Proceedings of International Joint Conference on Neural Networks, San Jose, California, USA, July 31 – August 5, 2011

Traffic Flow Breakdown Prediction using Feature Reduction through Rough-Neuro Fuzzy Networks

C. Affonso
Universidade Estadual Paulista
São Paulo, Brazil
affonso@itapeva.unesp.br

R. J. Sassi (IEEE Member)
Nove de Julho University
São Paulo, Brazil
renato.sassi@ieee.org

R. P. Ferreira
Nove de Julho University
São Paulo, Brazil
kasparov@uninove.edu.br

Abstract— **The prediction of the traffic behavior could help to make decision about the routing process, as well as enables gains on effectiveness and productivity on the physical distribution. This need motivated the search for technological improvements in the Routing performance in metropolitan areas. The purpose of this paper is to present computational evidences that Artificial Neural Network ANN could be use to predict the traffic behavior in a metropolitan area such São Paulo (around 16 million inhabitants). The proposed methodology involves the application of Rough-Fuzzy Sets to define inference morphology for insertion of the behavior of Dynamic Routing into a structured rule basis, without human expert aid. The dynamics of the traffic parameters are described through membership functions. Rough Sets Theory identifies the attributes that are important, and suggest Fuzzy relations to be inserted on a Rough Neuro Fuzzy Network (RNFN) type Multilayer Perceptron (MLP) and type Radial Basis Function (RBF), in order to get an optimal surface response. To measure the performance of the proposed RNFN, the responses of the unreduced rule basis are compared with the reduced rule one. The results show that by making use of the Feature Reduction through RNFN, it is possible to reduce the need for human expert in the construction of the Fuzzy inference mechanism in such flow process like traffic breakdown.**

Keywords: Artificial Neural Network; Feature Reduction; Fuzzy Sets; Rough Sets; Traffic Breakdown.

I. INTRODUCTION

The dynamic routing forecasting could increase significantly the large cities traffic efficiency. The ability to predict the traffic flow variances could aid to choose the best way, in order to avoid lost of time on deliveries, mainly when the forecasts point to traffic flow breakdown. It is possible to find these parameters analytically [1,2]; however, solving this problem by applying classical theories of transport phenomena requires accurate information about the flow process.

Considering the above points, it is possible to think that the use of Fuzzy Sets is necessary, once this type of approach is especially applicable to systems where information is inaccurate and the strategies used by human experts in dealing with these systems are expressed through linguistic terms [3,4].

A sequential architecture is used in this work, once the Rough Sets and the Neuro Fuzzy Network have distinct functions: Rough Sets selects the critical features, while the Neuro Fuzzy Network generates the surface response input / output, considering that the Neuro Fuzzy Network has learnability and can adapt itself to the real world [5].

The great advantage of the RNFN approach consists in the synergy achieved by combining two or more technical capabilities to obtain a more powerful system with respect to learning and generalization process. [1,2].

The paper is organized as follows: section 2 introduces the fundamental concepts of Rough Sets. Section 3 presents the application of Fuzzy Sets to Traffic Flow. Section 4 presents a proposal for an application of RNFN. The methodology and experiments are discussed in section 5, and the conclusion and discussion in section 6.

II. ROUGH SET THEORY

The Rough Sets Theory (RS) was proposed by Zdzislaw Pawlak in 1982 as a mathematical model to represent knowledge and to treatment of uncertainty. An important concept in RS is the *reduct* concept. A reduct is a minimal set of attributes that can represent an object with the same accuracy as the original set of attributes. Elimination of redundant attributes can help in identification of strong non-redundant classification rules [9, 14].

A reduct of B – RED(B) – on information system (IS) is a set of attributes $B' \subseteq B$ assign all attributes $a \in (B - B')$, which are dispensable. Thus, $U/IND_S(B') = U/IND_S(B)$, where $IND_S(B)$ is, called the of Indiscernibility Relation [8].

The reducts computation is a n-p hard problem and the reducts processing in large databases require high computational processing. The reducts is generated by discernibility from Discernibility Matrix.

