Journal of KONBiN 2019
Volume 49 Issue 3
DOI 10.2478/jok-2019-0065

Olesya AFANASYEVA[1], Olga KOROSTIL[2]
[1] *Pedagogical University in Krakow (Uniwersytet Pedagogiczny w Krakowie)*
[2] *Ukrainian Academy of Printing*

FORECASTING AND PREDICTING IN ENGINEERING TASKS

Prognozy i przewidywania w zadaniach inżynierskich

***Abstract**: The work analyzes the tasks of solving problems, which consist in determining the events that may occur through some time after the completion of the process of solving the problem. One of the possible classifications of such tasks is proposed. The analysis of differences between different types of tasks is carried out, features of implementing the processes of their resolution are revealed. The paper considers in detail such types of tasks as prognosis and prediction. Differences are described between these processes with each other and the characteristics that determine each of the processes. The comparison of various types of processes in the overall forecasting process is presented.*
Keywords: prognosis, prediction, prediction models, hypotheses, heuristics

Streszczenie: *Tematem publikacji jest rozwiązywanie problemów polegających na określaniu zdarzeń, które mogą mieć miejsce po zakończeniu procesu rozwiązywania problemu. W artykule zaproponowano jedną z możliwych klasyfikacji takich zadań. Ponadto, przeprowadzono analizę różnic rozmaitych zadań. Przedstawiono także cechy wdrażania procesów ich rozwiązywania. W artykule szczegółowo rozważano takie rodzaje zadań, jak prognozowanie i przewidywanie. Opisano różnice między tymi procesami oraz cechy charakterystyczne, które określają każdy z wymienionych procesów. Przedstawiono porównanie różnych rodzajów procesów w całościowym procesie prognozowania.*
Słowa kluczowe: prognozowanie, przewidywanie, modele przewidywania, heurystyka

1. Introduction

Prognosis is a common process for identifying a random event (Ev_i),which is comprehensively studied, but due to its universality, there is a fairly wide range of tasks that should be considered. The general prognosis process is called a set of all possible types of processes to determine the occurrence of random events Ev_i after some time at the end of the process of its determination. One of the tasks of studying these processes is the possible classification of various prognosis processes, which can be attributed to the general idea of prognosis. Based on this classification, one can accept the differences between processes that can be attributed to representations about different types of prognosis of the occurrence of one or another Ev_i. As part of a possible classification, it is expedient to distinguish the following types of general prognosis process:

- prognosis,
- prediction.

The introduction of the given types of processes of the general idea of prognosis requires a justification of the corresponding division. In connection with this, it is necessary to analyze each of the types of the resulted processes and the need to identify the key differences between the processes that implement them. Such differences relate to ways of implementing processes that make up a general idea of prognosis ($Pr(PR)$), where PR is prognosis and Pr is the process of implementing the prognosis. In addition, the differences relate to the ways in which input data is formed for $Pr(PR)$ as well as the differences between the parameters of the goals of these processes. The parameters of the objectives of the investigated processes can differ from each other not only through the use of various modifications of the methods of implementing the corresponding processes, but also from the use of different volumes of input data and from their extensions that can be used within the general organization of these methods [5].

Also, an important problem is the definition of task classes, which must be used to solve the above types of processes. Such tasks and, accordingly, the processes of their resolution, are directly related to the characteristics of the object, or the process in respect to which the solution to the selected task is implemented.

The least investigated are the processes of prediction.

An example of the feasibility of using prognosis can be the problem of extending the resource of a complex technical object (CTO). Because the value of the resource ($Tr_i(CTO)$) is determined by the time during which CTO will function in accordance with the technical requirements, then it can be argued that the exact value of the resource (Tr_i) depends on various random factors that can affect the CTO. It is known that in practice the value of Tr_i is determined with some stock, which is provided by the interval of time for which Tr_i decreases. Obviously, this stock of Tr_i leads to a decrease in the efficiency of using CTO and is a value determined by a certain approximation.

