RELIABLE OPERATION OF CYBER-PHYSICAL SYSTEM WITH ACCOMPANIED BY A DIGITAL TWIN

The reliability of cyber-physical systems (CPS) operation is based on the general problem minimization of multi-factor risks, the margin of permissible risk, the forecast of the destabilizing dynamics of risk factors, principles, hypotheses, and axioms that are directly related to the analysis of abnormal situations, accidents, and disasters. To ensure the reliable operation of the CPS, a digital twin is created, which accompanies the operation of the CPS throughout its life cycle. The design of digital twins is based on simulation modelling methods that provide the most realistic representation of a physical environment or object in the virtual world.

Keywords: cyber-physical system, multi-factor risks, principles, hypotheses, axioms, digital twin.
оцінки адекватності динамічної моделі в процесі нагрівання повітря. Для адаптації динамічної моделі розроблено алгоритм пасивної ідентифікації невизначених параметрів, де мінімізуються відхилення невизначених коефіцієнтів моделі в процесі нагрівання повітря. Чисельні результати підтвердили ефективність запропонованої методики моделювання цифрового двійника електрокалорифера для кібер-фізичних систем промислових підприємств.

Ключові слова: кібер-фізична система, багатофакторні ризики, принципи, гіпотези, аксіоми, цифровий двійник.

Introduction.
Many governments have placed cyber-physical systems (CPS) on their priority list of innovations, deeming them critical to protecting their interests. Nowadays, CPS is spreading to almost all human activities, including a variety of technical systems, industrial, energy, military, transportation and many others.

The creation of modern technologies determines new requirements for ensuring the technological and environmental safety of the functioning of the CPS. The rapid development of the Internet of Things market leads, in turn, to an increase in demand for systems capable of processing large amounts of data. There is a need to search for new principles and approaches to assessing the guaranteed functioning of CPS with priority areas of safety and survivability, and, above all, complex technogenic and environmentally hazardous technical systems, based on the methodology of system diagnostics. In addition, it is necessary to consider the heterogeneity and diversity of data from different applications and devices, analyze the results obtained and make decisions in a timely manner.

It is predicted that the active technologies development of the third and fourth industrial revolutions will change not only industry, transport, and the economy as a whole, but will also radically affect human living conditions [1]. These technologies create a new environment of intellectual space, the management of which requires the development of new approaches, initiate new management problems not only of production processes, but also of all social processes, including culture and education. According to K. Schwab's forecasts, “Many of the innovations are just emerging, but soon the turning point will come when technologies begin to develop, layering and reinforcing each other, representing an interweaving of technologies from the world of physics, biology, and digital realities” [1]. It is predicted that technologies can begin to interact independently of a person: unmanned vehicles; artificial intelligence systems that will be able to make their own decisions about adjusting production processes, which will allow to quickly eliminate malfunctions, unforeseen deviations from the course of production processes, etc. Innovative technologies radically change not only the production structure and lead to a new organization of management of enterprises and organizations as a whole but will affect all spheres of human life. It is already realized that technological innovations can have both positive and negative effects, and it is necessary to make decisions on the feasibility of introducing new technologies, considering the usefulness and consequences of their introduction.
To ensure the reliable operation of the CPS, a digital twin (DT) is created, which accompanies the operation of the CPS throughout its life cycle [2]. When adopting a control strategy, DT allows for adequate displaying of the dynamics of the physical process, predicting the behavior, detecting system malfunctions, finding modifications in the structure of the physical process by observable effects, and ensuring efficient and uninterrupted operation of CPS.

