UDC 004.934.2: 534.442.6
DOI: 10.15587/1729-4061.2018.144146

THE OBJECTIFIED PROCEDURE AND A TECHNOLOGY FOR ASSESSING THE STATE OF COMPLEX NOISE SPEECH INFORMATION PROTECTION

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*
E-mail: s.nuzhniy@gmail.com
L. Parkhuts
Doctor of Technical Sciences, Professor
Department of information security
Lviv Polytechnic National University
Stepan Bandera str., 12, Lviv, Ukraine, 79013
E-mail: par7@i.ua
Yu. Kasianov
Senior Lecturer*
E-mail: yukas.nik.ua@gmail.com
*Department of Computer Technologies and Information Security**
**Admiral Makarov National University of Shipbuilding
Heroiv Ukrainy ave., 9, Mykolayiv, Ukraine, 54025

Зазвичай при побудові систем захисту мовної інформації для систем постановки активної завади використовують генератори білого шуму. Визначення рівня захищеності мовної інформації від витоку акустичними та вібраційними каналами відбувається за відповідними нормативними методом та технологією. Але використання зловмисником багатоканальних методів перехоплення мовної інформації та сучасних методів обробки цифрових фонограм (вейвлет-перетворення, кореляційного аналізу та інше) дозволяє отримати несанкціонований доступ. Спроби використання генераторів мовоподібних завад, що засновані на використанні білого шуму (та його кольорових клонів), методів реверберації, методу «мовний хор» та інших не вирішують поставленої задачі.

В рамках дослідження запропоновано спосіб подолання цих труднощів. Він заснований на використанні генераторів мовоподібної завади скремблерного типу та застосуванні об'єктивізованого методу та технології оцінки ступеню захищеності мовної інформації на межі контрольованої зони. Об'єктивізований метод поєднує методи визначення критеріїв залишкової розбірливості мови (методи Покровського та Speech Intelligibility Index), методи фільтрації складнозашумлених акустичних сигналів (вейвлет-перетворення, фонемно-кореляційний аналіз та інші) та метод порівняння тест-сигналу в точці розміщення джерела сигналу і на межі контрольованої зони. Це дозволяє підвищити достовірність отриманої оцінки рівня захищеності мовного сигналу від витоку акустичними та вібраційними каналами за межі контрольованої зони.

Для дослідження рівня захищеності мовної інформації при різних типах завади та різних співвідношеннях сигнал/завада розроблено імітаційну модель експерименту.

Запропоновано технологію синтезу тест-сигналів на основі випадкових фонограм та/або фонограм озвучення дикторами артикуляційних таблиць і мікшування з різними типами завадового шуму при заданих співвідношеннях сигнал/завада.

Дослідження проведені в блоці «Wavelet 1-D» середовища Matlab. Встановлено, що при використанні шумової завади типу білий шум та співвідношеннях сигнал/завада –20…24 дБ запропонована методологія збільшує залишкову розбірливість тест-сигналу з $W \leq 10$ % до $W \approx 40…60$ %

Ключові слова: захист мовної інформації, акустичні завади, розбірливість мови, очищення мовного сигналу

1. Introduction

Speech information is the object of special attention from criminals wishing to operatively obtain fresh information about the activities of government agencies, various businesses, and organizations, as well as individuals. Thus, the summer of 2013 saw the scandal associated with the former CIA officer Edward Snowden. Over the period of 201–2014, the world's leading mass media were publishing revelatory materials related to the activities of secret services from the United States of America – National Security Agency (NSA) and the Central Intelligence Agency (CIA). The included information on listening to higher officials of foreign states, political and public figures. In other words, even the most protected representatives of society cannot be protected.

White noise generators (or of its colored clones) are typically applied for the systems of active interference when designing systems for protecting speech information. The level of speech information protection against leaking through acoustic and vibrational channels is determined by employing the appropriate normative method and technology. However, use of the multi-channel methods for intercepting speech information or modern methods for processing digital phonograms (wavelet transform, correlation analysis, etc.) can allow the intruder to partially or even completely restore speech information.

Thus, there is an issue on creating new approaches to the tasks of building the systems to protection verbal information implying the development of the objectivized methodology and a technology to assess the degree of speech information protection at the border of controlled zone.

2. Literature review and problem statement

Research into speech information protection has been developing continuously, however, specialized reports related to findings in this field were first revealed in the first half of the 20th century. Thus, paper [1], when examining the impact of acoustic interference in the form of periodic and non-periodic meanders on the speech signal intelligibility, analyzed research results by several other scientists who had reported their studies over 1937–1949, including «Communication in the presence of noise» [2]. Authors of the same paper investigated the impact of the frequency switch acoustic signal and a noise interference in different combinations and at different ratios of signal to interference. That allowed the authors to determine the optimal ratios of the speech signal and the noise interference, the intensity of signals, temporal and spectral composition, etc. Their recommendations made it possible to design systems of speech information protection both during transmission of signals along the line of communication and in order to protect separate premises.

Part of the results reported by the authors of the work was described as an «unusual effect» and led to forming a new direction of research, bioacoustics. Its purpose was to study and to optimize parameters of the technical means of measurement, to restore and process sound oscillations considering the peculiarities of the structure of the human body, as well as psychological aspects. Bioacoustics achieved considerable progress in the development of specialized complexes for carrying out research under «field conditions». These include in the first place the means to investigate acoustic sources at a considerable distance, such as mobile complexes for wildlife researchers, hydro-acoustic tools, etc. However, its main achievements were recommendations for developers of musical equipment (microphones, speakers, synthesizers, etc.), tools for analog and digital processing and reproduction of acoustic signals. Considerable success was achieved in the development of medical technologies (research into damage and abnormalities in the organs of speech and hearing, conducting psychophysical studies of the brain, etc.) and means to improve the hearing of humans.

