4th WSEAS Int. Conf. on COMPUTATIONAL INTELLIGENCE, MAN-MACHINE SYSTEMS and CYBERNETICS Miami, Florida, USA, November 17-19, 2005 (pp210-216)

Optimizing neuronal complexity using wavelet based multiresolution analysis for Type-I fuzzy neural networks

S. M. AQIL BURNEY, TAHSEEN AHMED JILANI, MAJOR AFZAL SALEEMI
Department of Computer Science
University of Karachi
University Road, Karachi, 75270
PAKISTAN

Abstract: - In neural network the connection strength of each neuron is updated through learning. Through repeated simulations of crisp neural network, we propose the idea that for each neuron in the network, we can obtain reduced model with more efficiency using wavelet based multiresolution analysis (MRA) to form wavelet based quasi fuzzy weight sets (WBQFWS). Such type of WBQFWS provides good initial solution for training in type-I fuzzy neural networks thus the search space for synoptic connections is reduced significantly, resulting in fast and confident learning of fuzzy neural networks. As real data is subjected to noise and uncertainty, therefore, WBQFWS may be helpful in the simplification of complex problems using low dimensional data sets.

Key-Words: - crisp neural networks, wavelets based multiresolution analysis (MRA), extraction of fuzzy rules, fuzzy neural network (FNN), wavelet based quasi fuzzy weight sets (WBQFWS), density estimation techniques.)

1 Introduction

Fission of artificial neural networks and fuzzy inference systems have attracted the growing interest of researchers in various scientific and engineering areas due to the growing need of adaptive intelligent systems to solve the real world problems. A crisp or fuzzified neural network can be viewed as a mathematical model for brain-like systems. The learning process increases the sum of knowledge of the neural network by improving the configuration of weight factors. FNN are generalization of crisp neural networks to process both numerical information from measuring instruments and linguistic information from human experts, see [3], [18], and [21]. Thus, fuzzy inference systems can be used to emulate human expert knowledge and experience. An overview of different FNN architectures is discussed by [7], [10] and [19]. It is much more difficult to develop the learning algorithms for the FNN than for the crisp neural networks; this is because the inputs, connections weights and bias terms related to a regular FNN are fuzzy sets, see [21]. The new technique in mathematical sciences called wavelets can be introduced to reduce the problem complexity as well as the dimensions so that a FNN may provide a fast track for optimization. Wavelet based MRA provides better analysis of complex signals than Fourier based MRA as in [9] and [17].

The paper is organized as follows. In section II, we made a short study of learning procedures in crisp neural networks. In section III, we present concepts of fuzzy logic as our target work and later in section IV wavelet based MRA is introduced. In section V, simulation experiments are presented. These sets provide the initial design for type-I neuro-fuzzy networks as discussed by [1] and [6]. To our knowledge, the concept of obtaining WBQFWS through crisp neural networks has not been investigated in the literature.

2 Neural Networks

A neural network can be regarded as representation of a function determined by its weight factors and networks architecture [5]. The overall mapping is thus characterized by a composite function relating feedforward network inputs to output. For p-layer feedforward network, we have

$$\mathbf{O} = \mathbf{f}_{composite}(\mathbf{x}) = \mathbf{f}^{L_p}\left(\mathbf{f}^{L_{p-1}} \ldots \left(\mathbf{f}^{L_1}(\mathbf{x}) \ldots\ldots\right)\right)$$

Usually, we train a neural network with a training set, present inputs to the neural networks, and interpret the outputs according to the logical rules in the training set see [2], [4] and [22]. The most commonly used technique to adjust weight parameters of a neural network is backpropagation method based on LMS learning defined as

$$J = E\left[\sum_k e_k^2(n)\right]$$

where k= number of output neurons. Where

$$w_{ji}^{l}(n+1) = w_{ji}^{l}(n) + \eta\, \delta_j^{l}(n)\, y_i^{l-1}(n),$$

η is the learning rate and $\delta_j^l(n)$ is the local change made at each neuron in the learning, see [15]. But to deal with noisy and uncertain information, a crisp neural network has to use concepts of fuzzy inference systems [24] that will be discussed in the next section.

