Mathematical Fundamentals of Structural And Entropic Analysis of Digital Data Flows

Nataliia Vozna[a], Andriy Segin[a], Ihor Pitukh[a], Artur Voronych[b], Lyubov Nykolaychuk[c]

[a] *West Ukrainian National University, 11, Lvivska Str.,Ternopil, 46009, Ukraine*
[b] *Ivano-Frankivsk National Technical University of Oil and Gas, 15, Karpatska Str., Ivano-Frankivsk, 76019, Ukraine*
[c] *Nadvirna Vocational College by National Transport University, 177, Soborna str., Nadvirna, 78400, Ukraine*

Abstract

Areas of entropy application and structural analysis for solving a wide range of information problems in the field of states monitoring for control objects are identified. Mathematical bases of existing algorithms for entropy estimation of stationary random processes are presented. Criteria of structural complexity are systematized for microelectronic tools, which allow to compare the system characteristics of different structures for operating devices and specialized processors in the computer architecture. The most priority modern architectures of interactive CPS i nterms of the emergence and parallelization coefficient for data flows are defined. The principle of data encryption based on the entropy method of signals manipulation method is proposed. On its basis, the priority structures of crypto protection of data are offered. These structures are used for the reception and decoding of crypto-protected entropy-manipulated signals. The proposed structures are characterized by the limit characteristics of maximum speed and minimum time and structural complexity.

Keywords

Entropy, structures, Specialized processor for entropy estimation, cryptographic protection, entropy-manipulated signals**.**

1. Introduction

In modern cyberphysical systems, the volume of digital data streams is growing significantly and algorithms for their processing are being developed. One of the effective ways of data processing for a wide range, such as digital data research, encoding and encrypting data, transmitting information, etc. became entropic analysis. Hartley and Shannon formulas are most often used to estimate the entropy of digitized processes [1-2]. However, entropy analysis needs further development in terms of improving the theoretical foundations, practical implementation algorithms and specialized processors for their calculation [3-5]. In addition, it is necessary to improve the criteria for determining the complexity of cyberphysical systems using certain algorithms and technical means.

2. Justification of the relevance of entropy and structural analysis

There are a number of approaches and algorithms for entropy characteristics estimation of data flows. Based on them, appropriate specialized processors have been developed to calculate entropy estimates [6].

IntelITSIS'2022: 3rd International Workshop on Intelligent Information Technologies and Systems of Information Security, March 23–25, 2022, Khmelnytskyi, Ukraine
EMAIL: nvozna@ukr.net (N. Vozna); andriy.segin@gmail.com (A. Segin); pirom75@ukr.net (I. Pitukh); archy.bear@gmail.com (A. Voronych); lmnykolaychuk@gmail.com (L. Nykolaychuk)
ORCID: 0000-0002-8856-1720 (N. Vozna); 0000-0002-3556-248X (A. Segin); 0000-0002-3329-4901 (I. Pitukh); 0000-0003-0701-917X (A. Voronych); 0000-0002-7733-4573 (L. Nykolaychuk)

These specialized processors have different parameters of hardware complexity of their implementation, the time of calculating the final result, the accuracy of entropy estimation and others.

1. The method of estimating entropy using a centred autocorrelation function, takes into account statistical relationships between data. It is described by the expression as follows:

$$I_x(R) = n \times \hat{E}\left[\frac{1}{2}\log_2 \frac{1}{m}\sum_{j=1}^{m}\left(D_x^2 - R_{xx}^2(j)\right)\right], \quad (1)$$

where, n is the sample volume; $\hat{E}[\bullet]$ is the integer function with rounding to a larger whole; $j = \overline{1,m}$ are shifting parameters of time counts; m is a number of autocorrelation reference points; D_X is a dispersion; $R_{XX}(j)$ - is an autocorrelation function.

The centred autocorrelation function $R_{xx}(j)$ is bounded by asymptotics given by expressions (2) [6].

$$R_{xx}(j) = \frac{1}{n}\sum_{i=1}^{n} \overset{\circ}{x}_i \cdot \overset{\circ}{x}_{i-j}; \qquad \overset{\circ}{x}_i = x_i - M_x; \quad j = \overline{0,m}. \quad (2)$$

where, $M_x = \frac{1}{n}\sum_{i=1}^{n} x_i$ is a selective mathematical expectation at intervals [1, n], $D_x = \frac{1}{n}\sum_{i=1}^{n}\left(\overset{\circ}{x}_i\right)^2$ is a dispersion, the graph of which is shown in the Fig.1 [6].

The probability entropy function $I_X(R)$, which is calculated on the basis of the autocorrelation function $R_{XX}(j)$, is shown in Fig.2.

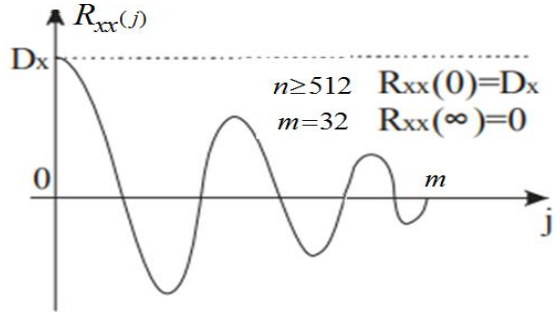


Figure 1: Asymptotes of autocorrelation function $R_{XX}(j)$

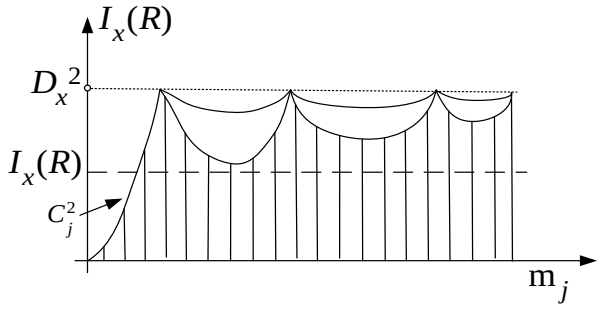


Figure 2: Estimate of correlation entropy $I_X(R)$, where $C_j^2 = D_X^2 - R_{XX}^2(j)$.