In the field of bioacoustics, *ASA* compiled recommendations and standards, such as [3–5] and others, which have acquired the status of *ANSI* adopted at international level. Standard [5] introduces the term «*Speech Intelligibility Index*» (*SII*) and the procedure for its calculation. SSI is the coefficient that characterizes speech intelligibility under various adverse conditions of listening, such as masking with noise (interference), filtering, reverberation, etc. The coefficient varies from 0 to 1 and is calculated from dependence:

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

where n is the number of bands based on which one calculates a speech intelligibility index (typically $n=6$ or 7, when using octave bands, or $n=18$ or 21, when using 1/3-octave bands); I is the information significance coefficient in the band; A is the coefficient of information recognition quality in the band.

The introduction of SSI is a further development in the procedure for determining the level of ability of the listener to recognize words in a message voiced by an announcer. The coefficient is introduced as a name change for the factor (index) of articulation. In this case, some improvements were made. In *ANSI/ASA* S3.5-1997, coefficient of information recognition quality was replaced with a function that depends on the difference between the peak level of a speech signal in the band and the effective level of noise (interference):

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

With respect to:

$$\Delta L = L'_S - L^e_n = 20 \cdot \lg \frac{P'_S}{P^e_n},$$

where L'_S is the peak level of the speech signal and the effective level of noise (interference) in dB, respectively; P'_S and P^e_n are the power of the specified signals in W; we can pass over to the signal/interference ratio, which is usual in Ukraine.

However, such a transition in determining the coefficient/function of information recognition quality has led to a certain incorrectness, replacing the peak level of signal in the band with the root-mean-square level. Correction of the results reported in [7–10] involved the introduction of an additional shift of the results obtained for a certain normalized magnitude. Thus, in the analysis shown in [12], the following limits for such a shift are specified, from 5 to 20 dB. In this case, the shift is applied to the entire frequency spectrum, rather than for individual bands.

An analysis of the methods for determining the level of speech signal protection reveals that there are three groups at present:

1) subjective or articulation methods, based on determining an index (or function) of speech intelligibility;

2) objective (or instrumental) methods, which determine the signal/interference ratio;

3) objectivized methods, which are trying, through instrumental methods, to determine the impact of a noise interference on the speech signal.

At the same time, it should be noted that underlying all the above-mentioned methods are the studies reported in [1, 2, 6], as well as others, which substantiate the application of SSI.

3. The aim and objectives of the study

The aim of this study is to explore a possibility to construct a system of speech information protection, capable to meet the requirements in terms of protection from leaking via acoustic and vibration channels, while using modern methods and technologies for the filtration of noise interference.

To accomplish the aim, the following tasks have been set:

– to define the priority method of forming an interference signal;

– to define the criteria for estimation of the level of speech signals protection;

– to propose a simulation model of the experiment to study the level of speech information protection at different types of interferences and for different signal/interference ratios.

4. Adaptation of SSI to the objectivized estimation method for determining the level of speech information protection

4. 1. Analysis of methods for forming the interference signals

The idea to employ specialized acoustic signals (noises) and systems (devices) for the protection of separate facilities (areas, buildings, etc.; hereafter referred to as «controlled zones» (CZ)) first appeared in the first half of the 20th century [1, 2, 6, 13], and by the mid-century it helped form the field of «Acoustic noise control systems» [14]. In this case, the basic method of studying the level of impact of a noise interference on the speech signal is the articulation method [5, 6, 11–14].

In our time, the specified operations are regulated in the United States, Canada, and the EU by a number of normative documents, the main being [5, 11, 15].

In the post-Soviet space, the use of the speech intelligibility index for determining the level of its protection was proposed in [16–18]. The basis of the studies was a procedure by Pokrovski [10], which is associated with the articulation peculiarities of the Russian language. The main idea of research is to determine the level of residual intelligibility of speech signal after its exposure to passive and active interferences, basically the same speech intelligibility index, only the working zone of the criterion is substantially shifted towards the region of negative values for the signal/interference ratio.

The idea gained further momentum in papers [19–23] and others. However, to prevent the recognition of a speech signal, the basic principle of protection has remained the use of an interference signal (white, colored, or speech-like types). It should also be noted that the term «speech-like» implies the modification of the form of a white noise signal; it is adapted to the characteristics of even volume.

Papers [22, 23] proposed methods that started considering the spectra of speech signal and structural units (phonemes) along the 1/3-octave bands. In this case, the protection method remains, again, the same; setting an interference based on white noise in order to hide unsafe frequencies.

The main disadvantages of the specified methods are:

– subjectivity of the procedure;

– significant amount of work to be done in the examination of OIA (object of information activity OIA);

– vulnerability of the method of speech information protection based on the use of systems for setting active acoustic and vibration interferences based on white noise generators and its colored clones and the «speech-like» type, as well as reverberation.

The listed disadvantages of the methods are generalized for all the sources considered and thus require further investigation.

At the same time, the development of mathematical methods of signal processing in telecommunications systems has led to the understanding of the necessity of creating new approaches to the formation of an interference noise signal. The end of the 20th century was marked by the emergence of two directions in the formation of speech-like interference [20, 21]:

1) the signals that are synthesized based on a speech signal, which must be protected (in the scientific literature they are often called the correlated interferences, such a name in most cases now is not correct any more). This type includes, first and foremost, the interferences of reverberation type, interferences with permutation and/or inversion of frequency bands, and others. Their main feature is the possibility of dynamic change in the volume of interference noise depending on volume of the speaker's speech;

2) the interference signals that use third-party speech sources. The most common are such interferences as the Language Choir (aka «sound of the crowd»).

