Оцінка рівня захищеності мовної інформації від витоку акустичними та вібраційними каналами виконується згідно міжнародних та національних стандартів, а також у відповідності з нормативними документами. Для оцінки рівня її захищеності, в багатьох країнах, нормативними документами передбачено використання коефіцієнту сигнал/завада. Однак метод має ряд значних недоліків, що не дозволяють визначити реальний стан рівня захищеності.

Запропоновано вдосконалений об'єктизований метод оцінювання, який ґрунтується на визначенні коефіцієнту залишкової розбірливості для тест-сигналу, після його відновлення методами математичного аналізу (адаптивної фільтрації, кореляційного та спектрального аналізів, вейвлет-перетворення та інше). Коефіцієнт залишкової розбірливості визначається для кожного слова, що входить до короткої фрази, тест-сигналу.

Проведено аналіз частоти вживання фонем в українській мові. Показано, що враховуючи визначення терміну «алофон» та кількість носіїв мови, можна вважати, що загальна кількість алофонів прагне до нескінченості. Для зменшення розрахункової складності використано формалізований підхід на основі спрощеної лінгвістичної моделі – фонема (буква), дифон (дві букви) та трифон (три букви). В якості джерела інформації, в такому випадку, можна використовувати текстові документи.

Запропоновано аналітичні залежності для розрахунку коефіцієнта залишкової розбірливості мови та його складових – коефіцієнтів частоти вживання алофонів в словах української мови та важливості розпізнавання алофону для розпізнавання слава.

Показано взаємопов'язаність коефіцієнту (класу) захищеності мови SPC та словесної розбірливості W. На їх базі запропоновано шкалу об'єктизованого оцінювання ступеню захищеності мовної інформації на межі контрольованої зони за критерієм залишкової розбірливості мови

Ключові слова: цифрова фонограма, коефіцієнт розбірливості мови, деструктивні зміни фонемної структури, вейвлет-перетворення

IMPROVEMENT OF THE METHOD FOR ASSESSING THE LEVEL OF SPEECH INFORMATION SECURITY

V. Blintsov
Doctor of Technical Sciences, Professor
Department of Electronic Engineering of Ship and Robotic Complexes*
E-mail: volodymyr.blintsov@nuos.edu.ua

S. Nuzhniy
PhD, Associate Professor
Department of Computer Technologies and Information Security*
E-mail: s.nuzhniy@gmail.com
*Admiral Makarov National University of Shipbuilding
Heroiv Ukrainy ave., 9,
Mykolayiv, Ukraine, 54025

1. Introduction

Determining the level of speech information security is one of the main problems for the systems of technical information protection and security services of an enterprise/organization (hereinafter referred to as the object) Confidence in its security provides the possibility of free communication when discussing issues that are critical to the economic, technological and innovative policies of the object, contain state and commercial secrets.

Despite all the variety of protection objects (their purpose, functioning peculiarities, number of staff), protection of speech information is provided by passive and active systems. However, to obtain confidence in the quality of protection, it is necessary to conduct specialized research, which would establish if the actual protection level meets the set requirements.

There are currently a significant number of methods for assessing the level of speech signal security – speech Intelligibility Index, speech transmission index, Articulation Index, speech Privacy Class, Signal-to-noise ratio, and others. However, as the analysis shows, these methods have one common drawback – they do not take into consideration the capabilities of an intruder to process an intercepted signal in order to clean it from extraneous noises.

The use of modern methods for digital processing of phonograms (spectral-correlation analysis, wavelet-transformation, adaptive filtration and others) make it possible to restore the linguistic component of the speech message (test-signal) even at significant levels of noise interference ($SNR \leq -10$ dBA). This leads to incorrect assessment of speech information security, obtained by these methods, and, accordingly, to underestimation of the possibility of interception and recognition of confidential speech information by an intruder.

Thus, there arises an urgent need for the development of the improved method for assessing the speech information privacy at the sites of information activities. The method should be independent on the structure and the principles of operation of the systems of technical protection of speech information with limited access and be able to take into consideration modern methods of digital processing of phonograms.

2. Literature review and problem statement

All the systems of assessing the level of speech information security at present can be divided into the following groups:

– control of energy parameters of a speech signal;
– control of frequency spectrum;
– control of the speech signal intelligibility.

Control of energy parameters of a speech signal on the boundary of the controlled zone in Ukraine is performed in accordance with [1, 2]. The estimation criterion is the Signal-to-noise ratio – SNR, which is measured in the frequency range from 100 Hz to 10 kHz by ⅓-octave bands. Calculation and comparison of the derived values are performed according to [3].

A similar method for assessing speech information privacy for closed spaces is used in Canada. Measurements are carried out in accordance with *ASTM E*2638 [4]. The evaluation criteria are the magnitudes of Speech Privacy Class (SPC) and uniform-weighted signal-to-noise ratio (SNRuni32).

Methods for assessing the level of speech signal privacy based on the study of frequency spectrum can be divided into several subgroups, which make it possible to determine:
– SII (speech Intelligibility Index) [5–8];
– STI (speech transmission index) [9–11];
– RASTI (Room Acoustics STI for premises) [10, 11];
– STIPA (STI for Public Address systems) [10, 11];
– STITEL (STI for Telecommunication Systems) [10, 11];
– %ALcons (Percentage Articulation Loss of Consonants).

Speech signal intelligibility is assessed by the criteria:
– W (word speech intelligibility (WSI) [12];
– AI (Articulation Index) [6, 12, 15].

The analysis shows that not all the above methods can be applied to the assessment of the level of threats of speech information interception by an intruder. Thus, the methods for assessing the speech signal privacy based on the study of frequency spectrum and determining indices STI, RASTI, STIPA, STITEL and %ALcons make it possible to establish the information integrity level. At the same time, this level will be almost the same for legal listeners and for an intruder (illegal listener). The difference will be caused by the parameters of the premises independent on a listener, such as sound insulation, parameters of the intruder's equipment and others.

Other methods, despite a significant difference in the principles of assessing the level of speech signal privacy and research technologies, have many similarities. The most essential of them is the assessment of the impact of noise interference on a speech signal on the boundary of the controlled area.

However, all above methods share one general drawback – they do not take into account the capability of an intruder to conduct specialized processing of an intercepted signal in order to clear it from outside noise. Such clearing methods can include, first of all, wavelet-transformation, phonetic-correlation analysis, adaptive and neuro network filtration. This leads, as shown in paper [14], to a significant increase in the level of speech signal recognition and, accordingly, to a critical increase in the values of such indices as SII, W (WII), AI (AI), SPC, SNR and SNRuni32.

The use of energy parameters has a series of important technological advantages – ease of research, availability of laboratory equipment and evidence of obtained results. At the same time, this method has a significant drawback – assessment of the level of speech information privacy by the *SNR* ratio means that the ratio of root mean square levels of the signal and noise intensity is found. In relation to an interference signal, which usually uses white noise

(or its "clones", such as "pink", "blue" and similar noises), this approach is justified. In relation to a speech signal, this approach, even with division of the frequency spectrum by ⅓-octave bands, is inadmissible.