1. Special properties of CPS and digital twins.

1.1 Special properties of CPS as open systems.

Considering that CPS are the driving force behind innovative transformations, many complex problems still need to be solved, for the solution of which it is advisable to consider the functioning of CPS as open systems under conditions of uncertainty. In open systems, in contrast to closed (isolated from the environment), thermodynamic laws appear that seem paradoxical and contradict the second law of thermodynamics. In accordance with this beginning, the general course of physical events in closed systems occurs in the direction of increasing entropy. At the same time, in open systems in which there is a transfer and transformation of matter, in accordance with the concept of L. von Bertalanffy "... it is quite possible to introduce negentropy", i.e., a decrease in entropy; and "... such systems can maintain their high level and even develop towards an increase in the order of complexity" [3].

One of the main fundamental characteristics of complex systems - the ability to resist entropic (system-destroying) tendencies and exhibit negentropic tendencies - in modern systems theory is attributed not to regularities, but to features. However, it can also be viewed as regularity. It is she who fundamentally distinguishes developing open systems with active elements from closed systems for which the law of physics is valid - the desire of the system to occupy a minimum energy state. According to Bertalanffy, the ability to resist entropic (system-destroying) tendencies and to exhibit negentropic tendencies is due to the openness of the system. But later it was realized that negentropic tendencies are initiated by active elements that stimulate the exchange of material, energy and information products with the environment and show their own "initiatives", an active principle. Due to this, in such systems, the regularity of the increase in entropy is violated (analogous to the second law of thermodynamics operating in closed systems, the so-called "second law") and negentropy tendencies are observed, i.e., self-organization itself, development, including "free will" [4]. On the one hand, there is an idea of entropy as chaos, a measure of the disorder of the system, and of negentropy as a measure of order. But, on the other hand, it is negentropic tendencies that are the basis for the development of any innovations, and they destabilize the system, introduce instability, such as disorder. And entropic ones, on the contrary, stabilize the state of the system, since the minimum energy state, to which entropic processes lead, is the most stable.

Thanks to the law of L. von Bertalanffy, the system exhibits the following abilities:
the ability to adapt to changing environmental conditions and interferences, both external and internal, not only in relation to interference, but also in relation to control actions;
the ability to develop options for behavior and change own structure (if necessary);
the ability to self-organize; (human);
the ability and desire for goal setting.

Open systems with active elements exhibit the following properties:
- uncertainty, non-stationarity of parameters, instability of the functioning of the system, unpredictability of behavior;
- uniqueness and unpredictability of the system’s behavior in specific conditions.

These properties are manifested in the system, due to the presence of active elements in it, as a result of which the system, as it were, manifests "free will", but at the same time there is also the presence of limiting possibilities determined by the available resources (elements, their properties) and characteristic a certain type of system structural.

1.2 Special properties of digital twin.

DT refers to a new innovative toolkit that helps exploit advanced scenarios of the Internet of Things (IoT) [5]. This toolkit is used to create digital copies of physical objects. These physical objects can be factories, power grids, transport systems, buildings, cities, and more.

One of the fundamental works on the standardization of DT development is the Industrial Internet Reference Architecture (IIRA) reference model proposed by IIC [6]. The document describes guidelines for the systems development, applications and solutions using IoT in industry and infrastructure solutions. This architecture is abstract, and provides general definitions for various stakeholders, system decomposition order, design patterns, and terms glossary. The IIRA model relate at least four stakeholder viewpoints (levels): business; usage; operation; implementation. Each level focuses on DT functional model implementation, the structure, interfaces and interactions between the DT components, and DT model’s system interaction with external elements of the environment to support CPS functioning. The DT technology includes (but is not limited) combinations of the program object: physical model and data; analytical model and data; temporal variable archives; transactional data; master data; and visual models and calculations. DT creation concept has a multifaceted architecture and correspondingly complex mathematical support for implementation.

The mathematical description of DT can be obtained using statistical modeling, machine learning, or analytical modeling techniques. Methods of statistical modeling can be divided into three groups [7]: regression analysis models; classification models; anomalies detection models. The method choice depends on the size, quality and data nature, as well as on the problems type and the process knowledge being modeled. For technological processes in CPS, analytical models are often used, which have valuable properties in engineering [8]. In [9] the CPS development using deterministic models, which have proven to be extremely useful, is discussed. Deterministic mathematical models of CPS are based on differential equations and
include synchronous digital logic and single-threaded imperative programs. However, CPS combines these models in such a way that determinism is not preserved.