An analysis of the above types of speech-like signals reveals the existence of certain limitations in their application. Thus, the first type of signals are characterized by such drawbacks as the presence of sufficient volume of information to synthesize an interference filter. This is predetermined by the following factors:

– when using the reverberation as a principle of forming an interference, a spectrogram clearly demonstrates the presence of the specified effect. It can be quite easily established by running a correlation analysis;

– when using the band-inverse and band-permutation methods of forming an interference, the difficulties are only related to using dynamic keys and the unconventional division of spectrum from a speech signal into bands. Using neural network filters makes it possible to restore speech of a speaker with a delay, in most cases, not longer than a few minutes;

– applying a dynamic change in the volume of an interference signal may substantially improve the bio-acoustical characteristics in premises. However, its application provides additional information for the neural network filtering; spectrograms clearly demonstrate the splitting of an announcer's mode into words.

Given that the above list of shortcomings is incomplete (including only the most important), we can consider a given method of forming an interference insufficiently reliable.

Using the method of forming an interference of the second type provides a better level of the announcer's speech signal protection. This is predetermined by that the restoring of speech requires a partial or complete restoration of speech by all the announcers on the basis of whom the interference signal was created. The level of recovery depends on the level of influence and the similarity of components of the interference signal to the announcer's speech. In this case, it is necessary to consider the following peculiarities:

– in most cases, a limited list of phonograms is used to create a signal of the interference (for the most part, it does not exceed 6–8);

– the length of phonograms is limited and does not exceed 10–15 minutes, which leads to applying the loop procedure loops;

– in most cases, the algorithm of forming an interference signal is static, it is formed during initialization of the interference generator early in the operational cycle, and it is unchanged over a given session;

– when using the dynamic code of change in phonograms, for example 4 active phonograms out of 8, a substantial increase in the interference generator's RAM is required. Phonograms are retrieved from the database of the device or available remotely (for example, from radio, the Internet, and other sources). Using them requires downloading them into the memory of the generator and control over recovery. In this case, it is essential that professionally-designed phonograms should be used; it is not recommended to employ phonograms from the announcers that are not characteristic of a given room. This, in turn, significantly restricts the list of third-party sources.

The specified list of comments somewhat limits the use of the indicated type of an interference generator. However, such interference generators, while taking into consideration the comments, are capable of creating quite a lasting protection, even for neural filters.

On the other hand, in the absence of strict restrictions on the time of processing a phonogram of the intercepted speech signal, an intruder can design appropriate neural-network filters to restore the announcer's speech. The first three items on the list might contribute to that.

It should also be noted that the past decades have witnessed the emergence of devices of the combined type. The principle of their action is the simultaneous application of the first and second methods for formation an interference signal. Thus, papers [24–26] considered a generator that employs a method of the Language Choir from 6 active sources from a database of 10 sources, the combinations of the temporary band-inverse and frequency band-inverse permutations. The main difference from analogs is the use of a feedback from the emitted interference signal. This generator is termed, as proposed by authors of [24–26], a generator of speech-like signal of the scrambler type.

The use of this type of a speech-like interference generator creates a very complicated combination of articulation elements, including those created using the announcer's speech, which makes the application of neural-network filters impractical; blocks of the filter must be trained for each interim step in the presence of considerable ambiguity.

Thus, paper [27] reports results of using a neural network (RBF Network) and a wavelet transform to filter the interference of different types. Authors show the possibility of detecting the presence of a speech signal when using a quasi-periodic interference signal that has the spectral power of density similar to speech (for example, music, interference of the Language Choir type, and others). Studies have shown that at a signal/interference ratio of 2:1 (3 dB), when using neural networks (RBF Network), and 1:1 (0 dB) when employing a wavelet transform, it still is possible to restore a speech message. However, the specified magnitudes are critical.

In recent years, there have been studies in which a value for the signal/interference ratio for neural networks was reduced to –3 dB, and some authors suggest even –6 dB. However, it is significantly less than the required level [12].

4. 2. Residual speech intelligibility index

In [15–17], it was suggested using SSI to determine the level of its protection. That was predetermined by the broad implementation of speech-like interference signals for systems of information protection, the development of new methods and technologies for phonogram processing. In this case, it was principal to understand that the use of a signal/interference ratio does not make it possible to correctly determine the level of speech protection, which is an integral coefficient. That is, it does not account for a significant amount of articulation factors of the human speech.

At the same time, a simple return to the articulation method and determining a speech intelligibility index is not an alternative, as it requires significant human and material costs. It is especially associated with the new methods of processing digital phonograms (wavelet transforms, neural-network filters, etc.), which make it possible to process complex noise signals. An important factor in carrying out such processing is introducing certain distortions to a speech signal, which are related to influencing its spectral composition due to the impossibility to perfectly separate the harmonics of an interference from the speech signal.

Thus, there is a need for the introduction of a new criterion in determining the level of speech signal protection – a residual speech intelligibility index (or a coefficient of destructive changes). This coefficient must take into consideration a curvature in the spectral composition of minimum speech units (it is proposed to apply allophones) along the 1/3-octave bands with respect to information weight coefficients in the band and the information recognition quality in the band.

The procedure also takes into consideration the factor of importance of allophone recognition for the recognition of a word (phrase, sentence) and the frequency coefficient of using a given allophone in speech (it is possible to replace with the frequency of allophone usage in articulation tables).

Thus, expression (1) to calculate *SII* from [5] takes the form:

$$S^R = \frac{1}{m}\sum_{y=1}^{m}\left[R_y^a H_y \sum_{i=1}^{n} I_i A_i\right], \qquad (2)$$

where S^R is the residual intelligibility index (RII); m is the total number of allophones in a word; R^a is the coefficient of importance of an allophone recognition for the recognition of a word (phrase, sentence); H is the frequency coefficient of using a given allophone in speech.

The proposed dependence defines a possibility of transition to the objectivized estimation method for determining the level of protection of the complex noise speech information.

An important element of the proposed method is the development of a map for splitting a test speech signal (hereafter, a test-signal), read by an announcer directly during the experiment or as a phonogram, into allophones. To create the test-signals, we refer to articulation tables [5, 10, 22–24, 28].

The maps for splitting a test-signal into an allophone contain information about the number and type of allophones (m), which comprise words in the test-signal.