When replacing a speech signal having a random form with its root mean square levels, even when using the division of frequency spectrum by ⅓-octave bands, there is a significant divergence between peak values and root mean square level [13, 14].

In [4], there is an attempt to fix this drawback by introducing the SPC and SNRuni32 coefficients.

SPC is determined from expression:

$$\text{SPC} = LD(\text{avg}) + L_B(\text{avg}), \tag{1}$$

where $LD(\text{avg})$ is the average value of sound intensity, measured at some points of the closed premises (distant premises, hereinafter referred to as DP), calculated in the average value for control points (CP) of outer space (beyond the controlled zone); $L_B(\text{avg})$ is the level of natural interference in the CP.

According to [4], the CP are chosen on condition of maximum benefit for an intruder, that is, in places of the worst sound insulation level, without taking into consideration the possibility of installing by an intruder of devices for speech information interception. The distance from the outside of the enclosing structure of a DP (for example, walls, windows, doors, etc.) to the CP is 0.25 m, according to [4]. This makes it possible to increase the level of the speech signal in the CP by 11–12 dB in comparison with measurements at the distance of 1 m.

The SNRuni32 coefficient, according to [4], is determined as the difference of levels in the CP of a speech signal $L_{TS}(f)$ and noise levels $L_B(f)$. These differences are cut off so that their values could not be less than 32 dB, in this case, speech will not be heard

$$\text{SNR}_{\text{UNI32}} = \frac{1}{16} \sum_{f=160}^{5,000} \left\{ L_{TS}(f) - L_B(f) \right\}_{-32}. \tag{2}$$

By neglecting the values, for which the difference is less than 32 dB, $L_{TS}(f)$ can be found as

$$L_{TS}(f) \approx L_{SP}(f) - LD(\text{avg}), \tag{3}$$

where $L_{SP}(f)$ is the sound intensity, measured at some points in the closed premises (controlled zone).

Then, (2) will take the form

$$\text{SNR}_{\text{UNI32}} \approx \frac{1}{16} \sum_{f=160}^{5,000} \left\{ L_{SP}(f) - LD(\text{avg}) - L_B(f) \right\}_{-32}. \tag{4}$$

During transition to average by ⅓-octave bands, we will obtain

$$\text{SNR}_{\text{UNI32}} \approx L_{SP}(\text{avg}) - LD(\text{avg}) - L_B(\text{avg}). \tag{5}$$

Taking into consideration (1), we obtain

$$\text{SNR}_{\text{UNI32}} \approx L_{SP}(\text{avg}) - \text{SPC}. \tag{6}$$

Hence,

$$\text{SPC} \approx L_{SP}(\text{avg}) - \text{SNR}_{\text{UNI32}}. \tag{7}$$

The expression above shows the relationships between SPC and SNRuni32 coefficients.

Papers [20, 21] present the values of intelligibility threshold in the rooms for SNRuni32, depending on the necessary level of speech information privacy. Three levels were considered: "intelligibility in free space", "intelligibility in the premises" and "intelligibility of the fact of talk", which correspond to values –11 dB, –16 dB and –22 dB [21]. The features of designing security systems based on the coefficient were shown and the scale of conformity of coefficient values to the level of the object security (its category) and frequency of intelligibility of words (short phrases) within the normative period of time was introduced. The shortcomings of this paper include the limited coefficient SPC (linguistic peculiarities of speech information and the influence of noise interference were not taken into account), and limitation of the proposed scale only to coefficient SPC.

In [15], it is recommended to select SNRuni32 for (7) of "–11 dB", and in [21] – "–16 dB".

Thus, if we accept for $L_{SP}(\text{avg})$ typical value of 64 dB (according to [12, 22] "quiet speech"), SPC takes the form

$$\text{SPC} \approx L_{SP}(\text{avg}) + 11\,\text{dB} = 75\,\text{dB},$$

which corresponds, according [15], to the level of "speech privacy" ("confidential talks"), or

$$SPC \approx L_{SP}(\text{avg}) + 16\,\text{dB} = 80\,\text{dB},$$

which corresponds to the level of "speech security".

Paper [15] deals with the analysis of the features of using AI and SII coefficients, article [6] contains the recommendations on their application. It was shown that the SII coefficient is the improved AI and it is intended to determine the level of intelligibility of speech information. The scope of application includes bioacoustics, occupational safety and protection of places where there are people (offices, dwelling, public places and others), as well as speech information privacy (estimation of the level of speech information intelligibility on the boundary of the controlled area). The calculation method makes it possible to explore and calculate the SII coefficient on condition of the influence of acoustic noise of different types and intensity on a speech signal. In addition, in [15] and [6], they imposed restrictions on the signal/noise ratio, in which the comprising magnitudes are accepted as equal to zero (for ⅓-octave bands or for specialized areas). This leads to the possibility of incorrect evaluation of information availability in the band (area). On the other hand, in the case of signal filtering, at significant levels of noise interference, there is a significant distortion of the spectrum, which leads to the possibility of mistaking the introduced distortions for the information signal.

Paper [16] explores the use of the AI coefficient for different structures and intensity of noise interference, considers their influence on the possibility of signal recovery – spectral composition and signal levels on the bands. However, the paper does not take into consideration the linguistic component of the information signal.

The specific drawback was considered in [17]. The conducted studies revealed the possibility of substitution of vowels with consonants. This occurs due to the appearance of spectrum distortion in the process of the interaction of a speech signal with noise interference. The paper considered the characteristics of noise interference and estimated the magnitude of their influence on a speech signal. The studies were conducted with the involvement of presenters who are native speakers of English. The impact of noise on the adequacy of the speech signal perception for other languages was not explored in the paper.

The studies of the influence of interferences of combined noise on speech transmission were performed in [18]. The studies were performed based on the model of open public space. Sound fields for dominant noises were implied using the typical model of a city square, surrounded by houses. The traffic noise and two types of construction noises, corresponding to stationary and pulsed noises, were selected as background noises. Listening tests were conducted on a group of adults, and the speech transmission quality was evaluated according to the criteria of the difficulty of presenter's speech recognition, as well as intelligibility indicators. However, instrumental parameters of speech signals and noise were not controlled in this case, only subjective assessment of the possibility of presenter's speech recognition was carried out.

In article [19], the evaluation of the speech intelligibility level was conducted by the indicator of short-term objective intelligibility (STOI). In the studies, the speech that was mixed with white noise at low values of the SNR coefficient was supplemented by the ideal binary mask (IBM). The STOI was used for prediction of intelligibility of both loud and frequency-weighted speeches. Native speakers of English (British) participated in the studies. The possibility of restoring the signals at a significant influence of noise interference, sufficient for its recognition by instrumental methods, was shown. Other types of noise interference and their influence on objective and subjective characteristics of a speech signal were not considered in the paper.

In studies [24–26], it is stated that the minimum number of recognized formants for confident recognition of a phoneme should by not less than two. This is due to the complex spectral composition of phonemes, caused by the peculiarities of their formation in the voice tract of a person. Paper [24] contains the results of the formation of phoneme [a] by allophone transformations from short-term fragments under conditions of low intensity of a sound (whisper). The studies showed that only 70–80 % of short-term fragments are able to form the corresponding phoneme with an assigned reliability of identification. This is caused by the existence in short-term fragments of a significant number of additional non-stationary sections that do not bear any information about the appropriate phoneme.

