HYBRID MODULAR MODEL FOR TIME SERIES FORECASTING BASED ON NEURO-FUZZY NETWORK AND FUZZY COGNITIVE MAPS

Sergey Yarushev, Alexey Averkin

***Abstract**: In this paper, we consider a hybrid approach to forecasting time series using neuron-fuzzy prediction models and Fuzzy Cognitive Maps. Main idea of proposed approach is hybridization two different ways for time series forecasting. We can make quantitative and qualitative forecast. In addition, we describe the different approaches to learning and optimization of the network, such as the methods of particle swarm, evolutionary methods, as well as variants of the hybridization of these methods. Also, in a comparison of the results of the prediction for example, one of the indicators of the State Program of Development of Science and Technology is forecast schedule.*

***Keywords**: forecasting, fuzzy cognitive maps, time series, hybrid-forecasting models.*

***ITHEA Keywords**: I. Computing Methodologies: I.2 Artificial Intelligence: I.2.1 Applications and Expert Systems*

Introduction

Nowadays modelling and forecasting time series are among the most active areas of research. For example, depending on the historical data, situation on sales market, changes in prices for shares of population growth and banks deposits are forecast. Forecasting time series affects the lives of people around the world, so it has great practical value and perspectives of research in all areas of the modern society, which is also an important area in the field of computer application.

The solution of problems of identification of dynamic objects should be used in a variety of fields: it can simplify temperature controllers, or complex management and forecasting. It can also solve the forecasting problem, along with a number of different methods, for example, statistical analysis, neural networks [Haykin, 1994]. Identification of the object may be difficult if the exact structure of the model of the object is unknown, some of the parameters of the object change due to obscure principles, or the exact number of parameters of the object is unknown. In such cases, the hybrid neural network can be used for identification of dynamic objects. There are many types of neural networks that are used for identification of dynamic objects. Despite the large number of neural network methods for identification of dynamic objects, most of these algorithms have some limits, or do not provide the required accuracy.

Among all kinds of neural networks, architectures that can be used for identification of dynamic objects allocated a class of neural networks based on self-organizing maps of Kohonen with hybrid architecture. Hybrid neural networks of this type will get special attention in this article because they are becoming more widespread and successful applications for solving various problems of recognition [Efremova, 2012], identification [Trofimov, 2010], and forecasting. We will also consider a number of biomorphic neural networks applicable for solving identification problems and management.

In the development and future changes in time series, there is reflexivity between events, their participants and the actual prognosized process (time series), between the researcher and the process being studied [Lefevr, 1965]. The theory of reflexivity in the economic world suggests that the situation that has arisen affects the behavior of the participants in the process, and their thinking and behavior act on the development of the situation to which they are participants [Soros, 2003]. It is clear that using only a tool for forecasting time series, it would not be as powerful as any, it is impossible to reflect and take into account the situation and events affecting the process under study, since a neural network is allowed to work with historical data. The practical way out of the situation is to develop such methods that could operate both with a cause-effect relationship between events and the projected process, as well as with the numerical values of the time series, its historical data. Therefore, it is expedient to develop a hybrid forecasting system capable of operating both qualitative data and quantitative ones.

In this paper, we propose a new hybrid time-series forecasting model based on fuzzy relational cognitive maps and a hybrid neural-fuzzy network with regression analysis.

Modular Neural Networks

The core of the modular neural networks is based on the principle of decomposition of com-plex tasks into simpler ones. Separate modules make simple tasks. More simple subtasks are then carried through a series of special models. Each local model performs its own version of the prob-lem according to its characteristics. The decision of the integrated object is achieved by combin-ing the individual results of specialized local computer systems in a dependent task. The expan-sion of the overall problem into simpler subtasks can be either soft or hard-unit subdivision. In the first case, two or more subtasks of local computer systems can simultaneously assigned while in the latter case, only one local computing model is responsible for each of the tasks crushed.

