Probability and Information Entropy in the Analysis of Reliability of Technical Systems

A. S. Dulesov[1] [1], D. Y. Karandeev[1] [2], A. A. Konovalov[1] [3], D. B. Bayyr[1] [4], N. V. Dulesova[2] [5]
[1]*Katanov Khakass State University, 92, Lenina ave., Abakan, 655017, Russia*
[2]*Khakas Technical Institute of Siberian Federal University, 15, Komarova ave., Abakan, 655017, Russia*

Keywords: probability measure, information entropy, reliability analysis.

Abstract: The paper outlines the issues of reliability analysis of complex technical systems. The justification of the need to perform reliability analysis of complex technical systems, the structure of which is a network, is given. We propose to apply for this purpose, in addition to probabilistic characteristics, information entropy calculated based on the K. Shannon model. A classical assessment of a random event in reliability analysis as a possible outcome (test result) is given. Following the theories of probability and information, the classical measures of probability and their relationship to the measure of uncertainty of information (entropy) are considered. The role of the private entropy obtained as a result of the tests and its participation in the classical Shannon model are indicated. A distinctive feature of this model is the differentiation of entropy into components related to two incompatible events with the probability of failure-free operation and the probability of failure. This avoids mixing entropies in the additive event accounting process. According to the classical theorems of addition and multiplication of probabilities, as well as consideration of compatible, incompatible and overlapping events, mathematical expressions of entropy are presented. They allow us to assess the level of information uncertainty in the task of reliability analysis for subsequent comparison and selection of system structures.

1 INTRODUCTION

The development of information technology provides for the implementation of analysis not only on the basis of the application of information measures, but also on the possibility of using information theory. In it, the key place is occupied by the issues of measuring information and its uncertainty when considering problems about the state and choice of systems.

Any technical system, especially a complex one (Bar-Yam, 2002; Kudzh, 2014), involves the creation of a model that, in the modeling process, due to the presence of uncertainty, doesn't fully reflect the variety of relationships between the elements of the system and the environment.

A complex technical system is classified to as an information technology system endowed, in particular, with the property of reliability. This property involves the study of approaches to the analysis of the relationship between subsystems (elements), the violation and restoration of which changes the behavior and state of the entire system.

At the same time, it is important to understand how the system interacts with the external environment (Monakhov, Savinykh, Tsvetkov, 2005). Therefore, it should be considered as open, for which interaction is considered through information and energy.

In the process of studying the state and behavior of complex systems, structural analysis (Butko, 2017) and structural modeling (Tsvetkov, 2017) are in demand. Taking into account the system analysis,

[1] https://orcid.org/0000-0001-6371-0171
[2] https://orcid.org/0000-0001-6202-5228
[3] https://orcid.org/0009-0001-2685-1503
[4] https://orcid.org/0009-0000-6904-7286
[5] https://orcid.org/0000-0001-8445-2095

structural reliability analysis is an integral part of it. These studies relate to the rejection of a number of system characteristics and the transition to the implementation of selective analysis. These studies concern the abandonment of a number of system characteristics and the transition to selective analysis. Thus, the system has to be replaced by an idealized theoretical abstract object (Tsvetkov, 2017). In fact, such a concept as "complexity" doesn't provide for consideration of all the properties of the system, so we propose a way to simplify, understanding that one should not expect to get a one hundred percent result.

Reliability analysis concerns not only a complex technical system, but also a complex network. This network has a simpler structure of connections between the elements. Each element is endowed with the function of converting and transferring energy, as well as resources. The network structurally characterizes the state of the object. The transition of an object from one state to another depends on the external random influence of nature and purposeful human control.

Since the network is a simpler object in comparison with a complex system, it is therefore often depicted as a graph (Borisenko, Lakhno, Chepovsky, 2010). Despite the fact that the representation of a complex system and a complex network are not equivalent, they have common features. These include elements or blocks that have both external and internal connections. Considering the network structure as a graph, each branch of it is an element attached to some vertex. The whole set of formed vertices of the graph is characterized by a high density of internal connections and a low density of external connections, which is typical for closed systems. However, the graph representation of a network has the property of grouping networks. This allows you to move from a simple nodal representation of the network to a complex one, that is, in the form of blocks. Each of the blocks is an association of elements according to functional characteristics, representing small physical systems with a common input and output, which facilitates the analysis of structural reliability. Therefore, by presenting the network as a graph, the task of analyzing reliability is facilitated without losing the properties of the system.