The Discernibility Matrix of information systems S, stands by DM(B) is a symmetric matrix n x n with: $mD_{(i,j)} = \{a \in B \mid a(E_i) \neq a(E_j)\}$ for i,j=1,2,…,n. with $1 \leq i, j \leq n$ e $n = |U / IND_S(B)|$. Thus, the elements of Discernibility Matrix mD(i,j) is a set of conditional attributes of the B that differentiate the elements of classes in relation their nominal values.

$$F_S(\mathbf{a}_1^*, \mathbf{a}_2^*, \ldots, \mathbf{a}_m^*) = \wedge\{\vee\, m_D^*(i,j) \mid i,j = 1,2,\ldots,n, \quad m_D(i,j) \neq 0\}$$
$$\text{With: } m_D^*(i,j) = \{\mathbf{a}^* \mid \mathbf{a} \in m_D(i,j)\} \tag{1}$$

With the simplifications methods of booleans functions in Fs(B) function, the reducts of S are generated. This simplification is an algebric handle of the logics functions with the goal of reduct the number of attributes and needs operation.

The discernibility function Fs(B) is obtained as following: For all attributes stands into each element from Discernibility Matrix MD(B) is applied the sum operator, “or” or “$\vee$” and, among the cells of this matrix is used the “product” element, “and” or “$\wedge$”, which result in a Boolean expression of “sum of products”.

The discernibility function Fs(B) is defined from the discernibility matrix, as defined on the TABLE I.

TABLE I
DISCERNILITY MATRIX

	E1	E2	E3	E4	E5	E6
E1	∅					
E2	EV	Æ				
E3	EV	Æ	Æ			
E4	EV, QL	EV, QL	EV, QL	Æ		
E5	EV,QL,BL	QL,BL	QL,BL	EV, QL	Æ	
E6	QL,BL	EV,QL,BL	EV,QL,BL	EV, QL	EV	∅

The Fuzzy Sets evaluates the membership among the elements from the same class, while RS is concerning to the relationship among groups of elements in different classes. Therefore, RS does not compete with the Fuzzy Sets Theory but complements it. In fact, the RS theory and the Fuzzy Sets theory are two independent approaches for the treatment of imprecise knowledge [7].

The knowledge acquisition bottleneck is a significant problem that hinders the building of intelligent monitoring systems. The generation of a good knowledge bases for this task is notoriously difficult. This is particularly the case when experts are not readily available [6]. Machine learning techniques (especially rule induction methods) can be of great benefit to this area by providing strategies to automatically extract useful knowledge, given large enough historical datasets [10].

Rough Selection provides a means by which discrete or real valued noisy data (or a mixture of both) can be effectively reduced without the need for user-supplied information. Additionally, this technique can be applied to data with continuous or nominal decision attributes, and as such can be applied to regression as well as classification datasets. The only additional information required is in the form of Fuzzy partitions for each feature which can be automatically derived from the data. This corresponds to the case where only the decision attribute values are Fuzzy, once the conditional values are crisp [11].

III. FUZZY SETS

In 1965, Zadeh [4] assigned a number to every element in the universe, which indicates the degree to which the element belongs to a Fuzzy set.

To formulate this concept of Fuzzy set mathematically; we present the following definition [12]. Let X be the universe. A mapping $A: X \to [0,1]$ is called a Fuzzy set on X. The value $\mu(x)$ of A, at $x \in X$ stands for the degree of membership of x in A. The set of all Fuzzy sets on X will be denoted by F(X). $\mu(x) = 1$ means full membership, $\mu(x) = 0$ means non-membership and intermediate values between 0 and 1 mean partial membership, $\mu(x)$ is referred to as a membership function as x varies in X. Considering $\mu(x)^c = 1-\mu(x)$. Like ([0,1], $\vee$, $\wedge$) ,(F(X),$\cup$,$\cap$) is not a Boolean algebra since it is not complemented, i.e. $A \cup A$ c = X and $A \cap A$ c = $\varnothing$ do not hold generally.

Importantly, the choice of a membership function is context-dependent. For instance, it is clearly different that the temperature of a steel-smelting furnace is high and the temperature of a human body is high. Even in a same context, the choice is dependent on the observer [12].

A relation is a subset of the Cartesian product of two sets. It is automatically fuzzified while a subset is fuzzified. A Fuzzy relation could be defined as follow: Let A and B be two non-empty sets.