Each modular system has a number of special modules that are working in small main tasks. Each module has the following characteristics:

— The domain modules are specific and have specialized computational architectures to recognize and respond to certain subsets of the overall task;

— Each module is typically independent of other modules in its functioning and does not influence or become influenced by other modules;

— The modules generally have a simpler architecture as compared to the system as a whole. Thus, a module can respond to given input faster than a complex monolithic system;

— The responses of the individual modules are simple and have to combine by some integrating mechanism in order to generate the complex overall system response.

The best example of modular system is human visual system. In this system, different modules are responsible for special tasks, like a motion detection, color recognition and shape. The central nervous system, upon receiving responses of the individual modules, develops a complete realization of the object which was processed by the visual system.

Review of Hybrid ANFIS Models for Time Series Forecasting

At the moment, there are many different learning algorithms ANFIS networks, each of them has its own advantages. Consider some of the studies that examined various hybrid-learning methods.

Chinese scientists, [Wang, 2015] presented its own model of forecasting of financial flows in the banking sector using a modified algorithm swarm optimization, which is called APAPSO (Adaptive Population Activity PSO). In view of the fact that the use of traditional methods do not give stable prediction results, researchers have proposed hybrid learning algorithm based APAPSO algorithm in combination with the method of least squares. In comparative experiments, the algorithm developed compared with the standard backpropagation technique in combination with the method of least squares (LMS), and with the traditional method swarm optimization-LMS. Results showed an increase in speed optimization, in comparison with conventional algorithms, as well as increase the accuracy of the prediction.

Indian scientists [Gunasekaran, 2011] offered a hybrid-forecasting model based on the integration of ANFIS and immune algorithm for the prediction of the Indian stock market. To create an effective model of prediction, the researchers decided to use an artificial immune algorithm to adjust the parameters of fuzzy system output functions. The data daily close of trading on the National Stock Exchange of India (NSE) were used as input data for system testing, as well as well-known technical indicators. The output is a forecast of the future value of NSE index. The experimental results were compared with other models on the basis of soft computing and actual data from the auction. As a result, the experimental results have shown that the prediction model proposed gave much more accurate prediction results, compared to conventional models.

Artificial neural network (ANN) is a very good approximation method, which has the characteristics of adaptability and self-study [Dong, 2006]. However, using ANNs easily fall into a local minimum. By combining with fuzzy inference system has been proposed a new kind of non-linear prediction method, namely, adaptive neural fuzzy inference system (ANFIS) [Catalo, 2011]. This method can be used as the fuzzy rules and a neural network structure for implementing adaptive learning, so the prediction accuracy is higher than the one of the artificial neural network. In order to further improve the accuracy of predicting adaptive ANFIS system, you can use a variety of teaching methods, for example, the PSO algorithm (particle swarm optimization) or the method of particle swarm is used to optimize the structure of the network parameters. For example, a new hybrid approach combining particle swarm and ANFIS network is used for short-term forecasting of wind power in Portugal, as a result, it is possible to achieve the required accuracy of prediction using the proposed approach [Pousinho, 2011]. The radial basis function neural network (RBFNN) with non-linear evolutionary method swarm cha-particles, (NTVE-PSO) is proposed, and the simulation results show that the proposed NTVE-PSORBFNN has higher prediction accuracy and computational efficiency for predicting electricity consumption [Meng, 2012]. Improved PSO based on artificial neural network (ANN) was proposed by the researchers, results show that the proposed SAPSO based on ANN has a better ability to escape from local optimum and is more efficient than conventional PSO based on ANN [Cai, 2007]. Algorithm training based on a hybrid method of optimization of particle swarm (PSO) and the evolutionary algorithm (EA) for the prediction of 100 missing values of VRE-alternating series of 5000 data points that show the experimental D results that PSO-EA algorithm also proven effective in study [Wang, 2012].

Modular Hybrid System for Time Series Forecasting with ANFIS and Fuzzy Cognitive Maps

The developed forecasting system is based on a modular architecture that betrays the system additional stability when even if one of the modules crashed the remaining modules continue to perform their work.