2 THE CLASSICAL DETERMINATION OF PROBABILITY AND ENTROPY

Let's further consider the relationship between the probability of events and information entropy in the analysis of the reliability of technical systems. In this process, we will take into account the author's scientific results presented in (Dulesov, Khrustalev, Dulesova, 2016; Dulesov, Eremeeva, Karandeev, Dulesova, 2018; Dulesov, Karandeev, Eremeeva, Khrustalev, Dulesova, 2019).

Testing (experience, experiment) is distinguished among the basic concepts from the theory of probability and reliability. It refers to the fulfillment of a number of conditions to determine the achieved level of reliability. Tests, for example, include control of equipment operating modes, exposure to loads on the test object, etc. Testing can also be attributed to the operation of the facility, where operating conditions, unintended impacts, environmental influences, etc. are considered as loads.

A random event (a possible event or just an event) is understood as any fact or factor that may or may not occur during testing. For example, during laboratory testing, equipment may break down, and during operation, various kinds of failures and accidents may occur. A random event in the reliability analysis is considered as a possible outcome, the result of the test. Considering the state of the system, the issues of randomness are related to the need to improve process control systems aimed at consolidating efforts to combat the elements (nature, unintended and intentional (dangerous) human actions, etc.).

From the point of view of identifying cause-and-effect relationships between objects, event A occurs because the cause is the appearance of the preceding event B. These events can be considered incompatible, the occurrence of one of them excludes the appearance of the other in the same test. For example, if an object fails during the testing process, then it cannot be considered as working. Events A and B are considered joint if they can occur together in the same trial. Two incompatible (opposite) events – when one of them must necessarily happen. The event opposite to event A is denoted as $\overline{A}$.

Let's start considering the classical probability and the measure of information uncertainty. At the same time, we will follow the theories of probability and information.

In the practice of reliability analysis, it is important to be able to compare events according to the degree of possibility of their occurrence. For example, during the testing process, the technical system should be much more efficient than in a state of failure. At the same time, it should be characterized by a certain level of information. Therefore, to compare events, informed decisions on the application of probability and information measures are required. The probability of a random event is a numerical measure of the degree of objective possibility of an event occurring. In addition, information entropy is taken into account – measure of the uncertainty of the system, measuring the degree of unpredictability of the occurrence of a probabilistic event, as well as a measure of information, meaning (according to the Shannon approach) the degree of knowledge about the possible state of the object. In addition to the above, we highlight the concept of a state, which denotes a set of stable values of variable parameters that characterize an object from the occurrence of an event to its transition to another event. The state is characterized by describing the variable properties of an object.

Most information technology systems are endowed with statistical data on their states. Then it becomes possible to determine the probability of an event and the amount of information in the analysis process. To do this, you can use the relative frequencies (frequencies) of event *A*:

$$W(A) = M / N \qquad (1)$$

where *N* is the total number of experiments, *M* is the number of experiments in which event *A*.

In this case, the amount of information obtained is calculated in bits as:

$$h = -\log 2\ W(A) \qquad (2)$$

The value of *h* is called partial entropy. For example, it is a measure of the uncertainty of random events that occur during tests (experiments) conducted in factory laboratories. Since the number of experiments is quite large and if the experiments are performed under the same conditions, (then according to the conditions of Lyapunov's theorem on the normality of the distribution of random variables with an unlimited increase in their number). We can testify to the following: when performing tests, the relative frequency W(A)) changes little, fluctuating around a certain constant number $p^*(A)$. This fulfills the condition: $p^*(A)=W(A)$. Therefore, $p^*(A)$ can be considered a statistical probability of the event in question.