Each allophone, based on its type, is given a coefficient of the recognition importance (R^a). The coefficient indicates the degree of influence of the recognition accuracy of a specific allophone on content of the test-signal depending on the distribution of its frequency spectrum along the 1/3-octave bands. This is important in the recognition of assonant allophones and when considering the peculiarities of the speaker's speech (a dialect, speech defects, etc.). In analyzing the complex noise speech information, this factor takes into consideration the impact of an interference and/or distortion in the frequency spectrum that occur during transforms and filtering.

Coefficient of allophone speech usage frequency (H) indicates the probability of the occurrence of a specific allophone in the analyzed phonogram. Introduction of the coefficient is predetermined by the influence of an interference and/or distortion in the frequency spectrum that occur during transformation and filtration on the quality of a specific allophone recognition.

Specific values for the coefficients of allophone recognition importance and the frequency of its use are determined from a special chapter in articulation tables.

Procedure for determining the R^a and H coefficients for the English language is given in [3–5] and several other documents.

For the Russian language, determining the R^a and H coefficients is regulated by GOST R 50840-95. However, it does not make it possible to account for the use of slang, dialect, languages of national minorities, and so on.

A significant difference in phonetics between the Ukrainian and Russian languages leads to an incorrect application of articulation tables based on GOST R 50840-95. Attempts to design such tables for the Ukrainian language were described in papers by Arkhipova ([28] and others), however, the study was not completed. The indicated task is relevant in Ukraine.

That introduces the limitation in the use of the Ukrainian language when investigating the level of speech information protection at an object using the SSI (1) criterion and/or a ratio of residual speech intelligibility (2).

5. Simulation model of experiment

5. 1. The purpose and main tasks solved with the simulation model

Conducting a field experiment aimed at studying the effect of different types of an interference and at different signal/interference ratios on the level of speech information protection, as most other studies, is too expensive. For such experiments, the prerequisite is to employ, during experiments, actual objects, which limits their use for their purposes, as well as several types of interference generators.

Paper [26] proposed a laboratory complex to conduct a study, which greatly reduces the cost of conducting experiments. Such a complex makes it possible to simulate any premises. However, in this case, it requires considerable time to prepare the experiment; it is necessary to ensure compliance of the integral and 1/3-octave model parameters with the acoustic parameters (specifications) of an actual room and its structural elements.

The authors' proposed procedure for conducting computer modelling of the study process makes it possible to significantly reduce the costs of time, to extend the range of the examined objects and the types of interference generators.

The main tasks of simulation model of the experiment are:

1) development of test-signals:

a) development of specialized announcer's test-signals based on the voice-over text and random phonograms of varying complexity (voice, voice + music, voice + random thematic signals, etc.);

b) development of an interference test-signals (interference generator, natural and technogenic backgrounds);

c) synthesis of the examined test-signal – mixing the announcer and interference test-signals at the predefined signal/interference ratio based on the integral and 1/3-octave parameters;

d) application of specialized filters to simulate the effects of acoustic parameters of an actual room and its structural elements on the examined test-signal;

2) running a filtration procedure of the examined test-signal;

3) design of the map for splitting the announcer and examined test-signal before and after filtering for allophones;

4) running a correlation and spectral analysis of the maps of allophone splitting;

5) calculation of the speech signal protection level based on RII for (1) and (2).

5. 2. Sequence of research

Our research was conducted in the programming environment MatLab R2015b, block «wavelet 1-D». The test-signals were prepared in the environment Sony Sound Forge Pro 11, which prevents the possibility of using announcer's test-signals during filtration procedure.

The study was conducted in 7 stages:

1) preparation of test-signals (voice-over, interference, and examined);

2) design of the map of splitting the announcer's and examined (prior to filtration procedure) test-signals into allophones;

3) preliminary analysis of the examined test-signal (function *Analyze*, block «wavelet 1-*D*»);

4) filtration of the examined test-signal (function *Denoise*, block «*wavelet* 1-*D*») and analysis of its result (function *Analyze*, block «*wavelet* 1-*D*»);

5) design of the map of splitting the examined test-signal into allophones after a filtering procedure;

6) running a correlation and spectral analysis of allophones splitting maps;

7) calculation of the speech signal protection level based on RII for (1) and (2); analysis of the derived values.

5. 3. Design of test-signals

The main feature in the preparation of test-signals is the possibility for their software setting. At such setting, it is necessary to take into consideration that the sound editors operate the scales that are not within a classic acoustic range (dBA), but the scale dBfs (dB – decibels, FS=Full Scale). A dBfs scale's feature is the match of 0 in dBfs to the level of a full signal at ADC of the audio device (all unities). This is related both to the need to prevent the occurrence of the clipping and dithering phenomena and the necessity to protect the source registers of ADC from overload.

The synthesis of the examined test-signals was performed for the typical interference/signal ratios used in the systems of technical protection [5–11], (in dB): –3; –6; –10; –15; –18 and –25.

Fig. 1 shows the phonogram (announcer's test-signal) with a length of 5.5 s and with selected 10 blocks of 0.55 a (*a*), the spectrogram of a long-range spectrum for blocks (*b*), and the spectrogram of a long-term spectrum for a section (*c*). Basic statistical characteristics of the phonograms are given in Table 1.

