Non-Digital Information Processing in Biotechnical Systems with Biofeedback

Vladimir Mescheryakov[1], Dmitry Mescheryakov [2], Anna Gnatovskaya [3], Dmitry Kondratyuk [4], Aleksandr Salabash [5]

[1,3,4,5] *Odessa State Environmental University, Address, Lvovskaya st., 15, Odessa, 65016, Ukraine*
[2] *JSC PETROSOFT, Address, M. Govorova st., 18a, Odessa, 65032, Ukraine*

Abstract

The possibility of using statistical non-digital data processing in homeostatic biotechnical systems with biofeedback is considered. It is shown that the non-digital representation of the primary data of the connection between the input action and the output response of the organism is more consistent with the model of the physiological system. The impossibility of using arithmetic transformations when processing non-digital data for sliding windows with a small number of samples has shown the advantages of using the Kemeny median. Comparative analysis of data processing with abnormal emissions by sliding linear and nonlinear filters operating in real time mode of biological object functioning is carried out. The advantage of using nonlinear filtering when working with time sequences containing anomalous sample values has been substantiated. For biotechnical regulatory systems for medical purposes with biofeedback, an option for making decisions on the criterion of signs is presented, which increases the stability of the system.

Keywords

Feature space, non-digital data, median of Kemeny, anomalous outliers, filtering, fuzzy transformations.

1. Introduction

Methods of non-digital statistics [1] are used in expert systems in decision-making, sociology, political science [2], psychology, in areas where there are no or difficult opportunities for unambiguous decision-making [3]. Consider the possibility of using non-numerical statistics to convert the original human signals as a reaction to the intensity of infrared radiation[2] [4].

The original feature space of a biological object is probable due to the low level of output signals, the application of heterogeneous signals, noise and external influences [5]. This involves the use of statistical methods of data processing and taking into account the fact that the resulting sequence is non-stationary [6]. Nonstationary leads to the need to allocate quasi-stationary sections, where you can select the moments of the stationary sequence [7].

In addition, sign signals, such as the resistance of the skin, obtained in different parts of the body, have a significant variance [8]. This is due to the uneven location on the surface of the skin of the sweat glands, different thickness of the epidermis, etc. Moreover, there is no direct relationship between skin resistance and sweat gland activation, and a change in resistance twice does not mean that their activity has changed proportionally. It follows that the original data

ISIT 2021: II International Scientific and Practical Conference «Intellectual Systems and Information Technologies», September 13–19, 2021, Odesa, Ukraine
EMAIL: gradan@ua.fm (A.1); Dmescheryakov1135@gmail.com (A.2); aninfo2000@gmail.com (A.3); dumitrukondratiuk@gmail.com (A.4); aleksandr1996@gmail.com (A.5);
ORCID: 0000-0003-0499-827X (A.1); 0000-0001-7224-749X (A.2); 0000-0002-0018-5696 (A.3); 0000-0002-7411-8509 (A.4); 0000-0002-3708-0505 (A.5).

carry information about physiological processes in the body, but these data are qualitative rather than quantitative [9]. The resistance of the skin is not equivalent to the activity of the sweat glands, so the binding of the physical readings of the device is also qualitative rather than quantitative.

A fundamentally different principle of information processing in biological and technical objects is a significant problem of data processing in homeostatic biotechnical systems [10]. Where decision-making technical component is made on the basis of physiological reactions of the organism [11]. The aim of this work is to improve the quality of pre-processing of data in the biotechnical system by using statistical non-digital methods of time series processing.

2. Non-digital approach to processing the original feature information

The qualitative nature of the samples in contrast to the quantitative representation has its own characteristics. Thus, the values of the samples cannot be made, because the obtained values lose their meaning. If we assume that the sequence of samples are elements $x_1, x_2, \ldots, x_n$ of a nonlinear set X, then under these restrictions it becomes clear that the determination of the average value of the sample requires other approaches compared to those adopted. Even when analyzing the samples of a series, the arithmetic mean value is acceptable only for the case of a sufficiently uniform value of the members of the series. If there is an anomalous value in the sample, the value of the arithmetic mean does not always adequately characterize the average, because the influence of this component is much more significant than others.

