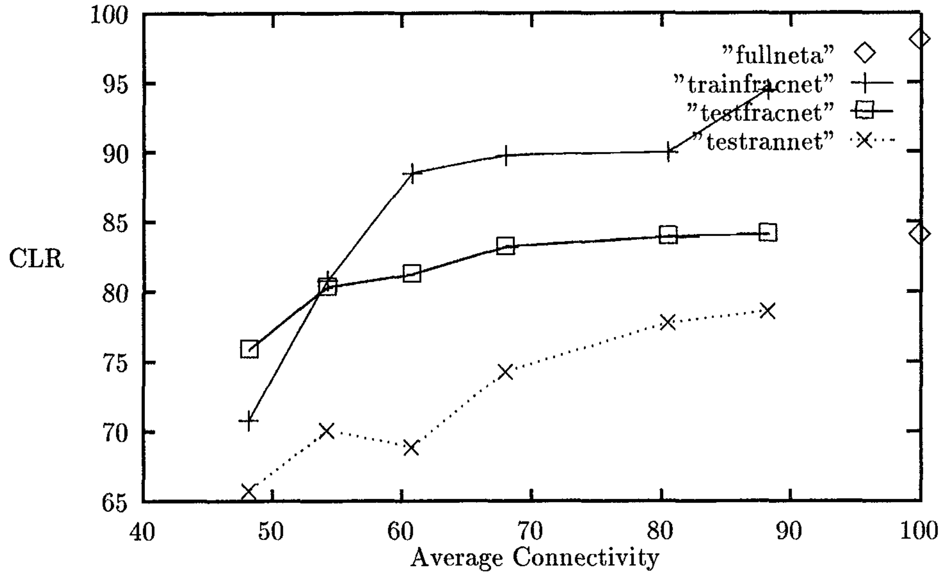

Figure 4 : Classification rates (CLR) for Full net, Fractal net (for test and train samples) and Randam net (test samples) with different average connectivity for Sonar data

Chapter 4. Generalization Behaviour of Proposed Model

Generalization is the most fundamental and desirable quality of a learning system. It is the ability of the system to achive equal performance with respect to the training patterns used for the design of the system as well as the unseen patterns outside the training set. The generalization capability of our proposed model compared to randomly connected and fully connected net in classification problem has been studied and the simulation results of sonar data has been presented in this chapter.

The generalization Measure (GM) has been defined as

$$GM = \frac{CLR_{Train} - CLR_{Test}}{CLR_{Train}} \tag{4}$$

where CLR stands for Classification rate measured under identical net condition. For ideal generalization the train set and the test set classification rate should be same. Thus the best and the worst value of GM should be 0 and 1. Low value of GM corresponds to better generalization.

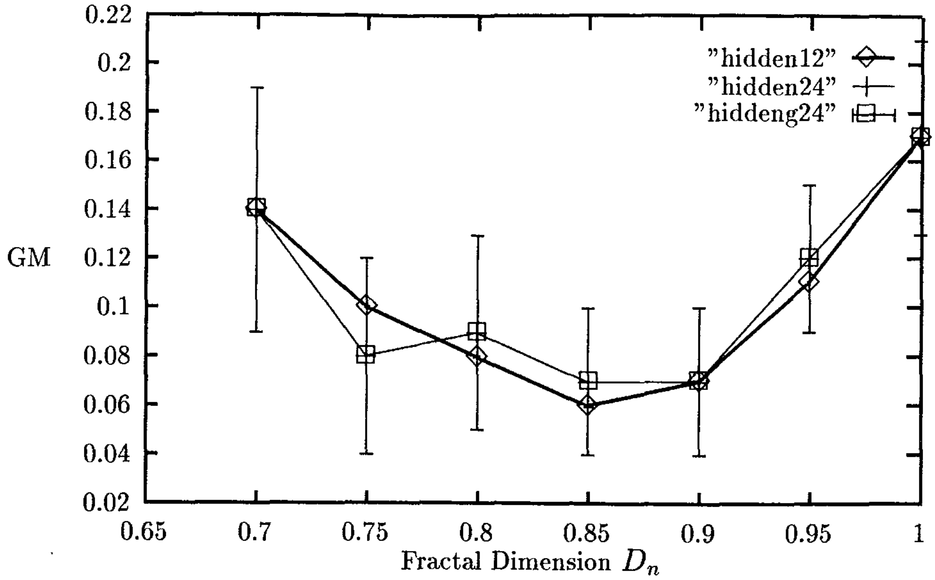

Figure 5 : Generalization measures of the Full connection and Fractal net with different fractaldimension