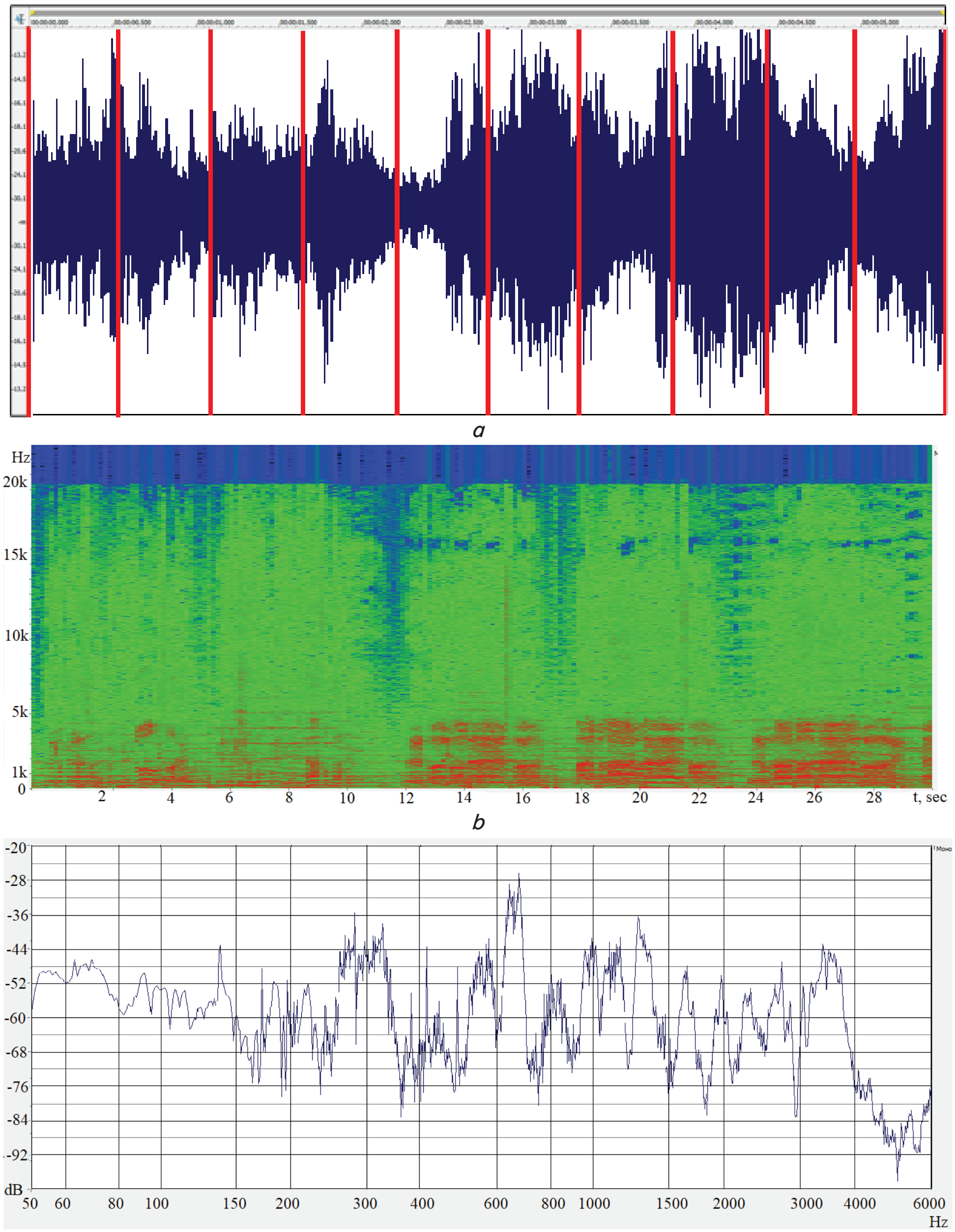
Table 1

Statistical characteristics of a phonogram

Cursor position (Time)	00:00:05.500
Position of sample with a mimimal value (Time)	00:00:03.685
Sample minimal value (dB)	–11.014
Sample maximal value (dB)	–9.980
Root-mean-square level (dB)	–24.998

In Fig. 2, lines show the selection of a long-range spectrum (Fig. 1, *c*) of the 1/3-octave bands (green color) in the spectrogram, values for the root-mean-square level (red color) and the root-mean-square level of the 1/3-octave bands (blue color). Fig. 2 demonstrates that values for the root-mean-square level at individual 1/3-octave bands and the actual level of a speech signal peak may differ by a magnitude exceeding 16...18 dB.

Fig. 3 shows spectrograms of the examined test-signal, which was synthesized from the announcer's test-signal and the signal of interference, the type of «white noise», with a ratio of –21...–24 dB, which correspond, according to [16, Table 9], to the ratio of speech intelligibility W=10 %.



c

Fig. 1. The phonograms of the announcer's test-signal: *a* — announcer's test-signal (phonogram «In memory of Caruso») split into 10 blocks (1/3-oktave); *b* — sonogram of the long-range spectrum for the blocks of the announcer's test-signal; *c* — integrated spectrogram of the long-term spectrum of the announcer's test-signal

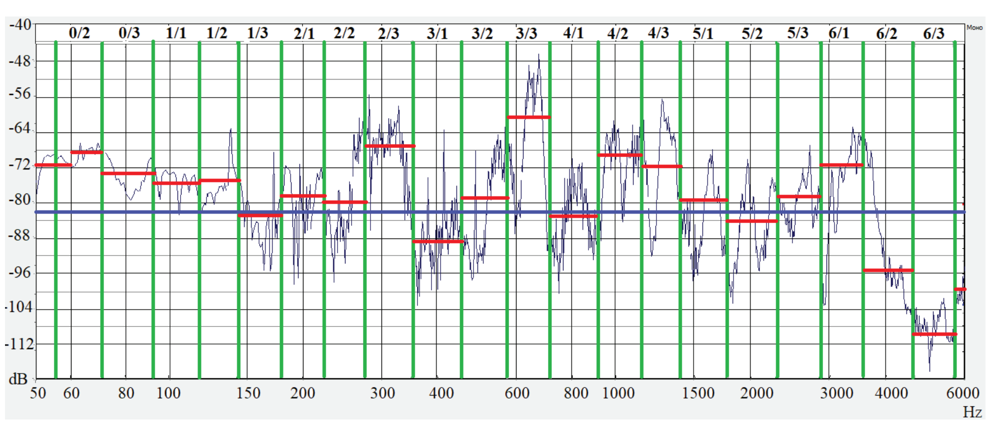


Fig. 2. Analysis of the selection of root-mean-square levels in the spectrogram of the announcer's test-signal (block 1)