The practice of using analytical models to describe the functioning of technological processes indicates that ready-to-use models are extremely rare, because such models are developed under conditions of conceptual uncertainty that need to be disclosed for a particular physical process. Conceptual uncertainty arises from the knowledge incompleteness of the physical environment, process or system. Conceptual uncertainty is complex [10], conceptual uncertainty examples are an uncertainties combination: objectives; operation of process; structure modeled system; system elements interaction, or interaction with the external environment, and others. For analytical models, the above uncertainties are complicated by information uncertainty, which is caused by: methodological uncertainty (complex processes are linearized when modeling); measurement distortion (due to inaccuracy and inertia of sensors and the presence of disturbance)

2. **The reliable operation of Cyber-physical Systems with Guaranteed Survivability and Safety.**

The reliable operation of CPS is based on the general problem minimization of multi-factor risks, the margin of permissible risk, forecast of the destabilizing dynamics of risk factors, principles, hypotheses, and axioms that are directly related to the analysis of abnormal situations, accidents, and disasters. The key idea of the strategy is based on the main principle: to provide timely and reliable detection and estimation of risk factors, prediction of their development during a certain period of operation, and timely identification and elimination of the causes of abnormal situations before failures and other undesirable consequences occur and prevention of the transition from normal to an abnormal mode [11]. The fundamentally important peculiarities are the following: sets of risk factors and sets of situations are largely unlimited; a set of risk situations is in principle not a complete group of random events; a threshold restriction of time for decision forming is a top priority; the problem is not completely formalized; indicators of a multifactor risk estimation are not determined; criteria of a multipurpose risk minimization are not determined. The communication with computational systems and different types of sensors is implemented online in real-time. Joint actions of CPS components determine the properties and special features of the mode of functioning of a complex system at any moment of time.

The software for guaranteed survivability and safety of the functioning of CPS is implemented as an informational platform for technical diagnostics [12]. This tool allows timely and reliably determine, evaluate and forecast risk factors and based on that taking into account the feedback timely identify causes of emergency situations before failures can occur and other undesirable consequences.

The main idea of the strategy is to ensure, in real conditions of the functioning of a complex system, the timely and reliable detection, assessment of risk factors, forecasting of their development over a given time of operation, and on this basis, the timely elimination of the causes of emergency situations before failures and other undesirable consequences [13].
The diagnostic unit, constituting the basis of the algorithm for managing the security of cyber-physical systems in abnormal situations is implemented as an information platform, containing the following modules [14]:

- Receiving and processing of initial information in the process of functioning CPS.
- Restoring functional dependencies (FD) and identifying patterns based on empirical discretely specified samples.
- Quantizing the original variables.
- Forecasting of non-stationary processes.
- Reliability of information transmitted from sensors.
- Construction of the process of technical diagnostics.

Let us consider in more detail the just mentioned 5 module of the information platform for technical diagnostics (IPTD).

Evaluation of the reliability of the information transmitted from the sensors. To obtain information, the CPS uses sensors that continuously provide computing systems with data received from the environment. Both wired and wireless data transmission methods can be used to exchange data from the CPS sensors. On the basis of the obtained data, computing systems can control the physical elements of the system, or supporting their process of functioning. One of the important issues in this interconnection is the reliability of the transmitted information from various sensors and transmitting devices. One of the types of unreliability of the transmitted information during the functioning of the CPS is the failure or malfunction of the sensors. The fundamental complexity of this problem lies in the fact that a priori it is difficult to identify this failure directly without cross-checking sensors, the installation of which is usually unprofitable. Moreover, each investigated CPS has its own peculiarities of recording critical parameters and their tracking.