In non-digital statistics, the measure of difference is an indicator $d : X^2 \to [0, +\infty]$, the essence of which is to capture the fact that the more $d(x, y)$, the more different x and y [12]. In relation to the empirical mean, this means minimizing the expression:

$$E_n(d) = Arg\min\left\{\sum_{1 \le i \le n} d(x_i, x), x \in X\right\}, \quad (1)$$

where the mean $E_n(d)$ represents the set $x \in X$ for which the function

$$f_n(x) = \frac{1}{n}\sum_{1 \le i \le n} d(x_i, x), \quad (2)$$

reaches the minimum value on the set X and is the median or mean for a sample of rankings by Kemeny. In [12] it is shown that for qualitative values for the ordinal scale as a mean it is possible to use only the median, and not the arithmetic mean or geometric mean. Proof of the convergence of theoretical and empirical averages is based on the law of large numbers. With a limited sample, the concept of a ε-heel f is introduced, which is a neighborhood in terms $Arg\min(f)$ of a function that is minimized. This, in particular, removes the question of choosing metrics in space X. The size ε of the area is determined both by the accuracy of determining the values and by the sensitivity thresholds used, if the modulus of the difference between the samples is less than or equal to the sensitivity threshold.

It also follows from the peculiarities of qualitative representation that the increase in the sample size may not lead to an increase in the reliability of the assessment, as it is impossible to talk about the stationary and centeredness of the analyzed process, along with the negative consequences of such a statement samples. This fact is critical for real-time systems, as it introduces a delay of at least half the sampling time. The small sample size leads to a significant variation of the indicators relative to the average, because, for example, for control systems, the indicator of stability is important.

If we consider the stability as the absence of control effects on the tolerances, the reaction of a biological object of the type "cold-warm", or "comfortable-uncomfortable" is more stable than the perception of the values of ambient temperature. The feeling of warmth is perceived by each person individually, the physiological reaction of the organism is primary, and the quantitative description of conditions is secondary.

The scales of qualitative features are the ordinal scale and the scale of names [2a], the first of which corresponds to the problem to be solved. Comparison of the two samples Y and Z can be done by their average values:

$$f(Y_1, Y_2, \ldots, Y_n) < f(Z_1, Z_2, \ldots, Z_n). \quad (3)$$

If the transformation in the ordinal scale ρ, such Y_i as Z_i normalization, is allowed, then $\rho(Y_i)$ and $\rho(Z_i)$ change to and.

To form the average of the data set, you can use the sign of the distance from a given point to the points of the neighborhood, and the degree of proximity are smaller distances. Since it is not possible to use the summation operation for qualitative values, we use the difference indicator. For problems with a limited sample, it is necessary to determine the empirical average, which under certain conditions provides convergence with the theoretical average.

For a space of arbitrary form X with elements $x_1, x_2, \dots, x_n$ of a real-valued function $f(x, y)$ with value in X, the values of the difference function differ $f(x, y)$ the more, the more x and y differ. The average value $\bar{x}$ relative to the degree of difference $f(x, y)$ is the solution of the optimization problem [12]:

$$\sum_{i=1}^{n} f(x, y) \to \min, \quad y \in X . \qquad (4)$$

The theoretical average does not differ from the classical average for the law of large numbers when $n \to \infty$, in accordance with Hinchin's theorem, tends to a mathematical expectation:

$$\frac{1}{n}\sum_{i=1}^{n} f(x_i, y) \to Mf(x, y) . \qquad (5)$$

When $f(x, y) = |x - y|$ and with an odd number of samples $n = 2k + 1$, the value of the mean is equal to $\bar{x} = x_{k+1}$, i.e. we obtain a sample median. With an even number of sample members, we obtain the half-sum of the sample values x_k and x_{k+1}. To exclude arithmetic operations, you can limit the odd number of samples.

To determine the average of Kemeny, it is necessary to rank the data. The filter delay for real-time systems is determined by half of the sample, so it cannot be large. With a limited sample, the ranking operation consists in arranging the data in a non-killing order, ie increasing with the possibility of the existence of elements with the same values. Algorithms for implementing this function are known and consist in a sequential comparison of the current element of the sample with the elements constructed in ascending order. Next, the median is determined, which is the average of Kemeny [13].