Article [25] presents the research into the phenomenon of formant splitting, when a group of sub-formants emerge in the place of one simple (scalar) formant. Such phenomenon is characteristic of sibilants and hushing consonants [c], [з], [ш], [ж], [с'], [з'], [ш'], [ж'] and results in appearance of additional resonant frequencies.

Nasal sounds [м], [н], [м'], [н'] have even more complex spectral composition. These sounds have not two, but three characteristic frequencies – the first basic nasal formant $Fp1(n)$, the first basic mouth formant $Fp1(m)$ and the second constant formant $Fp2$. This led to the appearance of additional spectral components. For example, according to [25], for sound [н], apart from the specified $Fp1(n)$, $Fp1(m)$ and $Fp2$, the additional spectral composition is determined from dependence: $Fp2 + \mathbf{n} \cdot Fp1(n)$, where $n = \{1, 2, 3, 4\}$. Sound [л] in the Ukrainian language is characterized by the existence of additional combinatory formants that occur at the frequencies close $Fp2 - Fp1$, $Fp2 + n \cdot Fp1$, where $n = \{1, 2, 3, 4, 5\}$.

In addition, it is necessary to take into account the fall by 3 dB on the octave of signal intensity that is characteristic of most of the phonemes, which leads to an increase in the impact of noise interference [12, 22].

Thus, analysis [1–21] shows the lack of scientifically substantiated method assessing the level of speech information privacy that can take into account the use of modern methods for speech information recovery by an intruder. The method should take into consideration objective (instrumental) and subjective (linguistic) characteristics of a speech signal and assess the privacy level, depending on the category of information, requirements to the selected premises and temporal parameters of observation.

3. The aim and objectives of the study

The aim of this study is to improve the method for assessing the level of speech information privacy, under conditions of the possibility of using modern methods and technologies of noise interference filtration by an intruder.

To achieve the aim, the following tasks were set:

– to perform analysis of the methods for assessment of the level of speech information privacy used in Ukraine, EU, USA and Canada and to conduct their comparative analysis;

– to substantiate the need for improvement of the procedure for determining the level of speech information privacy based on the coefficient (index) of speech;

– intelligibility by taking into consideration the features of allophones and introduction of the coefficient of the residual speech intelligibility;

– to study the possibility of using diphones and triphones as a simplified linguistic model of allophones and to analyze the frequency of their use in the Ukrainian language;

– to analyze the impact of the quality of recognition of the basic tone frequency and formants for the reliability of assessment of the level of speech information privacy;

– to propose the scale of objective evaluation of the degree of speech information privacy on the boundary of the controlled area by the criterion of residual speech intelligibility.

4. Analysis of methods for assessing the level of speech signals security

Comparison of coefficients SPC (according to [4, 21]) and SNR (according to [12, 22]), in the definition that is accepted on the post-Soviet territory and the coefficient of word intelligibility of speech W [22] is shown in Table 1.

Thus, the coefficient of word intelligibility of speech W makes it possible to link the methods for evaluation of speech information privacy, based on the principles of measuring the sound intensity, such as [4, 22].

It should be noted that the SPC sets up the requirements for DP at the design phase of safety systems – sound insulation systems. The value of the coefficient essentially depends on

the maximal level of probable average value of speech sound intensity in the DP. In case of inspection of premises and determining the actual level of information privacy, the SPC is determined based on the SNRuni32 coefficient, according to (7).

More information about the effectiveness of security systems is given by the AI and SII coefficients. It should be noted that the SII, according to [6], is a formal substitution for the AI articulation coefficient, which was in 1997. However, [4] was put into operation in Canada in 2010, where the use of the AI coefficient was resumed.

Table 1

Relationship of coefficients *SPC* and *W* and *SNR*

Level	SPC			Word intelligibility of speech		*SNR*
	Description of the level of speech privacy	Numerical value, dB	Frequency of recognition	Description of level of speech privacy	*W*, %	Numerical value, dB
Standard speech privacy	Seldom comprehensible and often heard	≤70	1 short phrase in 2 min	Partial concealing the talks content in DP	40	–6
Standard speech security	Sometimes comprehensible and seldom heard	70–80	1 short phrase in 15 min – 1 h	Considerable concealing the talks content in DP	30	–8
High speech security	Not comprehensible by essence, sometimes heard	80–85	1 short phrase in 2 – 4.5 h	Concealing the talks subject in DP	20	–10
Very high speech security	Incomprehensible and not heard in essence	>85	1 short phrase in 16 – 20 h	Concealing the fact of the talks in the DP	10	–13

At significant levels of noise interference (SNR= –10...–22 dB) in spectral ⅓-octave bands, a signal of interference starts to increase the levels of formants of the second and third groups for most phonemes. The level of formants of the fourth and the following groups is usually hardly comprehensible at the ratios SNR≈–0 dB. Thus, for example, during forensic phonoscopic examinations [23, 24], it is a mandatory condition in order to determine the authenticity of a presenter to ensure the minimal impact of the influence of noise interference on the presenter's speech. This is only possible at SNR≥0 Db (3... 5 dB).

The main advantage of SII in comparison with SPC and SNR is taking into consideration the informativeness of spectral ⅓-octave bands. That is, not the signal/noise ratio, but rather the presence the formants of the phoneme in a specific band of frequencies is essential in calculations. According to [6], SII is determined as:

$$S = \sum_{i=1}^{n} I_i A_i, \qquad (8)$$

where *n* is the number of bands, according to which the speech intelligibility coefficient is calculated (usually *n*=6 or 7 – when using octave bands, or *n*=18 or 21 – when using ⅓-octave bands); *I* is the coefficient of information weight in the band; *A* is the coefficient of information recognition quality in the band.

Coefficient of information weight in band *I* is determined for each speech and depends on the likelihood of appearance of sound formats in a particular octave or ⅓-octave band and existence of additional spectral components [24, 25].

The coefficient of information recognition quality in band A reflects the likelihood of correct recognition of a formant or its additional spectral components in the studied octave or ⅓-octave band and, according to [6], is determined as:

$$A(\Delta L)_i = \begin{cases} 0, & \Delta L_i \le 0 \text{ dB.} \\ \dfrac{\Delta L_i}{30}, & 0 < \Delta L_i \le 30 \text{ dB,} \\ 1, & \Delta L_i > 30 \text{ dB.} \end{cases} \qquad (9)$$

where ΔL is the difference between the peak level of a speech signal and effective noise level. That is,

$$\Delta L = L'_S - L_n(\text{avg}) \approx$$
$$\approx \{L_S(\text{avg}) + 15 \text{ dB}\} - L_n(\text{avg}) = \text{SNR} + 15 \text{ dB.}$$

Summand "15 dB", according to [6], ensures the transition from the peak signal level in the band to its root mean square (effective) level and further to coefficient SNR that is typical in Ukraine.