Under these conditions, the application of the information entropy of K. Shannon (11), calculated by the expression:

$$H(X) = -\sum_i^m p_i(x) \log p_i(x) \qquad (3)$$

subject to consideration of independent events:

$$\sum_{i=1}^m p_i(x) = 1 \qquad (4)$$

where the base of the logarithm is arbitrary.

The mathematical expression of Shannon (3) is valid both for calculating information I and for entropy H. Entropy is endowed with some properties, which can be found in (12). One of them is: entropy is additive, that is, the total entropy consists of the entropies of events i:

$$H(X) = H(X_1) + H(X_2) + \ldots + H(X_i) + \ldots + H(X_n) \qquad (5)$$

The application of (3) in reliability analysis has its own specifics. Complex technical systems have a high (close to one) probability of trouble-free operation. It means that the failure of the object doesn’t occur within the specified operating time or a specified time interval. The value of $p^*(A)$ is the probability of failure-free operation obtained as a result of testing and it, together with the failure rate, characterizes the reliability of the object. In turn, the private entropy, measured in bits, according to (2):

$$h = -\log_2 p^*(A) \qquad (6)$$

It will be close to zero, meaning that entropy as a factor of information uncertainty means the following: during the testing process, the information received (entropy) indicates a small number of events leading the test object to failures. The search for undesirable events (leading to failures) doesn’t require a lot of additional information and time.

Considering the events as incompatible, from the point of view of reliability, the probability of failure-free operation p(t) and the probability of failure q(t) are distinguished (in most situations, coinciding with the probability of restoring an operable state). Then the function (3) will have the following form:

$$H = H(p) + H(q) = -\sum_i^n p_i(t) \log p_i - \sum_j^m q_j(t) \log q_j \qquad (7)$$

subject to consideration of independent events:

$$\sum_{i=1}^n p_i(t) + \sum_{j=1}^m q_j(t) = 1 \qquad (8)$$

Here n is the number of events related to the probability of failure-free operation $p_i(t)$, m is the probability of failure $q_j(t)$. Having proposed these expressions, we note the following: they are valid for reliability analysis, since random events relate to two qualitatively different indicators of object reliability (operation and failure). Expressions (3) and (4) will be valid without taking into account the qualitative separation of events.

Let's turn to the consideration of geometric probability and entropy. Since the number of outcomes in the analysis of the reliability of systems is limited, the geometric probability and entropy can be interpreted as follows.

Let's highlight the time interval for the entire period of operation (testing) of the object – the MN segment. This segment has its own length and is characteristic of only one state of the object. Suppose that a segment CD ≤ MN randomly falls on this segment, then this will be an event that covers part of the original segment with its length and its own state. At the same time, we don’t take into account which specific section of MN the event will appear on. Then the probability of occurrence of event A on any part of the segment MN can be calculated using the formula:

$$p(A) = \frac{t_{CD}}{T_{MN}} \quad (9)$$

where tCD is the length of the segment CD related to a random event, TMN is the length of the segment MN.

Information entropy, as a measure of uncertainty (unpredictability) of the occurrence of event A on the considered segment MN:

$$h(A) = \log p\,(A) \quad (10)$$

From expression (10) it can be seen that with increasing length of the segment CD, followed by coincidence with MN, $\lim_{p(A)\to 1} h\,(A) = 0$. The partial entropy of h(A)=0 means that there is no uncertainty when the event fully covers the segment MN. For very small segments CD, the partial entropy has a large amount of uncertainty and it becomes difficult to determine the location of the CD segment on MN.

By analogy, the task can be completed when an area g ≤ G falls on flat area G, then the probability of such an event A is determined by the expression:

$$p(A) = \frac{s_g}{S_G} \quad (11)$$

where s is the area of part g of the region G, and S is the area of the entire region G.

Information entropy, as a measure of uncertainty (unpredictability), is characterized by the event A about the presence of area s on area S. Calculation of h(A) is performed by expression (10).