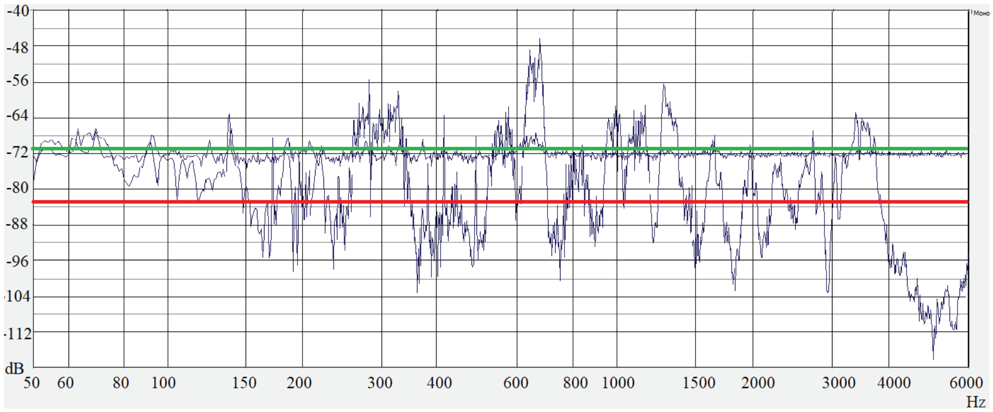


Fig. 3. Analysis of the spectrograms of the test-signal and the signal of an interference, the type of «white noise», with a ratio of –15...–18 dB

5. 4. Analysis of research results

The procedure of filtering and noise purification using a wavelet analysis (software MatLab, block «wavelet 1-D», functions Analyze and DE-NOISE) results in a decrease in the signal-interference ratio.

Fig. 4 shows results of research into a simulation model based on the announcer's test-signal «Vulnerable Zone». The synthesis of the examined test-signal was carried out at the signal/interference ratio –21 dB (Fig. 4, top), which corresponds, in line with [16, Table 9], to the speech recognition level $W<10$ %.

Based on the results of filtering and noise purification, the signal/interference ratio increased to the level –8...–5 dB (Fig. 4, bottom) with the speech recognition level $W>60$ %.

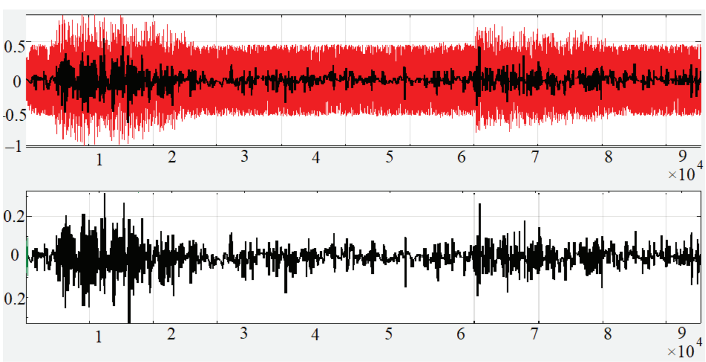


Fig. 4. Analysis of the results of noise purification of test-signal at the signal/interference ratio –21 dB, based on the method of a wavelet analysis (software MatLab, block «wavelet 1-D»)

An analysis of the results obtained has revealed the improvement in the quality of a phonogram to the level at which it is possible to not only recognize unambiguously an information message, but to run an analysis of allophones in the announcer's speech.

For comparison, we split into allophone units the announcer's (Fig. 5, *a*) and the examined test-signal, which had undergone filtering and noise purification (Fig. 5, *b*).

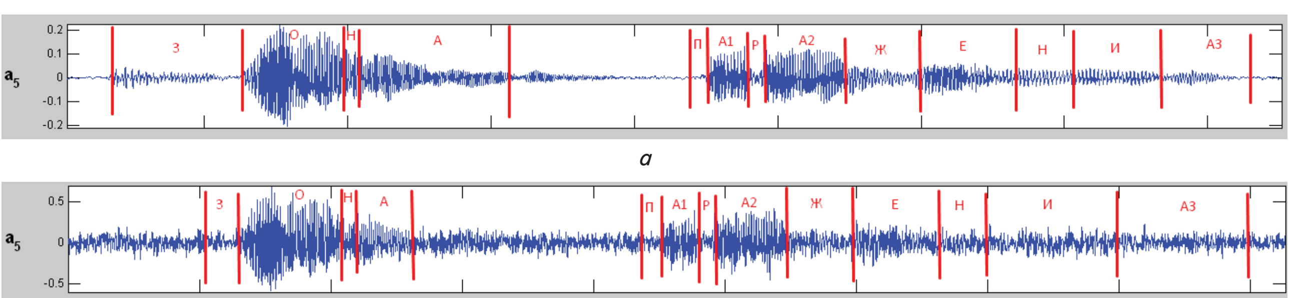
The Figures show that in most cases the characteristic attributes of allophones, despite the significant impact of a noise interference and a large number of transformations (filtering), were retained. The differences are characterized by two factors:

1) objective – the impact of a noise interference on wavelet transforms predetermined the emergence of wavelet shifts along the temporal sweep;

2) subjective – the splitting into allophones was done under a manual mode, that is, there is a presence of the significant influence of noise signal residuals on defining the beginning/end of an allophone.

An analysis of the results obtained (Fig. 4, 5) reveals that such an increase in intelligibility is critical and allows the termination of investigating a given examined test-signal.

