Prediction of Stock Market Indices using Hybrid Genetic Algorithm/ Particle Swarm Optimization with Perturbation Term

Tarek Aboueldahab* and Mahumod Fakhreldin**

Cairo Metro Company, Ministry of transport, Cairo, Egypt,*

*Computers and Systems Department, Electronic Research Institute, Cairo, Egypt***

heshoaboda@ hotmail. com, mafakhr@ mcit. gov. eg

Abstract

Stock market indices prediction is one of the most important issues in the financial field. Although many prediction models have been developed during the last decade, they suffer a poor performance because indices movement is highly non stationary and volatile dynamic process. As improving the prediction accuracy becomes an important issue, we propose a new hybrid genetic Algorithm / Particle Swarm Optimization (GA/ PSO) model with perturbation term inspired by the passive congregation biological mechanism to overcome the problem of local search restriction in standard hybrid (GA/ PSO) models. This perturbation term is based on the cooperation between different particles in determining new positions rather than depending on the particles selfish thinking which enables all particles to perform the global search in the whole search space to find new regions with better performance. Experiment study carried out on the most famous stock market indices in both long term and short term prediction shows significantly the influence of the perturbation term in improving the performance accuracy compared to standard hybrid (GA/ PSO) models.

Key words

Stock markets, Particle Swarm Optimization, Genetic Algorithm, Perturbation term, Global search.

1 Introduction

In the recent years, a lot of attention has been devoted to the analysis and predication of future values and trends of stock market indices which are affected by many complex factors such as political events, monetary policies, and traders' expectation causing volatility and non-stationary in stock market data. Also, stock market prediction is more difficult than other time series prediction problems because it is regarded as a highly dynamical nonlinear task rather than a random one [6, 9].

Particle Swarm Optimization (PSO) algorithm was used to train the weights of different neural networks architectures to predict famous stock market indices [16-17] but the problem of poor prediction accuracy remains unsolved due to the PSO major disadvantage of premature convergence to local minimum [8,12]. To overcome this problem, many researches introduced Genetic Algorithms (GA) operators such as selection, crossover, and mutation to the PSO algorithm to form hybrid (GA/PSO) model to increase the diversity and to give particles the ability to fly to new regions in the search space. Selection is the process of selecting highly performance particles as parents in the creation of a new generation, the crossover operator produces two new children particles by recombining the information of the two selected parents, and the mutation operator is used to mutate the new generated particles [1, 3-4, 10-11]. For example, in Hybrid Genetic Algorithm Particle Swarm Optimization (HGAPSO), the upper-half of the best-performing particles are updated using PSO while the other particles are updated using GA [3]. Genetical Swarm Optimization (GSO) model integrates the two algorithms and divides the population randomly into two parts, each part run one algorithm and combines them again to produce a new generation [4]. Another hybrid model integrated the GA mutation operator into PSO algorithm by randomly selecting a particle at the end of each generation to be mutated producing a new position in the search space to be used in the new generation [1]. The Breeding Swarm hybrid model introduced the Velocity Propelled Averaged Crossover (VPAC) in which the GA crossover and mutation operators were used in the worst performance particles according to predefined breeding ratio [10-11].

However, in all these models, each particle in PSO algorithm is restricted in doing the local search around its best position and the global best position without any cooperation with other particles in the swarm while the GA operators are responsible in doing the global search by making particles fly to new regions in the search space [11, 15] leading to poor model performance in terms of accuracy, convergence speed and robustness [8, 12]. To overcome this problem and improve the model performance, each particle in the swarm should be capable of performing the global search and should not be only restricted in local search [8, 15]. Thus, we introduce our new proposed hybrid (GA/PSO) model with perturbation inspired by the passive congregation biological mechanisms of school of fishes and flock of birds by using a perturbation term based on inserting the position difference between other two particles into the particle position update equation. This term enables each particle to compare between its selfish thinking and cooperation with other particles to adjust its position in better regions the search space. We analyze three of the most famous stock market indices which are DJIA, NASDAQ100, and S&P500 and compare the performance between our proposed hybrid

(GA/PSO) model with perturbation term and the standard one. Simulation results show significantly that the perturbation term positively enhance the prediction accuracy and keep the fast convergence speed and robustness. This paper is organized as follows: section 2 reviews main concepts of the hybrid (GA/PSO) model and section 3 introduces our proposed hybrid GA/PSO model with perturbation. Experimental study is presented in section 4 and finally the conclusion and future work are presented in section 5.

