The Method of Detection of Speech Process Signs in the Structure of Electroencephalographic Signals

Vasil Dozorskyi[a], Oksana Dozorska[a], Evhenia Yavorska[a], Leonid Dediv[a] and Andrii Kubashok[a]

[a] *Ternopil Ivan Puluj National Technical University, Ruska str., 56, Ternopil, 46001, Ukraine*

Abstract

The article is devoted to the issues of mathematical modeling of electroencephalographic signals for the problem of impaired human communicative function compensation. In particular, the choice of a mathematical model in the form of a piecewise stationary random process, which is adequate to both the physical nature of such signals and the problem, was made, and the method of processing of electroencephalographic signals was developed to identify signs of speech process in their structure. The developed method is based on the application of methods of spectral-correlation analysis of stationary random processes and the sliding window method. According to the developed method, within each broadcast of the sliding window the calculation of the average estimates of the power spectral density distribution is performed. The values of these assessments are compared with the threshold values for the state of rest and the state when the patient tries to say something. These threshold values are obtained at the preparatory stage, which is described in the article. The experimentally selected electroencephalographic signals were processed and it was found that according to the obtained average estimates of power spectral density distribution it is possible to distinguish between state of rest and the state when patient's tries to say something, i.e. the method allows to detect signs of speech process in electroencephalographic signals. Fisher's criterion was used to assess the reliability of the obtained results, which confirmed the high agreement of the experimental results with theoretical assumptions.

Keywords [1]

Communicative function, electroencephalographic signal, mathematical model, piecewise stacionary random process

1. Introduction

The ability to exchange information between people is provided through communicative function, in the process of which a significant number of organs are involved that form a complex system [1, 2]. There is an increase in the number of people with limited or lost communicative function, in particular due to dysfunction of organs or systems of the human body that perform this function. Therefore, it is important for medicine to find ways to indirectly compensate the impaired human communicative function.

To solve such problem, technical means of speech correction (speech therapy simulators and visualizers) or technical means of partial compensation of the impaired communicative function can be used. However, the disadvantages of such systems are the lack of medical equipment on the market, the high cost of individual orders, and the long time it takes to adapt the software to an individual patient.

ITTAP'2022: 2nd International Workshop on Information Technologies: Theoretical and Applied Problems, November 22–24, 2022, Ternopil, Ukraine

EMAIL: vasildozorskij1985@gmail.com (A. 1); oksana4elka@gmail.com (A. 2); yavorska_eb@yahoo.com (A. 3); dediv@ukr.net (A. 4); andriy.kubashok@gmail.com (A. 5);
ORCID: 0000-0001-6744-3015 (A. 1); 0000-0001-7053-863X (A. 2); 0000-0001-6341-1710 (A. 3); 0000-0002-2963-6948 (A. 4); 0000-0002-7504-461X (A. 5);

Accordingly, the development of new technical means of impaired communicative function compensation is urgent.

It is known that the main source of information about the operation of the system is the signal that is formed during the functioning of this system. Accordingly, it is possible to solve such problem by properly processing the biosignals that arise during the speech process. The processing methods will determine the algorithms of the functioning of the software of technical means for impaired communicative function compensation. With this approach, it is possible to use for subsequent processing the electroencephalographic (EEG) signals [3-5]; electromyographic (EMG) signals of facial muscles [6-8] or EMG signals registered from the surface of the patient's neck [9]. However, the processing of such signals independently of each other has significant drawbacks, which limited their practical application.

A promising method of impaired communicative function compensation is based on the selection and processing of two groups of biosignals: the first are electromyographic signals taken from the surface of the human neck near the vocal folds; the other group is electroencephalographic signals, locally selected from the areas of the patient's head surface, located near the speech centers. In the structure of the last group of signals, electrical images of nerve impulses will be displayed, which these centers will send to the corresponding organs of the vocal apparatus in the process of implementing the communicative function.

This principle is based on the method, proposed in the works [10-12]. Such method is based on synchronous selection and processing of EEG and EMG signals. In this case, EEG signals should be taken from the patient's head surface near the speech centers - Brock Center, Vernicke, and associative center. EMG signals should be selected from the neck surface near the vocal folds. At the same time, in accordance with the proposed method, in the structure of EEG signals it is necessary to identify the time intervals of the appearance of signs of the process of human communicative function implementing.