Evaluation of the reliability of the information transmitted from the sensors can be implemented using various techniques, in particular, the Chauvenet’s criterion is used. Based on this criterion on the test sample, as well as at each later step when receiving a new element from the set of incoming indicators, for each coordinate, estimates of the expected value and variance are calculated separately from the set of all previous values of these indicators. Then these estimates are substituted into the inverse distribution function of the Gaussian random variable and the probability is determined for each coordinate. Then multiplication is performed by the size of the sample considered up to the current moment, and if the obtained value is greater than 0.5 for at least one coordinate of the indicators by which the sensor is monitored, then it is considered that the transmitted information from the corresponding sensor is incorrect.

Detection of random sensor failures can be based on the construction of Bollinger bands [15] and step functions of the 1st and 2nd level. The use of both step functions and Bollinger bands, which characterize a technical analysis tool and a technical indicator reflecting the current deviations of the observed value, provides a decrease in the level of dependence on the error of the measured indicators.
Bollinger Bands are plotted as upper (moving average plus 2 standard deviations) and lower (moving average minus 2 standard deviations) boundaries around the moving average, but the band width is proportional to the standard deviation from the moving average over the analyzed time period. The moving average period can be selected arbitrarily, but it should be noted that: the longer the period of the moving average, the less sensitive it will be to changes in the observed value; a moving average with a very small period will generate a large number of false signals; a moving average with a very long period will be constantly late.

Taking into account these factors, by empirical considerations, the period of the moving average can be equal to 10. Thus, for 10 dimensions of a given sample, a moving exponentially smoothed average and average deviation are calculated for this interval. The upper and lower Bollinger bands are formed: an exponentially smoothed mean plus two standard deviations and an exponentially smoothed mean minus two standard deviations.

A genetic algorithm can be used to parameterize the model. The procedure for detecting possible sensor malfunctions is based on the following considerations. If the sensor functions are normal, each of its readings does not go beyond the threshold level. Any indication can be confirmed by the previous and subsequent values. First, this is due to the nature of the monitored processes: most changes in the status of a process do not occur instantly. Therefore, an abrupt change in sensor readings can be taken as evidence of failure of the measuring instruments. This approach is implemented as follows. At each step, the arithmetic average of the previous and subsequent measurements is calculated. Then this value is compared with the current value. If the deviation exceeds the threshold level, the operator displays a message about a possible sensor malfunction. Sensor failure can also be tracked by comparing predicted and actual measurements. Since the prediction follows the general behavior of the system based on the latest measurements, a deviation in the actual value could indicate a sensor malfunction. Therefore, the system performs a regular comparison of forecasts and their respective recovered values. As in the previous case, a deviation exceeding the threshold level gives a message about a possible sensor malfunction.

There are general problems with the operation of sensors, that is, deviations of the recorded values from the true ones. An exponential smoothing method can be used to smooth out these deviations. This method is well suited for working with dynamically changing quantities, since the most recent measurements have the greatest weight. Thus, if at some stage there is a significant deviation of the value of a certain parameter from the previous one, and at the next step the value of this parameter returns to its previous level, then, most likely, the sensor has failed and the exponential smoothing method will ensure the elimination of this obstacle. If at some stage there really was a leap, then this will be reflected in all subsequent measurements and the model will quickly pass into a new dynamic state.
3. Approach to developing a digital twin of CPS under conceptual uncertainty.

Generalized structural scheme of DT development procedure based on analytical model of physical process is considered. This takes into account the parametric uncertainty of the physical process mathematical description.

In the first stage of development, it is necessary to conduct a literature analysis in the applied field of research for the physical process. This will help to determine the model structure, existing advantages and disadvantages. As a rule, the analytical model of the studied process has the system form of differential, difference or algebraic equations. The practice of using analytical models shows that ready-to-use models are very rare. Even tested models require adjustment of parameters in order to adapt them to specific conditions of use. Thus, when DT is developed for a particular physical process, the researcher needs to determine the uncertainty "physical limits of that process" in the numerical values form of mathematical model parameters. To do this, the researcher needs to perform passive identification of the mathematical model parameters. A very important role at this stage is played by the data quality for the model identification, so the formation of the database should be guided by the known requirements of informativeness, synchrony and correctness.