2 Hybrid GA/PSO

Particle Swarm Optimization (PSO) is a population based stochastic optimization technique developed by Kennedy and Eberhart and each particle in the swarm represents a potential solution of the optimization problem in *D*-dimensional space [5]. At the k^{th} generation the i^{th} particle in the swarm of size *S* particles is represented by its current position $X_i(k) = (x_{i1}(k), x_{i2}(k), .., x_{iD}(k))$, its current velocity $V_i(k) = (v_{i1}(k), v_{i2}(k), .., v_{iD}(k))$, and its current fitness function $F_i(k)$, the particle position and velocity are updated in the next generation by the following equations

$$V_i(k+1) = w * V_i(k) + C_1 * \varphi_1 * \left(X_g - X_i(k)\right) + C_2 * \varphi_2 * \left(X_{ibest} - X_i(k)\right) \quad (1)$$

$$X_i(k+1) = X_i(k) + V_i(k+1) \quad (2)$$

Also the i^{th} particle autographical memory remembering its best previous position $X_{ibest}(k) = (x_{ibest1}(k), x_{ibest2}(k), .., x_{ibestD}(k))$ associated with its current best fitness function F_{best} and X_g is the published knowledge of the best current position found by all particles corresponding to the global best fitness function F_g. From equation (1), the first term represents the inertia of the particle pervious velocity and w is inertia weight; the second term is the social term representing the cooperation among all particles where C_1 is the social constants and the third term is the cognition term which represents the private thinking and the selfish behavior of the particle itself where C_2 is cognitive constants and both φ_1 and φ_1 are random variables in the range [0, 1]

The GA operators are used after the fitness evaluation of all the particles in each generation, the worst (S*Ψ) particles performance are discarded and removed from the population where Ψ is the breeding ratio determining the discarded proportion of the swarm and its value is arbitrary selected in the range [0.0, 1.0] to provide better results [10-11]. From the remaining (S*(1-Ψ)) particles, parents particles are selected randomly to undergo crossover operator producing new child particles using the Velocity Propelled Averaged Crossover (VPAC) [10-11] to accelerate the new particles away from their parents' direction according to the following equation:-

$$X_1^c(k+1) = \frac{X_1(k) + X_2(k)}{2} - \beta_1 * V_1(k) \quad (3-a)$$

$$X_2^c(k+1) = \frac{X_1(k) + X_2(k)}{2} - \beta_1 * V_2(k) \quad (3-b)$$

Where $X_1^C(k+1)$ and $X_2^C(k+1)$ are the positions of child 1 and 2 respectively, $X_1(k)$ and $X_2(k)$ are the positions of the parents particles and $V_1(k)$ and $V_2(k)$ are their associated velocities and β_1 and β_2 are two uniformly distributed random variables in the range [0, 1]. For the mutation process, all variables in the new created child particles are equally probable to be mutated [1].

3 Proposed hybrid (GA/PSO) with perturbation

Biologists found out that in a spatially well-defined group such as school of fishes or flock of birds, each individual can monitor both the environment and its immediate surrounding such as the position and the speed of neighbors and proposed two types of biological mechanisms to realize these requirements which are active aggregation and passive congregation [7, 12].

Active aggregation is the transfer of necessary and active information among different individuals in the entire group such as the place with the most food and displays their social behaviors, this mechanism can be represented by the social term in equation (1) because it transfers the best position found so far to all particles in the swarm. Passive congregation is an attraction from one individual to others but does not display social behaviors. Individuals may have law fidelity to other group members if they have little or no genetic relation to each others, also each individual has multitude of potential information from other group members helping in reducing the possibility of miss detection and incorrect interpretations. Thus, the passive congregation helps individuals to make global search in the whole search space and find new regions with better solutions [7, 12].