A mapping R: Ax$B \to$ [0,1] is called a Fuzzy (binary) relation from A to B. For $(x,y) \in A$xB, R$(x,y) \in [0,1]$ is referred to as the degree of relationship between x and y.

IV. FUZZY SETS APPLIED TO TRAFFIC BEHAVIOR

The analysis of the traffic flow involves a series of notable occurrences that affect its behavior. TABLE II shows the types of occurrences that were used in the RNFN.

TABLE II
NOTABLE OCCURRENCES THAT OPERATE IN THE TRAFFIC FLOW.

Notable Occurrences	
1. Accident victim	9. Manifestations
2. Flooding	10. Occurrence with loaded truck
3. Running over	11. Occurrence with dangerous goods
4. Truck broken	12. Bus asset towards
5. Defect in the trolleybus	13. Falling Tree
6. Lack of electricity	14. Light off
7. Fire	15. Light flagged
8. Fire Vehicle	16. Vehicle excess

Based on the database were defined membership functions for all the variables involved in the process, the criterion for defining the deviation and center of each of these functions was achieved with support from the experts[19]. Were considered 3 linguistic labels α for each membership functions: low, medium, high; with their respective centers c_α, and standard deviation σ.

To model the membership functions we used the Gaussian function.

$$\mu^{\alpha}(x) = e^{-\frac{1}{\sigma}(x-c_{\alpha})^{2}} \quad (2)$$

To establish an inference mechanism to simulate a system, are required two stages, as suggested in [13]:

- The premises of all rules are compared with controlled entries to determine which rules apply to a situation.
- The outputs are compared with the established rules that have been determined.

In this paper, we used the T-norm applied to input $x = (x_1, x_2,..,x_n)$ as suggested by [13], where the value of $\Phi_{j,\, j=1,\dots,m}$ rules of inference be calculated as follows:

$$\Phi_j(x, y) = T(R(x), \mu_B^{\alpha y}(y)) \quad (3)$$

To obtain the network output value at the generalization phase is necessary to perform the clustering of the values from inference rules [12].

Where Fuzzy relation $R_{j,\ j=1,\dots,m}$ for each inference rule, are given as:

$$R_j(x) = \mu_{x1}^{\alpha 1}(x_1) \wedge \mu_{x2}^{\alpha 2}(x_2) \wedge \dots \wedge \mu_{xn}^{\alpha n}(x_n) \quad (4)$$

V. A ROUGH-NEURO FUZZY NETWORK (RNFN)

The Fig. 1. illustrates the full process of Rough-Neuro Fuzzy Network. In the initial fase, the RS preprocesses the datasets and the input vectors with minimal attributes are generated.

Through the inference mechanism established for the Traffic flow was determined the values of the rules Φ_j as shown in equation (3). This reduced rule base is used to training the RBF/MLP and RBF ANN.

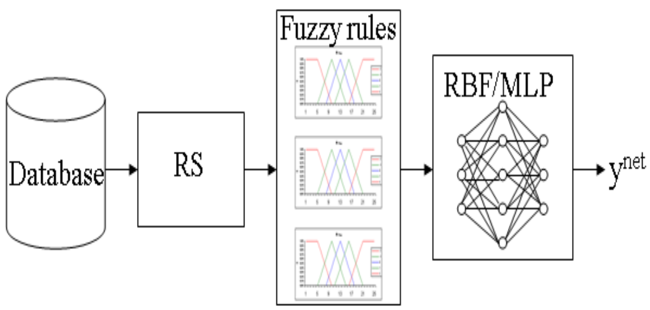

Figure 1. A Rough-Neuro Fuzzy Network (RNFN).

This paper presents a supervised feedback ANN architecture in 3 layers: the input layer represents the values of the reduced rule base. The neurons of network hidden layer are trained from the set of inference rules (reduced).

Will be considered initially RBF architecture. In this architecture are used, in order to approximate the surface response, a combination of Radial Basis Functions, in this particular case the Green function [15];

$$G(\Phi - c_i) = e^{-\frac{(\Phi - c_i)2}{\sigma}} \quad (5)$$

The learning algorithm chosen for this application was the K-means, where the synaptic weights W, and the centers of the basis functions C are recalculated from the parameters of learning η_1 and η_2, in order to minimize the error ε, as follows[15];

$$\begin{gathered} w(n+1) = w(n) + \eta_1 \sum \varepsilon(n) G_{ij}(x_j - c_i) \\ c_k(n+1) = c_k(n) + \eta_2(x - c) \end{gathered} \quad (6)$$

A Multilayer Perceptrons is one of the most used ANN in classification. The ANN architecture MLP typically consist of a specification of the number of layers, type of activation function of each unit and weights of connections between the different units should be established for the construction of the neural architecture [15].