The last step in DT model development is the identified model discretization. Here it is necessary to set correctly the sampling time for the mathematical model. On the one hand, the sampling time should not be small in order to ensure the information distribution over the CPS network. On the other hand, a large sampling time will lead to the loss of intermediate information for short-term forecasts. The obtained numerical model, even a sufficiently adequacy high degree, does not yet guarantee a prediction estimates high quality if the basic uncertainties for the mathematical model of the physical process are not taken into account. Therefore, after designing DT, it is necessary to check the possibility of using it for solving the assigned forecasting tasks.

An important characteristic for DT is to determine the received prediction quality. Often, the quality of forecast estimates is determined with the help of LSM. However, LSM is one of many possible statistics that depends on the data scale. Therefore, only this characteristic is not enough for the analysis of a qualitative prediction. The quality of linear and pseudo-linear models is assessed using several statistical quality criteria [16], since each criterion has its own specific purpose and characterizes one property of a prediction evaluation. Therefore, DT developer must comprehensively study a physical process, existing mathematical models, possible perturbing effects on a physical process and justify a use of adequacy criteria for DT in conditions of conceptual uncertainty.

In the proposed approach, the DT development will be carried out on the basis of an analytical model in the state space. In order to obtain adequate computational data at the first stage of DT synthesis, it is necessary to identify the continuous model, using the passive identification algorithm [17]. Also, it is highly desirable to reduce the computational resources of DT. It is known from modeling theory
that computational resources for simulation differential equations are more than for their discrete analogs based on difference equations. Therefore, let us consider a discrete representation of continuous model.

The continuous mathematical model can be represented in a discrete form [18]

$$\bar{X}_{k+1} = \bar{A}_d \bar{X}_k + \bar{B}_d U_k,$$

(1)

here $\bar{A}_d = e^{X_T V_k}$, $\bar{B}_d = \int_0^{T_k} e^{X(T_k - t)} \bar{B} \, dt$, $T_k$ is sampling period.

Thus, DT synthesis methodology for the electric heater consists of steps:
1) the uncertain parameters identification ($\alpha$ and $G_A$) of mathematical model by the considered algorithm;
2) transition from the continuous model in the state space to the discrete model (1), which is DT;
3) if during operation the DT accuracy has deteriorated (due to non-stationarity of the physical process) then go to step 1 to identify the parameters of the model.

4. Results of the digital twin simulation.

The proposed methodology was used to develop and simulate DT of an electric heater using the MatLAB software package. Let’s consider the example of DT simulation using model in state space [18]:

$$X' = AX + BU$$

(2)

$$X = \begin{bmatrix} \Delta \theta_A \\ \Delta d_A \\ \Delta \theta_E \end{bmatrix}, \quad A = \begin{bmatrix} -1/T_A & 0 & k_2/T_A \\ 0 & -1/T_d & 0 \\ k_i/T_E & 0 & -1/T_E \end{bmatrix}, \quad B = \begin{bmatrix} k_3/T_A & 0 & k_4/T_A \\ 0 & k_5/T_d & k_6/T_d \\ 0 & 0 & 0 \end{bmatrix}, \quad U = \begin{bmatrix} \Delta \theta_{A0} \\ \Delta d_{A0} \\ \Delta G_A \\ \Delta N_E \end{bmatrix},$$