The cognitive term in equation (1) is not enough to represent the passive congregation mechanism because it displays only its best previous position without thinking of other particles positions [12]. To fully represent this mechanism, a perturbation term is included and described as the position difference between two other randomly selected particles in the swarm. Thus in each generation, for the i^{th} particle in the swarm, the position difference $\delta_i(k+1)$ can be written as follows:-

$$\delta_i(k+1) = X_l(k) - X_m(k) \tag{4}$$

Where $X_l(k)$ and $X_m(k)$ are the positions of particles l and m respectively and $i \neq l \neq m$. Using this position difference, the i^{th} particle perturbation term $P_i(k+1)$ is written as follows:-

$$P_i(k+1) = C_3 * \varphi_3 * \delta_i(k+1) \tag{5}$$

Where C_3 is the perturbation constant and φ_3 is random variable in the range [0, 1] and the new velocity with perturbation $V_i^P(k+1)$ is constructed by inserting this perturbation term instead of the cognitive term and is written as follows:-

$$V_i^P(k+1) = w * V_i(k) + C_1 * \varphi_1 * \left(X_g - X_i(k)\right) + P_i(k+1) \tag{6}$$

The above velocity with perturbation term is based on other cooperative thinking between different particles while the standard velocity in equation (1) is based on the selfish thinking of the particle without any cooperation with other particles. Consequently, the position with perturbation $X_i^P(k+1)$ associated with this new velocity is calculated as follows:-

$$X_i^P(k+1) = X_i(k) + V_i^P(k+1) \tag{7}$$

The fitness function $F_i^P(k+1)$ associated with this perturbation term is compared with the standard one $F_i(k+1)$ associated with the cognitive term to determine the new particle position and velocity and adjust accordingly the best particle position and the global best position. Thus all particles in the swarm can perform the global search by comparing the model performance resulting from both the original cognitive term and the new added perturbation term.

If the perturbation term doesn't improve the model performance, so there is no genetic or little genetic relation between this particle and the other randomly selected particles. On the other hand if it yields better performance, so these particles help in making correct interpretation through the global search in the whole search space to find new position with better performance. The proposed algorithm can be summarized in the following steps

Step1 (Initialization): randomly create an initial swarm of particles and setup the required inertia, social, cognitive, and perturbation constants. At each generation, for each particle in the swarm do the following

Step2 (Cognitive term utilization): update the velocity $V_i(k+1)$ and position $X_i(k+1)$using the cognitive term according to equations (1) and (2) respectively and compute its associated fitness function $F_i(k+1)$.

Step3 (Perturbation term construction): randomly select two other particles to construct the position difference between them $\delta_i(k+1)$ according to equation (4) and calculate the perturbation term $P_i(k+1)$ using equation (5).

Step4 (Perturbation term utilization): compute the velocity $V_i^P(k+1)$ and position $X_i^P(k+1)$ using the perturbation term according to equations (6) and (7) respectively and compute its associated fitness function $F_i^P(k+1)$

Step5 (Fitness Check): Select the best one of these two fitness functions to be assigned to the particle fitness and modify accordingly both the local best and the global best

Step6 (selection): after modifying all particles positions, velocity and fitness in the current generation, select the particles with the worst performance part according to the breeding ratio Ψ to be replaced with other particles selected randomly as parents particles from the remaining part of the swarm

Step7 (crossover): construct the new child particles $X^C(k)$ to replace the worst particles according to equations (3-a), and (3-b).

Step8 (mutation): select randomly with equal probability one variable in the search space from the new children particles to be mutated. Return to step 2 to start a new generation.

4 Experimental Study

Our experiments will be carried out in both long term and short term for NASDAQ100, DJIA and S&P500 stock market indices according to preliminary researches suggested that input data comprise daily open, maximum, and closing values while the output data is the next day closing value [6, 16-17].

The neural network architecture used as the predictor to the stock market indices is the Sigmoid Diagonal Recurrent Neural Network (SDRNN) because it was proved in our previous work that this architecture is better than other different architectures for reducing the error and increasing the accuracy in many applications [13-14]. Both the standard hybrid (GA/PSO) model and the hybrid (GA/PSO) with perturbation model are used to train the SDRNN predictor weights set and the accuracy of these two models are compared to show the influence of our proposed model in both long term and short term data set. In our experiment, we choose a swarm size of 40 particles, breeding ratio 0.1 and all constants and inertia are set to one