Prognosis in the tasks of extending the resource Tr_i, or $Tr^P - Tr = \Delta Tr_i$ where Tr^P is the value of the extended resource is to determine an adequate value ΔTr_i. Value ΔTr_i cannot be set accurately enough and, therefore, in practice is taken down. Values ΔTr_i are determined on the basis of the data obtained when solving engineering and design problems arising in the course of work on the continuation of the resource. Confirmation of this state of affairs is well-known cases when CTO operates in accordance with technical requirements in cases that are significantly higher than Tr^P, or Tr_i. This means that the inaccuracy of the definition ΔTr_i, when performing work on the continuation of the resource of CTO, when its reduction leads to its reduction in relation to technical capabilities of CTO, and its inadmissible increase, in terms of technical capabilities, can lead to emergency situations in the process of CTO functioning.

2. Basic concepts

Definition 1. Prognosis is the process of solving a task in a selected algorithm that does not conform to the model of the natural process of occurrence Ev_i, but allows you to determine Ev_i. Such an algorithm uses initial data that is characterized by a certain degree of uncertainty with respect to Ev_i [6].

The uncertainty of an event that needs to be predicted is due to the following reasons:

1. Lack of necessary information about the natural processes of occurrence Ev_i.
2. Absence of fragments of algorithms in the system of prognosis process.
3. Absence of a number of necessary initial data x_i, which may affect the appearance Ev_i.

Measure of uncertainty $\eta(PR)$ is a relative value and depends on the needs of the problem that needs to be solved, when implementing the process PR. To determine it in the form of a certain numerical value, it is necessary to establish a method for measuring the amount of information which, for each type of cause of uncertainty, should be established separately.

In the first case, the method of measurement $\eta(PR)$ is based on the use of the corresponding value of the amount of information that can be described as:

$\Im(Ie) = \sum_{i=1}^{m}[x_i * j(x_i)]$ where x_i is a data element, $j(x_i)$ is a description of the interpretation of the relevant data, which is known in relation to the task

$Za_i(PR)$, $Ie_i = x_i * j(x_i)$ is an information element related to the description of the problem to be solved.

In the second case, the method of measuring information uses the value that is the number of elementary algorithms Ae_i, each of which is a description of a certain set of components that allow you to define a separate element of the goal. The set of such algorithms allows you to get a solution to a task that is determined by the general purpose $C_i(Za_i) = \sum_j^m C_{ij}(Za_i)$ of the task Za_i. If the target $C_i(Za_i)$ is formed in such a way that it cannot be divided into a set of final components of the goals, then the corresponding components $C_i(Za_i)$ represent intermediate goals $Cp_i(Za_i)$ which can be described as: $Cp_1(Za_i) \rightarrow Cp_2(Za_i) \rightarrow \cdots \rightarrow Cp_m(Za_i) \rightarrow C_i(Za_i)$.

In the third case, the method of measuring information uses a value that is the number of data elements x_i, which is not enough to implement all the necessary elementary algorithms Ae_i, each of which ensures the achievement of sub-targets $C_i(Za_i)$.

Definition 2. Prediction (Pp) is a process $[\Pr(Pp)]$ determining the possibility of occurrence of some event on the basis of analysis of incomplete initial data and using elementary algorithms Ae_i, which cannot be sufficiently substantiated.

One of the differences of prediction from prognosis is that in the first case there is a well-formulated task and goal that needs to be achieved, and in prediction there is no clearly formulated problem, there is the goal of the task to be achieved. Such a goal in the tasks of prediction is not sufficiently determined by the initial data.

The next difference is that prognosis uses a mathematical apparatus, mainly probability theory, which is sufficiently tested and, in most cases, allows for sufficiently adequate results. Examples of such models can be: regression models, models using Bayes affirmations, models based on the use of numerical series and others. In the case of prediction, heuristic models are used based on the use of those or other hypotheses formulated on the basis of the results of the analysis of the purpose of the prediction and the known initial data. An example of such models can be: a model of analogies, a model of successive approximations, models that use different types of non-classical logic, etc. [4, 7].