Taking into account the fact that the structure of EEG signals will reflect the changes in the electrical activity of the brain that is the result of the necessity of providing mechanisms of self-regulation and functioning of all organs and systems of the human body, implementation of the function of adaptation to the variables of external and internal factors affecting the human body, etc., as well as the representation of the EEG signal in the form of a superposition of all potentials of the action of all neurons with varying degrees of influence on the resulting EEG, the article focuses on the evaluation of the results of this group of biosignals processing.

The processing of signals and the functioning of information systems is implemented at the links of the triad "model - algorithm - software realization" triads [13]. In this case, the choice of model is a decisive factor: the model should combine important properties of the research object in its structure in accordance with the solvable problem, replace it with theoretical researches, be the basis for the organization of experimental researches and the basis for the processing and interpretation of their results. [13]. Accordingly, the mathematical model and methods for processing EEG signals should take into account the physical nature of such signals and have the means of detecting the signs of speech process in their structure.

2. Materials and Methods

The structure of EEG signals should show signs of realization of the communicative function. To substantiate the mathematical model of these signals, the following assumptions are made: 1) areas of EEG signals at rest - in the absence of speech process with constant additional factors (emotional state, human position in space, closed eyes, external conditions) - will be stationary; areas of EEG signals when patient tries to implement the communicative function will be stationary, but with different parameters from similar areas for rest (estimates of mathematical expectation, variance, etc.). The task of detecting the signs manifestations of the process of communicative function realization in the structure of EEG signals will be reduced to the task of detecting the changes in the properties of these biosignals characteristic of such a process.

In this case, it is advisable to use a piecewise stationary random process [11,14] as a mathematical model of EEG signals. This model is used to describe the basic dynamics of information activity of the brain and segmentation of EEG signals on such stationary areas. Thus, in 1977, Bodenstein and

Pretorius proposed a generalized concept of EEG structure, according to which the EEG signal consists of practically stationary segments, which are connected by rapid transients.

If we extend the model of a piecewise stationary random process to the class of EEG signals, so such biosignals can be given in the form [11, 12, 14]: $\Xi_n(t)=(\xi_1(t),\xi_2(t),\ldots,\xi_n(t))$, where: $\Xi_n(t)$ – random vector-process, specified on the interval $t\in[a,b]$; sequence of sets B_k, $k=\overline{1,n}$ is a breakdown of the interval $[a,b]$ by points $t_1,t_2,\ldots,t_k$; $I_{B_k}(t)$ – indicator function of the set B_k: $I_{B_k}(t)=\begin{cases}1, & t\notin B_k;\\ P, & t\in B_k,\end{cases}$, where P – coefficient, characterizing the change in the parameters of the process. Accordingly, the task is reduced to the task of detecting time points $t_1,t_2,\ldots,t_n$ of the occurrence of changes in parameters of stationary sections of EEG signals. Conditionally this is shown in Figure 1.

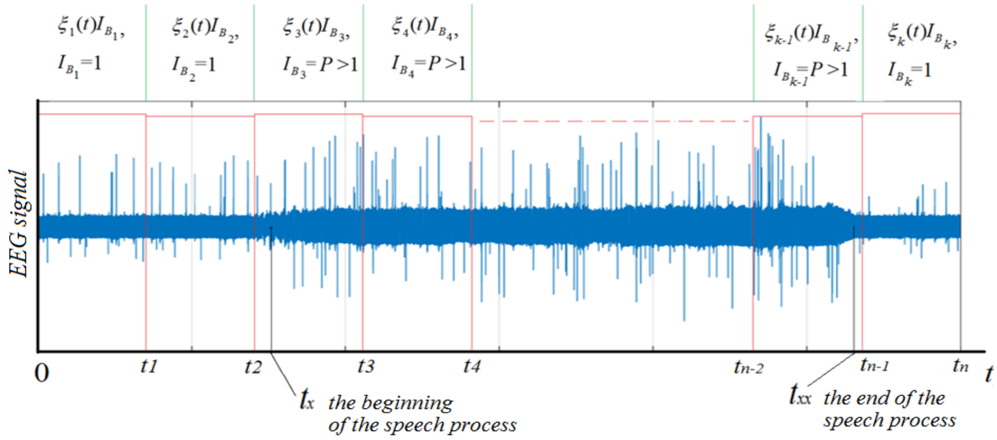

Figure 1: Detection of EEG signal changes, which characterizes the increase of brain activity – beginning and the end of the speech process

