Neural Network Ensembles Based Approach for Mineral Prospectivity Prediction

Vanaja Iyer, Chun Che Fung, Warick Brown and Kok Wai Wong

***Abstract*— In mining industry, accurate identification of new geographic locations that are favourable for mineral exploration is very important. However, definitive prediction of such locations is not an easy task. In recent years, the use of neural networks ensemble approach to the classification problem has gained much attention. This paper discusses the results obtained from using different neural network (NN) ensemble techniques for the mineral prospectivtity prediction problem. The proposed model uses the Geographic Information Systems (GIS) data of the location. The method is tested on the GIS data for the Kalgoorlie region of Western Australia. The results obtained are compared to some of the commonly known techniques: the majority combination rule, averaging technique, weighted averaging method tuned by Genetic Algorithm (GA) and a newly proposed rule based method. The results obtained using the different techniques are discussed.**

***Index Terms*— Neural network ensemble; Mineral prospectivity; Geographical Information System.**

I. Introduction

The primary objective of research in mining industry is aimed towards finding new mineral deposit locations. A definite method to do this is by drilling the location of interest and to carry out extensive analysis of the composition of the retrieved samples. As this depends on trial and error, it is a costly procedure. The other step involves developing new methods that can predict with reasonable accuracy of locations where new mineral deposits can be found. During the late 1980s, with the advancement in computing software and hardware capacity, commercial geographic information systems (GIS) that could handle large spatial data have become available on a range of platforms. GIS is defined [2] as computer systems that have functional capability to bring together spatial data from a large variety of sources into a single data base as a series of data layers that overlap correctly at all locations. Due to these features, they are emerging as an important technology in mineral exploration. A typical application of GIS in the geological communication is for mineral potential mapping.

In this study, the problem is now defined as the prediction of potential mineral deposit and barren regions of a particular geographic area by analyzing the regional exploration data set contained in the GIS data. This data comes from multiple sources that reflect the geology, geochemistry, geophysics etc of the region. This is a developing area of research with a great potential use of soft computing techniques.

Some of the mineral prospectivity methods that use geosciences data sets include Boolean algebra [7] and the index overlay method [2, 13, 14]. While the binary and index overlay methods are simple, they have disadvantages such as limited information in the map and not being suitable to model complex non-linear relationships. Statistical methods such as the multiple linear regression was one of the earliest methods used in the mineral prospectivity mapping [5, 6, 16]. But the method is based on assumptions that the relationship between the input and output variables is linear. It is obvious that in most cases, such assumption is not true.

Artificial neural networks (ANN) have been extensively used in other fields of research but are still not used exhaustively in the area of mineral exploration. Until recently, backpropagation neural network (BPNN) has made up for a majority of the neural network applications. However, there are inherent issues in using this approach. Lately, further studies are involved in investigating the performance of other neural networks. This paper is pointing to such a direction in deploying alternative ANN architecture to the problem of mineral prospectivity analysis.

Brown et al have published [3, 4] the results of using Back propagation neural network (BPNN) for mineral prospectivity prediction. Probabilistic neural network is used by Singer and Kouda [18, 19] to classify deposits into deposit types based on the presence or absence of 58 ore and alteration mineral. Singer and Kouda [17] compared the performance of probabilistic neural network (PrNN) with weights-of-evidence methods for the prediction of mineral potential. His report has found probabilistic neural network performance to be better. Harris and Pan [9] compared the performance of probabilistic neural network with that of General regression neural network (GRNN) and have found probabilistic neural network gave slightly better results.

From a survey of the methods used previously for mineral favourability prediction problem, it is evident that where neural network has been used, the common practice is to

V. Iyer is a PhD research student at the School of Information Technology, Murdoch University, Murdoch W.A., Australia (e-mail: v.iyer@murdoch.edu.au).

C. C. Fung is the Director of Centre for Enterprise Collaboration in Innovative Systems, Murdoch University, Murdoch, W.A., Australia (e-mail: L.Fung@murdoch.edu.au).

W. Brown is with the Centre for Global Metallogeny The University of Western Australia W.A., Australia (e-mail: wbrown@segs.uwa.edu.au).

K. W. Wong is with the School of Information Technology, Murdoch University, WA, AUSTRALIA (email: k.wong@murdoch.edu.au).