Coefficients I and A can be determined in another way – the coefficients are normed in many languages. Thus, in [12, 22], these coefficients are specified for the Russian language – weight coefficient of band k_i and formant speech parameter in band ΔA_i.

Analysis of coefficient of information intelligibility quality shows that according to (9), SII is determined in the dynamic range of 30 dB, shifted to the right by 15 dB. This indicates that it is appropriate to use coefficient SII to assess the level of speech information privacy at SNR≥−15 dB. That is, in conformity with the "Very high speech security" level in Table 1.

However, the coefficient has a considerable drawback. According to (9) SII, it is inappropriate to use it in automated systems for assessment of the level of speech information privacy at considerable noise pollution level equal (SNR<−15 dB). In papers [12, 22], it was shown that at coefficient W=20 % for 5 and 6 octave bands (root mean square frequencies of bands of 2,000 Hz and 4,000 Hz), coefficient SNR is 18.2 dB and 24.7 dB. Measurements were performed using white noise at interference. Weight coefficient of band k_i for the specified bands was found to be 0.30 and 0.26, respectively. In this case, these values, according to [12, 22], are the highest, that is, more than a half of all formants are concentrated in them.

5. Improvement of the method for determining the level of speech information privacy by introducing the coefficient of residual speech intelligibility

Determining the level of speech information privacy in the systems, which use the means of setting active interference, should include the preparatory stage. The main requirements for it are:

– the phonograms of specialized articulation tables, formed on the basis of short common use phrases and by professional orientation of the studied protection object (including terminology, slang, proper names, abbreviations, etc.) are used as test signals;

– the studied phonogram is obtained according to the typical technology for determining the level of speech information privacy based on the SNR coefficient (or any of SPC, AI, SII, SNRuni32 coefficients);

– clearing the studied phonograms is performed by modern methods and technologies of noise interference filtration.

Assessment of the level of speech information privacy based on the cleared phonogram can be performed by the following methods:

1. Linguistic examination of oral speech.

2. Determining the SII coefficient based on the correlation analysis of the original of a test-signal and a "cleared" signal.

3. Determining the coefficient of residual speech intelligibility.

The first method implies the involvement of specially prepared auditor (listener). Its advantage is determining the level of speech information privacy based on the direct method – listening. The shortcomings of the method are subjectivism and considerable labor intensity.

The second method is based on the preparation of the specialized database for each articulation table. The database specifies the coefficients of information weight for the bands for each phoneme in each of the words that are included in the articulation table, based on which the phonogram of a test-signal was recorded. The base is formed relying on analysis of a test-signal (phonogram) at the preparatory stage. During the base generation, it is necessary to take into consideration the peculiarities of the speech, professional focus of the staff of an object, homonyms and others. Thus, for example, the word "compass", depending on the professional orientation, can change the stress – "ко́мпас" or "компа́с". It is possible to use the words-homonyms "на бе́резі" and "на бере́зі", and so on. This causes a change in the frequency of basic formants, their splitting, appearance of additional formants [24, 25].

The disadvantages of this method are:

– significant overstatement of the level of speech information privacy due to the use of the correlation method and databases;

– existence of residual noise interference at frequencies of the upper octaves;

– distortion of the spectrum of formants at frequencies of upper octaves.

Method 3.

It is proposed to introduce a new criterion of assessment of the level of speech information privacy – RII (Residual intelligibility index) – S^R, determined from expression:

$$S^R = \frac{1}{m} \sum_{y=1}^{m} \left[R_y^a H_y \sum_{i=1}^{n} I_i A_i \right], \qquad (10)$$

where m is the total number of allophones in a word; R^a is the coefficient of the importance of allophone recognition for recognition of a word (phrase, sentence); H is the coefficient of the frequency of use of this allophone in speech, n is the number of ⅓-octave bands, by which coefficient is calculated (n=18 or 21); I is the coefficient of information weight in the band for allophone recognition; A is the coefficient of information recognition quality in the band.

If we accept that

$$R_y^a H_y \sum_{i=1}^{n} I_i A_i = K^R,$$

then K^R is the coefficient of destructive changes of the allophone that expresses the influence of noise interference on the form and frequency-temporal spectrum of the allophone. The coefficient structurally takes into account two components of speech recognition: linguistic $\left(R_y^a H_y\right)$ and formant $\left(\sum_{i=1}^{n} I_i A\right)$.

Fig. 1 shows the example of research in the MATLAB environment (version R2015a) into the influence of noise interference (white noise) at SNR=−10 dBA on the word "ship" from the expression "coast guard ship". Fig. 1, *a* shows the phonograms of the original (Original), noise signal (Signal & Noises) and residual (Residual Intelligibility) signals. Their spectral analysis is shown in Fig 1, *b*. As Fig. 1, *b* shows, at SNR=−10 dBA, the frequency of the basic tone is unambiguously determined for noises signal and residual signal. For the residual signal, cleared from interference, it is possible to recognize formants *Fp*1, *Fp*2 and *Fp*3, using correlation analysis for splitting the words into allophones and correlation analysis of the spectrum of allophones by "falls".
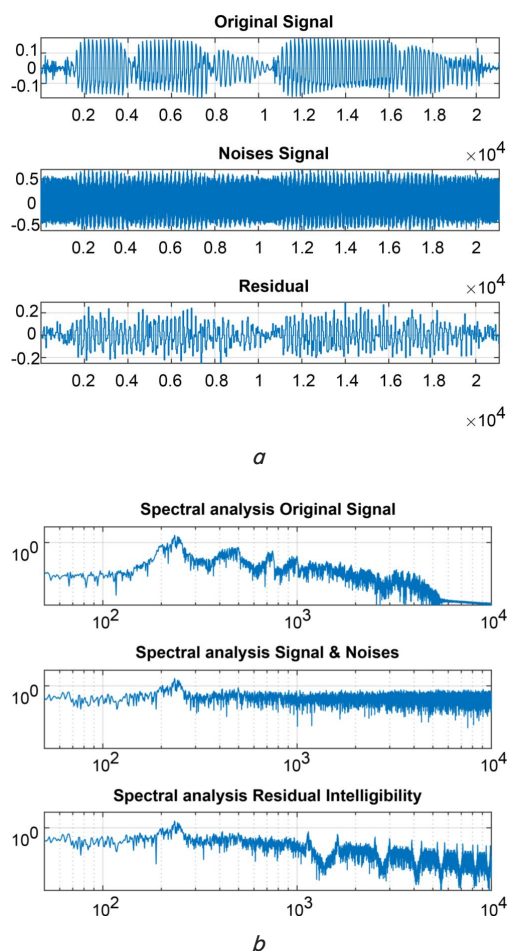

a

b

Fig. 1. Studying the influence of noise interference (white noise) at *SNR*=−10 dBA on word "ship" in the MATLAB environment: *a* — noise interference filtration; *b* — spectral analysis on the long-term spectrum

The linguistic component considers the peculiarities of formation and use of the language in the state (region), dialects, words-parasites, slang, etc. That is, this component considers the factors that affect the formation of words from allophones – the importance of allophone recognition

by a listener to understand the essence of the word and the frequency of its use in speech. The importance of these factors is that:

– the duration of sounding of the formed phoneme in the equal tempo text is normally 50–60 % of the time of its sounding (depending on the type of phonemes – vowel/consonant, sonorant/noisy, etc.). That is, the process has a discrete nature by time and readiness of a phoneme for analysis. Given that the length of the phoneme in the word is 0.1–0.25 s, time for the formation of the tangent in the frequency-temporal distribution for the formed phoneme at a significant influence of the noise interference is insufficient. This results in a decrease in the likelihood of the reliable recognition;

– the process of forming a word from allophones is continuous by nature. This is due to the essence of the concept of allophone as a realization of a phoneme, its variant, caused by specific phonetic surroundings in the word, the stress in a word and in a sentence. Thus, the process of allophone formation actually begins and ends in the structure of the previous or the posterior allophone (or a pause between the words) – there occurs the process of mutual allophone overlapping. Thus, the period of time to analyze the formation of the tangent to the formants significantly increases. In this case, the use of the accumulation method can significantly reduce the influence of noise interference and, accordingly, increase the probability of reliable recognition of allophones and a word in general;

– different frequency of using allophones in words makes it possible to significantly increase the level of allophone recognition in a word. Given the definition of the term "allophone" and the number of native speakers, we can assume that the total number of allophones tends to infinity. Therefore, there is a need to formalize the approach based on a simplified linguistic model – a phoneme (a letter), a diphone (two letters) and a triphone (three letters). Text documents are used as an information source [27–32].

An increase in recognition reliability is possible in the transition from analysis of frequency of using phonemes to frequency of using allophones. For the word "ship", the frequency of using phonemes [к], [о] and [р] can be found from Fig. 2 or [28, 29, 32] to be 4.01; 8.23 and 4.12, respectively.

Conducted studies of diphones show that for the Ukrainian language in the texts with the total number of letters of 276 thousand, phoneme [а] occurs 27,318 times. Table 2 shows the data on the frequency of using some diphones with background [а].

As can be seen from research results shown in Table 2, there is a low probability of occurrence in the studied phonogram of a large number of diphones, and accordingly, the allophones associated with them:

Analysis of frequency of using triphones, created based of vowel [о], shows that they can be divided into four groups:

1. Triphones that have the highest frequency of use – more than 2 %. These triphones include [кор], [мов], [нов], [ног], [ної], [ном], [нос], [пов], [пор], [пос], [роб], [ров], [тор] and [фон].

2. Triphones that have medium frequency of use – in the range from 0.5 % to 2 %. They include [бор], [вод], [гою], [доб], [дов], [дод], [док], [доп], [зон], [йог], [йол], [коï], [ком], [кон], [кос], [лог], [лож], [лою], [мог], [нок], [нор], [пож], [поз], [пол], [пом], [пох], [роз], [ром], [рот], [соб], [тов], [тод], [ток], [тот], [точ], [фор], [ход] and [хом].

3. Triphones that have low usage frequency – in the range from 0.1 % to 0.5 %.

4. Triphones that are based on diphones [ао], [ґо], [єо], [ио], [їо], [уо], [цо], [юо], [яо], [оґ], [єо], [ио], [іо], [ой], [оу], [оф], [оц], [ощ], [оь] and [оя] have the frequency of use of less than 0.1 %.
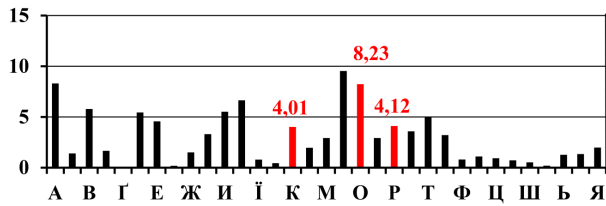

Fig. 2. Averaged values of frequency of using the phonemes in the Ukrainian language, according to [28, 29, 32]

Table 2

Frequency of using some diphones with background [a]

Diphone type	[ац]	[ца]	[ас]	[са]	[ає]	[єа]	[аї]	[їа]	[аі]	[іа]	[ао]	[оа]
Frequency of use, number	2,250	0	823	247	330	0	641	0	0	494	3	7
Relative frequency of use, %	8.23	0	3.01	0.9	1.21	0	2.34	0	0	1.81	0.01	0.03

Fig. 3 shows the results of research into frequency of using triphones based on phoneme [o]. According to the obtained results, the frequency of using triphone [кор] in the Ukrainian language is 2.21 %.
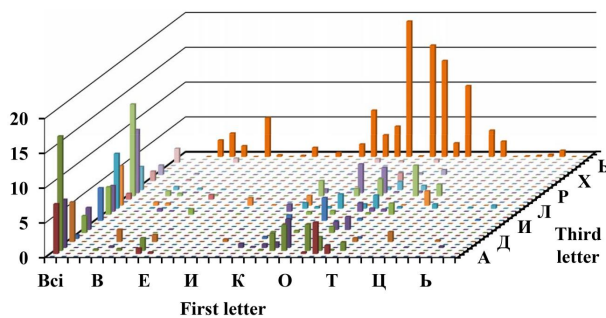

Fig. 3. Averaged values of frequency of using triphones based on phoneme [o] in the Ukrainian language

At the same time, it is necessary to take into consideration that the phonemes that are close by the way of formation due to the destructive influence of noise interference or filtration means and/or phonogram processing can change the formant frequencies.

At the same time, out of 1,444 triphones, formed based of vowel phoneme [o], only 117 are used– their frequency of use is not zero.

Thus, the prevalence of the RII coefficient over the SII coefficient manifests itself by the transition from phonemes to the allophones, which have the envelope form that is more resistant to outside influences and take into consideration frequency-temporal distribution of information in the formants. The distribution, in general, depends on the features of the speech formation of both a particular presenter, and the speech of the given region (state). This allows, even at critical values of signal/noise ratio (SNR≤−15 dB), substantial recovery of speech information.

Analysis (10) shows that the RII coefficient is the average of residual speech intelligibility for this test-signal under specific conditions of noise. This makes it possible to neutralize the influence of separate significant deviations of intelligibility by words when long-term changes are taken into account.

Technologies of determining the RII coefficient involves the development of the specialized allophone base, which can take into consideration not only the characteristics of a particular speech, but also the dialect peculiarities of the speech of the regional presenters and the influence of other speech. Such base will be much larger than the base that is developed by the second method considered above. However, its advantage will be universality (independence on a certain articulation table) and the possibility of more distinct recognition of diphones, affricates, sibilants and hushing consonants and others.

6. Analysis of frequency of using of diphones and triphones (linguistic models of allophones) in the Ukrainian language

Objective evaluation of the degree of speech information privacy on the boundary of the controlled area should be based on the likelihood of recognition of speech information in the intercepted signal by an intruder with the reliability that is sufficient to recover a certain volume of information. Given that an allophone is accepted in (10) as a speech unit, it is logic to link the evaluation scale to the averaged level of allophone recognition, and accordingly word recognition – the use of the dual ordinate scale. If an expert (or a decision support system) has the originals of test-signals, by which it is possible to establish the correspondence of a certain group of allophones to a particular word, this approach is correct. The dependence argument is the number by order of an allophone in the word of test-signal.

