A Hybrid Method for Tuning Neural Network for Time Series Forecasting

Aranildo R. L. Junior
Department of Physics
Federal University of Pernambuco
Recife, Pernambuco, Brazil
proj.brain@gmail.com

Tiago A. E. Ferreira
Statistics and Informatics Department
Rural Federal University of Pernambuco
Recife, Pernambuco, Brazil
taef.first@gmail.com

ABSTRACT

This paper presents an study about a new Hybrid method - GRASPES - for time series prediction, inspired in F. Takens theorem and based on a multi-start metaheuristic for combinatorial problems - Greedy Randomized Adaptive Search Procedure(GRASP) - and Evolutionary Strategies (ES) concepts. The GRAPES tuning and evolve the Artificial Neural Network parameters configuration, the weights and the minimum number of (and their specific) relevant time lags, searching an optimal or sub-optimal forecasting model for a correct time series representation. An experimental investigation is conducted with the GRASPES with some time series and the results achieved are discussed and compared, according to five well-known performance measures, to other works reported in the literature.

Categories and Subject Descriptors

I.6.5 [**Model Development**]: Modeling methodologies; I.2.8 [**Problem Solving, Control Methods, and Search**]: Heuristic methods

General Terms

Experimentation

Keywords

Evolutionary Strategies, Neural Network, Time Series, Forecasting, GRASP method

1. INTRODUCTION

In this paper a systematic procedure based on a hybrid intelligent system is proposed, Greedy Randomized Adaptive Search Procedure with Evolutionary Strategies (GRASPES), where this define automatically the weights, architecture and inputs (relevante lags) of the Artificial Neural Network (ANN) applied to time series forecasting problem [5, 6]. The GRASPES combine a MultiLayer Perceptron (MLP) network, Greedy Randomized Adaptive Search Procedures (GRASP) [3] and Evolutionary Strategies (ES) [2], which searches and defines the best evolved neural network structure in terms of the processing units number, network weights and the minimum number of (and the particular [7]) lags necessary to solve the forecast problem.

2. THE GRASPES METHOD

The main GRASPES characteristic is based on the GRASP idea, an multi-start procedure, where each iteration is made up of construction phase, where a randomized greedy solution is constructed, and a local search phase which starts at the constructed solution and applies iterative improvement until a locally optimal solution is found [3]. The optimal solution on all the GRASP iterations is kept as the final result and can be trivially implemented in parallel computers [3] (GRASPES inherit this behavior). The general expectative is that, close to sub-optimal solution, the method will find other sub-optimal (or optimal) solution with high probability. Consequently, the search will tend to look around of such solution.

Evolutionary Algorithms (EA) are based on the evolutionary ideas and they are composed by a set of trial solutions of the problem (population), with each solution (individual) is coded by a data structure. Then, the genetic operations (crossover and mutation) can be applied to this set of individuals in order to create the offspring (new generation). The Evolutionary Strategies (ES) are a particular class of EAs, where the population is just one individual and only mutation operation is applied [2]. Based on these affirmations, each individual I codifies an three layer ANN (MLP).

The individual will be evaluated by a defined fitness function, given by, $fitness = f(I)$, where the better individuals will return higher fitness values. The population P_n is a set of I, where n is the quantity stored by the proposed method of the best individuals generated by the parent in all iterations. After that, the parent's chromosomes I_p are cloned and they will go to undergo a mutation operation. This operation will change the genes of the chromosome (some features will be changed).

For tuning the ANN structure [6], two integer random numbers $l \in [1..10]$ are generated, one to define the lags (processing units in input layer i) and another to define the number of processing units in hidden layer (sigmoidal units j). For each weight of the best individual I, a gaussian random number (zero mean value and standard deviation σ, $N(0, \sigma)$) is generated, where mutation step size σ is governed by the law described by the $1/5$ success rule [2]. The mutation step size is co-evolving with the solutions, where the initial value is 1. The next step is add $N(0, \sigma)$ to the weight: $x'_i = x_i + N(0, \sigma);\ i = 1, 2, \ldots, p,$.

The new individual is evaluated and it will be included in the population $P_n = [I_1, I_2, ..., I_n]; n = pop\_size$, if and if only, its solution quality is better than the actual father. Then the process will continue with the procedure repetition: the parent's chromosomes is cloned and the operation of mutation is executed. This steps will be repeated until the mutated individuals number criterium or the size of the population n is reached, defining a Parent's Generation (PG). Consequently, when it is reached, the best individual of the current population P_n is selected, substituting the parent, as described by $I_{best} = max(f(P_n))$.

The stopping criterium is researched when occurs a defined iteration number of the PG without better individual generation, where a new individual is considered "bet-

GECCO'08, July 12–16, 2008, Atlanta, Georgia, USA.
ACM 978-1-60558-130-9/08/07.

ter" when your fitness is great than parent's fitness plus a constant u, i.e., $Better\ Individual : f(Individual) > f(Parente) + u$

The basic steps of the GRASPES are described in the Algorithm 1.

```
begin GRASPES
    initialize PG, u, n, σ, parent, γ ← 0;
    evaluate f(parent);   // f(·): fitness function
    while γ < PG do
        clone parent, τ ← 0;
        for i=1 to 30000 do
            define the input layer i and hidden layer j;
            perform the mutation operation on the sons I_τ;
            evaluate f(I_τ);
            if f(I_τ) > parent's fitness value then
                move I_τ for P_n;
                τ ← τ + 1;
                if τ > n then
                    exit for, the size of P_n was reached;
                end
            end
        end
        if P_n = ∅ then
            γ ← γ + 1;
        else
            choose the greatest fitness value by I_best = max(f(P_n));
            if (f(parent) - f(I_n)) > r ) then
                the individual I_best will be the new parent;
                γ ← 0;
            else
                γ ← γ + 1;
            end
        end
        adjust the σ value;
    end
end
```

Algorithm 1: The GRASPES procedure.