The assessment of the prediction performance and accuracy of reducing this fitness function as much as possible is done in terms of the maximum absolute difference (MAX), the Mean Absolute Percentage Error (MAPE) and the Root Mean Square Error (RMSE) [2, 16-17] and they are defined as follows:

$$MAX = \max_n\left(\left|Y_{actual}(n) - Y_{predict}(n)\right|\right) \tag{8}$$

$$MAPE = \frac{1}{\mathrm{N}}\sum_{n=1}^{N}\left(\frac{\left|Y_{actual}(n) - Y_{predict}(n)\right|}{Y_{acutal}(n)} * 100\right) \tag{9}$$

$$.RMSE = \sqrt{\frac{\sum_{n=1}^{N}\left(Y_{actual}(n) - Y_{predict}(n)\right)^2}{N}} \tag{10}$$

Where N is the size of the training data set, and $n = 1,2, \dots, N$ and $Y_{actual}(n)$ is the actual stock market closing price and $Y_{predict}(n)$ is the predictor output at the n^{th} sample

In the long term stock market indices prediction, the training data were chosen starting form 3/1/2000 to 1/1/2006 while the testing data was chosen from3/1/2006 to 14/2/2011 and the best fitness obtained through all generations using both standard hybrid (GA/PSO) and our proposed hybrid (GA/PSO) with perturbation for NASDAQ100, DJIA and S&P500 are shown in Figure 1, 2, and 3 respectively.