Example. Let there be a piece on the chessboard. We don’t know which platform (cell) s it is located on. It is necessary to find a figure. The entire area of the board S consists of 64 platforms s. The probability that the figure is on the square (according to (11)) will be 1/64. By expression (10) information entropy h(A)=log2(1/64) = 6. After completing the action: dividing the board in half, the probability of finding a piece on the remaining half will be 1/32, h(A)=5. This action allowed us to remove the uncertainty of the shape search by 1 bit. Each further division in half allows you to remove the uncertainty by 1 bit until the search is completed.

3 A SET-THEORETIC INTERPRETATION OF THE BASIC CONCEPTS OF ENTROPY

Monitoring of tests (equipment failures during operation of a technical system), as a factor in the use of digital technologies, includes consideration of the set of all mutually exclusive outcomes, that is, the space of elementary events related to the elements of the system.

Let’s introduce operations on events that are equivalent to operations on the corresponding sets. We will consider the events from the perspective of reliability analysis, taking into account the work (Dulesov, Kondrat, 2014; Dulesov, Kondrat, 2014; Karandeev, Dulesov, Bazhenov, Karandeeva, Dulesova, 2022).

Since the probabilities of events are directly related to entropy, indirect methods are used to determine it, not direct methods (represented, for example, by expressions (3) and (7)). They allow us to determine the probabilities and entropy of other events related to them based on the already known probabilities and entropy of some events. At the same time, the main theorems of probability theory are taken into account. The mathematical expressions presented below are consistent with the probability theorems and mathematical calculations of A. Khinchin (16), as well as the geometric interpretation of the main actions on events using Venn diagrams.

The product of two events A and B (denoted AB or $A \bigcap B$) is an event consisting of those outcomes that are included in both A and B. Such an event AB provides for the simultaneous occurrence of events A and B. As an example, we can note the simultaneous presence of operable and non-operable states and their combinations only for two or more elements of the system.

According to the property of independence of events and the multiplication theorem, the calculation of probability has the form:

$$p(AB)=p(A)\cdot p(B) \quad (12)$$

The product of two events is the result of obtaining a joint entropy, the average amount of information for a pair of joint events A and B:

$$H(AB) = p(B)\cdot H(A) + p(A)\cdot H(B) \quad (13)$$

where H(A) = p(A)·log p(A) и H(B) = p(B)·log p(B) – the entropy of events A and B.

The sum of two events A and B (denoted A+B or $A \bigcup B$) is an event consisting of all outcomes: a operable and non-operable state, included in either A or B. In other words, A+B is understood as the following event: either event A or event B. The appearance of these events simultaneously (operable and non-operable) is not possible for the system element in question.

The probability of two incompatible events is:

$$p(A+B) = p(A) + p(B) \quad (14)$$

and the entropy for two incompatible events is calculated by the formula:

$$H(A+B) = H(A) + H(B) = -p(A) \log p(A) - p(B) \log p(B) \quad (15)$$

According to the probability addition theorem, the probability p(A+B) of the sum of events A and B for two elements is:

$$H(A+B) = H(A) + H(B) - H(AB) = p(B)\cdot H(A) + p(A)\cdot H(B) \quad (16)$$

The sum of the probabilities of opposite events is 1:

$$p(A) + q(A) = 1 \quad (17)$$

and the entropy of these events is calculated by the formula (7).

The difference between two events A and B (denoted A–B or A\B) is an event consisting of the outcomes included in A, but not included in B. The meaning of this difference of events is that event A occurs, but event B doesn't occur. In this case, we mean the conditional probability p(B\A) = pA (B) of event B, provided that event A has occurred. Similarly, p(A\B) = pB (A) can be written.

Accordingly, conditional entropy is the amount of information about event A, provided that its own information is subtracted from it in the event of an event with probability p(B):

$$H(A\backslash B) = H(A) - p(B)\cdot H(A) \quad (18)$$

$$H(B\backslash A) = H(B) - p(A)\cdot H(B) \quad (19)$$

Based on the multiplication theorem, the probability of the product of two events is equal to the product of the probability of event A by the conditional probability B, provided that the first event occurred:

$$p(AB) = p(A) \cdot p_A(B) \quad (20)$$

and vice versa:

$$p(BA) = p(B) \cdot p_B(A) \quad (21)$$

Then the entropy of the joint events A and B can be calculated using the formulas:

$$H(AB) = p(A)\cdot H(B) \quad (22)$$

$$H(BA) = p(B)\cdot H(A) \quad (23)$$

Expressions (22) and (23) in total correspond to expression (13).