The qualitative nature of the samples leads to nonparametric models of process description. The parametric probability-statistical model is represented by a vector of fixed dimension, which does not depend on the sample size. In nonparametric models, the notion of distribution density is unacceptable, so it can be replaced by the probability of ε-hitting the -region. The formation of the ε-area in the simplest case can be the setting of the noise level and the signal being processed, the required sensitivity or other criteria. In the initial stage it is possible to provide a variant of asymptotic approach to the purpose of regulation, and then to specify depending on existing restrictions. This solution will allow us to talk about the ability of the proposed approach.

For the task of controlling the intensity of human infrared radiation on its physiological characteristics, it is important that the stability of the determination of the physiological response is higher than the numerical values of the devices, because it is primary. It is obvious that the use in addition to the resistance of the skin, other signs of human response to infrared radiation allows you to maintain the nature of these signals. Thus, heart rate and respiration only indirectly reflect the fact of increased heat extraction by the peripheral vascular and respiratory systems.

In a biological object, the response to each reaction is accompanied by the formation of an elementary goal, the implementation of an active act, checking its achievement, adjusting the elementary goal, and so on. These actions take place within the framework of a higher level goal, such as maintaining the temperature conditions of the body's functioning. Naturally, this is a very simplified model, but it can serve as a basis for reconciling mathematical and functional models.

Sampling $x_1, x_2, \dots, x_n$ of the initial probable size X of a biological object due to the above reasons cannot have a known distribution function. As the sample size increases $F(x)$, according to the central limit theorem, the distribution function tends to the normal distribution law. For a non-parametric model, the most appropriate solution to the decision-making problem are two criteria: the criterion of signs and the criterion of sign ranks [13]. For the sign criterion $F(m) = 0{,}5$, i.e. each of the random variables is probably more than the second sample:

$$R_j = \begin{cases} -1, & если \quad x_j < m_0 \\ +1, & если \quad x_j > m_0 \end{cases}. \qquad (6)$$

If the value m_0 corresponds to the response of the sensor in the active part of the process, the presented connection reflects the decision in the vicinity of the active therapeutic zone. The decision "-1" indicates that the radiation intensity must be increased, "+1" – the radiation intensity must be reduced, "" $x_j = m_0$ – to remain unchanged. This approach is known as the principle of follow-up balancing. The main advantage of the following balance is the high stability of the conversion at a low signal-to-noise ratio, the disadvantage is the low speed of entering the mode. If the latter disadvantage is not fundamental, for example, due to the preheating of the emitters before the procedure, the decision-making on the control of infrared emitters may be limited to the criterion of signs in this case. Thus, the qualitative representation of the initial features of the biological object, which is in the feedback circuit of the biotechnical system, allows to form requests for control of the intensity of infrared radiation by the physiological response of the organism. According to the initial physiological information, both reactions to external thermal influence, and to internal adaptation of an organism to own purposes and the executed processes, change of external influence taking into account ambiguity of transformation is carried out. The implementation of the presented variant of non-digital data processing for controlling the intensity of heating of the patient in the infrared peloidotherapy chamber is made on the basis of ARDUINO technology. Limitations of radiation intensity of infrared emitters were chosen on condition of discomfort of stay indoors. This condition is met by the value of the resistance of the leather above 400 Koh. From the point of view of carrying out procedure the range is not of interest as dry epithelium signals absence of the remains of heat in an organism. The value of skin resistance is less than 100 ohms with significant sweating close to pain and can serve as the upper limit of the heating intensity range.

2.1. Processing of time series with anomalous

Modern methods of filtering anomalous emissions are based on the calculation of a sample variance followed by data substitution, in which the deviation exceeds the threshold for noise-tolerant estimation of mathematical expectation, or on rejection of emissions using statistical hypothesis testing, which provides a sufficient number of members in biotechnical systems [14].

We compare the processing of the time series of a nonstationary process, which is the characteristic information of a biological object (heart rate read with a period of 20 seconds), the sliding linear window in determining the sampling center as the arithmetic mean and the median of Kemeny. From the given fragment of time sequence of samples it follows that process cannot be carried to a stationary series (Figure 1). The obtained data are influenced by obstacles, cyclic processes, trendy long-term processes that take place in the body during its functioning, so in the short term the time series is non-stationary. It is impossible to increase the size of the sample window due to the increase of the delay time on the analyzed effect.