The algorithm used in training MLP is the error backpropagation that works as follows: first a standard is presented in this work a standard will be a prototype vector and its label - the input layer of the network.

This pattern is processed layer by layer until the output layer provides the response rendered, fMLP, calculated as shown below:

$$f_{MLP}(x) = \varphi\left(\sum_{i=1}^{n} v_l \varphi\left(\sum_{j=1}^{m} w_{lj} x_l + b_{l0}\right) + b_0\right) \quad (7)$$

Where v_l and w_{lj} are synaptic weights; b_{l0} and b_0 are the biases, and φ the activation function, usually specified as the sigmoid function.

VI. CONDUCT OF EXPERIMENTS

This work should be viewed as an application of artificial intelligence techniques, in particular, RNFN to determine the flow systems behavior. In accordance with this approach [16], Neuro Fuzzy Networks have been used to solve such problems [20].

The literature [17,18] shows that flow process are phenomenon where direct mathematical modeling is not feasible due to the large number of variables. Moreover, as is typical in studies that use Fuzzy Logic, it is required indispensable tacit knowledge of experts in the field.

To perform an analysis about the traffic behavior, especially in large cities, the main difficulty is to determine a reasonable number of parameters to be studied. Thus the use of Rough Sets is intended to identify the variables that have greater relevance to the proceedings and eliminate those which have little significance.

An inference mechanism for the learning and generalization phases was established from the membership functions of variable, according to the main attributes selected by the Rough Sets.

The algorithm for construction of the MLP and RBF network was entirely developed by the authors. Scilab 5.1 was used as a programming language. The hardware platform used in the experiments was a Pentium Dual Core with 2.4 MHZ, 512 MB RAM and 40 GB hard drive.

A. Rough Neuro-Fuzzy Network (RNFN) type MLP and RBF applied to traffic flow breakdown in the city of São Paulo

Data were collected from notable occurrences of traffic in the metropolitan region of São Paulo on December 2009.

The main goal was to analyze the impact of such occurrences in the flow and on the traffic behavior. Experiments were conducted analyzing figures from the São Paulo traffic behavior, using the same inference mechanism discussed above.

Fig. 2 below shows the value of the average error, $\varepsilon = 1/ne \sum | Y_{NFN} - Y_{REAL} |$, n_e = number of events.

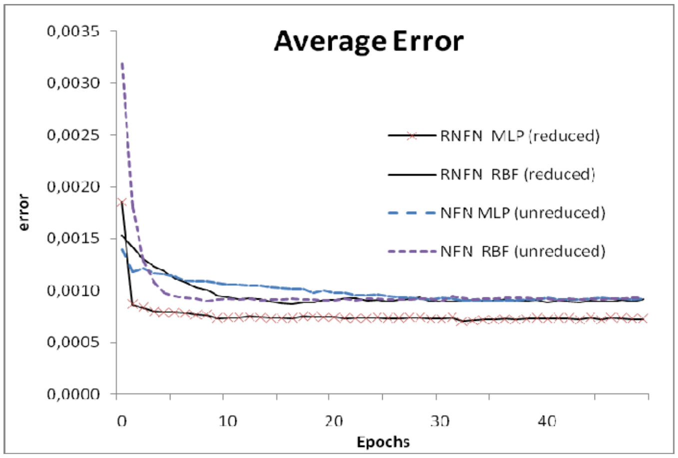

Figure 2. Average error: Rough-Neuro Fuzzy Network type MLP (RNFN-MLP) and type RBF (RNFN-RBF) reduced and unreduced by Rough Sets.

The choice of parameters for both RNFN was: number of hidden neurons 10; and learning rate constant η = 0.01; momentum factor = 0.75; and the maximum epochs 150.

RNFN MLP has initially higher average error, which decreases asymptotically with the increase in the epochs. Both rule bases reduced by the RS exhibit good behavior during the generalization phase.