The basis of the proposed method is the use of spectral-correlation analysis methods, and to process the EEG and EMG signals at intervals of the specified duration (t_1-t_2, t_2-t_3,…,t_{n-1}-t_n) - within the sliding window. The developed method for processing of EEG signals to detect the time moments of appearance of signs of the speech process includes the following stages:

1) the formation of a sliding window of a given width, which is a broadcast of EEG signal in time;

2) within each sliding window, the power spectral density distribution (PSDD) of the EEG signal is estimated;

3) calculation of averaged PSDD estimates;

4) on the basis of the calculated averaged PSDD estimations the formation of a criterion for decision-making on existence of signs of communicative function realization is carried out. As a criterion, it is proposed to use the variation of the averaged PSDD estimates, assuming that the values of variation of these estimates for the rest state will be significantly different for the state of attempt to implement the communicative function and will not overlap.

The proposed method includes two stages: preparatory and basic. The purpose of the preparatory stage is the registration of EEG signals, when the patient tries to mentally pronounce certain test sounds and words at certain intervals. Through a set of test statistics and its subsequent analysis, ranges of numerical values of the averaged PSDD estimates for EEG areas during the mental pronunciation of test sounds and words and areas for rest are formed.

During the main stage, the EEG signals are constantly registered and the average PSDD estimates within the sliding window broadcasts are calculated. The numerical values of these estimates are used to calculate the time intervals of the presence or absence of speech process signs (trying to mentally pronounce arbitrary sounds or words or silence) based on the results of plasing these values in the appropriate ranges of numerical values of these estimates.

In the case of a discrete sequence of EEG values, the power spectral density (PSD) of the stationary sections of each signal can be estimated by performing a Fourier transform from the autocorrelation function of such sections. For this purpose, Wiener-Hinchin transformations were used, which connect the autocorrelation function of the signal and the PSD:

$$G_n(f) = \sum_{\tau=-\propto}^{\propto} R_n(\tau_x) e^{-i2\pi f \tau_x},$$
$$R_n(\tau_x) = M\big(x_n(k) \cdot x_n(k-\tau_x)\big), \tag{1}$$

where $M(\cdot)$ is the mathematical expectation

In the case of the sliding window method, the sample from the output signal is processed within each broadcast of the sliding window. To form such a sample, the following expression is proposed:

$$\xi_n = \xi[n\tau + 1,\ \ n\tau + T), \tag{2}$$

where ξ_n is n sample from the input signal ξ within n sliding window, $n = \overline{0, N}$, τ is shift of sliding window, T is the width of sliding window.

In the developed method of EEG signals processing, the estimations of the speech process implementation will be intervals (the method makes it possible to determine the time intervals within which these estimates will be present). The grounding of choosing the width of a sliding window is given in the papers [10-12].

3. Experiment and Results

The specialized Matlab software environment was used to process the experimental data. According to the proposed method of impaired human communicative function compensation, it is necessary to determine the time intervals of the beginning and the end of the speech process (time intervals of communicative function implementation) based on the results of EEG signal processing. To do this, EEG signals with signs of increased brain activity when patients mentally try to say something were loaded into the Matlab environment. Figure 2 shows the EEG signal with signs of increased brain activity after 18 seconds (increased signal amplitude). Rectangles conditionally denote a window that is broadcast in time and within which the signal is processed. The amount of shift of the sliding window in time is equal to the width of the window, i.e. the previous and next window do not overlap.

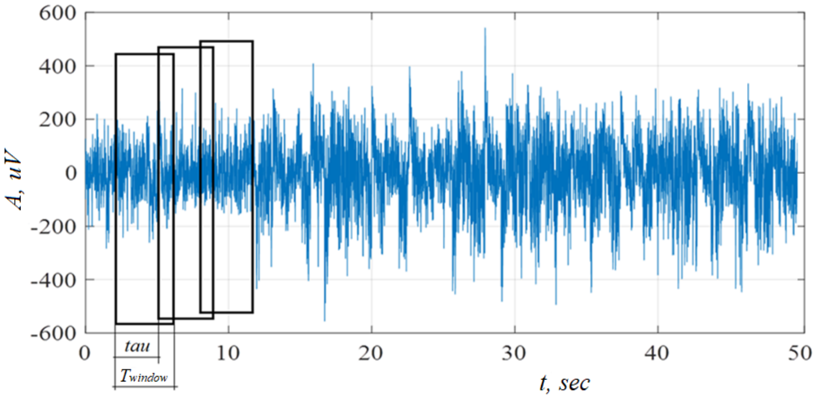

Figure 2: Broadcast of sliding window on the EEG signal

To establish an attempt to implement the speech process by the EEG signal, within each sliding window, the PSDD estimations and their subsequent averaging were performed:

$$\hat{M}_{\xi} = m_{\xi}(\hat{G}_n(\xi_{EEG})) \tag{3}$$

The assumption is made that the averaged PSDD estimates will be indicators of the beginning and the end of the speech process. At the same time, the value of these averaged estimates was postponed at one time axis together with the original EEG signal. The process itself for the formation of averaged PSDD estimates of the EEG signal within a sliding window is shown in Figure 2.

T_{window} are marked broadcasts of the sliding window, within which the EEG signal was processed (Figure 3, a). Within each window, the PSDD estimates of the signal were calculated (Figure 3, b). They were then averaged over frequency and power.

The obtained average estimates were plotted on the time axis in the middle of the time interval, which is equal to the width of the sliding window and the position of each individual broadcast of the window in time (Figure 3, c).

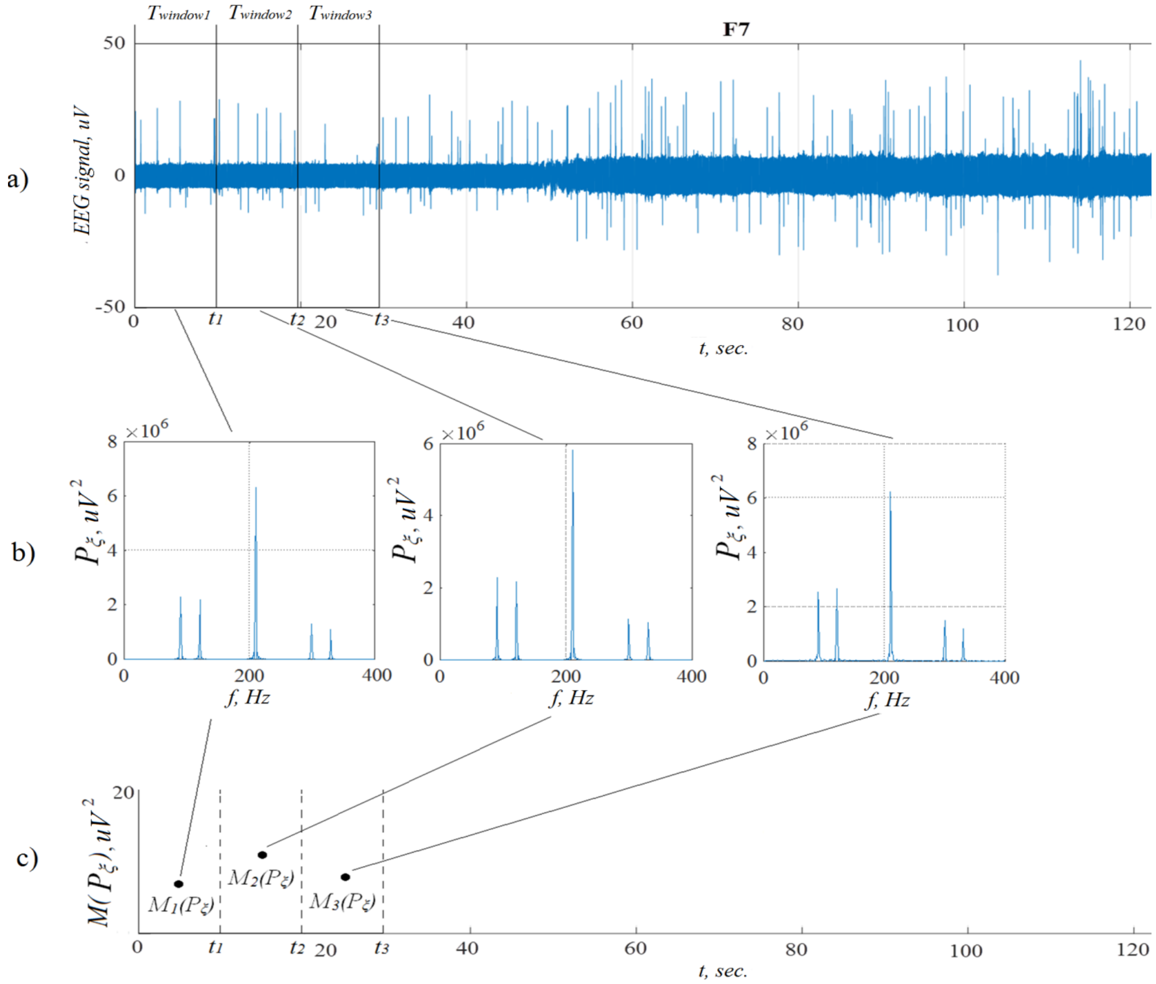

Figure 3: Formation of averaged PSDD estimates within a sliding window: a) - registered EEG signal with marked transmissions of the sliding window; b) - calculation of PSDD estimates of the EEG signal within the limits of each translation of the sliding window; c) - averaging calculated at the previous stage PSDD estimates and their deposition on one axis of time

Figure 4 shows a view of the EEG signal with signs of increased brain activity after 12 seconds and a graph of the averaged PSDD estimates, calculated within the broadcasts of the sliding window. The graphs show that the proposed averaged estimates are sensitive to the manifestations of changes in brain activity in the structure of the EEG signal.

As a criterion for determining the time intervals of attempts to implement the speech process, the variation of the averaged PSDD estimates $VAR(\hat{M}_\xi)$ is used.