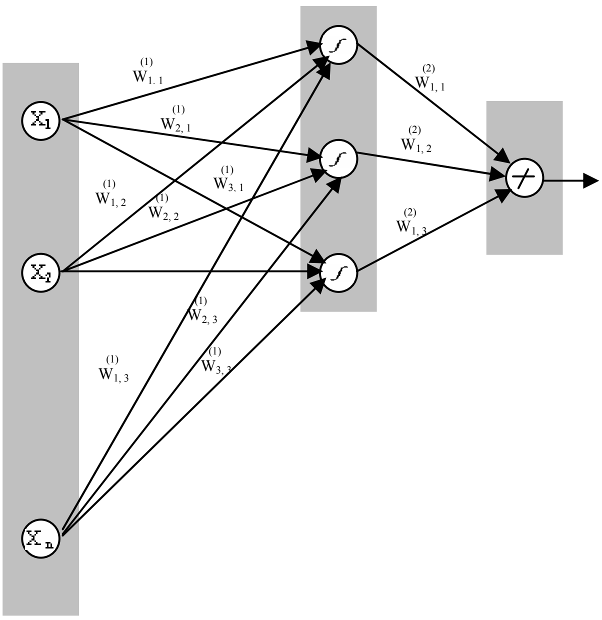

Fig.1. Structure of a crisp artificial neural network

3 Fuzzy Logic

Fuzzy logic was originally proposed by Prof. Lotfi A. Zadeh to quantitatively and effectively handle problems involving uncertainty, ambiguity and vagueness see [14], [17] and [25]. The theory which is now well-established was specifically designed to mathematically represent uncertainty and vagueness and provide formalized tools for dealing with the imprecision that is intrinsic to many real world problems. The ability of fuzzy logic is inherently robust since it does not require precision and noise-free inputs. Fuzzy inference systems are the most reliable alternative if the mathematical model of the system to be controlled is unavailable see [15]. The fuzzy sets and fuzzy rules can be formulated in terms of linguistic variables. Methods of fuzzy logic are commonly used to model a complex system by a set of rules provided by the experts. But fuzzy rules can also be applied in reverse problems: given the input-output behavior of a system, what are the rules which are governing the behavior. We cite definition of quasi fuzzy sets,

A quasi-fuzzy number **A** is a fuzzy set of the real line with a normal, fuzzy convex and continuous membership function satisfying the following conditions,

$$\begin{aligned} &\lim(t \to -\infty)\ \mathbf{A}(t) = 0,\quad \lim(t \to \infty)\ \mathbf{A}(t) = 0 \\ &a_l(\gamma) = \min\left(\mathbf{A}^\gamma\right),\quad a_r(\gamma) = \max\left(\mathbf{A}^\gamma\right) \\ &a_l : [0,1] \to \mathbf{R} \quad a_r : [0,1] \to \mathbf{R} \end{aligned} \tag{1}$$

Then $\mathbf{A}^\gamma = [a_l(\gamma), a_r(\gamma)]$. The support of A is the open interval $(a_l(\gamma), a_r(\gamma))$.

In order to reduce computational expense, we use triangular fuzzy numbers to define the fuzzy weights. The Wavelet based quasi fuzzy weight sets (WBQFWS) follow fuzzy arithmetic rules, and thus can be used for learning of fuzzy neural networks.