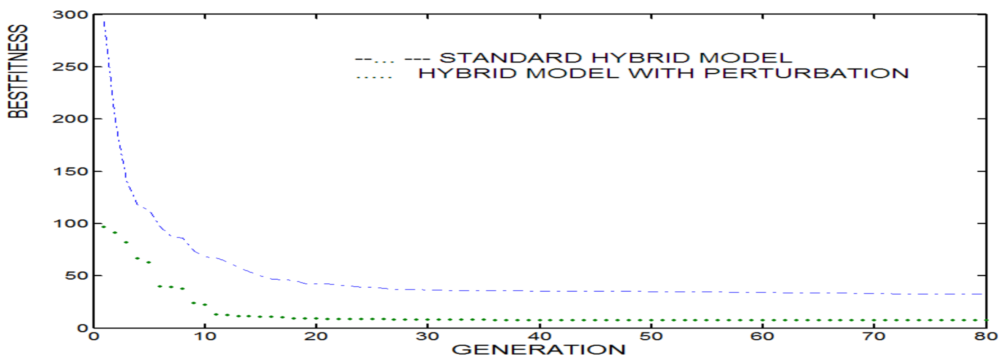

FIG1: DJIA best fitness

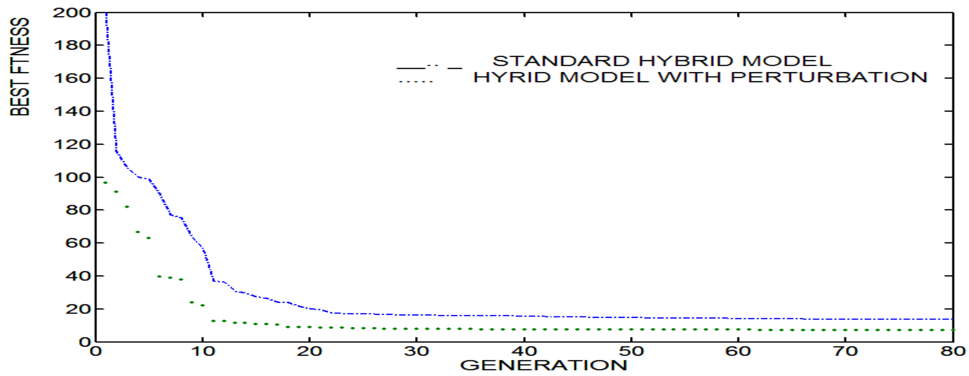

FIG2: NASDAQ100 best fitness

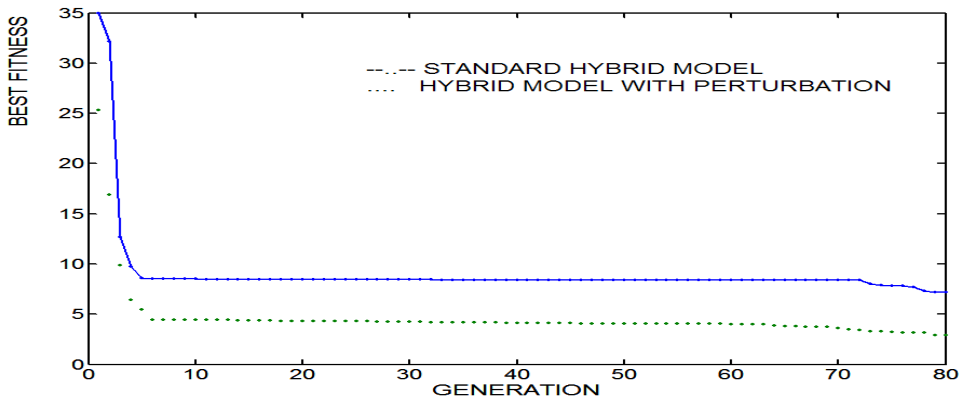

FIG3: S&P500 best fitness

The training and testing performance measurements using both standard hybrid (GA/PSO) and our proposed hybrid (GA/PSO) with perturbation for all these stock market indices are shown in Tables 1and 2 respectively.

	DJIA		Nasdaq100		S&P500	
	Standard Hybrid	Hybrid with Perturbation	Standard Hybrid	Hybrid with Perturbation	Standard Hybrid	Hybrid with Perturbation
MAX	893	١60	430	69	84	66
MAPE	0.30%	0.10%	0.56%	0.097%	0.44%	0.073%
RMSE	٦٨	15.2	١٠٧	١9.3	11.1	2.08

Table 1: long term training performance measurement using standard hybrid model and proposed hybrid model with perturbation

	DJIA		Nasdaq100		S&P500	
	Standard Hybrid	Hybrid with Perturbation	Standard Hybrid	Hybrid with Perturbation	Standard Hybrid	Hybrid with Perturbation
MAX	333	85	496	143	28	12
MAPE	0.٣١%	0.10%	0.52%	0.11%	0.48%	0.077%
RMSE	47	11.7	103	17.7	9.3	1.13

Table 2: long term testing performance measurements using standard hybrid model and proposed hybrid model with perturbation

In the short term indices prediction, the period is selected from 1/1/2008 till 14/2/2011 because at the start of this period, all indices began a major downtrend and reached their lowest values in 23/4/2009 before they start again the uptrend correction. The DJIA index dropped from 13043 point to 7957, the NASDAQ100 index dropped from 2609 point to 1652, and the S&P500 dropped from 1447 point to 851point. The performance measurements of both standard hybrid (GA/PSO) and our proposed hybrid (GA/PSO) with perturbation and the results of performance measurements are shown in Table 3.

	DJIA		Nasdaq100		S&P500	
	Standard Hybrid	Hybrid with Perturbation	Standard Hybrid	Hybrid with Perturbation	Standard Hybrid	Hybrid with Perturbation
MAX	333	82	481	135	26	8
MAPE	0.32%	0.11%	0.52%	0.11%	0.49%	0.079%
RMSE	58	14	109	18.6	11.5	1.7

Table 3: short term performance measurements using both standard hybrid model and proposed hybrid model with perturbation

The plots in Figs. 1-3 show that our proposed model reduces successfully the required fitness function within a small number of generations while the standard one failed throughout the whole generation, also the performance measurements shown in Tables 1-3 indicate obviously that our proposed model is superior to the standard one and yields better perdition accuracy in all the three different stock market indices for both the long term and short term prediction. This improvement is due to the existence of the perturbation term which enables all particles to perform the global search in the whole search space and to discover new position with better fitness function while the poor performance of standard model is due to the particles positions are dependent on their own thinking without knowing any information about other particles positions which could be useful for avoiding miss detection and increasing the correct interpretation.

5. Conclusion and Future work

A new hybrid (GA/PSO) model with perturbation is presented to increase the prediction accuracy for both short term and long term stock market indices prediction. This new perturbation term adequately represents the passive congregation biological mechanism which enables all particles in the swarm to perform the global search in the whole search space. The performance measurements of our proposed model is obviously better than the standard hybrid GA/PSO model in which only particles subjected to GA operators are capable in performing the global search.