The next difference is that in prognosis a result is obtained of solutions of the problem. In the case of prediction, a number of outcomes are obtained, among which the result is determined which is more appropriate, or, in accordance with the chosen criterion, is optimal in relation to the solvable problem. For prediction, it is characteristic that the choice of the optimal result is implemented within the framework of the overall process of predictions. This means that optimization is implemented on the basis of using not only the initial data but with using additional data obtained in different variants of the results of prediction. Therefore, prediction can be considered as a process implemented in several stages.

In the first stage, based on the initial data, different possible versions of the heuristic algorithm are implemented, the difference between which is determined by different permissible interpretations of the same input data.

At the second stage of prediction, the information expansion of the received and different results of the first stage is carried out. Such an extension is carried out by realizing the conclusions of the generalized interpretative descriptions of all the results obtained in the first stage.

In the third stage of prediction, the process of forming the end result is used, which uses expanded information about the subject of predictions.

The next difference is that one of the main parameters that characterizes prognosis is a measure of uncertainty in the subject of prognosis, which is determined by the amount of information that is not sufficient to obtain more accurate results. Therefore, prediction is characterized by the number of versions of heuristic algorithms used in this process.

3. Analysis of basic characteristics of the prognosis process

A general idea of the prognosis is an understanding of the process that should ensure the result of the calculations, which is as much as possible consistent with the parameters of the predicted event that would characterize it in the event of this event. The result that can be obtained in prognosis is compared with an event that actually occurs at a certain point in time. Therefore, one of the main parameters of an arbitrary method of prognosis is a measure of correspondence of the results obtained, using the model of prognosis, the real characteristics of the event Ev_i, which should come in accordance with the natural or technical preconditions of occurrence Ev_i. In many cases, prognosis is used to confirm or deny the basic hypothesis of the possibility of occurrence Ev_i. Therefore, let's look at the definition.

Definition 3. The basic hypothesis ($Bg(Ev_i)$) of prognosis is a hypothesis about the possibility of occurrence of some event Ev_i.

Thus, you can enter a parameter that determines the measure of probability that an event with expected characteristics will occur. Such a parameter is the likelihood of occurrence Ev_i will be noted as ($Pro(Ev_i)$) The peculiarity of this parameter lies in the fact that the possibility of occurrence of some event Ev_i is determined in relation to a certain moment of its origin t_i, regardless of which parameter is selected as the base. The base parameter defines a particular event as such, but it is always projected onto the time parameter. Therefore, we can assume that the next prognosis parameter is the prognosis time parameter Δt_i, the value of which is measured on the basis of the calculation of the probability of occurrence $Ev_i(\Delta t_i)$ through the time interval Δt_i whose initial count t_{ip} coincides with the moment of determining such an interval by the forecasting process, and the final counting interval t_{ik} is determined by the predicted moment of occurrence Ev_i. The value of this parameter, in the simplest case, is determined on the basis of numerical rows, if the latter can be constructed for the investigated events or processes. In other cases, the value Δt_i is determined on the basis of x_i parameter analysis and changes in their values, which determine the investigated event with their subsequent projection at the time of forecasting the process through which the event Ev_i should occur. An example can be an object for which one of the basic parameters is a change in its temperature. In this case, prognosis is conducted in relation to the occurrence of the event Ev_i, which is to achieve for the parameter x_i temperature Q_i, the value of which is established on the basis of the analysis of the object itself or its nature. Because the existence of x_i itself is considered in time, then projection of x_i to the parameter t_i is implemented by detecting the processes that use the parameter x_i. In this case, the environment of the parameter x_i expands, or $x_i \rightarrow U_j(x_i)$, where $U_j(x_i)$ is one of the sub-processes that besides the parameter x_i uses other parameters and, first of all, the time parameter and is implemented within the framework of the operation of the investigated object. If there is a situation where it is not possible to detect the processes associated with x_i and use the parameter t_i, then theoretically x_i cannot be projected on the time parameter. From the formal point of view, in this case, the value of the parameter x_i is not time-dependent, which may mean that it is not possible to observe the parameter x_i in the corresponding sub-process, or it may mean the absence of data on the laws of the operation of a separate sub-process U_j of the corresponding object in time related to the parameter x_i. This case will not be considered. Let us assume that any object characterized by the parameter x_i can cause the corresponding sub-process $U_j(x_i)$, and can be projected on the time parameter, which has an interpretation consistent with the operation of the whole facility and the interpretation agreed with the event Ev_i.