4 Wavelet Based Multiresolution Analysis

In recent years, researchers have developed powerful wavelet techniques for the multi scale representation and analysis of signals [8], [9] and [17. These new methods differ from the traditional Fourier techniques. Wavelets localize the information in the time-frequency space which makes them especially suitable for the analysis of non-stationary signals [16]. One important area of application where wavelets have been found to be relevant is fuzzy neural systems as discussed in [10] and [14]. This whole area of research is still relatively new but is evolving very rapidly. We examine the very important property of wavelet transformation i.e. maximization of signal energy using data compression for FNNs. There are essentially two types of wavelet decompositions, Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT), see [12], [16] and [20]. Continuous wavelets are usually preferred for signal analysis, feature extraction and detection tasks whereas the second type is obviously more adequate whenever it is desirable to perform some kind of data reduction or when the orthogonality of the representation is an important factor see [9]. However, the choice between them is optional depending upon the computational considerations. We will use the decomposition in terms of DWT using Mallat's pyramid algorithm which is faster than a CWT and obtained very satisfactory results see [17] and [20].

Let $f(t)$ be a signal defined in $L^2(R)$ space, which denotes a vector space for finite energy signals, where R is a real continuous number system. The WT of $f(t)$ in terms of continuous wavelets is then defined as

$$CWT_{\psi} f(a,b) = \langle f, \psi_{a,b} \rangle = \int_{\infty}^{\infty} f(t)\psi_{a,b}(t)dt \qquad (2)$$

where $\psi_{a,b}(t) = |a|^{\frac{1}{2}} \psi\left(\frac{t-b}{a}\right)$

$\psi(t)$ is the base function or the mother wavelet with $a, b \in R$ are the scale and translation parameters respectively. Instead of continuous dilation and translation, the mother wavelet may be dilated and translated discretely by selecting $a = a_0{}^m$ and $b = nb_0 a_0{}^m$, where a_0 and b_0 are fixed values with $a_0 > 1, b_0 > 0$, $m, n \in Z$ and Z is the set of positive integers. Then the discretized mother wavelet becomes

$$\psi_{m,n}(t) = a^{-\frac{m}{2}} \psi\left(\frac{t - nb_0 a_0{}^m}{a_0{}^m}\right) \qquad (3)$$

and the corresponding discrete wavelet transform is given by

$$DWT_{\psi} f(m,n) = \langle f, \psi_{m,n} \rangle = \int_{-\infty}^{\infty} f(t)\psi_{m,n}(t)dt \qquad (4)$$

DWT provides a decomposition and reconstruction structure of a signal using MRA through filter bank [9]. The roles of mother scaling and mother wavelet functions $\phi(t)$ and $\psi(t)$ are represented through a low pass filter L and a high pass filter H. Consequently, it is possible to obtain a signal f through analysis and synthesis by using wavelet based MRA, see [20].

$$f(t) = \sum_{n \in Z} c_{p,n} \phi_{p,n}(t) + \sum_{0 \le m \ge p} \sum_{n \in Z} d_{m,n} \psi_{m,n}(t)$$

where the sum with coefficients $c_{p,n}$ represents scaling or approximation coefficients and sums with coefficients $d_{m,n}$ represent wavelet or detail coefficients on all the scales between 0 and p. Data compression and energy storage in wavelets can be achieved by simply discarding certain coefficients that are insignificant. We combine this property of wavelets with neural networks and found a special class of mother wavelets db4, the most appropriate based on our data. We studied the effect of crisp weights on different neurons by reducing them using wavelets according to their energy preservation.

5 Experiment

The input/target pair presented to the network is $\{\mathbf{X}, \underline{t}\}$ where $\underline{\mathbf{X}} = [\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4, \underline{x}_5]$. A crisp neural network with 3 hidden and one output neuron is trained and repeated the simulations for first 128 successes. Through wavelet decomposition, we reduced dimensions by preserving 95% of the energy of original signal. The decomposed signal at level 5 using db4 wavelets for one of the input weight- vectors is shown in fig. 2.(a), and its compressed version along with original signal in fig. 2. (b).

