ISSN 1992-6502 (Print) Вестник УГАТУ ISSN 2225-2789 (Online)
Vol. 18, no. 5 (66), pp. 127-130, 2014 Vestnik UGATU http://journal.ugatu.ac.ru

UDC 004.7

TIME SERIES ANALYSIS TO PREPARE IT FOR FORECASTING TRAJECTORY BASED ON DATA MINING ONTOLOGY

I. P. VLADIMIROVA[1], **K. V. MIRONOV**[2]

[1]ooo-flaming@mail.ru, [2]mironov@ict.tuwien.ac.at

[1]Ufa State Aviation Technical University, Russia
[2]Vienna University of Technology, Austria

Submitted 2014, June 10

Abstract. In this article the time series analysis for predicting the state of dynamic objects is discussed. Application of the ontological model for storing the experimental results as the cases of solving the Data Mining task is proposed.

Key words: time series analysis; parameter estimation; search for precedents; prediction; the ontological model.

To study the complex system properties, including experimental studies, the approach based on intelligent analysis of the signals, produced by the system, is used. The classical problems of Data Mining (DM) are: cluster analysis and classification, forecasting, searching of association rules and identifying hidden dependencies in the data [1]. Prediction based on DM is relevant in the cases when it is difficult to describe mathematically the studying process, but there is a number of observations with the values of certain index (criterion), ordered in chronological sequence, i. e. in ascending order of variable t-time parameter.

For example, in seismology – recording the vibrations the earth crust, in good's delivery to the retail outlets – the data about changes in demand, in processes of the quality management – the number of defective products, the supply quality, etc. In this article the time series analysis to predict the state of dynamic objects is discussed.

STATE OF ART

Time series analysis – is a set of mathematical and statistical methods of analysis, used for detection the time series structure and its prediction. Identifying the structure of the time series is necessary in order to build a mathematical model of the phenomenon that is the source of the analyzed time series. Prediction of future values of the time series is used for the effective decision making. Generally, prediction comes from a certain given parametric model. Thus the standard parametric estimation methods (least-squares method, method of moments) are used. On the other hand, the methods of nonparametric estimation for the fuzzy defined models are sufficiently developed [2].

To describe the time series the following model can be used:

$$f_t = x_t + y_t, \qquad (1)$$

where f_t – current term of the time series in time moment t; x_t – random variable, that is generated by a deterministic function or a stochastic process; y_t – random variable, that is generated by random non-auto correlative process with zero mathematical expectation and constant variance.

Time series are studied for different purposes. In one problem it happens enough to get the description of time series characteristic features, in the other – is required not only to predict the future values of the time series, but also to control its behavior. A method for the time series analysis is determined, on the one hand, by the objectives of the analysis, and on the other hand, by the probabilistic nature of the values formation: correlation analysis, spectral analysis, smoothing and filtering, models of autoregression and moving average, forecasting.

CURVE FITTING PROBLEM DEFINITION AND SOLVING APPROACH

One of the most common ways to model the trend of the time series is constructing the analytic function, which characterizes the dependency between the time series levels and time, or trend. This method is called analytical alignment of the time

Research, presented in this article, was partially supported by RFBR grants 12-07-00377, 13-07-00273-a, and by MULTIC Erasmus Mundus program Action 2.

series. Since the time dependency can take different forms, for its formalization different types of functions can be used: linear trend; hyperbole; exponential trend; trend in the form of a power function; parabola of the second and more order, spline [3, 4].

In this paper the application of time series analysis for robotic catching of thrown object is considered. The robotic catching is considered as a stage of complex technological process called "Transport-by-throwing" (TbT). The concept of TbT was introduced in [5] and then developed in [6–11]. The concept consist in following: if there is a need to transport certain object from certain source point to certain destination point in space, this transportation may be reached via throwing the object by specific throwing device from source point towards the destination point. The object is caught at the destination point by specific catching device (gripper). This concept may be applied in factories of light industry for transportation of objects between machine tools that process these objects. Traditionally conveyor systems are applied for this task, however TbT has a potential ability to overcome conveyor systems in terms of flexibility, speed and energy consumption [8].