For noisy and cleared from noise signals in Fig. 1, only the section of frequency of basic tone ($Fp0$) gives high rates of coefficients of information weight and the quality of information recognition in the band. In other sections, the value of the coefficient of the quality of information recognition will be minimal, which leads to obtaining the false assessment of information privacy.

The consideration of the linguistic component through the correlation analysis of words for their splitting into allophones, in combination with correlation analysis of the frequency spectrum, makes it possible to recognize formants ($Fp1$, $Fp2$, $Fp3$ and $Fp4$), therefore, to significantly decrease the assessment of the level of speech information privacy.

Research results (Table 2, Fig. 2, 3) make it possible to determine the value of coefficient of frequency of using this allophone in speech (H) and coefficient of importance of allophone recognition for word recognition (R^a) in dependence (10).

According to the conducted research (Table 2, Fig. 2, 3), the maximum value of the frequency of using phonemes, diphones and triphones does not exceed 10 % (0.1 cond. u.). Thus, for the coefficient of frequency of using allophone (H) in a word of the test-signal, it is possible to introduce the dependence:

$$H = \begin{cases} 0.1, & h \le 0.01 \\ 10h, & 0.01 < h \le 0.1; \\ \frac{1}{2}(10h+1), & 0.1 < h \le 0.5; \\ 2(h+1), & 0.5 < h \le 2; \\ \frac{1}{2}(h+10), & h > 2, \end{cases} \quad (11)$$

where h is the frequency of using the triphone (cond. u.).

Dependence (11) is approximated by expression

$$H = 1.85 \cdot 10^{-4} \cdot h^5 - $$
$$-6.5 \cdot 10^{-3} \cdot h^4 \cdot 0.121 \cdot h^3 - $$
$$-1.03 \cdot h^2 + 4.4 \cdot h + 0.1. \quad (12)$$

Dependences (11) and (12) are designed both for manual calculation and for using in decision support systems for automated systems for assessing the level of speech information privacy.

7. Analysis of the quality of basic tone frequency recognition and formant on the quality of assessment of the level of speech information security

The quality of information recognition in the bands of frequency spectrum of a phoneme (a word) essentially depends on the level of destructive influence of noise interference and a curvature of frequency spectrum due to the use of digital methods of processing (Fig. 1). The simplest means for determining the recognition quality is the use of the spectral-correlation method.

However, at significant levels of destructive impact (at SNR $\le$ –10 dBA) of its use is ineffective. The solution to the problem is to use allophones and their main components – the frequency of the basic tone ($Fp0$) and formants ($Fp1$, $Fp2$, $Fp3$ and $Fp4$). In this case, in fact, it is actually necessary to assess the quality of recognition of envelopes and their reciprocal placement on the frequency spectrum. At significant influences, which lead to impossibility to determine the envelopes correctly (Fig. 1), the use of correlation coefficients is incorrect.

To assess the quality of recognition of the basic tone and the formants, it is proposed to use the coefficient of importance of allophone recognition for the recognition of word R^a, based on (10) and Fig. 1, is determined as follows:
– for Signal & Noises $\left\{ R^a \right\}^N$:

$$\left\{ R^a \right\}^N = \frac{1}{k} \sum_{j=0}^{k} \left[S_j \cdot \left(F_j - L_{n_j} \right) \right]; \quad (13)$$

– for Residual Intelligibility signal $\left\{ R^a \right\}^R$:

$$\left\{ R^a \right\}^R = \frac{1}{k} \sum_{j=0}^{k} \left[S_j \cdot \left(F_j - L_{R_j} \right) \right], \quad (14)$$

where k is the number of formants in an allophone (normally k=3...4, is determined for an allophone by a test-signal (Fig. 1)); S_j is the coefficient of weight of a formant in allophone recognition; F_j is the average value of the level of allophone formant signal in a test-signal (F_0 is the average value of the level of signal of the basic tone of an allophone in a test-signal); L_{n_j} is the average value of the noise level on the formant section; L_{R_j} is the average value of the level of residual noise and the noise caused by the spectrum curvature during mathematical treatment (filtration, wavelet-transformation, etc.).

8. The scale of objective evaluation of the level of speech information privacy

Taking into consideration (10) to (14), it is proposed to introduce the scale of objective evaluation of the degree of speech information privacy on the boundary of the controlled zone by the criterion of residual speech intelligibility. When forming the scale, the provisions used in [4] (Table 1) and correlated with the requirements [1–3], were taken into consideration.

When developing the scale, it is taken into consideration that at the average speech tempo, the average presenter pronounces 100–120 words/min. Taking into account that a word on average consist of 2–4 allophones, we obtain 200–480 allophones/min.

Under such original conditions, based on Table 1, it is possible to establish the scale of objective evaluation of the degree of speech information privacy on the boundary of the controlled zone by the criterion of residual speech intelligibility (Table 3).

Table 3 implies a uniform law of distribution of recognized words within the set period.

Table 3

Scale of objective assessment of the level of speech information privacy on the boundary of the controlled area by the criterion of residual speech intelligibility

Level	Description of speech privacy level	Recognition frequency		
		1 short phrase	3 words[1]	10 allophones[2]
Standard speech privacy	Seldom comprehensible and often heard	in 2 min		
Standard speech security	Sometimes comprehensible and seldom heard	in 15 – 60 min		
High speech security	Incomprehensible by essence and sometimes heard	in 2 – 4.5 hours		
Very high speech security	Unintelligible by essence and not heard	in 16 – 20 hours		

Note: [1] – on condition of uniform law of distribution in time; [2] – on condition of uniform distribution in time by groups of 2–3 allophones, which allow recovering completely or partially separate words of the test-signal with certain reliability

Assessment of the degree of speech information privacy by the frequency of allophone recognition implies their uniform distribution in groups of 2–3 allophones, which make it possible to recover certain words of a test-signal with certain reliability.

In Table 3, not only the requirements for the methods of assessing the level of speech information protection from leakage through acoustic and vibration channels were established, but also the distribution of premises taking into consideration the categories of information was introduced.

9. Discussion of the results of research into the improved method for assessment of speech data information privacy

The proposed mathematical dependencies (10)–(14) create the theoretical base for the method of determining the level of speech information privacy based on the residual intelligibility index (*RII*). The implementation of the method will make it possible:

– to improve the procedures and technologies of assessment of speech information privacy operating in Ukraine;

– to take into account the modern state of the methods of digital processing of phonograms;

– to ensure enhancement of reliability of assessment of its security level.

At the same time, the improved scale of objective evaluation of the degree of speech information privacy on the boundary of the controlled zone by the criterion of residual speech intelligibility (Table 3) was suggested, which allows:

– transition from the two-level assessment of the levels of speech information security that operates in Ukraine to the four-level assessment, which will enable constructing a security system depending on the object category and optimizing the material costs;

– establishment of the requirements for signal levels in the systems of setting active acoustic and vibration interferences in accordance with the object category;

– decrease in the level of signals of generators of acoustic and vibration interferences for levels "Standard speech privacy" (SSP) and "Standard speech security" (SSS), which makes it possible to significantly improve the working conditions of employees.