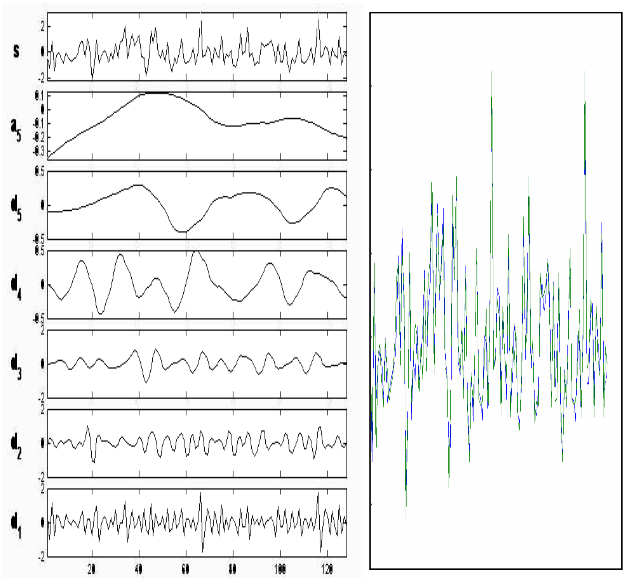

Fig. 2. (a) Signal decomposition (b) original and compressed signals

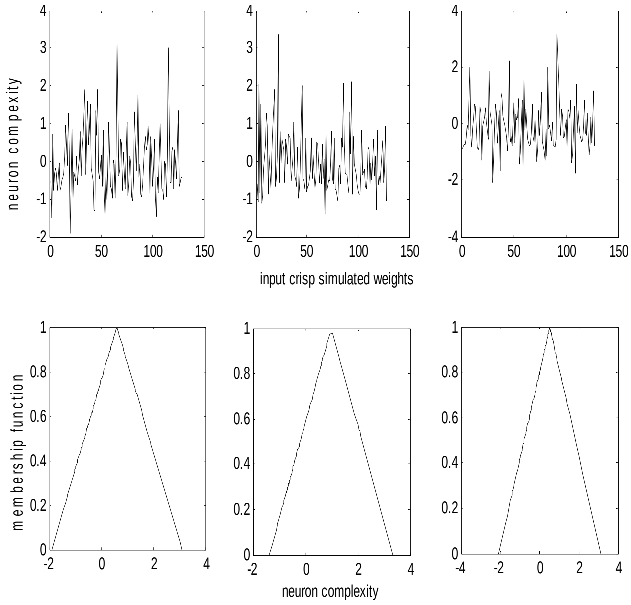

Fig.3. Wavelet based weights and corresponding triangular MF

We have used a threshold of 0.05 for the compression of signal, thus reducing the data dimensionality up to 50%. So that in place of high data requirements of QFWS in [1], we may have better performance using WBQFWS. The triangular membership function is constructed due to its reduced complexity see [23]. From fig. 1, the parameters of triangular-mf based WBQFWS using eq.1 will be,

$$a_l = \min(w^1_{1,1}),\quad a_r = \max(w^1_{1,1}),\quad a_m = (a_l + a_r)/2 \qquad (5)$$

In table.1, results of WBQFWS provide superior mapping of weight space obtained using repeated simulations of crisp neural network than 95% Gaussian confidence internal. Interesting to note that the first two moments are nearly similar in WBQFWS showing consistency of proposed interval sets but not in the case of Gaussian based intervals. When each of the 100 simulated values of weights are validated for significance then we observe considerable differences in the prediction capacity of two types of interval sets. The Gaussian based intervals are underestimating the values of table.1, but not in the case of WBQFWS. In validation process, the proposed WBQFWS provides more accurate bound (above 99% significant close sets) as compared to 95% Gaussian confidence intervals (less than 93% close sets) as show in table 3.

Conclusion

Learning with compressed wavelet neural networks using fuzzy weights will be efficient and higher level of generalization can be obtained with shorter computing time as compared to existing FNNs. We described the architecture of WBQFWS based FNNs that provide better initial search. neuro-fuzzy learning with fuzzy weights requires initialization of an interval based fuzzy sets, which require higher computing than for crisp learning to deal with uncertainty, vagueness and linguistic behaviors of some real life situations see [13], [24] and [26]. Results showed that less than 1% chance of bound independent values is possible, thus providing above 99% accurate mapping, in comparison with Gaussian bounds that are less than 93% accurate.