Based on the described approach of entropy estimation using the correlation function $R_{xx}(j)$, the structure of the special processor is developed. This structure is shown in Fig. 3 [6].

The specialized entropy estimation processor shown in Figure 3 consists of: $x(t)$ – input analog signal; 1 – synchronizer; 2 – ADC; 3 – digital data centring module, $\overset{\circ}{x}_i = x_i - M_x$; 4 – multiplication and

squaring module $C_j^2 = (\overset{\circ}{x}_i - \overset{\circ}{x}_{i-1})$; 5 – multi-bit shift register; 6 – generator of adjugate squares C_j^2; 7 – a group of adders; 8 – pyramid adder; 9 – binary logarithmic function encoder; $I_X(R)$ – output code.

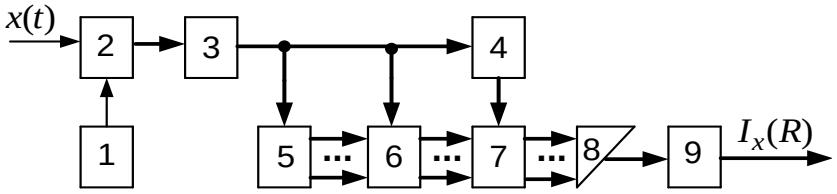


Figure 3: The structure of the special processor for determining entropy, taking into account statistical relationships using a correlation function

This approach to entropy estimation using the autocorrelation function $R_{xx}(j)$ has the following disadvantages:

i) the need to perform a data centring operation, which leads to an increase in computational time;

ii) the presence of the operation of accumulation of the products sum for squares of the centred values.

The consequence of these shortcomings is the considerable hardware complexity structure of the specialized processor for entropy estimation and significant time costs, which lead to low performance.

As a result, such structural implementation of specialized processor for entropy estimation is characterized by and low performance.

2. The next way for entropy estimation uses the equivalence correlation function $F_{XX}(j)$. This formula of entropy estimation has next form (3):

$$I_X(F) = n \cdot \hat{E}\left[\log_2 \frac{1}{m} \sum_{j=1}^{m} \left[M_X^2 - F_{XX}^2(j) \right] \right], \tag{3}$$

where, $\hat{E}[\bullet]$ is integer function with rounding to a larger integer number; $F_{XX}(j)$ – autocorrelation equivalence function.

Asymptotic characteristics of the equivalence function $F_{XX}(j)$ are described by expressions (4) [6].

$$F_{XX}(j) = \frac{1}{n} \sum_{i=1}^{n} \overset{\vee}{Z}(x_i, x_{i-j}), \quad j = \overline{1, m}, \tag{4}$$

Its graph is presented in Fig. 4.

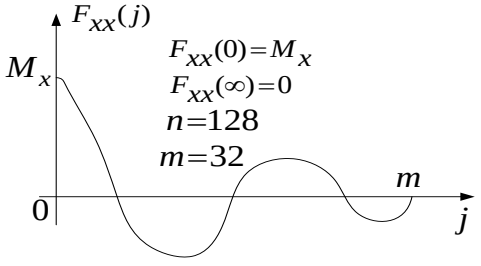


Figure 4: Graph of equivalence function $F_{XX}(j)$ and its asymptotes

The entropy estimation $I_X(F)$ based on the correlation equivalence function is displayed in Fig. 5.

Developed structure of the specialized processor based on the entropy estimation (3) using the equivalence function $F_{XX}(j)$ is presented in Fig. 6 [6].

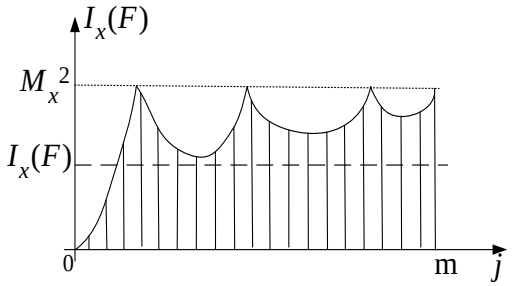


Figure 5: Entropy estimation $I_X(F)$ for correlation equivalence function.

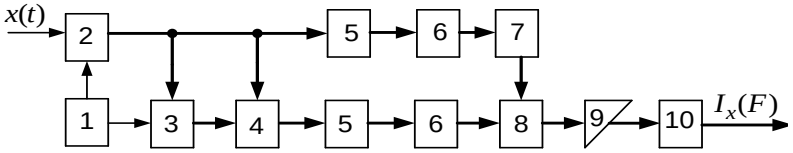


Figure 6: Structure of a specialized processor for entropy estimation based on function $F_{XX}(j)$

The following notations are used in Figure 6: $x(t)$ – input analog signal; 1– synchronizer; 2 – ADC; 3 – multi-bit shift register; 4 – a group of logical elements "AND"; 5 – counters; 6 – square generators; 7 – encoder; 8 – a group of adders; 9 – pyramid encoder; 10 – a binary logarithmic encoder; $I_X(F)$ – output code for entropy estimation.

The advantages of this method of entropy estimation and the corresponding specialized processor are:

i) a lack of centring and multiplication operations;

ii) Using the operation of comparing $Z(x_i, x_{i-j})$ of values x_i and x_{i-j};

iii) As a result of points i) and ii) the simpler algorithm and higher performance of the specialized processor of entropy estimation;

iiii) 4 times reduced the required sample volume $n \geq 128$ of input digital data with calculating the m-points of autocorrelation function.

The analysis of entropy estimation algorithms [6-8] and corresponding structural solutions allows to develop single-crystal specialized processor and widely use them in telecommunication systems and networks [9], as digital receivers of entropy-manipulated signals. It is also advisable to extend the functionality of such specialized processors by parallel outputting of entropy estimation codes and intermediate results of centred values calculations x_i, mathematical expectation M_x, dispersion D_x and estimated values of autocorrelation functions $R_{xx}(j)$, $F_{xx}(j)$, which are integral characteristics of entropy as it shown in Fig.7.