It is established that the values of variation increase by more than an order of magnitude in the presence of signs of increased brain activity:

$$VAR(\hat{M}_\xi)_{rest} = (2{,}9522 \pm 10\%)uV^2, \quad VAR(\hat{M}_\xi)_{speech} = (83{,}3164 \pm 10\%)uV^2 \tag{4}$$

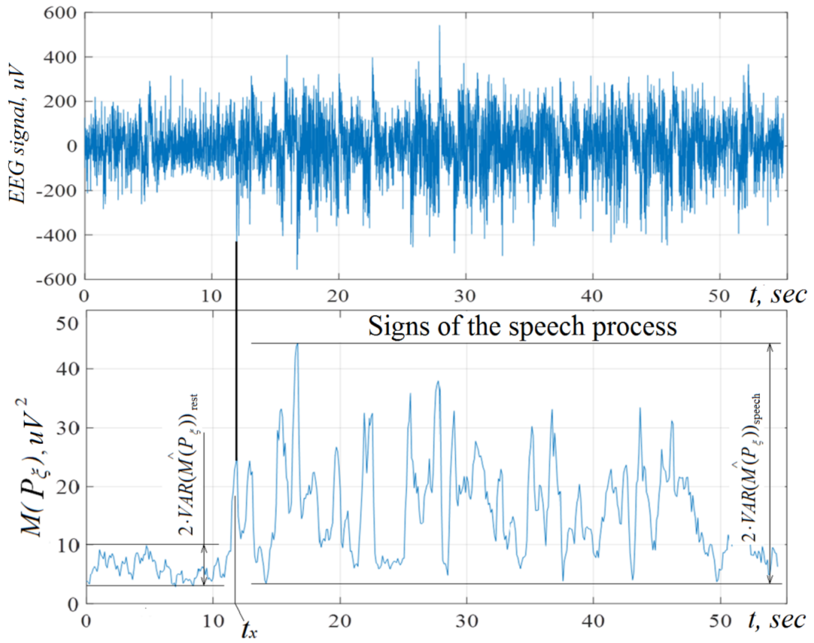

Figure 4: Registered EEG signal (upper figure) and averaged PSDD estimates (bottom figure), calculated within the sliding window, delayed at one time axis

Accordingly, the proposed criterion is sensitive and can be used to set the time intervals of attempts to implement the communicative function by the EEG signal.

4. Evaluation the reliability of results of EEG signals processing

As noted above, as a criterion for establishing time intervals of attempts to implement the communicative function, it is proposed to use it is proposed to use the values of the variation of averaged PSDD estimates the samples from the EEG signal. In this case, the zero hypothesis H_0 is put forward, that in a state of rest (the communicative function is not realized) the value of variation the averaged PSDD estimates will have a value of $d_{\xi 1}$. For the alternative hypothesis H_1, we made the assumption that at the time interval of attempts to implement the communicative function the values of variation will have a value of $d_{\xi 2}$, however $d_{\xi 1} \neq d_{\xi 2}$. That is, on segments of the stationary EEG signal, the sampled variations of the averaged PSDD estimates will be almost identical and differ for different pieces of stationary. To evaluate the statistical significance of the processing results and their reliability, the Fisher's criterion was used [15, 16], which makes it possible to compare the values of the sample variances of two series of observations. Let's denote by d_1 and d_2 the sample estimates of variations $d_{\xi 1}$ and $d_{\xi 2}$ respectively. Then the statistics of Fisher's criterion will be: $F = d_1 / d_2$.