Secondly, from [11] and [21], each hidden weight connection of neurons lies approximately in the interval

$$-\frac{1}{\sqrt{n_h}} < w_{ij} < \frac{1}{\sqrt{n_h}}$$

Table 1
A comparison of WBQFWS with 95% Gaussian confidence intervals

Weight position (from fig. 1)	(a) Wavelet based fuzzy weight sets				(b) 95% Gaussian confidence interval			
	Min	Max	Mean	S. E.	Confidence Bound		Mean	S. E.
W1(1,1)	-1.9205	3.0983	-0.0637	0.8606	-1.5256	1.4032	-0.0612	0.8903
W1(2,1)	-1.3789	3.3606	-0.0447	0.8025	-1.3914	1.3065	-0.0424	0.8201
W1(3,1)	-2.0946	3.1390	-0.1147	0.8544	-1.5654	1.3390	-0.1132	0.8829
W2(1,1)	-2.9017	3.2209	0.26124	1.2035	-1.8234	2.3362	0.2564	1.2644
W2(1,2)	-1.9873	3.1346	0.1987	1.0948	-1.6512	2.0429	0.1959	1.1229
W2(1,3)	-2.2610	2.9249	0.3128	1.0381	-1.4537	2.0798	0.3130	1.0741

Table 2
Validation of WBQFWS and 95% Gaussian confidence intervals

Weight position (from fig. 1)	(a) Wavelet based fuzzy weight sets				(b) 95% Gaussian confidence interval			
	Min	Max	Mean	S. E.	Confidence Bound		Mean	S. E.
W1(1,1)	-1.9205	3.0983	-0.0621	0.8921	-1.5290	1.4051	-0.0619	0.8921
W1(2,1)	-1.3789	3.3606	-0.1384	0.6248	-1.1659	0.8892	-0.1384	0.6248
W1(3,1)	-2.0946	3.1390	-0.0951	0.8352	-1.4685	1.2782	-0.0951	0.8351
W2(1,1)	-2.9017	3.2209	-0.0004	0.9267	-1.5243	1.5236	-0.0004	0.9267
W2(1,2)	-1.9873	3.1346	0.0991	1.1389	-1.7740	1.9721	0.0991	1.1389
W2(1,3)	-2.2610	2.9249	0.3918	1.1216	-1.4526	2.2362	0.3918	1.1215

Table 3
Number of insignificant WBQFWS and Gaussian confidence interval
(in 100 simulations for validation)

Weight position (from fig. 1)	WBQFWS	95% Gaussian C.I
W1(1,1)	1	5
W1(2,1)	1	2
W1(3,1)	1	7
W2(1,1)	0	7
W2(1,2)	0	13
W2(1,3)	0	13
Deficiency	0.43%	7.8%
Accuracy through validation	99.57%	92.2%

and our proposed set in eq. (5) provides little large interval to search for weights of hidden part of a FNN with above 99% accuracy.

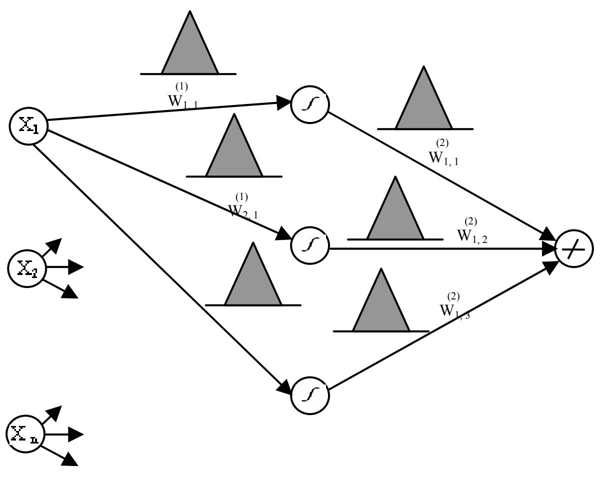

Fig.4. Proposed WBQFWS neural network

Future work

Further improved identification of suitable membership function is possible by determining the underlying probability structure of synaptic connections of a crisp neural network using non-parametric statistics. Thus based on this idea, we can form fuzzy inference systems with varying rules based on neuron complexity. This may provide new research directions to compare different WBQFWS based FNNs. A comparison for most suitable wavelet and optimization algorithms with varying learning parameters is also possible. As future work we will extend this concept on type-II fuzz logic systems as worked by [15].