The result indicates a promising way for to associate RS with Neuro Fuzzy networks to replace the human expert in the construction of inference rules.

The graph in the Fig. 3 shows the rate of slow transit logged every 30 minutes, since 6h to 18h. The database was obtained from Metropolitan Traffic Company of São Paulo (CET).

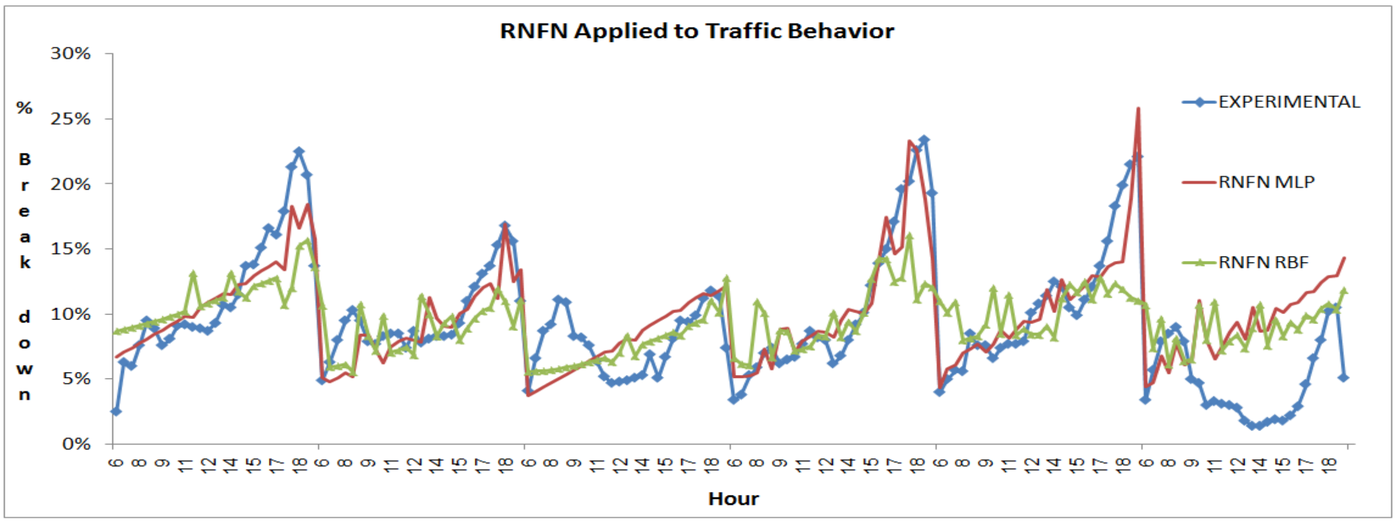

Figure 3. Rough Neuro Fuzzy Network applied to Traffic Behavior: Experimental, RNFN type MLP and RNFN type RBF

VII. Discussion and Conclusion

The application of RNFN showed a great ability to generalize, identify behavior patterns, and allow the creation of an inference mechanism in high complex systems, such as traffic behavior.

The aim is to continue with this early work using other data samples collected on different days of the week and in different months, and input them on the RNFN in order to predict in any day the traffic of the city of São Paulo.

Furthermore, by performing a comparative analysis between the reduced and unreduced rule basis, it is possible to conclude:

-The reduced rule basis has a better ability to generalize that an unreduced one.

-The reduced number of membership functions for each inference rule allows the creation of an automatic inference mechanism without the support of a human expert.

When applied to a real world dataset, the RNFN was able to identify the significant features like defined by a human expert. The main advantage of use the RNFN is the reduction of dependence on a human expert for the choice and construction of the rules of inference mechanism.

This gain is important, considering that one of the weaknesses of the approach using the Fuzzy Sets is its dependence on the human expert. If there is a reasonable number of attributes and a structured database, it may be even possible to eliminate the need for support from the human expert for the construction of inference rules, using his support only in the construction of membership functions.

Wavelet neural networks (WNN) have recently attracted great interest, because of their advantages over radial basis function (RBF) as it is an universal approximations; therefore we suggest the use of WNN as an alternative network in future works.

ACKNOWLEDGEMENT

This work was supported by Universidade Estadual Paulista and Nove de Julho University.