The main direction for solving the considered task (high-quality filtering of a noise interference, the type of «white noise») is the transition to methods that enable a destructive influence on a speech signal. Such properties are characteristic of the speech-like interferences that are formed from a speech signal, subject to protection. However, application of such methods only is insufficient. As was indicated above, they have significant systemic shortcomings.



a

b

Fig. 5. Splitting into allophones:
a – announcer's test-signal; *b* – examined test-signal that underwent filtering and noise purification

Thus, the most rational one is the generator of a speech-like signal of scrambler type proposed in [25]. Its use in the systems of speech information protection would make it possible to obtain the required level of protection even at a decrease in the signal/interference ratio to –12 ...15 dB.

6. Discussion of results of studying the level of protection of complex noise speech information using a simulation model

The insufficient level of speech information protection when employing the white noise generators in the systems for setting active interference, has been discussed for quite a long time – since the mid-1990's. However, there have been two problems:

1) how can one determine, with the predefined accuracy, the actual level of speech information protection?

2) what can replace white noise generators?

To address them, an extensive method has been used – increasing the level of an interference at the borders of controlled zones to the magnitude at which existing methods do not make it possible to recover information.

At the same time, the development of methods to determine the level of speech protection has been always evolving. The main result is the introduction of *SSI* in [5] and the language intelligibility of speech [16–18]. However, their main disadvantage is the dependence on the level of a signal/interference rate.

To separate a speech signal from an interference, a number of methods that are considered above have been proposed. Their common characteristic, as was shown above, is the limitation in the level of an interference – the ratio signal/interference should not be worse than –10 dB.

We have proposed an improved objectivized method and a technology for research, which provide the opportunity to determine the level of speech information protection at significantly greater levels of interference. To test their efficiency, a simulation model of the experiment has been suggested, which is maximally close to a field study and could be integrated into an actual research complex.

The advantages of the method include the application of standard filtering functions for complex noise phonograms (executed in block «wavelet 1-D» from the programming environment Matlab) and determining the impact of noise and filtering processes on allophone recognition. That necessitated the introduction of a residual speech intelligibility coefficient (refer to (2)). At present, however, its application is limited due to the absence of articulation tables for the Ukrainian language.

This temporary constraint has resulted in the use, as a criterion for the level of speech protection, at this stage of research, of the indicator of speech language intelligibility W [16–18]. Its application also makes it possible to compare the results obtained.

An analysis of our results reveals a significant improvement in the quality of speech information recognition (test-signals). Thus, at the signal/interference ratio at the level of -10 ± 0.2 dB there is an almost complete recovery of the test-signal, which is confirmed by a change in the indicator of speech language intelligibility W from 20 % (refer to [16]) to 90 %. At the signal/interference ratio of -21 ± 0.2 dB the application of the procedure makes it possible to increase the indicator from $W\leq1$ % to $W>60$ %, corresponding to the signal/interference ratio of -6 ± 1 dB.

Limitations in the application of the procedure for the time being is the absence of the specialized articulation tables and the lack of acquired statistical data for the generalization, construction of calculation diagrams, and synthesis of analytical dependences.

7. Conclusions

1. We recommend using a method for the formation of a speech-like interference of scrambler type for the active systems of speech information protection. This is predetermined by that the speech-like interferences of scrambler type are resistant to contemporary specialized methods for processing digital phonograms and methods of noise filtering in phonograms. At the same time, the application of such methods in order to filter interferences based on white noise and the speech-like interferences of the first and second types ensures improved quality of speech information recognition to critical levels. Therefore, it would make it possible to design systems of speech information protection capable to meet requirements for protection against leaking via acoustic and vibration channels.

2. It is shown that an increase in the reliability of the resulting estimate of speech signal protection level against leaking via acoustic and vibration channels beyond the borders of controlled zone now becomes possible owing to the combination of methods for determining the criteria of residual intelligibility, methods for the filtration of complex noise acoustic signals, and a method for comparing a test-signal at the point of a signal source location and at the border of controlled zone.

Thus, in particular, an analysis of methods for determining a speech intelligibility index (SII) and that by Pokrovski has allowed us to propose, as a criterion for estimation of the speech signal protection, a coefficient of residual speech intelligibility. That made it possible to correctly account for the distortion in the spectrum of a speech signal (test-signal) through the influence of an interference signal and filtration process.

By employing the filtering of complex noise acoustic signals, we achieve an almost complete restoration of the test-signal, which is confirmed by a change in the index of speech language intelligibility W from 20 % to 90 % at the level of signal/interference of -10 ± 0.2 dB.

Comparison of the test-signal at the point of a signal source location and at the border of controlled zone is carried out in accordance with the standard method and technology, taking into consideration the proposed amendments.

3. We have proposed a simulation model of the experiment to explore the level of speech information protection at different types of interference and at different signal/interference ratios. The procedure of the experiment consists of 7 stages: synthesis of test-signals based on random phonograms and/or phonograms of articulation tables, recorded by announcers, and mixing them with different types of interference noise at the assigned signal/interference ratios, and others. We conducted our study in the block «wavelet 1-D» from the programming environment Matlab. It was established that when using a noise interference of the white noise type and at the signal/interference ratio –20...24 dB the proposed procedure, based on the results of investigating the simulation model, improves residual intelligibility of the test-signal from $W=10$ % to $W=40...60$ %. Thus, we have brought to light the shortcomings of existing systems for speech information protection based on an interference of the white noise type.

