The Concept of Data Structuring Based on Entropy Forming, Transmission and Processing Methods of In....

Conference Paper · March 2014
DOI: 10.13140/RG.2.1.2221.5447

4 authors, including:

Artur Voronych
Ivano-Frankivsk National Technical University of Oil and Gas

9 PUBLICATIONS  18 CITATIONS
The Concept of Data Structuring Based on Entropy Forming, Transmission and Processing Methods of Information Flows

Nataliya Vozna, Artur Voronych, Taras Pastuh, Volodymyr Gladyuk

Abstract – the concept of data structuring based on different estimations of information entropy measure is described, it is shown that entropy approach could be widely used in information-measuring and telecommunication systems of data transmission.

Keywords – structuring, entropy, autocorrelation function.

I. INTRODUCTION

Structuring data [1] is the fundamental concept concerning methods, algorithms and architectural solutions for measurements, formation of data codes in various theoretical-numerical bases and specialized transformations on information flows.

Figure 1 shows generalized processes of information sources flow structuring, where IS – information sources, S - sensory system, # - sampling processes in the Hemminh's space, FS - functional structuring, F(U,\cap) - operations of structural unification and separation.

![Fig. 1. The process of structuring.](image)

One of the perspective directions of development and improvement of information theory and information processes in science, society and industry is data structuring that can approach to the own entropy of information sources based on entropy methods of forming, transmitting and processing of information flow in computer systems and networks. Therefore, the concept of structuring data based on the entropic approach is crucial for estimation of efficiency and effectiveness of use of the problem-oriented structured data.

II. ESTIMATIONS OF INFORMATION ENTROPY MEASURE

In developing technology of structural mapping results of measure of gas flow on the bases of the calculation of entropy noise in gas pipelines evaluation measures of C. Shannon's information entropy is effectively used:

\[ I_\sigma = \sum_{i=1}^{n} P_i \log_2 P_i, \quad (1) \]

where \( P_i \) is probability of discrete digital values and quantum noises.

As a result characteristics depending on the entropy of noise signals from the gas flow in laminar and turbulent conditions are received (Fig. 2).

![Fig. 2. Characteristics of noise entropy (I_H) change depending on the gas flow (Q).](image)

Another successful example of structuring data based on the entropy approach is entropy estimate that is based on the Hartley's information measure for a wide class of analytically specified signals:

\[ I_H = \hat{E}[\log_2 A], \quad (2) \]

where \( A \) is a digitization range of amplitude signals, \( \hat{E}[\cdot] \) is an integer function rounded to the whole number.

Moreover, in the particular case when the ADC output \( A = 2^k \), where \( k = 4,8,10,16 \) is the digit capacity of source codes, Hartley entropy is calculated according to the simplified expression (2):

\[ I_H = \log_2 A. \]

In the work [2] it is shown that Hartley’s entropy estimation is the upper estimation of the entropy of signals, but Shannon’s entropy estimation has significant functional limitations, as it does not react to the change of dynamic, correlation and spectral characteristics of the signals (Fig. 3).

On Fig. 3 the implementation of a random process of information source, with equal probabilities but different ordering of amplitudes values is shown.

Hartley’s entropy estimate (fig. 3) for two cases is the same and defined by the formula (2):

\[ I_H = 32 \cdot \log_2 10 = 32 \cdot 4 = 128 \text{ bits.} \]
Shannon's entropy estimation for two cases is calculated according to the formula (1) and for all cases, implementation of information source (Fig. 3) is the same as it is characterized by the same stochastic characteristics and is independent of the statistical amplitudes distribution:

\[ I_{sh} = 32 \cdot 2.96 = 94.88 \text{ bits.} \]

Information measure of entropy based on the calculation of logarithmic function of the difference of squares and variance of one of the autocorrelation function is proposed by the expression:

\[ I_N = \frac{1}{2} \log_2 \left( \frac{1}{m} \sum_{j=1}^{m} (D_x^2 - W_{xx}(j)); j \in [0, m] \right) \tag{3} \]

where \( D_x \) is a dispersion, \( W_{xx} \) is an autocorrelation function, \( m \) is the number of autocorrelation function’s points.