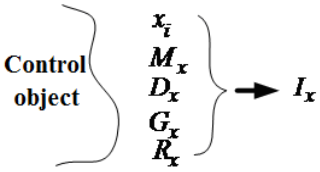


Figure 7: Entropy and her integral characteristics

Determination of entropy estimation $I_X(H)$ is carried out according to the formula of C. Shannon [1], which is based on the probability distribution of events:

$$I_x(H) = -\sum_{i=1}^{m} (p_i \cdot \log_2 p_i). \quad (5)$$

where $p_i = \frac{N_i}{N_0}$ – probability of appearance of i-event; m – a number of statistically independent events, N_0 – the total number of options.

It is more practically convenient to calculate the probability entropy according to the algorithm [6].

Since the $N_0 \geq N_i$ calculation of the logarithmic function is performed according to formula (6):

$$\log_2\left(\frac{N_i}{N_0}\right) = \left(\log_2 N_0 - Log_2 N_i\right), \tag{6}$$

Thus, the calculation of probabilistic entropy when $N = 256$ will be performed according to the expression:

$$I_x(H) = \frac{1}{N_0}\sum_{i=1}^{256} N_i\left(\log_2 N_0 - \log_2 N_i\right), \tag{7}$$

The graph of the entropy calculation results according to Shannon's formula in the decimal number system with the sample volume $n = 256$ and the total number of random messages $m = 256$ are shown in Fig.8.

Figure 8: The graph of the probabilistic entropy estimation in decimal number system

You can see the following properties of estimating the probability entropy according to Shannon's formula as a result of computer modelling of the corresponding calculations and from the graph shown in Figure 8:

i). The entropy value $I_x(H) = 0.5$ corresponding to the equal probability of independent events is observed in two cases when, for the given experimental conditions, the probabilities $p_i = 64$ and $p_i = 128$;

ii). The maximum value of the entropy estimate $I_x(H) = 0.530737$ is observed when $p_i = 94$;

iii). The characteristic of the estimated entropy $I_x(H)$ in the range of $1 \leq p_i \leq 255$ is asymmetric in contrast to the known traditional graphs of entropy estimates, which are symmetric in relation to the point of maximum entropy estimate.

For N_i values that correspond to the whole binary digits, there is symmetry of the same values of entropy estimates.

Thus, when $N_i = 16$ and $N_i = 192$ $I_x(H) = 0.25$;

when $N_i = 32$ and $N_i = 175$ $I_x(H) = 0.375$;

when $N_i = 64$ and $N_i = 128$ $I_x(H) = 0.5$.

It is obvious that the form of the graph of entropy estimation values according to Shannon's formula is conditioned by the graphical representation of the logarithmic function, with the argument defined in the range from 1 to 255.

Since modern digital electronics is based on binary codes, for convenience, Table 1 shows the results of calculating the logarithmic function in decimal and binary number system.

Table 1 shows the results of calculations of the logarithmic function of the products $p_i \cdot \log_2 p_i$ with the number of registered random events N_i corresponding to integers $N_i = 2^k$, $k = \overline{0,8}$.

Table 1

Value of entropy $N_i.(\log_2 N_0 - \log_2 N_i)$ for all $N_i = \overline{1 \ldots N}$

$\log_2 N_0$	N_i	$\log_2 N_i$	$N_i.(\log_2 N_0 - \log_2 N_i)$
8	1	0	1 (8- 0) = 8
8	2	1	2 (8- 1) = 14
8	4	2	4 (8 – 2) = 24
8	8	3	8 (8 - 3) =40
8	16	4	16 (8 – 4) = 64
8	32	5	32 (8 - 5) = 96
8	64	6	64 (8 – 6) = 128
8	128	7	128 (8 – 7) = 128
8	256	8	256 (8 – 8) = 0

As a result of the entropy calculation $I_X(H)$ in the decimal and binary number systems, the numerical values of estimation $I_X(H)$ are obtained, which are presented by the informative fragments in Table 2.

Table 2

Value of entropy in binary and decimal system for all $N_i = \overline{1 \ldots 256}$

$N_i(10)$	$I_x(H)(10)$	$N_i(2)$	$I_x(2)$
1	8,0	00000001	1000,000000000
2	14,0	00000010	1110,000000000
3	19,2451125	00000011	10011,011000000
4	24,0	00000100	11000,000000000
5	28,39035953	00000101	11100,011000000
6	32,490225	00000110	10000,011000000
7	36,348555	00000111	100100,010000000
8	40,0	00001000	101000,000000000
9	43,47067499	00001001	101011,011000000
…	…	…	…
12	52,98044999	00001100	110100,1110000000
15	61,39664107	00001111	111101,011000000
16	64,0	00010000	1000000,00000000
…	…	…	…
31	94,41991438	00011111	1011110,01000000
32	96,0	00100000	1100000,00000000
…	…	…	…
63	127,4313648	00111111	00111111,01000000
64	128,0	01000000	10000000,00000000
…	…	…	…
88	135,5700176	01011000	10000111,10000000
…	…	…	…
94	135,8637172	01011110	10000111,10000000

Continue of table 2

…	…	…	…
100	135,6143810	01100100	10000111,10000000
101	135,520602	01100101	10000111,10000000
128	128,0	10000000	1000000,00000000
…	…	…	…
192		11000000	1001110.10000000
…	…	…	…
224	43,15249746	11100000	101011.00100000
225	41,89923198	11100001	101001.11100000
…	…	…	…
234	30,33465562	11101010	11110,01010000
235	29,01851756	11101011	11101,00000000
…	…	…	…
240	22,34623705	11110000	10110,01010000
241	20,99366997	11110001	10100,11110000
…	…	…	…
248	11,35931502	11111000	1011,01011000
249	9,959518915	11111001	1001,11110000
250	8,553928834	11111010	1000,10001000
251	7,142567958	11111011	111,00100100
252	5,725459278	11111100	101,10111000
253	4,302625602	11111101	100,010011000
254	2,87408956	11111110	100,010011000
255	1,439	11111111	1,01110000000
256	1,0	100000000	1,00000000

Since the logarithmic function is irrational, it is clear that with a limited number of digits, its value can be calculated only with a certain accuracy, which is limited by the number of decimal places. Accordingly, in the decimal number system its value will be displayed more accurately than in binary with the same number of digits. Limiting the accuracy of only the integer part of number of the logarithmic function in the binary number system is quite sufficient, given the method of entropy estimation.