REFERENCES

[1] C. Affonso and R.J. Sassi. An inference Mechanism for Polymer Processing Using Rough-Neuro Fuzzy Network. K. Diamantaras, W. Duch, L.S. Iliadias (Eds.): ICANN 2010, Part III, LNCS 6354, pp. 441-450, Springer, 2010.

[2] R. P. Ferreira, C. Affonso and R.J. Sassi. Application of a neuro fuzzy network to forecast the behavior of the urban traffic of the city of São Paulo. IEEE Conference Information Systems and Technologies (CISTI), 2010 5th Iberian Conference on, Santiago de Compostela, 2010.

[3] L.A. Zadeh, “Fuzzy Sets: Information and Control”, v.8, pp 338-353, 1965.

[4] T. Takagi and M. Sugeno, “Derivation of Fuzzy control rules from human operators control action”, in Proc. IFAC Symp. Fuzzy Inform. Knowledge Representation and Decision Analysis, pp. 55–60, 1983.

[5] F. Gomide, M. Figueiredo, W. Pedrycz, “A neural Fuzzy network: Structure and learning, Fuzzy Logic and Its Applications, Information Sciences and Intelligent Systems”, Bien, Z. and Min, K., Kluwer Academic Publishers, Netherlands, pp. 177-186, 1998.

[6] L. M. F. Carvalho, “A neuro-fuzzy system to support in the diagnostic of epileptic events using different fuzzy arithmetical operations. Neuropsiquiatria”, 2008.

[7] Z. Pawlak, “Rough Sets”, International Journal of Computer and Information Sciences, pages: 341-356, 1982.

[8] R. J. Sassi, L.A. Silva, Del Moral Hernandez, Emilio , “A Methodology using Neural Networks to Cluster Validity Discovered from a Marketing Database”. In: 10th Brazilian Symposium on Neural Networks (SBRN), 2008, Salvador.

[9] R. Jensen, “Combining rough and Fuzzy set for feature selection”, University of Edinburgh, 2005.

[10] Y. Y. Yao, “Combination of Rough and Fuzzy Sets Based on □-Level Sets”. Rough Sets and Data Mining: Analysis for Imprecise Data, 1997.

[11] X. Wang, D. Ruan, E. Kerre, “Mathematics of Fuzziness”, Chapter 3, Springer, 2009.

[12] S. K. Halgamuge, M. Glesner, “Neural networks in designing Fuzzy systems for real world applications”, Fuzzy Sets and Systems 65, 1994.

[13] E. H. Mamdani and S. Assilian, “An experiment in linguistic synthesis with a fuzzy logic controller”. Int. J. Man Mach. Studies 7, 1–13 (1975).

[14] Z. Pawlak,”Why Rough Sets? Fuzzy Systems”. Proceedings of the Fifth IEEE International conference on, volume 2, 8-11, sept 1996, pages: 738-741.

[15] S. Haykin, “Neural Networks: A Comprehensive Foundation.” New York: Willey & Sons, 1999.

[16] T. I. Liu, X. M. Yang, G. J. Kalambur, 1995, ‘‘Design for Machining Using Expert System and Fuzzy Logic Approach,’’ ASME J. Mater. Eng. Perform. 4, No. 5, pp. 599–609.

[17] B. S. Kerner, Introduction to Modern Traffic Flow Theory and Control, DOI 10.1007/978-3-642-02605-8 11,pp.221-243 Springer-Verlag Berlin Heidelberg 2009.

[18] H. Reiss, A.D. Hammerich, and E.W. Montroll. Thermodynamic treatment of nonphysical systems: Formalism and an example (single-lane traffic). Journal of Statistical Physics, 42(3/4):647–687, 1986.

[19] T. I. Liu, X. M. Yang, and G. J. Kalambur, ‘‘Design for Machining Using Expert System and Fuzzy Logic Approach,’’ ASME J. Mater. Eng. Perform., 4, No. 5, pp. 599–609, 1995

[20] Y. Peng, G. Kou,Y. Shi, and Z. Chen, A Descriptive Framework for the Field of Data Mining and Knowledge Discovery , International Journal of Information Technology and Decision Making, Vol. 7, Issue: 4, Page 639 - 682, 2008