The system itself has three main modules responsible for the prediction task. A hybrid neural-fuzzy network performs the forecast of a time series based on numerical indicators and gives us a so-called quantitative prediction, the results of which pass through a verification system (estimates of the adequacy of the forecast), if the prognosis corresponds to the required accuracy, then it is passed on to the next module. In parallel with the neural-fuzzy network, a module with a fuzzy cognitive map operates, which receives data on the event-related effects on the time series, and constructs a cognitive map that takes into account all factors of influence on a specific predicted indicator. At the output, the cognitive map gives us a forecast with the probability of its implementation, that is, with the consonance of the factor that tells us whether the forecast will be fulfilled or not. Further, all data received from these modules is fed to the third module, which operates on the basis of the neural network, which aggregates the information obtained from the previous modules and outputs the final prognosis. In Figure 1 is a diagram of the forecasting system.

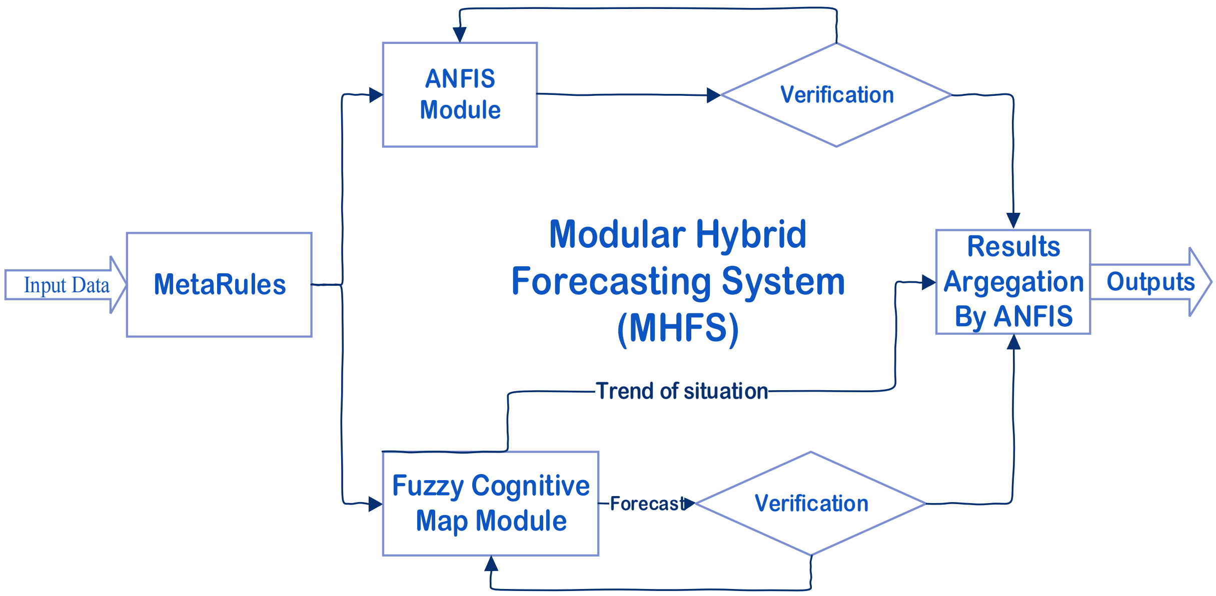


Figure 1. Modular Hybris Forecasting System

Further, we will dwell in more detail on forecasting based on fuzzy cognitive maps and their training, since a rather large number of studies are devoted to neural-fuzzy networks [Yarushev, 2016], with a slightly different situation with cognitive maps.