3. Theory and structural characteristics of wireless bus and 2D topologies cyber-physical systems

Estimates of hardware and time complexity are traditionally used to assess the system characteristics of cyber-physical systems (CPSs) components [10-12]. At the same time, these estimates do not take into account the current level of micro- and nano-electronics technologies in the crystal environment. The structural and technological complexity of such crystals, which contain transistors and the connections between them, is almost the same. There are many other estimates of the complexity of microelectronic computing modules in the CPS design process [11-15]. It is advisable to use the following more extensive estimates of the system characteristics of CPS components, among which the most important is the structural complexity [10]. Table 3 shows the criteria of structural and functional-informational complexity of microelectronic components and structures of CPS [10].

Table 3

Criteria for structural and functional-informational complexity of microelectronic components and structures of CPS

№	Analytical expression
1.	Petri. $$k_c = \frac{V_k}{b_n + b_m}$$ k, n, m is number of vertices, unidirectional and bidirectional edges
2.	Quine. $$S_K = \sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j$$ n, m is the number of inputs and outputs of the structure respectively
3.	M. Kartsev . single-level $A_C = \sum_{i=1}^{n} A_i$ A_C is general assessment of hardware complexity; i, j, k are types of components or levels of device structure.
4.	S. Mayorov. two-level $A_C = \sum_{j=1}^{m} \sum_{i=1}^{n} A_{ij}$, three-level $A_C = \sum_{j=1}^{m} \sum_{i=1}^{n} \sum_{k=1}^{l} A_{ijk}$ m, n, l is the appropriate number of different components types or levels of the structure of the device
5.	M. Cherkaskyi. Logarithmic structural complexity $$S = -E \log_2 \frac{E}{n(n-1)}$$ where E is the number of elements of the adjacency matrix of the system; n is the number of vertices of the graph
6.	M. Cherkaskyi. Software complexity $$P = -F \log_2 \frac{F}{n \cdot m}$$ F ; n, m is the corresponding number of control signals, control inputs and time samples of the time chart;
7.	V. Glukhov. $$L = \sum_{i=0}^{m-1} (g_i + v_i) \approx (1/2...3/4)m^2 \ ; \ g_i = x_i + 1 \ , \ v_j = m + d_i + 1$$ g_i are lengths of horizontal, v_i are the lengths of the vertical connections on the conditional FPGA.
8.	J. Martin. Structural complexity of the network structure $$K_d = \frac{N_i}{N_0} \ ; \ K_d = \frac{S}{G}$$ N_i are numbers of connections, N_0 is number of components; S is number of readings or requests, G is number of records or data updates
9.	Y. Nykolaychuk, I. Pitukh. Advanced assessment of network complexity $$K_{ed} = \frac{S_i \cdot G_0}{S_0 \cdot G_i}$$ S_i, S_0, G_i, G_0 are the actual number of requests, the maximum possible number of requests, the actual number of records or updates, the maximum possible number of records or updates in the node of the matrix model, respectively

№	Analytical expression
10.	N. Vozna. Criterion of complexity of multifunctional structure $k_c = \sum_{i=1}^{n} \alpha_i P_i \;\; P_i \in (l, P, x, d, r, h, z, b, c, i, n, a, f)$ P_i are informative parameters of structures attributes, α_i are weights of expert assessments of structural complexity, n is the number of microelectronic structure components.
11.	Y. Nykolaychuk, N. Vozna. Information and structural complexity $K_e = K \times \frac{F_C}{k_C} \Rightarrow \max ; \; F_C = \sum_{j=1}^{m} f_j$ *K* is data level identifier; F_C is information complexity of microelectronic structure.
12.	N. Vozna. Information and functional complexity of inputs and outputs $f_j = \sum_{i=1}^{n} \beta_i \times f_{input} + \sum_{i=1}^{m} \lambda_i \times f_{output}$ f_j are functional and informational characteristics; β_j, λ_j are coefficients of informativeness of input-output functions; m, n is number of inputs and outputs.
13.	Y. Nykolaychuk, V. Hryga. Additive criterion for estimating the complexity of the data ordering structure. $K_v = A + \tau; \; A = \sum_{1}^{n} A_i; \; A_i = A_M;$ A_i, τ_i - respectively hardware and time complexity of the *i*-th microelectronic component;
14.	A. Melnyk. Multiplicative normalized criterion of operating device complexity $K_M = 1/W_k \times T_k \Rightarrow \max;$ W_k are total equipment costs; T_k is duration of data processing;

It should be noted that the multiplicative normalized criterion of complexity of the operating device proposed by Professor A. Melnyk [16] is the most informative assessment of maximizing the efficiency of system characteristics of ADC components, vector, scalar and quantum supercomputers.

Systematized criteria (Table 3) for assessing the structural complexity of microelectronic components CPS can increase the efficiency of comparing the system characteristics of different structures of operating devices and specialized processors in the architecture of computing facilities.

This is especially true of the criteria presented (Table 3, No.11&14), which are the minimum characteristics of the efficiency of the equipment use for processor operating devices and computer memory.

An important criterion for the structural complexity of network 2D architectures CPS is the criterion of emergence proposed by J. Martin (Table 3, No.8).

2D network architectures CPS are classified: monopoly, hierarchical multilevel, ring, star-bus, interactive hierarchical, star-ring with open atmospheric optical channels communication, hierarchical one-level, bus, systolic, interactive monopoly, interactive multilevel hierarchical, ring-star, problem-oriented dialog.

The multilevel hierarchical, ring, systolic, star-bus and star-ring structures are the most perfect in the structure of CPS, which belong to the DCS [17, 18], in terms of functional and structural priority characteristics. The system characteristics of complexity for the specified network architectures CPS are calculated, according to the criterion of emergence of J. Martin [10] (Table 4-5).