Another common option of the prognosis process is a parameter that determines the degree of danger, which can be caused by the effect of an event Ev_i on the object or process, if it arises. We will write this parameter as Da_i. The importance of this parameter, for the characteristics of prognosis processes, is determined by the degree of negative influence (*Ni*) of an appropriate Ev_i on the environment, in which the object is functioning, which is written as Eo_i. It is advisable to note that in the case of using negative influence *Ni* we can talk about events that have a positive effect (Pi) on Eo_i [3] . As part of the prognosis problems, this case is not considered, because it is in most cases planned and predictable.

Measurement or evaluation of $Ni(Eo_i)$ or $Da_i(Eo_i)$ is very important, since it depends on the choice of one or other means of forecasting Ev_i, which can be described by the ratio $Ni(Eo_i) \rightarrow Pr[PR(Ev_i)]$, where $Pr[PR(Ev_i)]$ is the prognosis process of Ev_i. The concept of danger implies the existence of an object in the environment Eo_i, or a process in relation to which an action of Ev_i may be dangerous. It should be noted that the problems of prognosis, first of all, are considered in connection with the possibility of negative impact Ev_i on the Eo_i components and on Eo_i in general. To solve the problem of determining the degree of risk of Ev_i influence on the object or process, it is necessary to define, for example, the parameters of the technical object (TO), which will be considered an object of the environment, which can be influenced by $Ni(To_i)$. Such assessments, in most cases, relate to TO, and one of the known ways to implement them is approaches that use risk concepts [1]. A prerequisite for the existence of the parameter $Da_i(TO_i)$ or $Da_i(Eo_i)$ is the presence in Ev_i of factors that can interact with elements of TO, or Eo_i. This means that the factors $\varphi_i\big(Da_i(Ev_i)\big)$ must have parameters that are common to the component parameters TO_i, which are coordinated with each other in a range of values and measurement methods. Determining the extent of the negative impact and the assessment of such impact on TO_i is performed in diagnostics [2]. Therefore, we will accept that Ev_i can negatively act on TO_i, and the magnitude of such influence determines the degree of relevance $\mu[PR(To_i)]$ of solving the prognosis task that is written as $Zak(PR)$. The measure of relevance is directly related to the object, which can be influenced by a random event Ev_i, which is predicted.

The next important characteristic of the prognosis process is the assessment of initial or input data that determines the significance of the prerequisites necessary for the implementation of adequate prediction of the possibility of occurrence Ev_i. Prognosis of Ev_i apart from the prerequisites or from a source of Ev_i occurrence is not possible, since the initial data that the forecasting system uses are data including the corresponding source. Considering the source of Ev_i as a necessary component of the prognosis process allows us to expand the information about the input data and thus provide greater reliability of the relevant prognosis. Investigation of the preconditions of Ev_i occurrence which we will be expressed as Ps_i is closely associated with Eo_i not just because Ev_i can affect Eo_i , but also because, in many cases, the source of events Ev_i is a part of Eo_i. This means that you can write the dependence: $Ps_i \rightarrow Ev_i \rightarrow Eo_i$ where Eo_i includes TO_i.