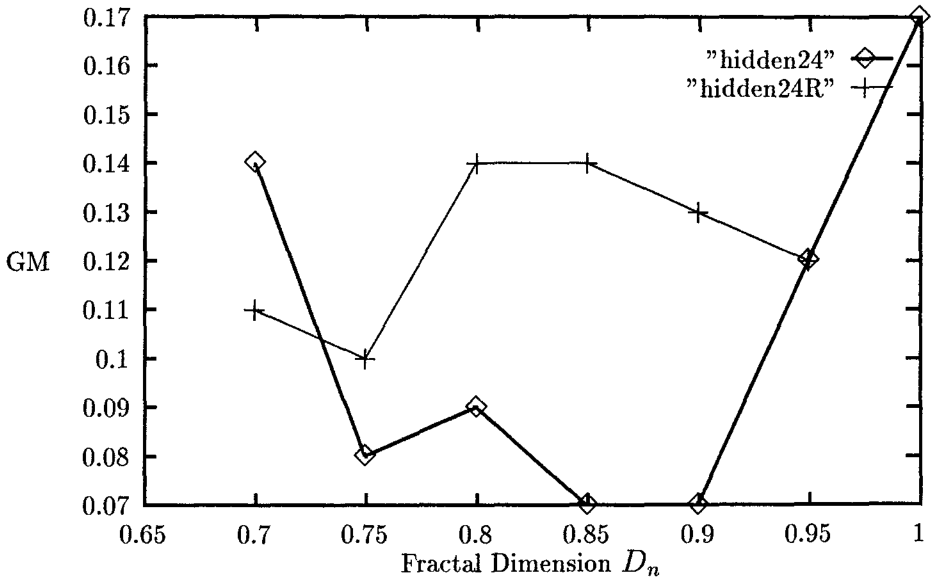

Figure 6 : Generalization measure of the Randomly connected net and Fractally connected net with same aveage connectivity

Simulation Results

Figure 5 represents the Generalization Measure of the full net and the fractal net with different fractal dimension. The deviation in Generalization Measure for different random partitions of the data has also been shown. The graph clearly indicates that the generalization capability of the fractally connected net with fractal dimension around. 85 is the best and greater than the full connection net.

Figure 6 represents the Generalization Measure for fractal net and randomly connected net of same average connectivity. The randomly connected net shows no improvement of generalization capabiliy over fullconnection net and generalizes poorly compared to fractal net of same average connectivity.

Chapter 5 . Conclusion

The final chapter concludes with summarizing the results of the study and pointing towards future scope of research. The main conclusion of the present study follows :

- Fractally connected hierarchical net with fractal dimension exceeding a limit performs equally well in terms of classification rate compared to fully connected net of same size in classification of correlated patterns thouth it has lower average connectivity.
- Fractally connected net performs better than randomly connected net of same size and average connectivity in correlated pattern classification.
- Fractally connected net have better generalization power compared to fully connected net and randomly connected net of same average connectivity when fractal dimension exceeds a certain limit.

The important points left to be studied follows :

- Fractal net has been found to perform better with correlated patterns, the perfomance for uncorelated pattern has to be investigated.
- The simulation results need to be explained on a theoretical basis.
- Backpropagation learning, though a successful invention for engineering application, seems to be biologically not plausible. A more biologically oriented learning process is represented by self-organization. The investigation of the behaviour of our proposed model with self-organizing learning principle may be interesting.

審査結果の要旨

全素子が並列に論理演算を行ない，その結合強度にメモリに分散している神経回路網は，柔軟な情報処理機能を持つ将来の脳型計算機の基本要素として期待され，研究が進んでいる。しかし，神経回路網の大規模集積化のためには，その情報処理機能を低減することなく，素子間の結合数を減少させることが要求されるが，これまでの結合の最適粗化の有効な設計論がなかった。

著者は，層間結合数が各素子からの距離のべき乗に比例して減少する階層型フラクタル神経回路に複数のパターン学習させ，その識別能力，汎化能力のフラクタル次元依存性を研究した。本論文は，これらの成果をまとめたもので全文5章よりなる。

第1章は序論で，本論文の目的と構成について述べている。

第2章では，従来の階層型及び部分結合型神経径路網の欠点について述べ，本研究の動機を明らかにしている。

第3章では，層内には結合がなく，隣接層間にフラクタル結合を持つフィードフォワード階層型フラクタル結合神経回路網を提案し，異なるカントール時系列と水中の異なる対象物からの反射音波標準時系列を200ノードの入力に対して学習させた場合の，それぞれの識別能力についてシミュレーションを行なった結果について述べている。どちらの場合も，フラクタル構成は，平均結合数を68%まで減少させても全結合構成の場合と同じ学習速度と識別能力を持つこと，同じ結合数でもランダム結合ではその能力が著しく低下すること，また，中間層数の増大によってフラクタル構成はランダム構成よりも識別能力の増加が著しいことなどを明らかにした。

フラクタル構成の臨海平均結合率は入力ノード数のべき乗に比例して減少することから，大規模集積化に対して有効であることを示したことは評価できる。

第4章では，階層型フラクタル結合神経回路網の汎化能力の指数と定義し，水中反射音波標準時系列に対して，この指数が結合フラクタル次元0.85の場合に最適値を取ることを明らかにした。また，ランダム結合の場合には結合数の変化に対して明らかな最適値を見いだすことはできなかった。このことは，過学習による汎化能力の低減に対する定量的研究の糸口を与えたもので高く評価できる。

第5章は結論である。

以上要するに，本論文は新しい階層型フラクタル結合神経回路網を提案し，大規模集積化に向けた設計指針を与えると共に，識別能力，汎化能力に対して新たな知見を加えたもので，情報物理学の発展に寄与するところが少なくない。

よって，本論文は博士（情報科学）の学位論文として合格と認める。