Note that the most priority modern architectures of interactive CPS are structures (Table 5, No.3,4), which are characterized by the highest level of emergence and parallelization of data streams and processing.

Such 2D CPS structures are used as information systems for background monitoring of natural protection areas.

The concept of the theory of formation and processing of interactive and dialog data in 2D architectures of DCS is shown on Fig.9.

Table 4

Emergence of 2D network non-interactive CPS architectures

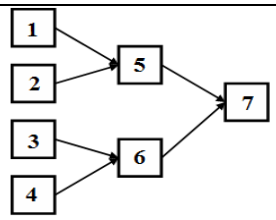
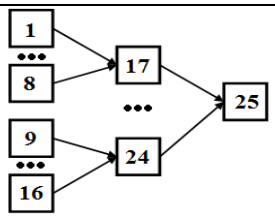
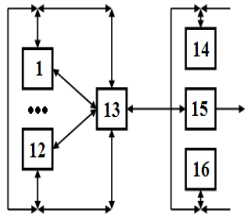
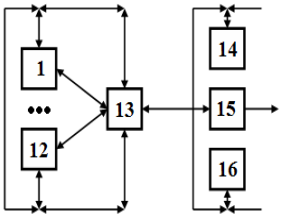
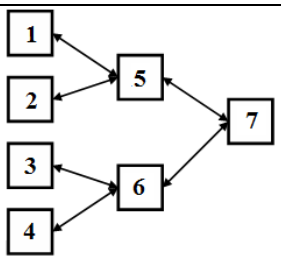
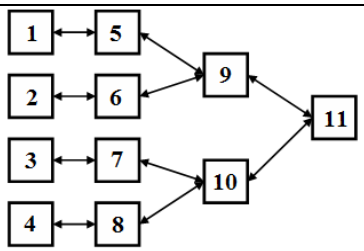
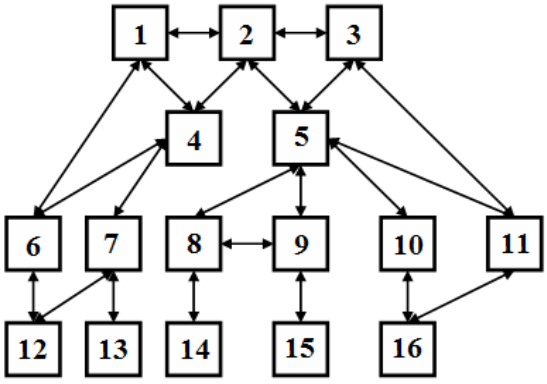
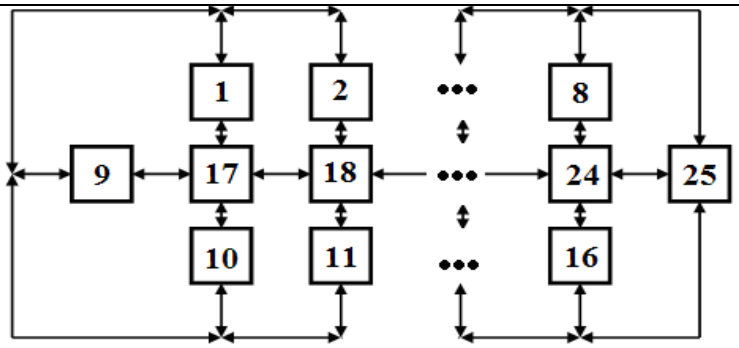
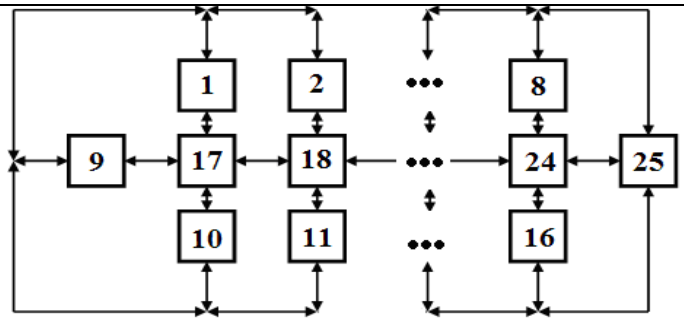
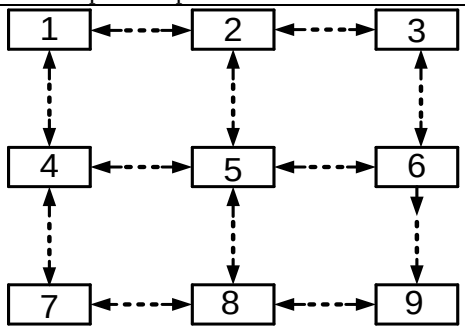
Modifications of DCS architectures			
Multilevel hierarchical		Star-bus	
			