Fuzzy Cognitive Maps in Time Series Forecasting Area

The time series is governed by two main forces - time and events that affect the change over time of the values of the time series. Most of these events are characterized by some uncertainty. Each value of the time series can be associated with a fuzzy variable with some membership function. In this connection, the most interesting for our research are methods based on the theory of fuzzy sets. Lotfi Zadeh in 1965 introduced the concept of fuzzy set, due to which it is possible to describe qualitative, fuzzy concepts and knowledge about the surrounding world, and then to operate them to obtain new information [Zadeh, 1976]. The application of this concept allows us to formalize linguistic information for constructing mathematical models [Rotshtein, 1997]. The notion of a fuzzy set is based on the proposition that the elements making up a given fuzzy set, as well as possessing common properties, can possess it and, consequently, belong to a given set in varying degrees. In this case, statements like "such and such an element belongs to a given set" lose their meaning, since it is still necessary to indicate the degree of belonging to a given set and its properties [Averkin, 1986].

To be able to operate with events that affect the time series, and events can be quite a lot and everyone can be related to each other, it makes sense to use fuzzy cognitive maps. They allow you to build a causal relationship between events and build a qualitative forecast of the development of the event, based on the strength of the influence of one event on another.

The cognitive map itself is an oriented graph, in which the vertices are the factors of the situation, and the weighted arcs are cause-effect relations, the weight of which reflects the force of the influence of the factors of the situation. Directional arcs of the graph are assigned the sign "+" or "-", i.e. they can be positive or negative. A positive relationship means that an increase in the value of the factor-cause leads to an increase in the value of the factor-effect, and a negative arc means that an increase in the value of the factor-cause leads to a decrease in the value of the factor-effect.

The tasks solved with the help of cognitive maps are to find and evaluate the influence of the factors of the situation, and to obtain, on the basis of the calculated influences, the forecasts of the development of the situation.

At present, for computation of the influences and forecasts of the development of the situation, fuzzy cognitive maps proposed by [Kosco, 1986] are widely used. In fuzzy cognitive maps, the force of influence between factors is given by means of linguistic meanings chosen from an ordered set of possible influence forces, and the values of the factors, their increments are also given in a linguistic form, and are chosen from the ordered sets of possible Values of the factor and its possible increments - scales of factors and incremental scales.

To construct a cognitive map that reflects the dynamic properties of the observed situation, it is necessary to determine the scales of the values of the factors and their increments.

To construct the scale of the factor, a lot of linguistic values of the factor are determined and structured. In determining the linguistic values, the absolute values of the factor are used, and not its evaluation of the type "large", "medium", "small". For example, the linguistic meaning of temperature can be the following: "so hot that you can barely put a hand on it" or the meaning "so cold that the hand immediately freezes," and not just "Hot" or "Cold." With this definition of the linguistic values of the situation factors, an objective standard of its value is set - a reference point. Setting an objective standard of the value of a factor facilitates the work of experts in determining the strength of the influence of factors and reduces expert errors.

The prediction problem is reduced to the matrix-matrix composition of the matrix of weights and the vector of initial increments of characteristics.

This algorithm works for positively defined matrices, while in our case the elements of the adjacency matrix and increment vectors can take negative and positive values.

Learning Algorithm for Fuzzy Cognitive Map

Suppose that we have a set of 3N historical data lines (hereinafter - training material) about the status of concepts in the system. From the point of view of the problem of forecasting based on increments of concepts (see "Method of obtaining a forecast"), increments of concepts from i-th iteration to (i + 1) iterations will constitute the initial increment vector. In this case, the fuzzy cognitive map should show that with such an initial increment vector, the values of the concepts will change in such a way that the resulting increments will lead to values on the (i + 2) iteration.

Let $A_i(t)$ be the value of the concept e_i at time t. Based on the specification of the learning material given above, we will consider triples of rows $A_i(t)$, $A_i(t+1)$, $A_i(t+2)$.

Define $x_i = \frac{A_i(t+1)-A_i(t)}{A_i(t)}$, $y_i = \frac{A_i(t+2)-A_i(t)}{A_i(t)}$. Here, x are the initial increment vectors, and y is the resultant increment vectors.

Let $o_i(t)$ be the increment e_i, obtained as a result of the prediction on the initial vector $x(t)$.

The learning task is to minimize the error of the fuzzy cognitive map, but with the values of x, y, o introduced in this paragraph.