4. Analysis of basic characteristics of prediction processes

The need to allocate predictions to a particular type of forecasting process is conditioned by the hypothesis that the processes that generate Ev_i are essentially nonlinear. This causes the incorrectness of the extrapolation of the processes of Ev_i generation to be implemented by known nonlinear functions and, moreover, linear functions, which are widely used in prognosis. Significant non-linearity, in contrast to nonlinearity of continuousness, is characterized by the presence of gaps of various degrees and other critical points in functions that could be used for extrapolation with the goal of solving tasks of predicting random events.

In this case, for the construction of prediction models, it is necessary to use the synthesis of various discrete simulation methods using logical means [8]. In order to correctly use this approach, it is necessary to implement it within the framework of heuristic methods based on the use of hypotheses about the processes of occurrence of events Ev_i [9]. In this case, it is possible to take into account significant nonlinearities. Thus, an important characteristic of prediction is the need to form hypotheses that could take into account significant or critical nonlinearities in processes generating events Ev_i. Prediction, like prognosis, approximates patterns that reflect processes occurring in one or another subject area and lead to the Ev_i emergence.

The presence of significant gaps of various kinds is conditioned by the nature of individual components of the general process, in relation to which it is necessary to use predictions of the occurrence of the event Ev_i. For example, such components can be thermal processes, electrical processes, chemical processes, and others. An example of the interconnected functioning of such processes can be the processes that are implemented in the system of electricity consumption, processes associated with the use of gas energy. As part of this example, significant non-linearity may occur, and there will be a spark in the grid, which will result in unauthorized ignition of the gas. For complex technical objects (CTO) there is a situation where a number of significant gaps in the overall implementation of processes in CTO can occur consistently, they can be interconnected, which ultimately leads to an event Ev_i that negatively affects the object.

The next feature of the prediction process is that the goal of predictions $C(Pp)$ is initially formed. This goal, to some extent, is expected by the user and plays a major role in the implementation of the prediction process. In addition to the goal, initial data is used that may not be complete. In the first step, the prediction process, based on the goal, builds an inverse procedure that could identify the known initial data and certain approximations of the individual initial data that is missing. The data obtained as a result of the implementation of the inverse procedure will not necessarily be sufficiently accurate to meet the necessary input data, even if the initial data is complete. The obtained data is compared with the initial data. Then, a situation may arise when the initial data obtained as a result of the implementation of the inverse procedure deviates from the data presented as an initial to an unacceptable extent. In this case, the hypothesis is modified, on the basis of which the next version of heuristics is selected. The given procedures are repeated until the deviation of the data groups, received as a result of the implementation of inverse procedures, from the data given at the beginning of the activation of the prediction process, or the input data obtained in the previous step of their definition, will be sufficiently close in magnitude. The prediction algorithm, in addition to the above processes, implements the procedures of a goal modification, if it is not possible to continue the above-stated iterative processes. Such modifications of the next target of predictions can be repeated. Due to this, as a result of the implementation of the general prediction algorithm, you can get a number of versions of the results of predictions. In the next step of the implementation of the prediction process, the procedures for selecting the optimal result of predictions are executed. There may be a negative target as a result of successive modifications to goals. Then, another feature of $Pr(Pp)$ will be the characteristic of the degree of positivity of the predicted goals. Let us introduce the following definition.

Definition 4. The ultimate goal of the prediction $C_i(Pp)$ is positive if it does not contradict the initial data that represent the initial input data of the general prediction process.

The idea of a positive prediction goal is used to formulate criteria that determine the acceptable modification of the goal and determine the allowable value of changes in the values of input data that arise in the various steps of the implementation of prediction. The idea of negative goals is associated with positive goals, since the idea of negative goals depends on their interpretation by the user.