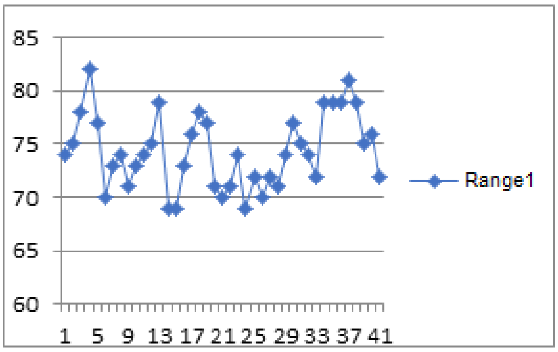

Figure 1: Non-stationary source feature of a biological object

We process this fragment with a 5-element sliding window (Figure 2).

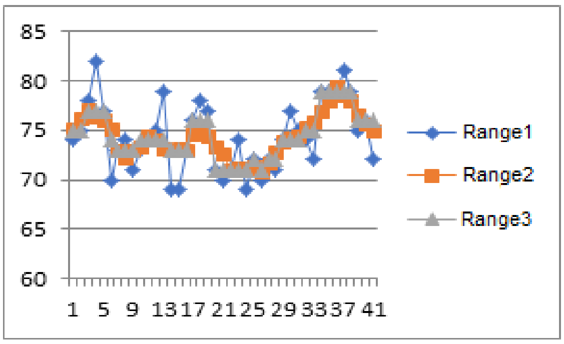

Figure 2: Processing a fragment of the original features (1) with a linear window (2) and a nonlinear filter Kemeny (3)

We make it with the replacement of the arithmetic mean value and the average Kemen average ranking of the sample values in the window and the selection of the median.

Analysis of curves 2 and 3 shows that there is no significant difference between them, except for sections 17–19 and 35–37, which requires a separate study and comparison of the reactions of linear and nonlinear filters. Methods of research of filters are worked out in detail in the literature on filtering of signals and time series therefore their detailed analysis to result in the given work can be considered inexpedient. One of the main problems of data processing, which is characterized by the ambiguity of the values due to the reaction of the biological object to the impact, is the rejection of abnormal values, or emissions. The most commonly used method of filtering anomalous emissions is to calculate sample variances with subsequent replacement of data in which the deviation of the mean exceeds a certain specified value of the calculated variance. The general approach to emission rejection is to use noise-tolerant assessment and test statistical hypotheses. The presented simulation experiments showed that the proposed filtering of anomalous measurements effectively works up to 18% of single emissions, and at 20% no longer works [15].

The reading of primary information by contact means from mobile patients is associated with problems with the conductivity of contact connections of the epidermis with electrodes, which leads to uncontrolled changes in the recorded data, ie the appearance of artifacts in the time series. When a person moves in a peloidotherapy chamber, the skin is bent in the places of reading the primary data, ie the appearance of various contact resistances is the basis of the physiotherapeutic method.

To assess the degree of influence of anomalous emissions on the possibility of using the results of primary features on the control capabilities in the experimentally obtained nonstationary series, we make anomalous emissions and process them with a linear and nonlinear filter (Figure 3–7).

Analysis of figure 3 shows that when processing a time series with single anomalous emissions by a linear filter, the effect of the anomalous component is significant because it is included in the arithmetic mean and shifts the filtered value towards the emission. When processing the time sequence by a nonlinear filter, the anomalous emissions have almost no effect on the results, because a single emission can only affect the displacement of the selected element on the neighboring ranked from the median.

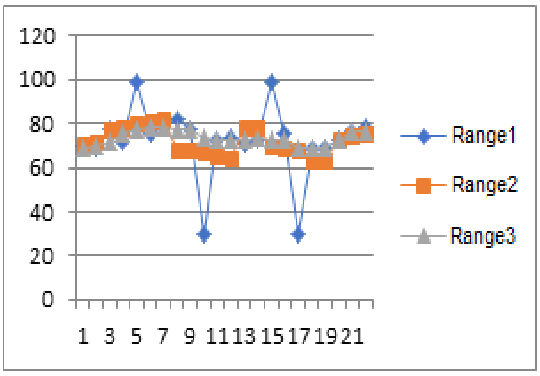

Figure 3: Time series processing with single anomalous emissions: 1 – primary series with anomalous emissions, 2 – linear filter treatment, 3 – nonlinear filter treatment