References

1. Miller G. A., Licklider J. C. R. The Intelligibility of Interrupted Speech // The Journal of the Acoustical Society of America. 1950. Vol. 22, Issue 2. P. 167–173. doi: https://doi.org/10.1121/1.1906584
2. Shannon C. E. Communication in the Presence of Noise // Proceedings of the IEEE. 1998. Vol. 86, Issue 2. P. 447–457. doi: https://doi.org/10.1109/proc.1984.13079
3. ANSI/ASA S3.42-1992 PART 1 (R2017) American National Standard Testing Hearing Aids with a Broad-Band Noise Signal. STANDARD by American National Standards of the Acoustical Society of America, 1992. 19 p.
4. 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. 6 p.
5. 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.
6. French N. R., Steinberg J. C. Factors Governing the Intelligibility of Speech Sounds // The Journal of the Acoustical Society of America. 1947. Vol. 19, Issue 1. P. 90–119. doi: https://doi.org/10.1121/1.1916407
7. Bykov Yu. S. Teoriya razborchivosti rechi i povyshenie effektivnosti radiotelefonnoy svyazi. Moscow; Leningrad: Gosenergoizdat, 1959. 350 p.
8. Ioffe V. K., Sapozhkov M. A. K voprosu o metodike rascheta razborchivosti rechi // Trudy komissii po akustike AN SSSR. 1951. Issue 6.
9. Sapozhkov M. A. Rechevoy signal v kibernetike i svyazi. Moscow: Svyaz'izdat, 1963. 472 p.
10. Pokrovskiy N. B. Raschet i izmerenie razborchivosti rechi. Moscow: Gos. izd-vo literatury po voprosam svyazi i radio, 1962. 392 p.
11. Technical Specifications for Construction and Management of Sensitive Compartmented Information Facilities. VERSION 1.4. IC Tech Spec – for ICD/ICS 705. An Intelligence Community Technical Specification Prepared by the National Counterintelligence and Security Center, 2017. 174 p. URL: https://www.dni.gov/files/NCSC/documents/Regulations/Technical-Specifications-SCIF-Construction.pdf
12. Gavrilenko A. V., Didkovskiy V. S., Prodeus A. N. Sravnitel'niy analiz nekotoryh metodov ocenki razborchivosti rechi // Elektronika i svyaz'. 2008. Vol. 1. P. 227–231.
13. Licklider J. C. R. The Influence of Interaural Phase Relations upon the Masking of Speech by White Noise // The Journal of the Acoustical Society of America. 1948. Vol. 20, Issue 2. P. 150–159. doi: https://doi.org/10.1121/1.1906358
14. Rosenblith W. A., Stevens K. N. Handbook of Acoustic Noise Control. Vol. I-II. 1953.
15. ISO 9921:2003. Ergonomics – Assessment of speech communication. International Organization for Standardization, 2003. 28 p. URL: https://www.iso.org/standard/33589.html
16. Zheleznyak V. K., Makarov Yu. K., Horev A. A. Nekotorye metodicheskie podhody k ocenke effektivnosti zashchity rechevoy informacii // Special'naya tekhnika. 2000. Issue 4. P. 39–45.
17. Horev A. A., Makarov Yu. K. Metody zashchity rechevoy informacii i ocenki ih effektivnosti // Special'naya tekhnika. 2001. Issue 4. P. 22–33.
18. Horev A. A. Sposoby i sredstva zashchity rechevoy (akusticheskoy) informacii ot utechki po tekhnicheskimi kanalami // Special'naya tekhnika. 2005. Issue 5. P. 54–60.
19. Didkovskiy V. S., Didkovskaya M. V., Prodeus A. N. Akusticheskaya ekspertiza kanalov rechevoy kommunikacii: monografiya. Kyiv: Imeks-LTD, 2008. 420 p.
20. Objective and subjective assessment of bandlimited signaling speech quality / Mytiai Y. O., Zamsha K. S., Lozinskyi B. V., Stepanovska O. S., Prodeus A. M. // Electronics and Communications. 2016. Vol. 21, Issue 1. P. 18–26. doi: https://doi.org/10.20535/2312-1807.2016.21.1.82250
21. Bogdanova N. V., Prodeus A. M. Objective quality evaluation of speech band-limited signals // Electronics and Communications. 2016. Vol. 19, Issue 6. P. 58–65. doi: https://doi.org/10.20535/2312-1807.2014.19.6.113479
22. Rezul'taty eksperimental'nyh issledovaniy ocenki vozmozhnostey perekhvata rechevoy informacii pri realizacii metodov dvuhkanal'nogo s'ema / Bortnikov A. N., Gubin S. V., Lobov V. A., Siromashenko A. V., Chernyshov P. V. // Voprosy zashchity informacii. 2007. Issue 1. P. 11–17.
23. Lobov V. A., Siromashenko A. V., Chernyshov P. V. Ocenka vozmozhnostey perekhvata rechevoy informacii pri realizacii metoda mnogokanal'nogo s'ema // Voprosy zashchity informacii. 2007. Issue 4. P. 27–35.
24. Nuzhnyi S. M., Kasianov Yu. I. Otsinka zakhyshchenosti movnoi informatsiyi pry zashumlenni i synkhronnomu vymiriuvanni z dvokh tochok // Zakhyst informatsiyi i bezpeka informatsiynykh system: materialy IV-oi Mizhnarodnoi naukovo-tekhnichnoi konferentsiyi. Lviv, 2015. P. 175–176.
25. Kasianov Yu. I., Nuzhnyi S. M. Otsiniuvannia efektyvnosti heneratora realnoi movopodibnoi zavady za kryteriem rozbirlyvist movy // Visnyk Natsionalnoho universytetu «Lvivska politekhnika». Seriya: Avtomatyka, vymiriuvannia ta keruvannia. 2016. Issue 852. P. 105–110. URL: http://ena.lp.edu.ua:8080/handle/ntb/36473
26. Nuzhnyy S. M. Udoskonalena tekhnolohiya otsinky stupenia zakhystu movnoi informatsiyi // Suchasnyi zakhyst informatsiyi. 2018. Issue 1 (33). P. 66–73. URL: http://journals.dut.edu.ua/index.php/dataprotect/article/view/1796
27. Hoyt J. D., Wechsler H. RBF models for detection of human speech in structured noise // Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94). 1994. doi: https://doi.org/10.1109/icnn.1994.374996
28. Arkhypova O. O., Zhuravlov V. M., Kumeiko V. M. Artykuliatsiyni tablytsi sliv ukrainskoi movy // Pravove, normatyvne ta metrolohichne zabezpechennia systemy zakhystu informatsiyi v Ukraini. 2009. Issue 2/19. P. 13–17.