For the future work, our proposed hybrid model can be used in other practical applications such as pattern classification and recognition and in other practical optimization problems.

References

[1] AHMED A.A.E., GERMINO L.T., and ANTONIO C.Z., A *Hybrid Particle Swarm Optimization Applied to Loss Power Minimization*. IEEE Transactions on Power Systems, Vol. 20, No. 2, pp. 859-866, May 2005.

[2] CHAOUACHI A., KAMEL R.M., ICHIKAWA R., HAYASHI H., and NAGASAKA K. Neural Network Ensemble-based Solar Power Generation Short-Term Forecasting, World Academy of Science, Engineering and Technology 54, 2009

[3] GRIMALDI E. A., GRIMACCIA F., MUSSETTA M., PIRINOLI P., and ZICH R.E., *A New Hybrid Evolutionary Algorithm for High Dimension Electromagnetic Problems*, IEEE Antennas and Propagation Society International Symposium Digest, Vol.2A, pp.61-64, 2005.

[4] JUANG C.F., *A Hybrid of Genetic Algorithm and Particle Swarm Optimization for Recurrent Network Design*. *IEEE* Transactions on systems, Man, Cybernetics- Part B: Vol. 34, No. 2, 2004.

[5] KENNEDY J. and EBERHART R.C. *Particle swarm optimization*, Proceedings of IEEE International Conference on Neural Networks, Piscataway, NJ, pp.1942-1948, 1995.

[6] MAJHI RITANJALI., PANDA G., SAHOO G., and PANDA ABHISHEK., *Prediction of S&P 500 and DJIA stock indices using particle swarm optimization technique. IEEE* Congress on Evolutionary (CEC), pp.1276-1282, 2008.

[7] PARRISH J. K., and HAMMER W. M., *Animal Groups in Three Dimensions*, Cambridge: Cambridge University Press, 1997.

[8] PREMALATHA K., and NATARAJAN A.M., *Hybrid PSO and GA for Global Maximization*, International Journal Open Problems Compt. Math., Vol. 2, No. 4,pp. 597 – 608, 2009.

[9] ROHIT C., and KUMKUM G., A Hybrid Machine Learning System for Stock Market Forecasting, World Academy of Science, Engineering and Technology, Vol. 39, pp 315-319, 2008.

[10] SETTLES M., and SOULE T., *Breeding Swarm: A ga/pso hybrid,* In Genetic and Evolutionary Computation Conference (GECCO). pp. 161–168. Washington, USA 2005.

[11] SETTLES M., NATHAN P., and SOULE T., *Breeding Swarms: A New Approach to Recurrent Neural Network Training.* In Proceeding of the Genetic and Evolutionary Computation Conference (GECCO-2005), pp. 185- 192, 2005.

[12] TANG W.H., HE S., WU Q.H., and RICHARDSON Z.J., *Winding Deformation Identification Using A Particle Swarm Optimizer With Passive Congregation for Power Transformer*. International Journal of Innovations in Energy systems and Power, Vol. 1, No.1, pp46-52, 2006.

[13] TAREK A. *Improved Design of Nonlinear Controllers Using Recurrent Neural Networks*. Master Dissertation, Cairo University, December 1997.

[14] TAREK A., and Mouhamed F., *Adaptive Control of Dynamic Nonlinear Systems Using Sigmoid Diagonal Recurrent Neural Network. IEEE* International Conference on Systems, Man, and Cybernetics (SMC), pp. 4341-4345, Istanbul. 2010.

[15] VRUGT J.A., ROBINSO B.A., and HYMAN J.M., *Self-Adaptive Multimethod Search for Global Optimization in Real-Parameter Spaces*, IEEE Transactions on Evolutionary Computation, Vol. 13, No. 2, pp. 243-259,APRIL 2009

[16] XUEGIN Z., YUEHUI C., and JACK Y.Y., *Stock Index Forecasting Using PSO Bases Selective Neural Network Ensemble*. International Conference on Artificial Intelligence (ICAI07), pp. 260-264. 2007.

[17] YUEHUI C., BO Y., and AJITH A., *Flexible Neural Trees Ensemble for Stock Index Modeling.* Neurocomputing, Vol.70, pp. 697–703. 2007.