$k_e = 0,8$	$k_e = 0,95$	$k_e = 8,3$	$k_e = 8,3$

Table 5

Emergence of 2D network interactive CPS architectures

Modifications of DCS architectures	
1. Multilevel hierarchical	
	
$k_e = 1,7$	$k_e = 1,8$
2. Problem-oriented dialog system	
	
$k_e = 2,8$	
3. Star-bus	
	
$k_e = 4,2$	

3. Star-bus
 $k_e = 4,2$
4. Star-ring with open atmospheric optical communication channels
 $k_e = 2,7$

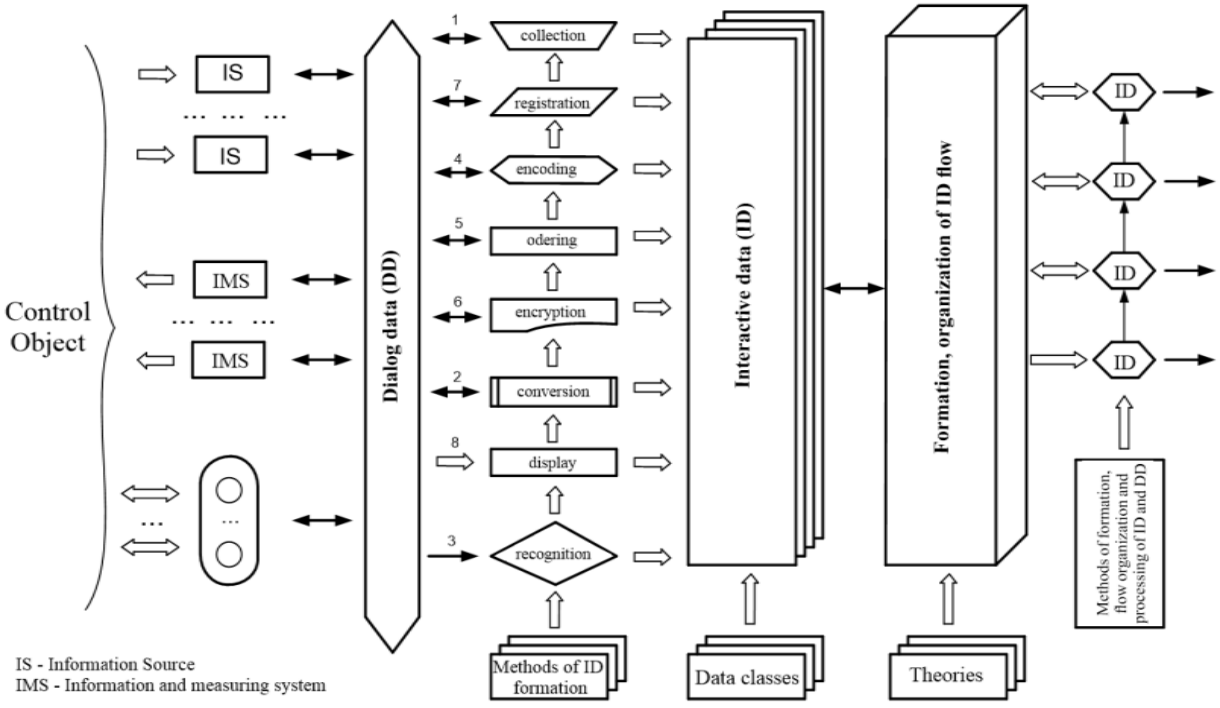


Figure 9: Structure and information functions of the formation and processing theory concept for interactive data

The main result of the using such concept in practice is the substantiation of methods of traffic organization and processing of information monitoring and dialog data in 2D structures of the CPS.

The developed concept is a basic tool for designing and improving the system characteristics of the components of monitoring, dialog, cyber-physical and interactive computer systems.

4. Crypto-protected transmission of information in cyber-physical systems based on entropy-manipulated signals

An important problem in the design of CPS for use in various industries, environmental and regime areas is the effective cryptographic protection of information data flows from unauthorized access.

There are known fundamental limitations of Shannon, which relate to the reliable receiving of manipulated signals against the background of noise [1, 3]. The essence of such restrictions is that the ratio of the sign of the manipulated signal (amplitude, frequency, phase, energy, etc.) must exceed the corresponding noise characteristic by 2 times according to the following statements:

$$\frac{P_s}{P_n} \geq 2;\ \frac{P_s(\Delta f)}{P_n(\Delta f)} \geq 2;\ \frac{P_s(f_i)}{P_n(f_i)} \geq 2;\ \frac{R_{xx}(j)_s}{R_{xx}(j)_n} \geq 2;\ \frac{H_s}{H_n} \geq 2;\ \frac{H_{cs}}{H_{cn}} \geq 2.\,, \tag{8}$$

where: P_S, $R_{XX}(j)_S$, H_S, H_{CS} – corresponding powers of amplitude, frequency, phase, autocorrelation, noise, entropy and crypto-protected entropy, P_n, $R_{XX}(j)_n$, H_n, H_{Cn} – corresponding powers of noise characteristics.

It is shown the characteristics (Fig. 10) of reliable signal extraction against the background of noise and interference depending on the distance of propagation according to the fundamental limitations of C.Shannon.

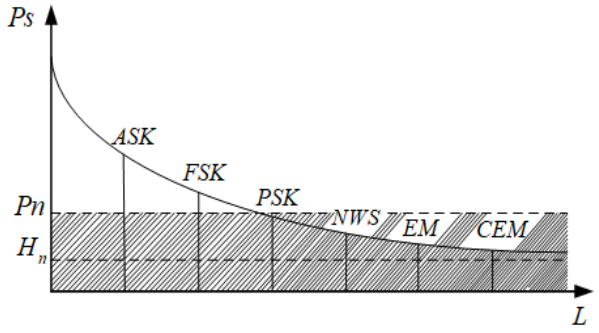


Figure 10: Methods of signal manipulation in conditions of intense interference

It is shown (Fig. 10) that the most promising methods of signal manipulation in modern CPS are CEM – crypto-protected multilevel entropic manipulation.

The structure of the device for determination of entropy according to the formula of probabilistic estimation of entropy of C. Shannon [1, 3] is offered in a work [9].

$$H^{<S>} = -k\sum_{j=0}^{S} p_j \log p_j\,, \tag{9}$$

where k is a positive coefficient that takes into account the basis of the logarithm; pj is the probability of the sj's state of information source; S is a number of independent states of information source.

The device is characterized by a high level of parallelization of information processing, has a regular microelectronic structure and contains: 1 – ADC; 2 – information input of the device, 1.1 – group of model resistors, 1.2 – comparators with paraphrase outputs (direct and inverse), 1.3 – logic elements AND-NOT, 3 – binary counters, 4 – synchronizer; 5 – encoders, 6 – pyramidal adder, 7 – device output.

In each channel of the device the counter (3) accumulates the sum of identical values of digital samples pj, and at the output of the tabular encoder (5) the product code pjlog2(pj) is formed. At the end of the cycle of sampling n-digital samples at the output of the pyramidal adder (6) the source code of the estimated entropy of the information source is formed.