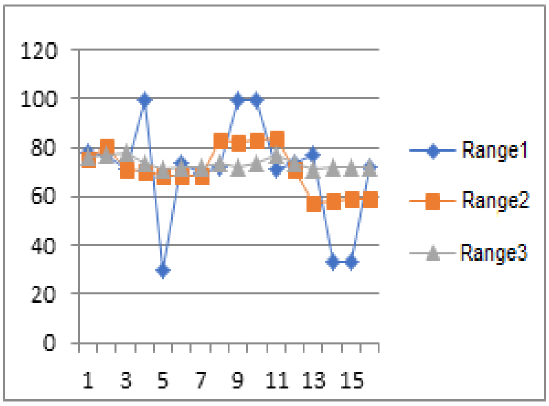

Figure 4: Time series processing with double anomalous emissions: 1 – primary series with anomalous emissions, 2 – linear filter treatment, 3 – nonlinear filter treatment

In figure 4 presents the results of time series processing of paired anomalous emissions by linear and nonlinear filters. The results of linear filtration show that anomalous emissions significantly affect the results, shifting the filtered curve towards anomalous emissions more than for single emissions. This is obvious, because the arithmetic mean value significantly depends on the anomalous emissions that are emitted at the level of other informative members of the series. When treated with a nonlinear 5-point filter, the effect of paired anomalous emissions is insignificant and is associated only with the homogeneity of the three remaining informative members of the series. Accordingly, in contrast to

[15], the limit of anomalous emissions is not 18%, 40% for a five-point filter.

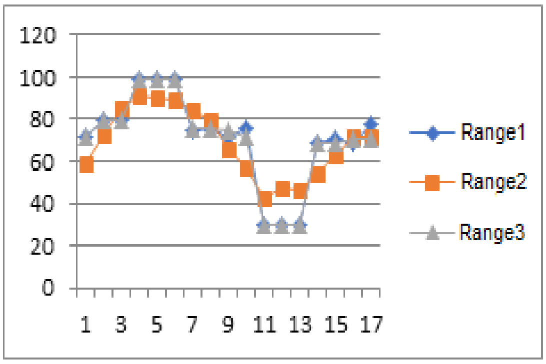

Figure 5: Time series processing with triple anomalous emissions: 1 – primary series with anomalous emissions, 2 – linear filter treatment, 3 – nonlinear filter treatment

In figure 5 presents the results of processing for three consecutive abnormal emissions of one sign relative to the signal. As the results of processing by a nonlinear window show, the filter does not cope with the task, because the median is anomalous value. That is, a filter with an odd number of elements (2n + 1) is operational provided that the number of consecutive anomalous emissions does not exceed n.

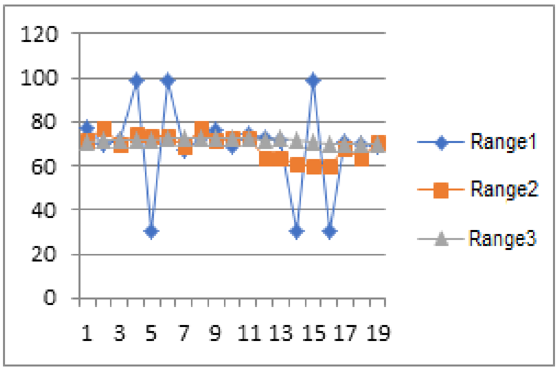

Figure 6: Time series processing with anomalous emissions alternating by sign: 1 – primary series with anomalous emissions, 2 – linear filter treatment, 3 – nonlinear filter treatment.

In figure 6 presents a variant of anomalous emissions consecutive on the sign, alternating. As shown by the results of treatment with a nonlinear filter, there is a mutual compensation of anomalous emissions, and the above condition for the number of anomalous emissions n can be exceeded.

In figure 7 presents the resulting processing of the time series, which contains anomalous emissions of different signs and durations.