According to [8] the task of providing the object transportation by throwing and catching may be divided into four subtasks (shown on Fig. 1):

• Throwing. The thrower must throw the object towards the destination area.

• Catching. The gripper must catch the object at the destination area.

• Prediction. This subtask is motivated by catching: prior knowledge about object trajectory in gripper workspace is needed to define the catching behavior of the gripper. Trajectory prediction is an application of time series forecasting. The time series in this case consist of measured coordinates of the flying object in time.

• Tracking. This subtask is motivated by prediction: the predictor needs the information about the object trajectory at the beginning of flight. Based on this data it can forecast the trajectory into the gripper workspace.

Forecasting of object trajectory may be done based on physical models of projectile motion [6-8] taking in mind the influence of gravitation and air drag on the airborne object. As the influence of the air flow near the flying object is hard to model (or the model become extremely computationally complicated) the performance of such systems is limited. As an alternative way of forecasting sample-based methods were proposed: neural network prediction [10] and nearest neighbor forecasting [11]. Pre-liminar results based on simulation show that nearest neighbors forecasting is a promising method for trajectory prediction [11] however in some cases it is sensitive to tracking errors which may lead to higher errors of trajectory prediction. The goal of this paper is to introduce the statistical method which decreases influence of observation errors.

An example set of trajectories produced by object throws from single point with varying velocity and direction is shown on Fig. 1.

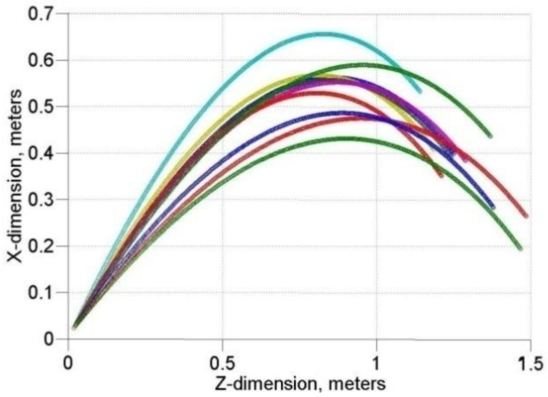

Fig. 1. 10 various trajectories of the object thrown from one point; variety of trajectories is provided by deviations in angle of throw and initial velocity

The object position is measured by the stereo camera set with fixed frame-rate (in the simulation it was set to 100 fps to increase clearness of the data; in reality the frame rate of 50–160 fps may be achieved using standard industrial cameras).

The error filtering is implemented by fitting the values of each coordinate with the approximation function in coordinate-time space. The analysis of the quality of approximation is done in order to define the most appropriate approximation function. The analysis was made using MATLAB Curve Fitting Toolbox (cftool). Example results for one trajectory are shown in table 1. If the approximation model is appropriate 1st and 4th criteria should have values near to zero, 2nd and 3rd criteria should have values near to one.

When selecting the parameters in cftool application of the working environment MATLAB [3, 4] is also possible to calculate the confidence intervals for the obtained values of the model parameters, corresponding to some given probability level, equal 95 %:

$$C = b \pm t\sqrt{S}\,, \qquad (2)$$

where b – obtained parameters values; t – inverse function for distribution function of Student; S – vector of the diagonal elements of the matrix sX^tX, where X – design matrix; s – mean-root-square error.