By this value, using the table data, we can assess the significance of the results of the EEG signals processing, construct the axis of significance, form the conclusion of accepting the hypothesis H_0 or rejecting it in favor of the hypothesis H_1 and, accordingly, assessing the reliability of the decision.

At the statistical processing of data is usually asked by some level of significance α, which characterizes the probability of appearance the error of first kind. Applying Fisher's criterion, we can estimate the level of significance of the processing results. If we denote the probability of error in the second kind through β, then the value of $1-\beta$ is called the power of the criterion [15], and in this respect, when comparing the variances of two numeric rows, which can be presented as two samples of EEG signals, Fisher's criterion is a powerful criterion [15].

To use Fischer's criterion, the sample variations of the sequence of averaged PSDD estimates was evaluated. In Figure 5 marked by $d_1 - d_5$ the sample variations with volume of 10 values, and: d_1=4,2465 uV^2, d_2=4,4063 uV^2, d_3=34,1038 uV^2, d_4=94,3297 uV^2, d_5=103,507 uV^2. By the requirement of the process of calculating the value of Fisher's criterion is the ratio of the greater value of the variation to the smaller.

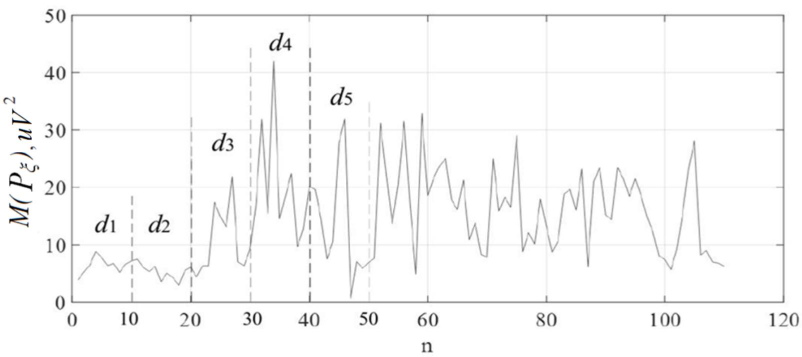

Figure 5: The chart of averaged PSDD estimates and the areas of sample variations calculation $d_1 - d_5$

Four values of Fischer's criterion were calculated:

$$F_1 = \frac{d_2}{d_1} = 1{,}0376; \quad F_2 = \frac{d_3}{d_2} = 7{,}7398; \quad F_3 = \frac{d_4}{d_3} = 2{,}7659; \quad F_4 = \frac{d_5}{d_4} = 1{,}0973;$$

The value of criterion F_1 was calculated for two sample variations, corresponding to the region of the EEG signal in the state of rest. In this case, the null hypothesis about the insignificant difference between the two variations must be confirmed.

The value of criterion F_2 was calculated for two sample variations, the first of which corresponds to the region of the EEG signal in the state of rest, and the second - in the state of implementation the communicative function. In this case, the null hypothesis should be rejected in favor of the alternative since the difference between the values of sample variances is significant.

The values of criterion F_3 and F_4 were calculated for sample variations corresponding to the region of the EEG signal in the state of implementation the communicative function. In this case, the null hypothesis about the insignificant difference between these variations must also be confirmed.

To confirm the made assumptions, it is necessary that the values F_1, F_3 and F_4 fell into the zone of insignificance of the Fischer's criterion, and value F_2 – into the zone of significance. For this purpose, an axis of significance of the Fischer's criterion was constructed (Figure 6).