The functional limitation of such device is the delay of the calculation process in the encoders (5) and the adder (6), which reduces the speed of the device. Therefore, the structure of the entropy estimation device (Fig. 11) is proposed [19], which is characterized by increased speed by parallelizing the processes of accumulation of the sum of probabilities pj and parallel encryption and estimating the initial sum of entropy according to the expression:

$$\tau_{H_x} = \begin{cases} \tau_{com} + \tau_c \\ \tau_t + \tau_e + \tau_{\sum} \end{cases}, \quad (10)$$

$\tau_{com}, \tau_c, \tau_t, \tau_b, \tau_{\sum}$ - respectively delays of the comparator, counter, trigger, encoder and adder.

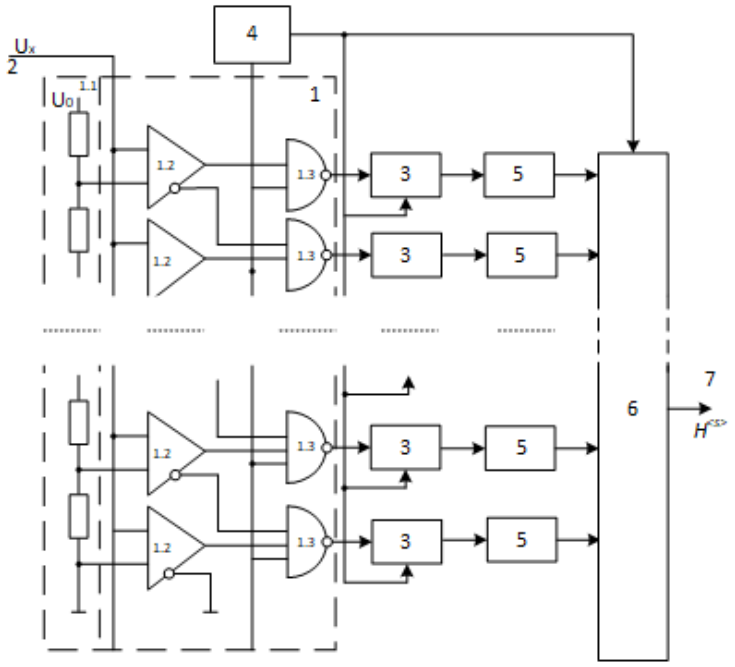


Figure 11: Device for entropy estimation

Structure of such a specialized processor [19] for receiving entropy-manipulated signals is proposed, which is shown in Fig.12.

Each channel of such device uses an n-bit jk-counter (3), the calculation results of which are registered by the memory register (5) on D-flip-flops. At the same time, in the process of calculating the product pjlog2(pj) and determining their sum by the pyramidal adder (7), the accumulation of new probability estimates pj in synchronous jk-counters (3) is carried out.

Patent [19] presents the results of comparing the hardware and time complexity of the two devices for entropy estimation at a sample size of m = 256, bit counts k = 8 and bit encoder codes h = 11.

Probability entropy detection devices are important components of telecommunication systems in the CPS structure, which provide an appropriate level of encryption of information data flows. The principle of data encryption based on the entropic method of signal manipulation, which provides noise-like formation of bit “0” and “1” bits is proposed. This modifies the structure of the entropy estimation device, which can receive and decode a bit-oriented stream of crypto-protected data with protection against unauthorized access.

It is shown an example (Fig. 13) of such a modified probability entropy determination structure [19], which is used to receive and decode crypto-protected entropy-manipulated signals.

The proposed method of crypto-protected entropy-manipulated is characterized by wide possibilities that require fundamental theoretical and experimental research, as well as a large amount of computer modelling.

Wide range of possibilities of methods of cryptographic protection of entropy-manipulated signals by hashing of streams {pi} and the possibility of their logical processing with logical elements "OR", delays and logical elements "AND".

In addition, multiplication by log2pi, log2pj, log2piz provides additional opportunities to increase cryptographic protection.

Then we can selectively summarize the individual $S_i \cdot \log_2 S_j$ to generate individual bits or quasi-ternary bits $H^{<S>}$.