Table 1

Approximation criteria

№	Criterion	Equation *	Value
1	Sum of squares due to error	$SSE = \sum_{k=1}^{n} w_k \left(y_k - \hat{y}_k\right)^2$	0.0013
2	R-square	$SSR = \sum_{k=1}^{n} w_k \left(\hat{y}_k - \bar{y}\right)^2$ $SST = \sum_{k=1}^{n} w_k \left(y_k - \bar{y}\right)^2$ $R(square) = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$	0.9984
3	Adjusted R-square	$Adjusted\_R(square) = 1 - \frac{SSE(n-1)}{SST(n-m)}$	0.9984
4	RMSE (Root mean Squared Error)	$RSME = \sqrt{\frac{SSE}{n-m}}$	0.0052

* w_k – weights, y_k – data in x_k, k – parameters values in x_k, , $\bar{y}$ – mean, n – number of data values, m – number of parameters.

In the example, when approaching the data by the second-degree polynomial the function

$$f(x) = -0.0005264\, x^2 + 0.03249\, x + 0.0142$$

with the best values of factors and confidence intervals is received.

Note that approximation is valid only for filtering errors on the tracking stage. It is not useful for prediction as polynomial model does not represent accurately ballistic motion in the air and will lead to systematic errors on a long scale.

Results of the analysis are the initial data for implementing the method of predicting the object motion trajectory based on the search of the closest case in the robotic TbT task. As a mean for storing the experimental results is proposed to develop the DM ontological model (ontology) with the use of Web Ontology Language (OWL). Ontology includes a set of classes of concepts from DM field, relationships between concepts, concept properties and individuals of classes [12].

Experimental results in the form of models describing the initial time series are presented in the ontology as a set of individuals of the class “Data analysis task”. The time series model together with the prediction result, obtained on its base, is presented as a case of solving the DM task. A set of cases is stored in the knowledge base, built in the DM ontology. Accumulation of the cases is a self-learning technology that allows realizing a method of forecasting the object trajectory based on the search of the closest case in the ontology.

CONCLUSION

This paper presents the approaches to the time series analysis with the use of analytic functions of the Curve Fitting Toolbox subsystem in MATLAB software environment. Approaching experimental example of one of the object trajectories when the object transportation by throwing from one robot to another, is shown.

The quality estimation analysis of the trajectory analytic function selection based on the eligibility criteria of the obtained approximations, which helped to determine the best parametric model with the obtained values of the parameters, was conducted. The result of the trajectories database analysis in Curve Fitting Toolbox subsystem in MATLAB software environment has shown that approximation by the second-degree polynomial with the use of the least squares method best of all describes the time series. Application of the ontological model for storing the experimental results as the cases of solving the DM task is proposed.

ACKNOWLEDGMENTS

The work is supported by the European Erasmus-Mundus Program and the grant of RFBR 13-08-00321 “Intellectual management of complex processes cooperation based on ontological analysis and knowledge processing under uncertainty”.