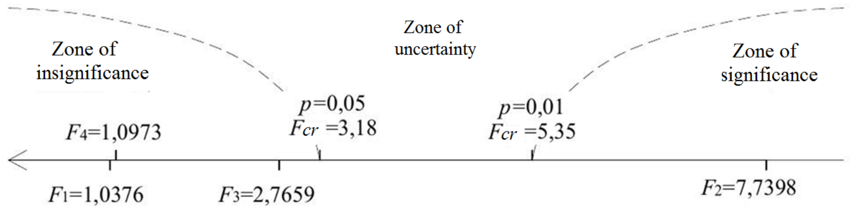

Figure 6: Axis of significance of Fisher's criterion

For the degrees of freedom n-1=9 and m-1=9 (n=m=10) of the samples according to the tabular data [15], were found the critical values of Fisher's criterion F_{cr} for probabilities of error $p = 0.05$ and $p = 0.01$. These values are deferred to one axis and indicate the zone of significance, the zone of insignificance and the zone of uncertainty. If the calculated value of Fisher's criterion falls into the zone of insignificance, then a zero hypothesis is accepted with a probability greater than 0,95; if the value of the criterion falls into the zone of significance, then the null hypothesis is rejected in favor of the alternative with the probability of 0,99; if the value of the criterion falls into the zone of uncertainty then we can not unambiguously accept or reject the null hypothesis.

In Figure 7 the calculated criteria values F_1-F_4 are delayed and it was found that the values of the criteria F_1, F_3 and F_4 fall into the zone of insignificance, and value F_2 – into the zone of significance of Fisher's criterion.

Accordingly, the use of the proposed criterion for determining the time intervals of implementation the communicative function by the variation of averaged PSDD estimates of the samples from the EEG signals gives the possibility with the reliability of $1\text{-}\alpha=1\text{-}0{,}01=0{,}99$ or in relative units - 99% increase brain activity in the process of communicative function implementing.

Previously, the assumption was made that the accuracy of determining the state of the process of communicative function implementation would increase if the next and previous windows, within which the averaged PSDD estimates are calculated, will be overlapped by the value *tau*. To confirm this assumption, an estimation of the decision-making reliability using Fisher's criterion was made, for which the estimation of sample variations was calculated for the averaged PSDD estimates at *tau*=0,2 T_{window}.

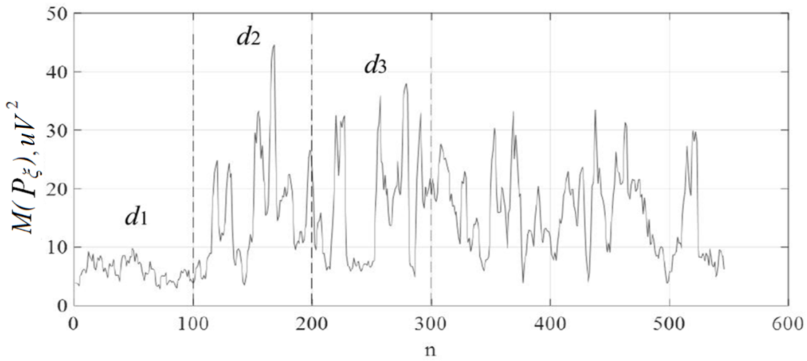

Figure 7: The chart of the averaged PSDD estimates and the area of calculation of the sample variations $d_1 - d_3$ at tau=0,2 T_{window}

In this case, the volume of samples was 100 values. For this number of freedom degrees by the tabular data [15] the critical values of Fisher's criterion F_{cr} were found for probabilities of error p=0,05 and p=0,01. The values of sample variations are calculated d_1=14,3310 uV^2, d_2=82,6879 uV^2, d_3=88,4858 uV^2 and the values F_1 and F_2:

$$F_1 = \frac{d_2}{d_1} = 5{,}7698; \qquad F_2 = \frac{d_3}{d_2} = 1{,}0701;$$

The obtained values are deferred to the significance axis (Figure 8).

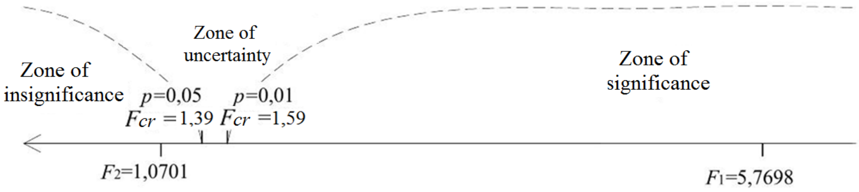

Figure 8: Axis of significance of Fisher's criterion

From Figure 8 it can be concluded that the value of F_1 is much further on the axis of the value F_{cr} at p=0,01 and accordingly, at larger volumes of samples, for which sample variations are calculated, the reliability of the decision is increasing.