Acknowledgements

We are very thankful for the support of Higher Education Commission (HEC) of Pakistan and Pakistan Army.

We also acknowledge the useful discussion with Prof. Dr. Mohsin Raza Naqvi of Department of Physics and Dean of Science, University of Balochistan, Quetta.

References

[1] Aqil Burney S. M., Tahseen Jilani and Cemal Ardil, Approximate Bounded Knowledge Extraction using Type-I Fuzzy Logic, in proc. of the 5th World Enformatika Conference, WEC'05, August 26-28, 2005 in Prague, Czech Republic.

[2] _____, A comparison of First and Second Order Learning algorithms in Artificial Neural Networks, International Journal of computational intelligence, vol. 1. no. 3, 2004, pp. 218-224.

[3] _____, Levenberg – Marquardt Algorithm for Karachi Stock Exchange Share Rates Forecasting, International Journal of

Computational Intelligence, vol. 1. no. 2, 2004. pp. 168-173.

[4] Aqil Burney S.M, Tahseen Jilani, Time Series forecasting using artificial neural network methods for Karachi Stock Exchange, (unpublished), 2002. A project in the Dept. of Computer Science, University of Karachi.

[5] Bishop M., Neural networks for pattern recognition, Clarendon Press, 1995.

[6] Buckley J. J., and Hayashi Y., Neural networks for fuzzy systems, Fuzzy Sets and Systems, 1995, pp. 265-276.

[7] Castellano G. and Fanelli A. M., Fuzzy inference and rule extraction using neural network, Neural Network World Journal, vol. 3,2000, pp. 361-371.

[8] Cohen A., Numerical analysis of wavelet methods, Studies in Mathematics and Its Applications, vol.32, 2003, Elsevier publishers.

[9] Daubechies I., Ten Lectures on Wavelets. CBMS-NSF Regional Conference Series, SIAM, Philadelphia, PA, U.S.A, 1992.

[10] Dunyak J. and Wunsch D., Training fuzzy numbers neural networks with alpha-cut refinement, in proc. IEEE Int. Conf. System, Man, Cybernetics, vol. 1, 1997, pp. 189-194.

[11] Haykin S., Neural Networks: A comprehensive foundation. Prentice Hall, 1999

[12] Hee. Y. Y., Chong F. L. and Zhong, B. L., A hierarchical evolutionary algorithm for constructing and training wavelet networks, Neural Computing & Applications, vol. 10, 2002, pp. 357-366.

[13] Ishibuchi H. and Nii M., Numerical analysis of the learning of fuzzified neural networks from if-then rules, Fuzzy Sets Syst., vol. 120, no. 2, 2001, pp. 281-307.

[14] Jang, Sun and Mizutani, Neuro-fuzzy logic and Soft Computing, a computational approach to learning and machine intelligence, Practice-Hall, 2003.

[15] Jerry M. Mendel, Uncertainly Rule-Based Fuzzy Logic Systems, Introduction and new Directions, Prentice Hall PTR, NJ.2001.

[16] Kaiser G., A friendly guide to wavelets, Birkhauser, Boston, MA, U.S.A, 1994.

[17] Mallat S. G., A theory of multiresolution signal decomposition: the wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, 1989, pp. 674-693.

[18] Mir F. Atiya, Suzan M. El-Shoura and Samir I. Shaken, A comparison between neural network forecasting techniques- case study: river flow forecasting, IEEE Trans. on Neural Networks. Vol. 10, No. 2. 1999.