This research work is supported by an Australian Research Council SPIRT Grant.

Authorized licensed use limited to: Murdoch University. Downloaded on June 15, 2009 at 05:06 from IEEE Xplore. Restrictions apply.

train one or more neural networks to find the best method based on the comparison of the performance results obtained on test data set. Bishop [1] describes two drawbacks in this method. Firstly, the waste of effort involved in using different methods and secondly, the possibility that the network chosen with the best performance on the validation set might not be the one with the best performance on a new test data set. This method also requires expertise in the field of neural computing so that it is possible to choose a suitable network architecture and structure. In recent years, neural network ensemble has been reported to have considerable success in classification problems. The technique involves combining the outputs of several separately trained neural networks to form one unified prediction. In this paper some of the commonly used neural network ensemble techniques are compared with a newly developed rule based technique.

One of the earliest reported works on neural network ensemble by Hansen and Salamon [8] concluded that generalization ability of neural network based system can be significantly improved by ensembling. Subsequently, such approach has been applied in various applications [20, 22, 23]. The neural network ensemble is built in two steps:

- training of the neural networks
- Combining the outputs from the trained networks.

In this paper, the problem of combining the predictions of the trained networks is addressed. On the issue of combining the outputs, some of the commonly used approaches are majority voting, averaging and weighted averaging. In this paper a new rule based method is proposed. The experimental results are compared with that obtained for the rule based method.

The rest of this paper is organized as follows. In section II various individual neural networks used for this study is discussed. In section III, various ensemble techniques that are commonly used and the new one proposed in this study are presented. The GIS data set used for testing is discussed in section IV. The experimental results are presented in section V and conclusion in section VI.

II. Individual Neural Networks in Ensemble

In this study four neural networks are trained for the problem of mineral prospectivity prediction. The basic framework of the neural network ensemble used is illustrated in Figure 1. As shown in the figure are the individual predictors: Probabilistic (PrNN), General Regression (GRNN), Polynomial (PNN) and Backpropogation (BPNN) neural networks.

Backpropogation neural network is the most commonly used neural network in many research studies and practical applications. Polynomial neural network is a flexible neural architecture whose topology is not predetermined but developed through learning. The design is based on Group Method of Data Handling (GMDH) which was invented by Prof. A. G. Ivankhnenko in the late 1960s [11] but enhanced by others. He developed GMDH as a means for identifying nonlinear relations between input and output variables. As described in [12], the GMDH generates successive layers with complex links that are individual terms of a polynomial equation.

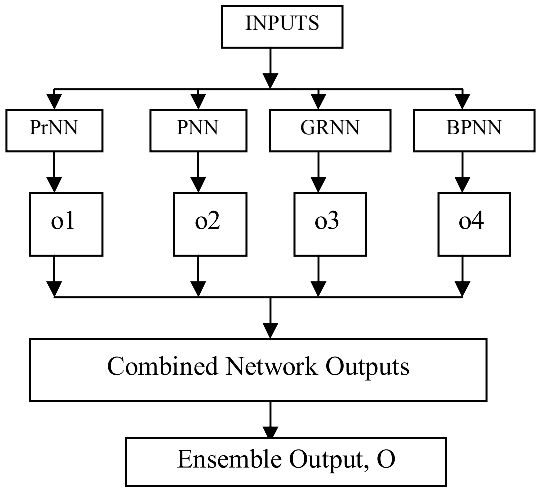

Figure 1: Ensemble Framework

GRNN are memory-based supervised feed-forward networks based on nonlinear regression theory for function estimation. GRNN was originally developed in statistics literature as Nadaraya-Watson kernel regression and was reinvented in 1990 as GRNN by Donald Specht [21]. GRNN is a 3-layer network that has an input layer, hidden layer consisting of at least one node for each pattern it is presented, and an output layer. The transfer function for this paradigm consists of a parameter called a smoothing factor, instead of a learning rate and momentum. Given a training data set and an independent data set, the transfer function is optimised by the selection of a single smoothing factor, which is the common spherical or radial basis function kernel band width. In most applications, there is a unique smoothing factor that produces the minimum Mean Square Error (MSE) between the network output and the desired output. This smoothing factor provides the same purpose in GRNN as the learning rate and momentum in BPNN whoch aims to determine how tightly the data will match the predictions or will fit the function.