5. Conclusion

In accordance with a grounded mathematical model of EEG signals in the form of a piecewise stationary random process, the method of statistical processing of such signals is developed for the task of detecting the time intervals of implementation the communicative function. The developed method is based on the application of methods of spectral correlation analysis and the method of a sliding window.

To establish the time intervals of implementation the communicative function by the EEG signal, a calculation of PSDD estimates was performed within each broadkast of a sliding window. Then the averaging of these estimates was calculated. It is established that the proposed averaged PSDD estimates are sensitive to manifestations of changes in brain activity in the structure of EEG signals in the implementation of the human communicative function. As a criterion for determining the time intervals of implementation the communicative function, the variation of the averaged PSDD estimates is used. It is established that the values of variation increase by more than an order of magnitude with the presence of signs of brain activity. The proposed criterion is sensitive.

The reliability of obtained results is evaluated on the basis of Fisher's criterion. It was established that the use of the proposed criterion for determining the time intervals of implementation the communicative function by the values of variation of the averaged PSDD estimates of the samples from the EEG signals gives an opportunity to establish with 99% reliability the state of increasing brain activity in the process of implementing the communicative function.

6. References

[1] Kashkin V.B.: Introduction to the theory of communication, PLINTA, 224 p. (2013).

[2] Remizov A.N., Maxina A.G., Potapenko A.Ya.: Medical and Biological Physics: Study. for high schools, Moskov, Drofa, 560 p. (2003).

[3] Porbadnigk A., Wester M., Schultz T.: EEG-Based Speech Recognition: Impact of Temporal Effects, 2nd International Conference on Bio-inspired Systems and Signal Processing, Porto, Portugal (2009).

[4] Geoffrey S. Meltzner, James T. Heaton, Yunbin Deng, Gianluca De Luca, Serge H. Roy, and Joshua C. Kline: Silent Speech Recognition as an Alternative Communication Device for Persons with Laryngectomy, IEEE/ACM Trans Audio Speech Lang Process, pp. 2386–2398 (2017).

[5] Herff C., Schultz T.: Automatic speech recognition from neural signals: A focused review. Frontiers in Neuroscience, 10, pp. 1-7 (2016).

[6] Wand M, Schmidhuber J.: Deep Neural Network Frontend for Continuous EMG-based Speech Recognition. Proc of the 17th Annual Conference of the International Speech Communication Association (Interspeech), pp. 3032-3036 (2016).

[7] Chuck Jorgensen, Diana D Lee, and Shane Agabon: Sub Auditory Speech Recognition Based on EMG/EPG Signals. Proceedings of the International Joint Conference on Neural Networks, pp. 1098-7576 (2003).

[8] Munna Khan, Mosarrat Jahan: Sub-vocal speech pattern recognition of Hindi alphabet with surface electromyography signal. Perspectives in Science, 8, pp. 558-560 (2016).

[9] Ambient Corporation. Buy Audeo Basic SDK. http://www.theaudeo.com/?action=buy.

[10] Dozorskyi V.G., Dozorska O.F., Yavorska E.B.: Selection and processing of biosignals for the task of Human Communicative Function Restoration, Kremenchug National University, 4(105), pp. 9-14 (2017).

[11] Vyacheslav Nykytyuk, Vasyl Dozorskyi, Oksana Dozorska: Detection of biomedical signals disruption using a sliding window, Scientific Journal of TNTU, Ternopil, 91(3), pp. 125–133 (2018).

[12] Oksana Dozorska: The mathematical model of electroencephalographic and electromyographic signals for the task of human communicative function restoration, Scientific Journal of TNTU, Ternopil, 92(4), pp. 126–132 (2018).

[13] Dragan Ya.P., Dozorsky V.G., Dediv I.Yu., Dediv L.E.: Principles and means of substantiation the methods of statistical processing of periodically correlated random process realizations, Lviv, Bulletin of the National University "Lviv Polytechnic", Computer Science and Information Technology, pp. 212-218 (2015).

[14] N.B. Marchenko, V.V. Nechiporuk, O.P. Nechiporuk, Yu.V. Pepa. Methodology of accuracy of information and information-viral systems of diagnostics. NAU, 377 p. (2014).

[15] Kobzar A.I. Applied mathematical statistics, Moscow: Fizmatlit, 816 p. (2006).

[16] J. Bendat, A. Pirsol. Applied analysis of random data. 540 p. (1989).