The presented set-theoretic interpretation of the basic concepts of entropy is limited, since statistical entropy concerns the consideration of discrete and continuous random variables and their numerical characteristics.

4 CONCLUSION

Considering the role of probability and entropy measures, the possibilities of their application in the task of monitoring and analyzing reliability are indicated by the needs for the use of digital technologies for the development of technical systems. Calculating not only probability according to statistical data, but also entropy is the result of choosing network structures that meet high reliability requirements. The presented theoretical calculations and mathematical expressions relate not only to the implementation of the initial stage of the solution search, but also to the reference point in the application of statistical estimates.

REFERENCES

Bar-Yam, Y., 2002. *General Features of Complex Systems. Encyclopedia of Life Support Systems*. London, UK: EOLSS UNESCO Publishers, Oxford.

Kudzh, S.A., 2014. The multidimensional nature of the consideration of complex systems. In *Prospects of science and education*, 1, 38-43.

Monakhov, S.V., Savinykh, V.P., Tsvetkov, V.Ya., 2005. *Methodology of analysis and design of complex information systems*. Moscow, Russia: Prosveshchenie.

Butko, E.Ya., 2017. A systematic approach in the formation of the structure. In *Slavic Forum*, 2(16), 25 - 31.

Tsvetkov, V.Ya., 2017. *Structural modeling: monograph*. Moscow, Russia: MAKS Press.

Tsvetkov, V.Ya., 2017. Complex technical systems. In *Educational resources and technologies*, 3(20), 86 - 92.

Borisenko, V.V., Lakhno, A.P., Chepovsky, A.M., 2010. Special representation of graphs and visualization of semantic networks. In *Fundamental and Applied Mathematics*, 8(16), 27 – 35.

Dulesov, A.S., Khrustalev, V.I., Dulesova, N.V., 2016. Quantification of information in the closed structure of a technical system. In *2nd International Conference on Industrial Engineering, Applications and Manufacturing*. ICIEAM.

Dulesov, A.S., Eremeeva, O.S., Karandeev, D.Yu., Dulesova, N.V., 2018. Approaches to information measurement of the structure state of technical systems. In *International Multi-Conference on Industrial Engineering and Modern Technologies*. FarEastCon.

Dulesov, A.S., Karandeev, D.J., Eremeeva, O.S., Khrustalev, V.I., Dulesova, N.V., 2019. The logarithmic basis to measure the amount of information related to the assessment of reliability of elements of the technical system. In *IOP Conference Series: Materials Science and Engineering. International Workshop "Advanced Technologies in Material Science, Mechanical and Automation Engineering – MIP: Engineering – 2019"*. Krasnoyarsk Science and Technology City Hall of the Russian Union of Scientific and Engineering Associations.

Shannon, C.E., 1949. Communication Theory of Secrecy. In *Systems Bell System Tech. Journal*. 28(4), 656 – 715.

Prokhorov, Y.V., 2022. Entropy in information theory. In *BigRussianEncyclopedia*.

Dulesov, A.S., Kondrat, N.N., 2014 The measure of uncertainty of information and its properties in relation to the assessment of the random behavior of a technical object. In *Journal "Scientific Review"*, 7(1), 258 - 264.

Dulesov, A.S., Kondrat, N.N., 2014. The amount of information at the overlap and intersection of elementary events. In *Journal "Scientific Review»,* 12(1), 146 - 150.

Karandeev, D.J., Dulesov, A.S., Bazhenov, R.I., Karandeeva, I.J., Dulesova, N.V., 2022. Entropy determination in the problem of the analyzing an electrical network state under the condition of independence of functioning of its elements. In *Journal of Physics: Conference Series*. IV International Conference on Applied Physics, Information Technologies and Engineering.

Khinchin, A.Ya., 1953. The concept of entropy in probability theory. In *Successes of mathematical sciences*, 8:3(55), 3 – 20.