Thus, the obtained results make it possible to indicate the achievement of the set goal. In addition, it is possible to indicate that:

1) The conducted analysis of the existing methods for assessing the level of information security showed the existence of three dominant methods – SNR (Ukraine and the state in the post-Soviet space), SII (USA and European countries) and SPC (Canada). However, each of these methods has its own peculiarities, which actually do not make it possible to compare transparently the results of their application. It was proposed to use the criterion of word intelligibility of speech (W), introduced in [12, 22] as an auxiliary section. The main comparison criterion was the requirements to the level of speech information security and the possibility of interception of its part by an intruder for a certain period. This allowed formalizing the approaches to assessing the level of speech information privacy by different methods and unifying the requirements for the systems of speech information security (Table 1).

2) The use of the complex methods that combine subjective (word intelligibility of speech) and objective (instrumental research into spectrum frequency) methods for assessing the level of speech information privacy is most promising. Such methods were called objective evaluation methods. The American method based on determining the speech intelligibility index (SII) is the most common. However, analysis of its use showed its limited applications – the capabilities of the modern methods for digital processing of phonograms, such as wavelet-transformation, adaptive filtration, spectral and correlation analyses, etc. are not taken into account.

Fig. 1, *b* shows the impact of the noise interference with the ratio SNR=−10 dBA on the test-signal and the spectrum distortion after the filtration procedure. As Fig. 1 and (9) show, when using the *SII* coefficient, a rather optimistic assessment of the security level will be obtained. The Signal & Noises analysis indicates that only on the spectrum section up to 300 Hz (section of basic tone), the value of coefficient of quality of information recognition in band (A) will differ from 0. According to (9):

– for sections from 50...300 Hz, we obtain

$$A(\Delta L)_i = \frac{\Delta L_i}{30} = \frac{L'_s - L_n(avg)}{30} = \frac{5}{30};$$

– for section from 300...5,600 Hz, we have $A(\Delta L)_i = 0.$ Thus, according to (8), obtain

$$S = \sum_{i=1}^{5} I_i A_i = 0.03 \cdot \frac{5}{30} = 0.005,$$

where $I = 0.03$ is the coefficient of information weight in the band, determined according to [12]. Such value corresponds to the level "High speech security".

However, after filtration (signal "Residual Intelligibility"), using wavelet-transformation, we obtain

$$S = \sum_{i=1}^{5} I_i A_i = 0.03 \cdot \frac{5}{30} + 0.12 \cdot \frac{3}{30} +$$

$$+0.2 \cdot \frac{1}{30} + 0.3 \cdot \frac{1}{30} + 0.26 \cdot \frac{2}{30} = 0,042.$$

That is the coefficient increased by an order of magnitude and corresponds to the level "Standard speech privacy".

The ways of solving the specified problem involve taking into account the linguistic component of speech information and allophone resistance to noise interferences and spectrum distortion. It was proposed to introduce the RII (Residual intelligibility index and the procedure of its calculation (10).

3) The use of diphones and triphones as a simplified linguistic model of allophones was proposed and analysis of the frequency of their use was performed (Fig. 2, 3, Table 2). The analysis was conducted for specialized texts in the area of "Information security", "Finance" and "Criminology". The texts of the total volume of more than 20 Mb of the text format were analyzed [32]. The results of the analysis were compared with the results of other authors [28–30]. While the results were generally the same, some divergences were identified, which can be attributed to the peculiarities of professionally oriented texts.

We proposed to divide the triphones into 4 groups, depending on the usage frequency, and derived the dependences for calculation of coefficient of the allophone usage frequency (11) and (12) as a component of (10).

4) The analysis of the quality of recognition of the basic tone frequency (Fp0) and of formants (*Fp*1, *Fp*2, *Fp*3 and *Fp*4) was performed to determine the quality of assessment of the level of speech information security. The dependences for calculation of coefficients of the importance of allophone recognition for recognition of words for speech and noise signals (13) and (14) residual intelligibility signals were proposed.

5) The scale of objective assessment of the degree of speech information privacy on the boundary of the controlled zone by the criterion of residual speech intelligibility

was proposed (Table 3). The requirements for the level of speech information privacy and the possibility of interception of its part by an intruder for a certain period were taken into account.

Thus, the results of the theoretical and applied research suggest the possibility of using the proposed method for assessment of the level of speech information privacy with the limited access on the information activity sites of Ukraine. The method takes into consideration the structure and the principles of operation of modern systems of speech information security and makes it possible to integrate into the current procedure of attestation of separated premises. Taking into consideration the capabilities of modern methods of digital processing of phonograms (wavelet-transformations, spectral-correlation analysis, adaptive filtration and others), together using articulation tables, ensures the reliability of the assessment of the level of speech information security.

10. Conclusions

1. A review of the methods for assessing the level of speech information privacy used in Ukraine, the EU, the USA, and Canada was carried out and their comparative analysis was performed. This allowed formalizing the approaches to the assessment of the level of speech information security by different methods and unifying the requirements for the systems of speech information privacy. The main criterion for comparison was the requirements to the level of speech information security and the possibility of interception of its parts by an intruder in a certain period.

2. We substantiated the necessity of improvement of the methods for determining the level of speech information security based on the speech intelligibility coefficient (index) by taking into consideration the peculiarities of allophones and the introduction of the coefficient of residual speech intelligibility This approach makes it possible to enhance the reliability of assessment of the level of speech information security during digital processing of phonograms – noise filtration based on wavelet-transformation, spectral and correlation analysis, application of adaptive filtration, etc.

3. The use of diphones and triphones as a simplified linguistic model of allophones was proposed. This allowed formalizing the structure of allophones and analyzing the frequency of their usage in the Ukrainian language. The triphones were divided into 4 groups, depending on the usage frequency, and analytical dependences for calculation of coefficient of frequency of usage of triphones (allophones) for the problems of technical protection of information.

4. The analysis of the destructive influence of noise interference and of distortions of spectral composition, which arise when using the methods of digital processing of phonograms, was performed to determine the quality of the recognition of the frequency spectrum of the basic tone and the formant. Their influence on the level of speech information security was estimated and analytical dependences for the coefficient of importance of allophone recognition for the recognition of words (phrases, sentences) for noise and sounds and residual intelligibility signals were proposed.

5. The scale of objective estimation of the degree of speech information security on the boundary of the controlled zone by the criteria of residual speech intelligibility and the possibility of interception of its parts by an intruder for a certain period was proposed. Its usage makes it possible to improve the quality of assessment of the level of speech information protection from leakage through acoustic and vibration channels, taking into consideration the information categories.

References

1. Normatyvnyi dokument systemy tekhnichnoho zakhystu informatsiyi ND TZI 2.4-010-2015.

2. Normatyvnyi dokument systemy tekhnichnoho zakhystu informatsiyi ND TZI 2.3-019-2015.

3. Normatyvnyi dokument systemy tekhnichnoho zakhystu informatsiyi ND TZI 2.2-008-2015.

4. ASTM E2638-10. ASTM Intl., West Conshohocken, PA.

5. ANSI/ASA S3.37-1987 (R2017) American National Standard Preferred Earhook Nozzle Thread for Postauricular Hearing Aids. STANDARD by American National Standards of the Acoustical Society of America, 1987.01.01.