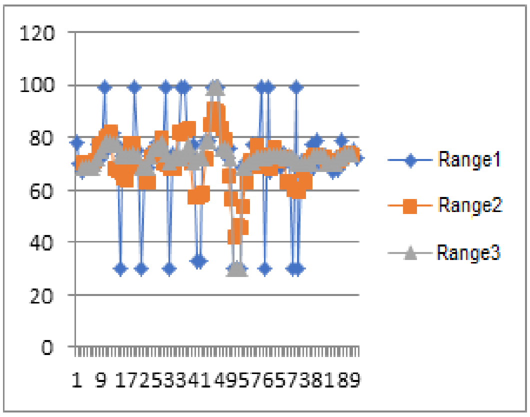

Figure 7: Time series processing with anomalous emissions of different duration: 1 – primary series with anomalous emissions, 2 – linear filter treatment, 3 – nonlinear filter treatment

In figure 8 presents the results of processing the experimentally obtained time series of values of the resistance of the skin under the influence of infrared radiation when moving the patient inside the chamber.

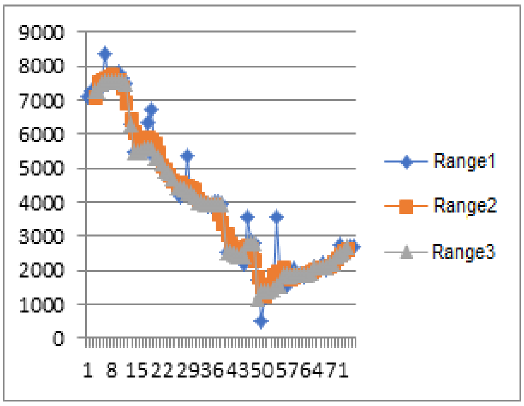

Figure 8: Processing of experimental data of a leather cover: 1 – primary signal, 2 – linear filtration, 3 – nonlinear filtration

The analysis of the primary data shows that, according to the criterion of stability of control, they are not suitable for direct use in a feedback system. Linear filter treatment showed that the effect of anomalous emissions is significant and significantly depends on the amplitude of the anomalous emissions. Non-linear filter treatments provide clear treatment for anomalous emissions, resulting in sustainable results that can be used for biological feedback systems.

It also follows from the above analysis that the proposed method of data processing allowed to process information that contains the uncertainty of the reaction of the biological object to the input effect.

A feature of nonlinear filtering compared to linear is the high dynamics of the process, because the median is not affected by neighboring ranked samples.

2.2. Fuzzy time series processing

The implementation of a fuzzy decision-making algorithm based on the resistance of the leather cover allows you to change the number of pulses of infrared emitters from ± 1 the central "zero" zone to the maximum, for example ± 8, in the upper and lower boundary zones. This allows you to quickly get out of uncomfortable areas and ensure the stability of the therapeutic area.

Fuzzy processing of information in non-digital form is important. Consider the function of belonging $\mu_A(x)$ to a fuzzy set A of elements x from a set X, in relation to the decision problem in the following interpretation. Define the membership function $\mu_A(x)$ as the degree x of proximity A to the prototype or similarity of affiliation $A = \{x, \mu_A(x)\}$. Then A represents a set of alternatives, and the $\mu_A(x)$ degree of preference and suitability of the choice as the value of the variable b. In this interpretation, the membership function plays the role of the ordering relation associated with the predicate A relation $x \geq_A x'$, which shows that x it corresponds more to another value x' of the same parameter in the current situation A. Continuing these considerations, we can show that inequality $\mu_Q(x, x') \geq \mu_Q(x, x'')$ describes a situation in which this expression means closer x' to x than $x' \geq_x x''$. Alternatives can be represented as fuzzy sets on a non-numerical scale, then a fuzzy set $B = \{(b, B(b)\}$ in the form where $(b, B(b))$ the set of fuzzy objects.

In static mode, it was possible to divide the core into components, which require a more detailed analysis of this area to improve the quality of decision-making. Thus, the use of fuzzy conversion of the original data allowed the use of biotechnical systems with biological feedback as a system for maintaining the intensity of infrared radiation for an individual patient according to the characteristic physiological response of the patient. The vague presentation of the characteristic information of the biological object and its processing by the technical component of the biotechnical system indicates a deeper overlap of functional and cybernetic models, ie the potential emergence of the emergence effect, for example in the form of new treatments [16].

3. Conclusions

1. The expediency of qualitative representation of the initial features is substantiated and the flow of the original features is processed by the methods of non – numerical statistics with the determination of the average in the sliding window as the medians of Kemeny.

2. A comparison of the results of processing non-stationary feature data with anomalous emissions typical of biological objects, linear and nonlinear filters showed that linear filters are inoperable. The use of nonlinear filters allowed to process time series with the number of anomalous emissions up to 40% of the number of samples in the window compared to 18% for existing anomalous emission filters.