[19] Mitra S., and Hayashi Y., Neuro-fuzzy rule generation: Survey in soft computing framework, IEEE Trans. Neural Networks., vol. 11, no. 3, 2000, pp. 748–768.

[20] Percival D. B., and Walden, A. T., Wavelet methods for time series analysis, Cambridge University Press, 2002.

[21] Puyin Liu and Hongxing, Efficient Learning Algorithms for Three-Layer Regular Feedforward Fuzzy Neural Networks, IEEE Trans. Fuzzy Syst., vol. 15, no. 3, 2004, pp. 545-558.

[22] Scarselli F., and Tosi A. C., Universal approximation using feedforward neural networks: A survey of some existing methods, and some new results, Neural Networks, vol. 11, no. 1, 1998, pp. 15-37.

[23] Sungwoo Park and Taisook Han, Iterative Inversion of Fuzzified Neural Networks, IEEE Trans. Fuzzy Syst., vol. 8, no. 3, 2000, pp. 266-280.

[24] Włodzisław Duch, Uncertainty of Data, Fuzzy Membership Functions, and Multilayer Perceptrons, IEEE, Trans. on Neural Network, vol. 16, no.1, 2005.

[25] Zadeh L. A., Outline of a new approach to the analysis of complex systems and decision processes, IEEE Trans. Systems, Man and Cybernetics, vol. 3, 1973, pp. 28-44.

[26] Zadeh L. A., The concept of linguistic variable and its applications to approximate reasoning, Parts I,II,III, Information Sciences, 8(1975) 199-251; 8(1975) 301-357; 9(1975) 43-80. 30.

S. M. Ail Burney received the B.Sc, first class first M.Sc. M.Phil. from Karachi University in 1970, 1972 and 1983 respectively. He received Ph.D. degree in Mathematics from Strathclyde University, Glasgow with specialization in estimation, modeling and simulation of multivariate Time series models using algorithmic approach with software development.

He is currently professor and approved supervisor in Computer Science and Statistics by the High Education Commission, Government of Pakistan. He is also member of various higher academic boards of different universities of Pakistan. His research interest includes AI, soft computing, neural networks, fuzzy logic, data mining, statistics, simulation and stochastic modeling of mobile

communication system and networks and network security. He is author of three books, various technical reports and supervised more than 100 software/Information technology projects of Masters level degree programs and project director of various projects funded by Government of Pakistan.
He is member of IEEE (USA), ACM (USA) and fellow of Royal Statistical Society United Kingdom and also a member of Islamic society of Statistical Sciences. He is teaching since 1973 in various universities in the field of Econometric, Bio-Statistics, Statistics, Mathematic and Computer Science. He has vast education management experience at the University level. Dr. Burney has received appreciations and awards for his research and as educationist including NCR-2002 award for Best National Educationist.

Tahseen A. Jilani received the B.Sc., first class second M.Sc. (Statistics) and M.A. (Economics) from Karachi University in 1998, 2001 and 2003 respectively. Since 2003, he is Ph.D. research fellow with Prof. Dr. S. M. Aqil Burney, in the Department of Computer Science, University of Karachi.
He is member of IEEE-Computational Intelligence Society. His research interest includes AI, neural networks, soft computing, fuzzy logic, Statistical data mining and simulation. He is teaching since 2002 in the fields of Statistics, Mathematic and Computer Science.

Major Afzal Saleemi received his bachelor's degree in mathematics and statistics from the University of Punjab Lahore, Pakistan in 1986.He received his master's degree in statistics also from Punjab University in 1988. He later joined Pakistan Army and served as directing staff (DS) in various prestigious institutions including Pakistan Military Academy. He is presently detailed by his department to do Ph.D. and carry out research in the fields of information technology. He is now research associate with Prof. Dr S. M Aqil Burney in the Department of Computer Science University of Karachi, Pakistan. His areas of research include data mining, wavelets, soft computing, statistical data simulation and modeling.