To solve the learning problem, a genetic algorithm is proposed. As a chromosome, a one-dimensional array of values is allocated, into which a two-dimensional array of weights of the fuzzy cognitive map is decomposed. Each value in this array is called a gene. Let's define the basic steps of the algorithm:

For all non-zero values of the weights of the initial map, a new non-zero weight value is defined, given by a small random number (the sign is not important). The initial non-zero values of the weights are determined by the expert (a non-zero value can be any, its only purpose is an indication that, according to the expert, there is a causal relationship between the two selected concepts).

Item 1 repeats PopulationSize times. Thus, the initial population of random solutions is formed.

The fitness function is defined for each chromosome (see below for the form of fitness function).

The pool of parents is determined by the method of "roulette".

In the pool of parents, "elite individuals" are added. The elite individuals in genetic algorithms are individuals who have shown the best value of fitness function on the last few generations (one individual per generation).

There is a crossing of chromosomes that fall into the parents' pool. The crossing of chromosomes A and B occurs as follows. The crossing boundary of l is randomly determined. Let A_{l+} be the part of chromosome A, consisting of genes located from l, and A_{l-}- part of the chromosome, located up to l. Then the result of crossing will be two chromosomes $A_{l-}B_{l+}$ and $B_{l-}A_{l+}$. The probability of crossing is determined in advance. If crosses do not occur, both parental chromosomes change without change into a population of offspring.

From the descendants obtained in step 6, a new population is formed (its size is exactly the same as the size of the population in the previous step of the algorithm).

There are mutations in the population of descendants. When mutating, a random gene is selected and replaced with a new random value. The probability of a mutation is determined in advance. If the mutation does not occur, the chromosome passes to the next iteration of the algorithm unchanged.

The following generation parameters are determined: an elite specimen (an individual with the best fitness value) to preserve its gene pool; the average fitness of the population (only relevant for evaluating the convergence of the algorithm); the value of fitness of an elite individual.

If the fitness value of an elite specimen is greater than a predetermined value of maximum fitness, the algorithm stops, and the selected chromosome is decomposed into the adjacency matrix of the fuzzy cognitive map (the training is considered complete). Otherwise, go to step 3.

The concept of elite individuals was introduced into the algorithm to accelerate the convergence of the algorithm. The number of elite individuals is taken equal to 60, while the size of the population is 100 (thus, at each step after the 60th generation only 40 chromosomes from the current population have chances of crossing - the rest is filled by an elite gene pool inherited from previous populations) .

The maximum fitness value is defined as 0.99. The results of the training are rounded to the nearest hundredths.

The probability of crossing is defined as 0.9, and the probability of mutation is 0.5. Such a high probability of a mutation (usually uncharacteristic for genetic algorithms) is justified in this case, since mutations introduce genetic diversity into the population. At the same time, since an elite gene pool is used, there is no risk of irretrievably "losing" useful genes from previous generations.

Conclusion

In this paper, we presented a modular time series prediction system, which is based on a neural-fuzzy network and fuzzy cognitive maps. Such a prediction system allows you to include all the factors that influence the development of the situation, this is the actual numerical time series that is predicted on a fuzzy neural network and events that directly affect the future development of the time series. Developed is a genetic algorithm for learning a cognitive map that allows you to speed up the process of developing and adjusting the links of fuzzy cognitive maps.

Acknowledgements

This work was supported by the Russian Foundation for Basic Research (Grant No. 17-07-01558).

Bibliography

[Averkin, 1986] Аверкин А.Н., Батыршин И.З., Блишун А.Ф., Силов В.Б., Тарасов В.Б. Нечет-кие множества в моделях управления и искусственного интеллекта / под ред. Д.А.Поспелова. - М.: Наука, 1986. - 312 с.

[Cai, 2007] Cai, X.; Zhang, N.; Venayagamoorthy, G.K.; Wunsch, D.C. Time series prediction with recurrent neural networks trained by a hybrid PSO-EA algorithm. Neurocomputing 2007, 70, 2342–2353.