6. ANSI/ASA S3.5-1997 (R2017) American National Standard Methods for Calculation of the Speech Intelligibility Index. STANDARD by American National Standards of the Acoustical Society of America, 1997.01.01.

7. ISO 9921:2003. Ergonomics – Assessment of speech communication (2003). International Organization for Standardization, 28. Available at: https://www.iso.org/standard/33589.html

8. IEC 60268-16:2011. Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index. Available at: https://webstore.iec.ch/publication/1214

9. ISO 7240-24:2010. Fire detection and fire alarm systems – Part 24: Sound-system loudspeakers (2010). International Organization for Standardization, 39.

10. UNE EN 60268-16:2011. Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index (Endorsed by AENOR in November of 2011).

11. ISO 3382 (2012). Acoustics – Measurement of room acoustics parameters – Part 3: Open plan offices.

12. Grigoriev, I. A., Kazanovski, A. I. (2010). Methodical approach to efficiency evaluation of voice data protection. Vestnik Voronezhskogo gosudarstvennogo tehnicheskogo universiteta, 5, 133–136.

13. Nuzhnyy, S. M. (2018). Improved technology of evaluation of stage of protection of license information. Modern Information Security, 1 (33), 66–73. Available at: http://journals.dut.edu.ua/index.php/dataprotect/article/view/1796

14. Blintsov, V., Nuzhniy, S., Parkhuts, L., Kasianov, Y. (2018). The objectified procedure and a technology for assessing the state of complex noise speech information protection. Eastern-European Journal of Enterprise Technologies, 5 (9 (95)), 26–34. doi: https://doi.org/10.15587/1729-4061.2018.144146

15. Hornsby, B. W. Y. (2004). The Speech Intelligibility Index. The Hearing Journal, 57 (10), 10–17. doi: https://doi.org/10.1097/00025572-200410000-00003

16. Allen, J. B. (2005). Consonant recognition and the articulation index. The Journal of the Acoustical Society of America, 117 (4), 2212–2223. doi: https://doi.org/10.1121/1.1856231

17. Phatak, S. A., Lovitt, A., Allen, J. B. (2008). Consonant confusions in white noise. The Journal of the Acoustical Society of America, 124 (2), 1220–1233. doi: https://doi.org/10.1121/1.2913251

18. Lee, P. J., Jeon, J. Y. (2011). Evaluation of speech transmission in open public spaces affected by combined noises. The Journal of the Acoustical Society of America, 130 (1), 219–227. doi: https://doi.org/10.1121/1.3598455

19. Graetzer, S., Hopkins, C. (2018). Evaluation of STOI for speech at low signal-to-noise ratios after enhancement with Ideal Binary Masks. Conference: 25th International Congress on Sound and Vibration.

20. Bradley, J. S., Gover, B. N. (2010). A new system of speech privacy criteria in terms of Speech Privacy Class (SPC) values. Proceedings of 20th International Congress on Acoustics. Sydney. Available at: https://nrc-publications.canada.ca/eng/view/accepted/?id=cf69d165-7fb1-46c9-b60f-e6c8020b0c11

21. Bradley, J., Gover, B. (2008). Speech privacy class for rating the speech privacy of meeting rooms. Canadian Acoustics, 36 (3), 22–23. Available at: https://jcaa.caa-aca.ca/index.php/jcaa/article/view/2018

22. Horev, A. A. (2009). Otsenka vozmozhnostey sredstv akusticheskoy (rechevoy) razvedki. Spetsial'naya tehnika, 4, 49–63.

23. Rybal'skiy, O. V., Solov'ev, V. I., Zhuravel', V. V. (2017). Fraktal'niy podhod k viyavleniyu sledov tsifrovoy obrabotki v analogovyh fonogrammah. Suchasna spetsialna tekhnika, 1, 4–9. Available at: http://nbuv.gov.ua/UJRN/sstt_2017_1_4

24. Solovyov, V., Rybalsky, O., Zheleznyak, V. (2014). Multifraktal structure of whisper and recognition of speech structures. Vestnik Polotskogo gosudarstvennogo universiteta. Seriya C, Fundamental'nye nauki, 12, 16–20. Available at: http://elib.psu.by:8080/handle/123456789/11215

25. Vakulenko, M. O. (2010). Akustychni invarianty ukrainskykh pryholosnykh. Naukovyi visnyk kafedry YuNESKO Kyivskoho natsionalnoho linhvistychnoho universytetu. Filolohiya, pedahohika, psykholohiia, 20, 4–16. Available at: http://nbuv.gov.ua/UJRN/Nvkyu_2010_20_3

26. Dobrushkin, H. O., Danylov, V. Ya. (2010). Osnovni pidkhody do rozpiznavannia movlennievoi informatsiyi (Chastyna 1). Visnyk Vinnytskoho politekhnichnoho instytutu, 4, 50–64. Available at: https://visnyk.vntu.edu.ua/index.php/visnyk/article/view/1746

27. Arkhypova, O. O., Zhuravlov, V. M., Kumeiko, V. M. (2009). Artykuliatsiyni tablytsi sliv ukrainskoi movy. Pravove, normatyvne ta metrolohichne zabezpechennia systemy zakhystu informatsiyi v Ukraini, 2 (19). 13–17. Available at: http://ela.kpi.ua/handle/123456789/9689

28. Arkhypova, O. O., Zhuravlov, V. M. (2009). Chastotnyi analiz vykorystannia bukv ukrainskoi movy. Radio Electronics, Computer Science, Control, 2, 53–56. Available at: http://ric.zntu.edu.ua/issue/viewIssue/1582/pdf_11

29. Sushko, S. O., Fomychova, L. Ya., Barsukov, Ye. S. (2010). Chastoty povtoriuvanosti bukv i bihram u vidkrytykh tekstakh ukrainskoiu movoiu. Ukrainian Information Security Research Journal, 12 (3 (48)). doi: https://doi.org/10.18372/2410-7840.12.1968

30. Babenko, T. V., Sushko, S. O. (2012). About an entropy of Ukrainian language. Ukrainian Information Security Research Journal, 14 (3 (56)), 104–107. doi: https://doi.org/10.18372/2410-7840.14.3397

31. Kulchytskyi, I. M., Shandruk, U. S. (2015). Vplyv orfohrafiyi na chastotu bukv u tekstakh. Visnyk Natsionalnoho universytetu «Lvivska politekhnika». Informatsiyni systemy ta merezhi, 814, 300–309. Available at: http://nbuv.gov.ua/ujrn/vnulpicm_2015_814_30

32. Nuzhnyi, S. M., Zanoskina, P. V. (2018). General Approaches to the Formation of Articulation Tables of the Ukrainian Language to Assess the State of Protection of Excessive-Noise Speech Information. Suchasna spetsialna tekhnika, 4 (55), 66–75. Available at: http://suchasnaspetstehnika.com/journal/ukr/2018_4.pdf