Probabilistic Neural Networks (PrNN) are known for their ability to train quickly on sparse data sets. They separate data into a specified number of output categories and hence they are deemed to be suitable for this application. These networks are three layer networks wherein the training patterns are presented to the input layer and the output layer has one neuron for each possible category. There are as many neurons in the hidden layer as there are training patterns. The network produces activations in the output layer corresponding to the probability density function estimate for that category. The highest output represents the most probable category.

III. Ensemble Techniques

Each network is trained with same set of input training data set. Except for the PrNN which gives an output of 0 (indicating barren cell) or 1 (indicating deposit cell), the other three networks outputs values ranging from 0 to 1. These values are classified as representing barren or deposit

Authorized licensed use limited to: Murdoch University. Downloaded on June 15, 2009 at 05:06 from IEEE Xplore. Restrictions apply.

cells by employing different thresholds or cut-off probability values. The outputs are calculated with each threshold (ranging from 0.1 to 0.9) in steps of 0.1. Following four types of ensemble methods are used in this study to combine the outputs obtained from these neural networks.

- ***Majority Voting**:* In this method, the ensemble output which indicates whether the output is barren or deposit is based on a simple majority combination rule. If the outputs from the four neural networks are assigned as o1, o2, o3 and 04, ensemble output O is obtained as follows for each threshold value from 0.1 to 0.9 in steps of 0.1.

  Ensemble Output O = 1
  if three or more of individual outputs (o1, o2, o3 and o4) are greater than threshold value.

- ***Averaging Method*:** In this method, the individual neural network outputs (o1, o2, o3 and o4) are averaged to get the ensemble output.

  Ensemble Output O = (o1+o2+o3+o4)/4

  The output O is then checked for each threshold value from 0.1 to 0.9 to obtain outputs corresponding to each threshold value.

- ***GA tuned Weighted Averaging Method*:** In this method, a weight value is assigned to each neural network output and the ensemble output is calculated using the weighted average formula as follows.

  O = (w1o1+w2o2+w3o3+w4o4)/(w1+w2+w3+w4)
  where w1,w2,w3,w4 are weights assigned to the neural network outputs.

  The output thus obtained is checked for each threshold value from 0.1 to 0.9 to obtain outputs corresponding to each threshold value. Initially a random weight factor is assigned to each neural network output. The GA developed to optimize the weight values has these weight values as chromosome and the sum of the errors of the training set as fitness function. The GA is run for each threshold value from 0.1 to 0.9 in steps of 0.1 to minimize the cumulative error.

- ***Rule Based Method*:** The output values of the PNN, BPNN and GRNN and all the previously discussed ensemble methods range from 0 to 1. So predicting if a region is a deposit or barren cell is again dependant on what is the threshold value above which the area can be considered to have deposit. Some of the earlier researchers have considered this value to be 0.5. So if the output value is 0.5 or greater the area is classified to be a deposit area. In reality this need not be true. To overcome the problem of having to choose the deciding threshold value, a rule based system is developed to determine the deposit cell. The rules are developed based on the individual NN network outputs. Following steps were adapted in the rule development.

  — Compute number of outputs greater than threshold 0.9.
  — The majority rule is then applied on these outputs. For example if all four of the NN outputs for the given input set are greater than 0.9, it is considered to be a deposit cell.
  — The remaining set is considered for further rule extraction. The remaining set has three or less outputs greater than 0.9. Input sets with none or one NN outputs greater than 0.9 is considered as barren.
  — The remaining data set has two outputs greater than 0.9. A set of rules were derived for this data set.
  — The rule set thus formulated is used on independent test data set to generate outputs.

The percentage correct deposit and barren cells is computed for each of the above method and compared. These results are given in Section V.

IV. GIS DATA SET

The GIS database is viewed as a collection of maps of a particular data type (such as solid geology, regional-scale faults etc) for a common geographic coordinate system. Within the map layers, two data structures are used to represent the spatial objects, namely the vector and raster structures. Most GIS support both structures and allow conversion from one structure to another.