The prediction process is characterized by the absence of an explicit prediction parameter Δt_i. The existence of such a parameter is a prerequisite for assigning the prediction process to the type of general prognosis processes. This means that the expected event should occur due to the time determined by the process $Pr(Pp)$. Since $Pr(Pp)$ is a model that describes the process of occurrence Ev_i, and an arbitrary process uses the time parameter t_i, the definition of Δt_i is based on the following. Parameter t_i in $Pr(Pp)$ can vary with the speed of real-time change for $Pr(Ev_i)$ which is commonly called a functioning $Pr(Pp)$ in real time $\tau(t_i)$, where $\tau(t_i)$ is the time interval during which real, natural or technical process of Ev_i occurrence is implemented. In this case, $Pr(Pp)$ does not correspond to the notions of prediction and is in a certain approximation only a model that describes the actual process of Ev_i occurrence. If $Pr(Pp)$ has a possibility to set the implementation of the process $P\,r(Pp)$ of the Ev_i emergence at a time interval $\delta(t_i)$ so that inequality $\tau(t_i) > \delta(t_i)$ is true, then the prediction time Δt_i will equal $\Delta t_i = \tau(t_i) - \delta(t_i)$ and the process $Pr(Pp)$ can be attributed to the class of general forecasting processes. It should be emphasized that the fundamental difference of $Pr(Pp)$ from the mathematical model of the process of Ev_i occurrence or $MM(Ev_i)$ is that $Pr(Pp)$ will always be less adequate to the natural or technical processes that determine the occurrence of the predicted event in relation to the model $MM(Ev_i)$, if the latter is possible to construct.

5. Summary

Different types of processes of general idea of prognosis for which their classification is proposed are investigated in this work. Analysis of the selected types of these processes, which collectively relate to the overall solution of the prognosis task, allows you to correctly choose one or the other method of solving the problem of determining key parameters, among which is the parameter time interval of the general process of prognosis of a random event, which may occur after the completion of the process of solving this task.

The given analysis of the parameters and features of each of the resulted process types allows you to choose the following:

- the most appropriate type of process of determining Ev_i from a plurality of processes that make up the overall prognosis process, with respect to the subject area and the applied task that is planned to be resolved in this area and which should use the results of the implementation of one of the types of the overall prognosis process,
- a method for implementing the process of determining the event, which may occur after some interval after completion of the selected type of prognosis process.

The analysis of various types of general prognosis in the work allows to construct models of corresponding processes that are adequate to applied problems, which use the results of calculations that are implemented by the relevant models.

6. References

1. Anagnostopoulos T., Anagnostopoulos Ch., Hadjiefthymiades S.: An adaptive location prediction model based on fuzzy control. Computer Communications. Vol. 34. Iss. 7, May 2011.

2. Ben-lin Dai, Yu-long He, Fei-hu Mu, Ning Xu, Zhen Wu: Development of a traffic noise prediction model on inland waterway of China using the FHWA. Science of The Total Environment. Vol. 482–483, June 2014.
3. Chu-Hui Lee, Yu-lung Lo, Yu-Hsiang Fu: A novel prediction model based on hierarchical characteristic of web site. Expert Systems with Applications. Vol. 38. Iss. 4, April 2011.
4. Guo-Dong Li, Shiro Masuda, Masatake Nagai: Predicting the subscribers of fixed-line and cellular phone in Japan by a novel prediction model. Technological Forecasting and Social Change. Volume 81, January 2014.
5. Leung M.T., An-Sing Chen, Hazem Daouk: Forecasting exchange rates using general regression neural networks. Computers & Operation Research. Vol. 27. Issues 11-12, September 2000.
6. Sobczyk M.: Forecasting. Theory, examples, tasks. PWN, 2008.
7. Sun H., Nguyen T., Luan Y., Jiang J.: Classified mixed logistic model prediction. Journal of Multivariate Analysis. Vol. 168. November 2018.
8. Wallach D., Thorburn P., Asseng S., Challinor A.J., Ewert F., Jones J.W., Rotter R., Ruane A.: Estimating model prediction error: Should you treat predictions as fixed or random? Environmental Modelling & Software. Vol. 84, October 2016.
9. Zhi Sun, Qinke Peng, Jia Lv, Jing Zhang: A prediction model of post subjects based on information lifecycle in forum. Information Sciences. Vol. 337–338, April 2016.