[Catalo, 2011] Catalao, J.P.S.; Pousinho, H.M.I.; Mendes, V.M.F. Hybrid wavelet-PSO-ANFIS approach for short-term electricity prices forecasting. IEEE Trans. Power Syst. 2011, 26, 137–144.

[Dong, 2006] Dong, Z. Study on the time-series modeling of China's per capita GDP. Contemp. Manag. 2006, 11, 15.

[Efremova, 2012] Efremova, N., Asakura N., Inui T.. Natural object recognition with the view-invariant neural network. In: 5th International Conference of Cognitive Science, 2012, pp.802-804.

[Gunasekaran, 2011] Gunasekaran, M., and K. S. Ramaswami. "A fusion model integrating ANFIS and artificial immune algorithm for forecasting Indian stock market." Journal of Applied Sciences 11.16 2011: 3028-3033.

[Haykin, 1994] Haykin S. Neural networks: a comprehensive foundation. New York: Macmillan, 1994.

[Kosco, 1986] Kosko B. Fuzzy cognitive maps //International journal of man-machine studies. – 1986. – Т. 24. – №. 1. – С. 65-75.

[Lefevr, 1965] Лефевр В. А. Исходные идеи логики рефлексивных игр //Материалы конфе-ренции «Проблемы исследования систем и структур». М.: Издание АН СССР. – 1965.

[Meng, 2012] Meng, Y.-B.; Zou, J.-H.; Gan, X.-S.; Zhao, L. Research on WNN aerodynamic modeling from flight data based on improved PSO algorithm. Neurocomputing 2012, 83, 212–221.

[Pousinho, 2011] Pousinho, H.M.I.; Mendes, V.M.F.; Catalão, J.P.S. A hybrid PSO-ANFIS approach for short-term wind power prediction in Portugal. Energy Convers. Manag. 2011, 52, 397–402.

[Rotshtein, 1997] Ротштейн А.П., Штовба С.Д. Нечеткая надежность алгоритмических про-цессов. - Винница: Континент-Прим, 1997. - 142 с.

[Soros, 2003] Soros G. The alchemy of finance. – John Wiley & Sons, 2003.

[Trofimov, 2010] Trofimov, A., Povidalo I. And Chernetsov S. Usage of the self-learning neural networks for the blood glucose level of patients with diabetes mellitus type 1 identification. Science and education, 2010, vol. 5. Availible at: http://technomag.edu.ru/doc/142908.html

[Wang, 2012] Wang, S.; Wu, L. An improved PSO for bankruptcy prediction. Adv. Comput. Math. Appl. 2012, 1, 1–6.

[Wang, 2015] Wang, J. S., & Ning, C. X. (2015). ANFIS Based Time Series Prediction Method of Bank Cash Flow Optimized by Adaptive Population Activity PSO Algorithm.Information, 6(3), 300-313.

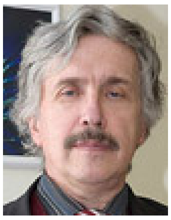
[Yarushev, 2016] Yarushev S. A., Averkin A. N. Review of studies on time series forecasting based on hybrid methods, neural networks and multiple regression //Программные продукты и системы. – 2016. – №. 1 (113).

[Zadeh, 1976] Беллман Р., Заде Л. Принятие решений в расплывчатых условиях. В кн.: Во-просы анализа и процедуры принятия решений. - М.: Мир, 1976. - С. 172-215.

Authors' Information

Sergey Yarushev *– is Postgraduate Student of Dubna State University, Dubna, Russia; e mail: sergey.yarushev@icloud.com*

Major Fields of Scientific Research: time series forecasting, neural networks, fuzzy systems, big data, deep learning, big data, artificial intelligence.



Alexey Averkin *– is Senior Researcher in Dorodnicyn Computing Centre, Federal Research Center "Computer Science and Control" of Russian Academy of Sciences;, Moscow, Russia; e-mail: , averkin2003@inbox.ru*

Major Fields of Scientific Research: data mining, big data, neural networks, time series forecasting, fuzzy systems, forecasting, deep learning, artificial intelligence.