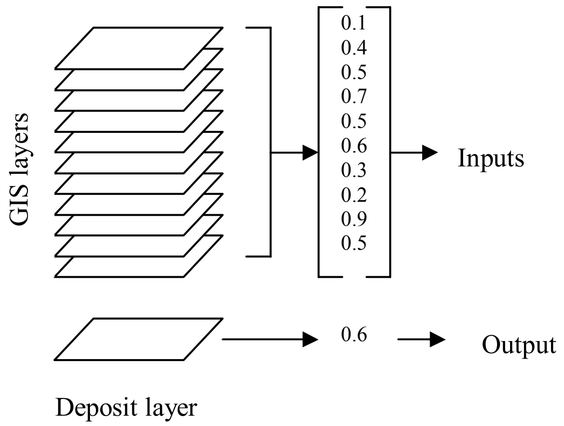

Figure 2. A typical input/output pattern for the neural network

For this study, GIS data set used is the one described in [3] to examine the prospectivity of orogenic Iode-gold deposits in an approximately 100 x 100 km area of the Archean Yilgarn Block, near Kalgoorlie region of Western Australia. In this study, 10 GIS layers in the raster data format are used to create the feature vectors. The GIS layers correspond to information such as solid geology, magnetic anomaly, gamma-ray survey and distance to faults are organized as shown in Figure 2. The thematic layers are divided into a grid of square cells of 100m side. Each cell is represented by the cell position and set of attributes within the two dimensional matrix of cells. The map area thus results in 1,254,000 cells. Out of these, only the 120 deposit cells with a total contained gold source greater than 1,000 kg along with 148 barren or non-deposit cells are used as

Authorized licensed use limited to: Murdoch University. Downloaded on June 15, 2009 at 05:06 from IEEE Xplore. Restrictions apply.

training and test data set.

V. RESULTS

The GIS data available is divided into training and test data set. Out of the data for the 268 cells, 147 cells were used for training and 81 cells for testing. Just like the training data set, the test data set has inputs corresponding to both deposit and barren cells. All the input values are scaled to [0, 1]. Table 1 shows the number of patterns in the test and training data set. All the four NNs are trained as 10-input single output models using the training data set to get the outputs. A typical example of inputs corresponding to 4 cells is given in the Table2.

Table 1. Number of patterns in the training and test data sets

Training Data Set			Test Data Set		
Deposit	**Barren**	**Total**	**Deposit**	**Barren**	**Total**
85	102	187	35	46	81

Table 2. Samples of input patterns

1	2	3	4	5	6	7	8	9	1 0
0.67	0.39	0.04	0.12	0.48	0.01	0.02	0.33	0.11	0.00
0.10	0.14	0.05	0.49	0.72	0.05	0.02	0.02	0.00	1
0.48	0.37	0.23	0.42	0 . 8	0.03	0.06	0.07	0.06	0.67
0.10	0.06	0.42	0.02	0.48	0.32	0.30	0.36	0.17	0
0.13	0.48	0.01	0.33	0.72	0.01	0.16	0.02	0.13	0.67

The NN outputs for the above inputs, along with actual output are given below in Table3. The ensemble output O for these outputs using the four methods described in section IV is computed. This is as given in Table 4 for the set of outputs in Table 4 for a threshold value of 0.9.

Table 3. Sample Network outputs

Cell Type	Actual Output	NN Outputs			
		PrNN o1	PNN o2	GRNN O3	BPNN o4
Deposit	1	1	0.88	1.00	1.00
Deposit	1	1	0.64	0.99	0.99
Deposit	1	0	0.75	1.00	0.99
Barren	0	0	0.07	0.00	0.07
Barren	0	0	0.26	0.13	0.38

Table 4. Samples of Ensemble outputs

o1	o2	o3	o4	m	ave	a	w-ave	w
1	0.88	1.00	1.00	1	0.97	1	0.99	1
1	0.64	0.99	0.99	1	0.90	1	0.96	1
0	0.75	1.00	0.99	0	0.68	0	0.63	0
0	0.07	0.00	0.07	0	0.09	0	0.005	0
0	0.26	0.13	0.38	0	0.107	0	0.039	0

where: **m**- Output using Majority Voting

ave -average of the outputs o1 to o4.

a - Output using Averaging Method for a threshold of 0.9 (a=1 if ave greater than or equal to 0.9)

w- ave-weighted average of the outputs o1 to o4.

w- Output using weighted averaging method for a threshold of 0.9 (w=1 if w-ave greater than or equal to 0.9).

After consultation with experts in the discipline, the rules extracted from the outputs for the training data set using the rule based method has the following format.