II. STRUCTURING DATA FOR ENTROPY APPROACH

Analysis of the most famous correlation functions and their asymptotic is described by the following expression:

covariance:

\[ K_{xx}(j) = \frac{1}{m} \sum_{j=1}^{m} x_j x_{j+j}; K_{xx}(0) = D_x + M_x; \quad K_{xx}(\infty) = M_x; \]

correlation:

\[ R_{xx}(j) = \frac{1}{m} \sum_{j=1}^{m} (x_j - M_x)(x_{j+j} - M_x); \quad R_{xx}(0) = D_x; \quad R_{xx}(\infty) = 0; \]

equivalence:

\[ F_{xx}(j) = \frac{1}{m} \sum_{j=1}^{m} (x_j - x_{j+j}; \quad F_{xx}(0) = M_x; \quad F_{xx}(\infty) = 0, \]

which shows that the autocorrelation function of equivalence is characterized by the lowest computational complexity and can be effectively used for the construction of the special entropy processors.

On Fig. 4 examples of structured data forming of analytic specified signals on the basis of the correlation entropy measure are shown (3).

Fig. 4. Estimation of correlation entropy measure of different signals.

On Fig. 4, the estimation of the correlation entropy measure of sinusoidal signal and Barker’s codes is shown, which shows that Barker’s codes are characterized by the maximum entropy in relation to the other signals.

Theoretical position and examples of entropy approach of structuring data adequately reflect schematic structures of appropriate special processors architectures [3].

It should be noted that in the process of the calculation of the correlation entropy measure the volume of the structured data decreases on one or two magnitude orders and there is an appropriate degree of approximation to own entropy information sources.
Special entropy processor is designed based on the analytical expression evaluation information measure of entropy (3) and the structure of the auto correlator [3]. This special processor is successfully applied in the forming and transmission system of entropy-manipulated signals in the wireless sensor systems (Fig. 5).

![Fig. 5](image)

Fig. 5. The architecture of a wireless sensor system based on the entropic noise signals manipulation (SU is a transmitter entropy-manipulated signals with antenna, M is device definition autocorrelation measure of entropy, D is demodulation unit, K is hub).

Method of quasi-three entropy noise signal manipulation is used in this architecture of wireless sensor systems used and is illustrated in Figure 6. The essence of it is that binary information message is placed in line with signal entropy distribution value [4]. A permanent component (Fig. 6, b) used to indicate the repetition - “sync”, beginning and end of the message can be found in the channel of communication. Information symbols “0” (Fig. 6,a) and “1” (Fig. 6,c) are placed in line with the noise signal with manipulable expectation.

![Fig. 6](image)

Fig. 6. Presentation of information messages with a manipulation of quasi-three signals with variable entropy: a) "1", b) "sync", c) "0".

An example of quasi-three signal with variable entropy for information message with a byte size is shown in the fig. 7, where: I - frame structure II - implementation of physical level of entropy manipulated signals, III - quasi-three code of manipulated signals, IV - Gauss distribution characteristics of signals with variable entropy. As seen from the fig. 7, with a help of quasi-three signal we can organize the "start" and "stop" bits without changing the signal structure, as well as eliminate repetition of information symbols, which in its turn provides for a high-quality bit synchronization.

![Fig. 7](image)

Fig. 7. Example of a quasi-three signal with variable entropy for informational messages up to 1 byte

### III. CONCLUSION

The concept of the theory and techniques of structuring data in information and telecommunication systems based on entropy approach of digital forming, transmitting and processing of information flows are described.

Wide demonstrated scope of the reduced methods of structuring data in the computerized and telecommunication systems determines the perspective for the further research in this area in order to improve the characteristics of the system software and hardware and special processor facilities for forming and digital data processing based on the information measure of entropy.

### REFERENCES