REFERENCES

1. **Ethem Alpaydin,** *Introduction to Machine Learning,* 2nd ed. Cambridge, Massachusetts: The MIT Press, 2010.
2. **Цыплаков А.** Введение в прогнозирование в классических моделях временных рядов // Квантиль. 2006. № 1. С. 3-19. **[** Ziplakov A. “Introduction to forecasting in classical models of time series,” *Kvantil*, no. 1, pp. 3–19, 2006. **]**
3. **http://matlab.ru**
4. **Anufriev I. E., Smirnov A. B., Smirnova E. N.** MATLAB 7.0. St. Petersburg: BHC-Petersburg, 2005.
5. **Frank H., Wellerdick-Wojtasik N., Hagebeuker B., Novak G., Mahlknecht S.** “Throwing objects – A bio-inspired approach for the transportation of parts,” in *Proc. IEEE Int. Conf. Robotics and Biomimetics, ROBIO '06, 2006*, pp. 91–96.
6. **Barteit D., Frank H., Kupzog F.** “Accurate prediction of interception positions for catching thrown objects,” in *Proc. IEEE Int. Conf. Industrial Informatics (INDIN)*, Daejeon, Korea, 2008, pp. 893–898.
7. **Pongratz M., Kupzog F., Frank H., Barteit D.** “Transport by throwing – a bioinspired approach,” in *Proc. IEEE Int. Conf. Industrial Informatics, INDIN 2010*, Osaka, Japan, 2010, pp. 685–689.
8. **Pongratz M., Pollhammer K., Szep A.** “KOROS initiative: automatized throwing and cathcing for material transportation,” *ISoLA 2011 Workshops*, pp. 136–143, 2012.
9. **Pongratz M., Mironov K. V., Bauer F.** A soft-catching strategy for transport by throwing and catching // Вестник УГАТУ. 2013, Т. 17, № 6 (59). С. 28–32. **[** Pongratz M., Mironov K. V., Bauer F. “A soft-catching strategy for transport by throwing and catching,” *Vestnik UGATU*, vol. 17, no. 6 (59), pp. 28–32*,* 2013. **]**
10. **Mironov K. V., Pongratz M.** Applying neural networks for prediction of flying objects trajectory // Вестник УГАТУ. 2013, Т. 17, № 6 (59). С. 33–37. **[** Pongratz M., Mironov K. V., “Applying neural networks for prediction of flying objects trajectory,” *Vestnik UGATU*, vol. 17, no. 6 (59), pp. 33–37*,* 2013. **]**
11. **Mironov K. V., Pongratz M., Dietrich D.** “Predicting the trajectory of a flying body based on weighted nearest neighbors,” in *Proc. Int. Work-Conference on Time Series*, Granada, Spain, June 2014, pp. 699–710.
12. **Черняховская Л. Р., Федорова Н. И., Малахова А. И., Владимирова И. П.** Аналитическая поддержка принятия решений на основе онтологического анализа прецедентов проблемных ситуаций // Информационные технологии интеллектуальной поддержки принятия решений: Тр. междунар. конф. Уфа: УГАТУ, 2013. С. 96–99. **[** Chernyakhovskaya L. R., Fedorova N. I., Malakhova A. I., Vladimirova I. P. “Analytical decision making support on the base of ontological analysis of problem situation cases,” in Proc. Information Technologies of Intellectual Decision Support: Int. Conf. Ufa: USATU, 2013. pp. 96–99. **]**

ABOUT AUTHORS

VLADIMIROVA, Irina Pavlovna, postgraduate student of the Department of technical Cybernetics. Diploma of economist-mathematician on mathematical methods in Economics (UGATU, 2009). Research in critical situation management, informatics and artificial intelligent.

MIRONOV, Konstantin Valerevich, Doctoral Stud. (Institute of Computer Technology). Dipl.-Eng., BSc (Ufa State Aviation Tech. Univ., 2012). Researches: material transportation by throwing and catching.

МЕТАДАННЫЕ

Название: Анализ временных рядов для прогнозирования траектории с использованием онтологической модели интеллектуального анализа данных.

Авторы: И. П. Владимирова, К. В. Миронов.

Организации:

ФГБОУ ВПО «Уфимский государственный авиационный технический университет», Россия;

Венский технологический университет, Австрия.

Email: ooo-flaming@mail.ru, mironov@ict.tuwien.ac.at

Язык: английский.

Источник: Вестник УГАТУ. 2014. Т. 18, № 5 (66). С. 127–130. ISSN 2225-2789 (Online), ISSN 1992-6502 (Print).

Аннотация: В статье рассматривается анализ временных рядов для прогнозирования состояния динамических объектов. Предложено применить онтологическую модель для хранения результатов экспериментов как прецедентов решения задачи интеллектуального анализа данных.

Ключевые слова: анализ временных рядов; оценка параметров; прогноз; поиск прецедентов; онтологическая модель.

Об авторах:

ВЛАДИМИРОВА Ирина Павловна, асп. каф. техн. кибернетики. Дипл. экономист-математик по матем. методам в экономике (УГАТУ, 2009). Иссл. в обл. поддержки принятия решений в соц. и экон. системах.

МИРОНОВ Константин Валерьевич, асп. каф. вычислительной техники и защиты информации, специалист по защите информации (УГАТУ, 2012). Иссл. в обл. транспортировки объектов перебросом.