Given cell is a deposit cell if any of the following conditions are true:

- If all the four NN outputs are greater than 0.9, the output is one.
- If three of the NN outputs are greater than 0.9 then
  - if these three are (GRNN, PNN and PrNN outputs) then output is one
  - elseif these three are (GRNN, BPNN and PrNN) then output is one
- If two of the NN outputs are greater than 0.9 then
  - if they are GRNN and PrNN and (PNN greater than or equal to 0.4) and (BPNN is greater than or equal to 0.6) then output is one;
  - elseif they are GRNN and BPNN and (PNN greater than or equal to 0.4) and (BPNN is greater than or equal to 0.6) then output is one;
  - elseif they are PrNN and BPNN and (PNN greater than or equal to 0.4) and (GRNN is greater than or equal to 0.6) then output is one.

The above are the main rules used to classify given set of cells into deposit or barren cells. These rules are then applied on the independent test data set. The results obtained for each method is given in Tables 5 to 8.

For all the ensemble methods described above except the rule based method values corresponding to different thresholds is obtained. Now it becomes necessary to choose the threshold that gives the best result. For the majority voting method, based on the training data results, 0.3 and 0.4 appears to give higher correct classification for deposit and barren cells. On the same basis, for the averaging method 0.4 and 0.5 gives a better training set result. For the weighted average method, 0.4 appears to give better results. However what is important is the result given for an independent test set. In the case of weighted average the results are much inferior to other methods. Comparing these results to the rule based method, it is found that for barren cell classification, rule based method is much superior to the other methods. For the deposit cell prediction, the percentage correct value of 77.1% is little less than the 82.9% for majority voting and averaging method. If the average correct value for both deposit and barren cells is considered, the rule based method is found to be better.

Authorized licensed use limited to: Murdoch University. Downloaded on June 15, 2009 at 05:06 from IEEE Xplore. Restrictions apply.

Table 5. Results for the Majority voting method

Threshold value	%correct Training set results		% correct Test set results	
	Deposit	Barren	Deposit	Barren
0.1	98.8	84.3	85.7	65.2
0.2	96.5	90.2	85.7	69.6
0.3	95.3	92.2	82.9	71.7
0.4	91.8	95.1	80.0	73.9
0.5	88.2	95.1	80.0	73.9
0.6	83.5	98.0	74.3	78.3
0.7	80.0	99.0	71.4	80.4
0.8	69.4	99.0	48.6	89.1
0.9	61.2	100.0	45.7	91.3

Table 6. Results of Averaging method

Threshold value	%correct Training set results		% correct Test set results	
	Deposit	Barren	Deposit	Barren
0.1	100	47.1	97.1	37.0
0.2	100	76.5	91.4	52.2
0.3	98.8	88.2	88.6	65.2
0.4	97.6	94.1	82.9	69.6
0.5	94.1	95.1	80.0	73.9
0.6	87.1	95.1	80.0	76.1
0.7	83.5	97.1	74.3	76.1
0.8	76.5	99.0	71.4	87.0
0.9	56.5	99.0	40.0	91.3

Table 7. Results of the Weighted Averaging Method

Threshold value	% correct Training set results		% correct Test set results	
	Deposit	Barren	Deposit	Barren
0.1	100.0	94.1	82.9	56.5
0.2	100.0	94.1	82.9	56.5
0.3	100.0	94.1	82.9	56.5
0.4	97.6	96.1	82.9	69.6
0.5	96.5	90.2	77.1	52.2
0.6	91.8	96.1	74.3	73.9
0.7	90.6	96.1	71.4	78.3
0.8	89.4	99	62.9	80.4
0.9	88.2	92.2	77.1	54.3

Table 8. Results for the Rule based method

% correct Training set results		% correct Test set results	
Deposit	Barren	Deposit	Barren
90.6	97.1	77.1	80.2

VI. CONCLUSION

The neural network ensembles are better way of solving real life problems without the necessity to have in-depth knowledge about the NN architecture and parameters. The proposed rule based method paves a new path to the application of mineral prospectivity prediction where the prediction of both deposit and barren cells are important to the mining industry. Further research is being conducted to automate the rule generation.

REFERENCES

[1] Bishop, C.M., *Neural Networks for Pattern Recognition*. 1995: Oxford University Press.