MatLAB was used to calculate the matrices $\bar{A}_d$ and $\bar{B}_d$ of DT (1). Simulation results are shown in Fig. 1. Fig. 1 (a) simulates the case with input influence

$$U(t) = \begin{bmatrix} 1 + 0.5 \sin(0.15t) \\ 1 + 0.2 \sin(0.15t) \\ -0.2 + 0.1 \sin(0.3t) \\ 0.5 + 0.2 \sin(0.05t) \end{bmatrix}^T,$$

and initial conditions for the physical model $X(0) = [2 \ 1 \ 15]^T$, $\alpha_0 = 161$, $G_A = 0.43$ and the identifiable model $\bar{X}(0) = [0 \ 0 \ 0]^T$, $\bar{\alpha}_0 = 250$, $\bar{G}_A = 0.8$. Fig. 1 (b) shows the surface isolines of criterion

$$I = M \left\{ \int_{t_0}^{t_f} (X - \bar{X})^T Q (X - \bar{X}) \, dt \right\} \rightarrow \min,$$

(3)

and the minimization trajectory of its parameters $\bar{\alpha}_0$, $\bar{G}_A$, which resulted in finding $\bar{\alpha}_0 = 160.3$, $\bar{G}_A = 0.41$. After identification of model (2) using MatLAB function c2d(...) numerical values of DT model matrices (1) are calculated for sampling period $T_{KV} = 2$:
Fig. 1 (c) shows the time characteristics of the state variables for the reference model \( X \) and the DT \( \hat{X}_k \) of the electric heater.

\[
\hat{A}_d = \begin{bmatrix}
0.0133 & 0 & 0.0829 \\
0 & 0.0125 & 0 \\
0.0552 & 0 & 0.7238
\end{bmatrix}; \quad \hat{B}_d = \begin{bmatrix}
0.9037 & 0 & -9.0374 & 0.547 \\
0 & 0.9875 & 0 & 0 \\
0.221 & 0 & -2.2097 & 6.1755
\end{bmatrix}.
\]

**Conclusions.** Technological development trends are increasingly faced with the creation of CPS for various innovative solutions. Such solutions must be able to handle a significant amount of data in an open systems environment. Open systems stand out for their ability to self-organize, while CPS, to the extent of its features, should be considered exactly as open systems. The article proposes an approach to building a CPS architecture with such properties, which includes an analytical level for the implementation of methods for assessing the guaranteed functioning of a CPS. Also, it allows you to take into account the heterogeneity and diversity of data from various applications and devices, analyze the results obtained and make decisions in a timely manner. The results of the analysis make it possible to judge that it is the approach to CPS as to an open system that will solve many complex existing problems in the evolution of the Internet of Things.
Digitalization of physical environments and processes implies the use of complex mathematical models with uncertain parameters, which need to be corrected in order to self-adapt to specific application conditions. As an example, the passive identification technique of the electric heater analytical model with subsequent synthesis of a digital twin for the CPS is considered. Based on system analysis methodology, the results obtained are summarized and an approach to the digital twin development using the analytical model under conditions of conceptual uncertainties is proposed.

The peculiarity of the proposed approach is the several key parameters identification of the analytical model, which are refined in the passive identification process. The use of a physical process analytical model makes it possible to abandon the search for all its parameters. It is known that in modern methods of system analysis the choice of structure and type of model plays an important role in further research and may require a lot of time and additional information for building an adequate model. In the proposed approach, the structure of the analytical model is known, for which only the key uncertain parameters are identified from the measured variables of the real physical process.

Example of parameter identification for the model (1) in the state space are given. It is shown that the uncertain parameters identification of the model in the state space belongs to the problem of single-extremal optimization. The procedure for synthesizing the model of the electric heater digital twin is proposed and numerically investigated. The simulation results confirmed the effectiveness of the proposed procedure for creating a digital twin using an analytical model.

The use of DT in CPS makes it possible to identify bottlenecks in technological processes, improve product quality, and reduce the risks of abnormal operation throughout the life cycle of equipment. DT are used for prediction of equipment operation modes and self-diagnostics, as well as optimization of the physical system structure. This approach provides a high-precision assessment of the plant's production capacity when drawing up the production.

References


Received 15.09.2023.