3. EXPERIMENTAL RESULTS

The series investigated were normalized to lie within the interval [0;1] and divided in three sets: training set (50% of data), validation set (25% of data) and test set (25% of data). The termination conditions are the constant $u = 10^{-3}$ and $PG = 150$. For each time series the maximum size of the set P is $n = 30000$ and the maximum value of σ is 10 to narrow the area of search.

In addition, experiments with the TAEF method (in-phase matching) found in [5, 4] and MMNN found in [1] are used for comparison with the proposed method.

3.1 Dow Jones Industrial Average (DJIA) Index Series

The Dow Jones Industrial Average (DJIA) Index series corresponds to daily records from January 1st 1998 to August 26th 2003 (1420 points).

For the prediction of the DJIA Index series (with 1 step ahead of prediction horizon), the GRASPES chose the 2-9-1 architecture as the its best individual. Table 1 shows the results for all performance measures and Figure 1 shows the comparison among the target values and the predicted values with the last 100 points of the test set.

Table 1: Experimental Results.

	TAEF		MMNN		GRASPES	
	S&P500	**DJIA**	**S&P500**	**DJIA**	**S&P500**	**DJIA**
ARV	0.0100	0.0346	7.4749e-3	3.4423e-2	7.3171e-3	3.3300e-02
MAPE(%)	1.0431	10.1529	0.92	9.67	0.8658	9.6800
MSE	7.4290e-4	8.4183e-4	9.7451e-5	8.3236e-4	9.6320e-5	8.2345e-04
Theil	7.24123	1.0006	0.9498	0.9945	0.9477	0.9871
POCID	50.54	47.57	81.31	50.85	84.78	54.0845

3.2 S&P500 Stock Index Series

The S&P500 series corresponds to the monthly records from January 1970 to August 2003 (369 points). In order to reduce exponential trend, the natural logarithm was applied to the original values of the series.

For the prediction of the S&P500 Stock Index series, the best individual has the architecture 3-5-1. Table 1 shows the results for all performance measures and Figure 1 shows the comparison among the target values and the predicted values.

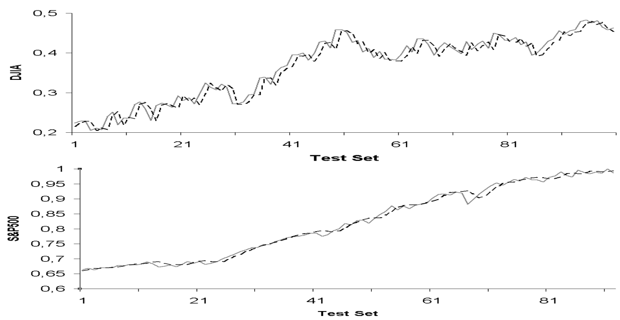

Figure 1: Prediction results - actual values (solid lines) and predicted values (dashed lines)

4. CONCLUSIONS

This paper has presented a intelligent hybrid system for time series forecasting problem, which combine ANN, ES and GRASP method, called GRASPES.

The experimental results using five different metrics (MSE, MAPE, U of Theil Statistics, POCID, ARV) suggest that GRASPES can be successfully used for financial time series forecasting and showed, according to Table **??**, that the proposed prediction model can obtain a superior performance than the other models displayed (TAEF [5, 4] and MMNN [1]). The experimental method validation was carried out with two complex and relevant financial time series.

Future works will consider: the phase prediction adjust procedure (proposed by Ferreira et al [5, 4]); the use of MLPs convectional training algorithms; and, other distinct forms of modeling the ANN. Other time series are being reaped for the efficiency confirmation of the proposed method and a study to determine the possible limitations of the method when dealing with other types of components such as trends, seasonality, impulses, steps, model exchange and other nonlinearities.

5. REFERENCES

[1] R. Araujo, F. Madeiro, R. de Sousa, L. Pessoa, and T. Ferreira. An evolutionary morphological approach for financial time series forecasting. *Evolutionary Computation, 2006. CEC 2006. IEEE Congress on*, pages 2467–2474, 16-21 July 2006.

[2] A. Eiben and J. Smith. *Introduction to Evoluionary Computing*. Springer, 2003.

[3] T. Feo and M. Resende. Greedy randomized adaptive search procedures. *Journal of Global Optimization*, 6(2):109–133, March 1995.

[4] T. Ferreira, G. Vasconcelos, and P. Adeodato. A new evolutionary method for time series forecasting. In *GECCO '05: Proceedings of the 2005 conference on Genetic and evolutionary computation*, pages 2221–2222, New York, NY, USA, 2005. ACM.

[5] T. A. E. Ferreira. *A new Intelligent Hybrid Methodology for the Time Serie Forecasting*. PhD thesis, Federal University of Pernambuco, 50732-970, Recife-PE, Brazil, Febuary 2006.

[6] F. Leung, H. Lam, S. Ling, and P. Tam. Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *Neural Networks, IEEE Transactions on*, 12(1):79–88, Jan 2003.

[7] F. Takens. Detecting strange attractor in turbulence. In A. Dold and B. Eckmann, editors, *Dynamical Systems and Turbulence*, volume 898 of *Lecture Notes in Mathematics*, pages 366–381, New York, 1980. Springer-Verlag.