[2] Bonham_Carter, G. F., Geographic Infomation Systems for scientists: Modelling with GIS. 1994, New York: Pergamon Press.

[3] Brown, W.M., et al., Artificial neural networks: a new method for mineral prospectivity mapping. Australian Journal of Earth Sciences, 2000(47): p. 757-770.

[4] Brown, W.M., Groves D.I., and Gedeon T., Use of noise to augment training data: a neural network method of mineral potential mapping in regions of limited known deposits. Natural Resources Research, 2003. 12(2).

[5] Geoffroy, J.D. and Wignall, T.K., A probabilistic appraisal of mineral resources in a portion of the Grenville province of the Canadian Shield. Economic Geology, 1971. 66: p. 466-479.

[6] Geoffroy, J.D. and Wignall, T.K., Statistical decision in regional exploration: application of regression and Bayesian classification analysis in the southwest Winconsin zinc area. Economic Geology, 1970. 65: p. 769-777.

[7] George, H. and Bonham_Carter, G.F., Spatial modelling of geological data for gold exploration Star Lake area. Statistical applications in the earth Sciences, 1990(157-169).

[8] Hansen, L.K. and P. Salmon, *Neural Network Ensembles*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990. **12**(10): p. 993-1001.

[9] Harris, D. and Pan, G.., Mineral favourability mapping: a comparison of artificial neural networks, logical regression and discriminant analysis. Natural Resources Research, 1999. 8: p. 93-109.

[10] Islam, M.M., X. Yao, and K. Murase, *A constructive A logorithm for Training Cooperative Neural Network Ensembles*. Ieee transactions of neural networks, 2003. **14**(4): p. 820-834.

[11] Ivakhnenko, A.G., Polynomial theory of complex systems. IEEE transactions on systems, man and cybernetics, 1971. 1(4): p. 364-378.

[12] Ivakhnenko, A.G., The Review of Problems Solvable by Algorithms of the Group Method of Data Handling (GMDH). Pattern Recognition and Image Analysis, 1995. 5(4): p. 527-535.

[13] Koch, G.S., Howarth, R.J., Scheunemeyer J.H. and Lowry R.K., Uranium resource assessment through statistical analysis of exploration geochemical and other data Economic Geology, 1981. 76: p. 1056-1066.

[14] Reddy, R.K.T. and G.S. Koch, A case study of silver exploration in parts of Idaho and Montana using mathematical modelling techniques. Economic Geology, 1988. 83(1008-1014).

[15] Sarycheva, L., Using GMDH for modeling economical indices of mine opening. Systems Analysis Modelling Simulation, 2003. 43(10): p. 1341-1349.

[16] Sinclair, A.J. and G.J. Woodsworth, Multiple regression as a method of estimating exploration potential in an area near Terrace, B. C. Economic Geology, 1970. **65**: p. 998-1003.

[17] Singer, D.A. and Kouda, R., A comparison of the weights-of-evidence method and probabilistic neural networks. Natural Resources Research, 1999. **8**(4): p. 287-298.

[18] Singer, D.A. and Kouda, R., Classification of mineral deposits into types using mineralogy with a probabilistic neural network. Nonrenewable Resources, 1997. **6**: p. 27-32.

[19] Singer, D.A. and Kouda, R., Use of a neural network to integrate geoscience information in the classification of mineral deposits and occurrences in Proceedings of Exploration 97: Fourth Decennial International Conference on Mineral exploration. 1997. Toronto, Canada.

[20] Shu, C. and D.H. Burn, Artificial Neural Network Ensembles and their Application in Pooled Flood Frequency Analysis. Water Resources Research, 2004. **40**.

[21] Specht, D.F., A General Regression Neural Network. IEEE transactions on Neural Networks, 1991. 2(6): p. 568-576.

[22] Zhou, Z.-H., et al., *Lung cancer cell Identification based on artificial neural network ensembles*. Artificial Intelligence in Medicine, 2002. **24**(1): p. 25-36.

[23] Zhou, Z.-H., Y. Jiang, and S.-F. Chen, *Extracting Symbolic Rules from Trained Neural Network Ensembles*. AI Communications, 2003. **16**(1): p. 3-15.

Authorized licensed use limited to: Murdoch University. Downloaded on June 15, 2009 at 05:06 from IEEE Xplore. Restrictions apply.