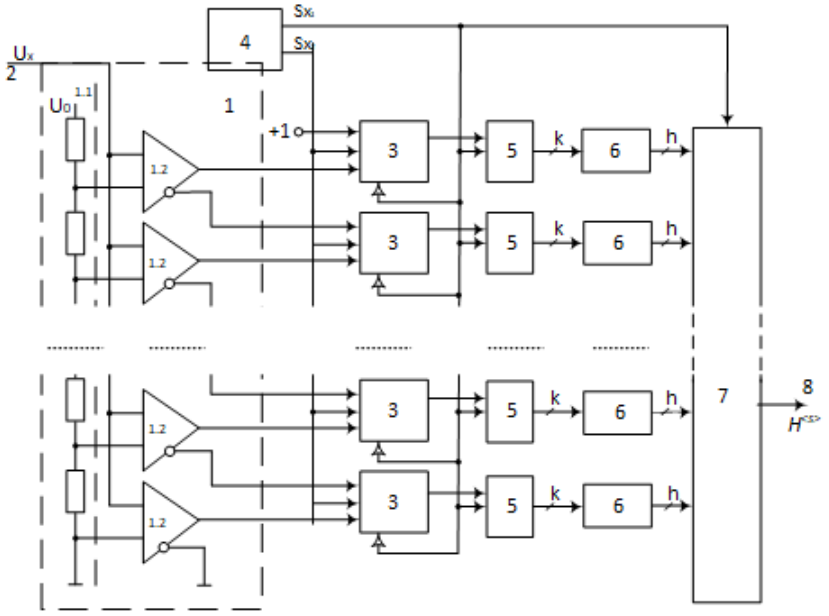


Figure 12: The structure of the specialized processor for entropy estimation

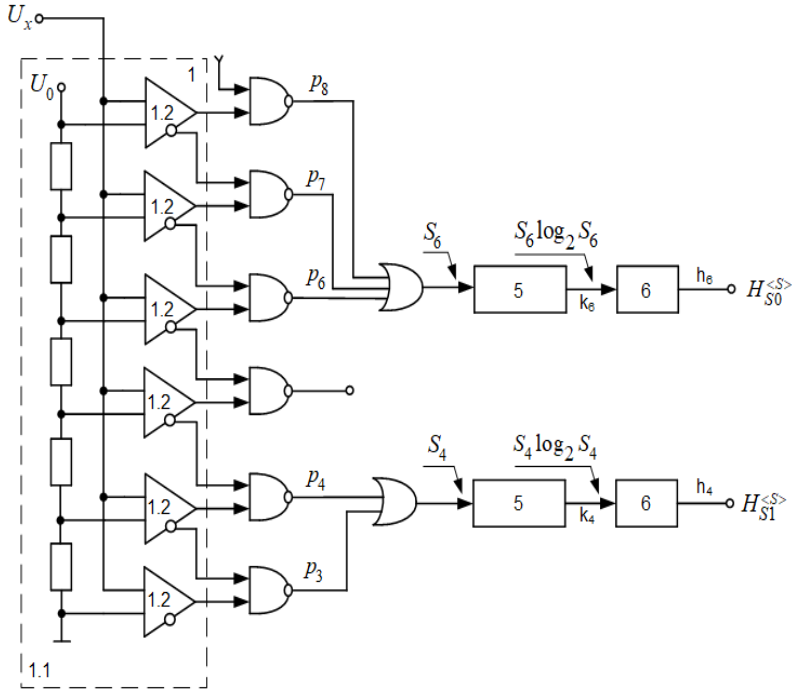


Figure 13: The structure of the modified device for entropy estimation

5. Acknowledgements

Thus, the analysis of existing entropy estimates is carried out and a new theoretically substantiated approach is proposed, taking into account correlation relationships. The results of entropy characteristics and properties of digital components of cyber physical systems are investigated and given. The prospects of entropy analysis and its use for the analysis of digital data flows are shown. High-performance structures of specialized processors for determining probability and correlation entropy are proposed. Improved structures of data cryptographic algorithms based on entropic signal manipulation. The generalizations of approaches of complexity estimations definition of cyber physical systems components are generalized and own criterion of structural estimation is offered and the mathematical apparatus of its definition is formalized.

6. References

[1]. C. Shannon A mathematical theory of communication ACM SIGMOBILE Mobile Computing and Communications Review, vol.5 n.1, Jan 2001 pp. 3-55.

[2]. Harré MS. Entropy, Economics, and Criticality. Entropy. 2022; 24(2):210. https://doi.org/10.3390/e24020210.

[3]. Pan Q, Zhou D, Tang Y, Li X, Huang J. A Novel Belief Entropy for Measuring Uncertainty in Dempster-Shafer Evidence Theory Framework Based on Plausibility Transformation and Weighted Hartley Entropy. Entropy. 2019; 21(2):163. https://doi.org/10.3390/e21020163.

[4]. J. Karmeshu. Entropy Measures, Maximum Entropy Principle and Emerging Applications. DOI https://doi.org/10.1007/978-3-540-36212-8.

[5]. Lee J, Lee K. A Method for Neutralizing Entropy Measurement-Based Ransomware Detection Technologies Using Encoding Algorithms. Entropy. 2022; 24(2):239. https://doi.org/10.3390/e24020239.

[6]. Artur Voronych, Lyubov Nyckolaychuk, Nataliia Vozna, Taras Pastukh. Methods and Specialized processors of Entropy Signal Processing. The Experience of Designing and Application of CADSM'2019, Feb., p. 3/59-3/62, 2019.

[7]. H. Zhang, Y. Deng. Entropy measure for orderable sets. Information Sciences. Vol. 561, 2021, Pages 141-151. https://doi.org/10.1016/j.ins.2021.01.073.

[8]. W. Xu, L. Jiang, C. Li. Improving data and model quality in crowdsourcing using cross-entropy-based noise correction. Information Sciences.Vol. 546, 2021, pp. 803-814, ISSN 0020-0255, https://doi.org/10.1016/j.ins.2020.08.117.

[9]. Patent of Ukraine 117037, July 11, 2018. (in Ukrainian).

[10]. N. Vozna. Structuring multifunctional data: theory, methods and tools: Monograph Ternopil: TNEU, 2018. 378 p. (in Ukrainian).

[11]. Dean, Walter. Computational Complexity Theory and the Philosophy of Mathematics†. Philosophia Mathematica, 2019, 27(3):381-439.

[12]. Miller, R. Computable Transformations of Structures. In: Kari, J., Manea, F., Petre, I. (eds) Unveiling Dynamics and Complexity. CiE 2017. Lecture Notes in Computer Science, vol 10307. Springer, Cham. https://doi.org/10.1007/978-3-319-58741-7_9.

[13]. B. Barak, P. Gopalan, J. Håstad, R. Mekha, P. Raghavendra, and D. Steurer Making the Long Code Short", SIAM Journal on Computing, 2015, Vol 44, pp 1287-1324.

[14]. Okhotin, A., Salomaa, K.: State complexity of operations on input-driven pushdown automata. J. Comput. Syst. Sci. 86, 207–228 (2017).

[15]. Dassow, J. (2017). Descriptional Complexity and Operations – Two Non-classical Cases. In: Pighizzini, G., Câmpeanu, C. (eds) Descriptional Complexity of Formal Systems. DCFS 2017. Lecture Notes in Computer Science, vol 10316. Springer, Cham. https://doi.org/10.1007/978-3-319-60252-3_3.

[16]. Melnyk, A., Melnyk, V. Specialized Processors Automatic Design Tools-the Basis of Self-Configurable Computer and Cyber-Physical Systems. IEEE International Conference on Advanced Trends in Information Theory, ATIT 2019. Ukraine, Kyiv 2019, pp. 326–331. https://doi.org/10.1109/ATIT49449.2019.9030481.

[17]. Milazzo M, Musciotto F, Miccichè S, Mantegna RN. Analysis of the Structure and Dynamics of European Flight Networks. Entropy. 2022; 24 (2):248. https://doi.org/10.3390/e24020248.

[18]. Suchecki K, Hołyst JA. Hierarchy Depth in Directed Networks. Entropy. 2022; 24(2):252. https://doi.org/10.3390/e24020252.

[19]. Declarative Patent of Ukraine, Patent 123920, June 23, 2021. (in Ukrainian).