3. It is shown that the levels of resistance signs in the central and peripheral zones differ more than 2 times, and the proposed methods of non-digital representation of information in conjunction with fuzzy logic provide information processing almost invariant to the scatter of the level of signs.

4. References

[1] Arroniz I. Extracting Quantitative Information from Nonnumeric Marketing Data: An Augmented latent Semantic Analysis Approach. (2007). Electronic Theses and Dissertations, 2004-2019. 3068. https://stars.library.ucf.edu/etd/3068.

[2] Leong, E. K. F., Ewing, M. T., & Pitt, L. F. (2004). Analyzing competitors' online persuasive themes with text mining. Marketing Intelligence and Planning, 22(2), 187–200.

[3] Rust, R. T., Lemon, K. N., & Zeithaml, V. A. (2004). Return on marketing: Using customer equity to focus marketing strategy. Journal of Marketing, 68(1), 109–127.

[4] Khudenko N.P., Mescheryakov D.V. Analysis of information of an object in infrared control system with biological feedback. Engineering sciences: Development prospects in countries of Europe at the Third Millennium. Collective monograph. Poland. V 2. 2018. 492 p. (p. 410–427).
[5] Khalifa Y., Coyle J.I. and Sejdic E. Non-invasive identification of swallows via deep learning in high resolution cervical auscultation recordings. Scientific Reports, volume 10, Article number: 8704 (2020). https://www.nature.com/articles/s41598-020-65492-1.
[6] Mescheryakov D. Usage of cluster analysis in a system management with a biological feedback. The scientific heritage. Hungary. № 22. 2018. P. 59–61.
[7] Burns, D., Leung, N.; Hardisty, M., Whyne, C.; Henry, P. and McLachlin, S. Shoulder Physiotherapy Exercise Recognition: Machine Learning the Inertial Signals from a Smartwatch. Physiol. Meas. 2018, 39, 075007. [Google Scholar] [CrossRef] [PubMed].
[8] Marzbani H., Marateb H. and Mansourian M. Neurofeedback: A Comprehensive Review on System Design, Methodology and Clinical Applications. Basic Clin Neurosci. 2016 Apr; 7(2): 143–158. doi: 10.15412/J.BCN.03070208.
[9] Richards, R.; van den Noort, J.C.; Dekker, J.; Harlaar, J. Gait Retraining With Real-Time Biofeedback to Reduce Knee Adduction Moment: Systematic Review of Effects and Methods Used. Arch. Phys. Med. Rehabil. 2017, 98, 137–150. [Google Scholar] [CrossRef] [PubMed].
[10] Voit E.O. Biochemical Systems Theory: A Review. International Scholarly Research Notices. 2013. URL: https://doi.org /10.1155/2013/897658.
[11] Blanquero, J.; Cortés-Vega, M.D.; García-Frasquet, M.Á.; Sánchez-Laulhé, P.R.; Nieto Díaz de los Bernardos, M.I.; Suero-Pineda, A. Exercises Using a Touchscreen Tablet Application Improved Functional Ability More than an Exercise Program Prescribed on Paper in People after Surgical Carpal Tunnel Release: A Randomised Trial. J. Physiother. 2019, 65, 81–87. [Google Scholar] [CrossRef].
[12] Orlov A.I. Non-numerical statistics. M.: MZ-Press. 2004. 513 p.
[13] William W.S. Wei. Multivariate Time Series Analysis and Applications. Wiley. 2019. 680 p. ISBN: 978-1-119-50285-2.
[14] Bhanu Chander, G.Kumaravelan. Outlier detection strategies for WSNs: A survey. ournal of King Saud University – Computer and Information Sciences. 2021. 23 p. URL: https://doi.org/10.1016 /j.jksuci. 2021.02.012.
[15] Serysheva I.A. Emission filtering in statistical and dynamic data processing problems in time and frequency standards. Bulletin of IrSTU. Volume 2. № 10. 2018. pp. 67 – 77.
[16] Lobachev I.V., Mescheryakov D.V. Information significance of the skin's response to infrared radiation. XII International scientific-practical conference "DYNAMICS OF THE DEVELOPMENT OF WORLD SCIENCE". August 5–7, 2020. Vancouver, Canada. Pp. 265–270.