Hardware-software complex of a speech-like interference generator for speech information protection

V V Alekseev, A V Yakovlev, M V Moiseeva and A A Tikhomirova
Tambov State Technical University, Sovetskaya street, 106, Tambov, 392000, Russia
E-mail: mariyamoiseeva@mail.ru

Abstract. The article describes a model of a speech-like interference generator based on a system analysis of ways to solve the problem of protecting speech information in office-type premises. The results of the effectiveness of the model application are presented.

1. Relevance

Many companies suffer significant financial losses every year due to leakage of confidential information. To combat this problem, they allocate significant financial resources, take various protective measures, but often do not pay due attention to the technical channels of speech information leakage, which is a significant mistake.

In acoustic information leakage channel, the technical unmasking (intelligence) features of the objects of protection are acoustic waves. Such leakage channels are typical for acoustic verbal intelligence (for intercepting speech information from places of human communication) and acoustic signal intelligence (for obtaining intelligence information about acoustic «portraits» of various acoustic devices, whose work is accompanied by acoustic fields).

Speech is a source of acoustic information, the carrier of acoustic information is acoustic signals. In air, the signal propagates in the form of a longitudinal elastic wave, which is the vibration of air particles along the direction of wave propagation. The acoustic signal can be intercepted by an intruder using a microphone, which converts the signal into electrical and allows to transmit the signal outside the monitored room.

The main means of protecting speech information from leakage through these channels are soundproofing of premises, search for «bugs» and active acoustic masking. The basis of masking means are audio jamming devices. In practice, noise generators have found the most widespread use. Application of this method makes it possible to reduce the signal-to-noise ratio at the input of the intelligence equipment by increasing the level of interference.

However, acoustic interference created by technical means of information security should not bring significant discomfort to the participants in the dialogue present in the room. The problem arises of bringing the power of acoustic interference to an optimal level that meets the requirements of both the room security and comfortable conversation.

2. Analysis of ways to solve the problem

The process of speech perception in noise is accompanied by the loss of the components of the speech message. The intelligibility of a speech message is characterized by the number of correctly received
words that reflect the qualitative area of intelligibility, which is expressed in the categories of the details of the certificate of the intercepted conversation compiled by the attacker.

As a solution, it was proposed to use an instrumental-computational method for assessing speech intelligibility, which allows you to select the level of noise interference of the protection means in such a way that the radiation of the generator is minimal, but ensures the protection of speech information from leakage. The power that meets these criteria is proposed to be considered optimal. Thus, it is necessary to measure the parameters set by the method, calculate the verbal intelligibility and adjust the masking tool to the optimal radiation level.

The method of forming speech-like interference correlated in level, spectrum and time with a hidden signal is the most effective way to actively protect speech information.

The use of speech-like interference in the information protection system against leaking over an acoustic channel is not only an effective method of preventing speech information leakage, but also is a «softer» method in relation to people in the room. A lower noise level required to achieve the set level of protection (relative to the use of other types of generators), a signal spectrum close to human speech – these factors make working in the space to be protected more comfortable and calm.

In systems of active speech information protection in rooms for negotiations, it is possible to use speech-like signals and speech sequences formed taking into account the linguistic features of the language and the statistical characteristics of the occurrence of phonemes in this language, as well as the length of words and sentences as masking signals. An overview of the active protection system against unauthorized eavesdropping is shown in figure 1.

The formation of speech-like signals can be performed by the compilation method based on the structural units of speech. As a result, the speech-like signals formed in this way retain all the shades of the speech of a certain speaker, and it is very difficult to distinguish them from the information signals of the same speaker.

To improve the quality of synthesized speech, you can use exponential spline functions at the boundaries of the transition from one phonemic structure to another and by superimposing the end of one phonemic structure on the beginning of the second phonemic structure. In this case, there will be some faster attenuation of the oscillation amplitudes of the end of one phonemic structure and an increase in the amplitude of the beginning of the second phonemic structure. This mechanism of compilation speech synthesis will eliminate signal jumps at the boundaries of phonemic structures. To perform such a compilation speech synthesis, you need a database of phonemic structural units of speech with slightly enlarged segments in the time domain, since during the synthesis the end of one phonemic structure is superimposed on the beginning of the second phonemic structure.

The analysis of speech segmentation methods showed that for the formation of the base of phonemic structural units of speech for speech synthesis by the compilation method, the segmentation method using dynamic programming is the most convenient. To do this, you must have a phonetic record of continuous speech, labeled manually into phonemic structural elements and containing all phonemic
structural units necessary for the base. Usually these are 300-400 allophones for Russian, Kazakh, Belarusian speech and about 1200 phonemic structural units.

Since the phonetic bases of the structural elements of speech at the beginning and at the end contain transition areas, splines for the transition areas should be used in speech synthesis. It is recommended to use the so-called «stitching» of allophones when compiling speech synthesis. The transition section of the end of the previous allophone, multiplied by a decreasing function varying from 1 to 0, is superimposed on the transition section of the subsequent allophone, multiplied by an increasing function from 0 to 1. If the lengths of the overlapping transition sections are not equal, then the length of the formed transition region is chosen equal to the length of the longer transition section.

Table 1. Values of signal-to-noise ratios at which the required efficiency of protection of acoustic (speech) information is provided.

| Type of interference                        | Verbal intelligibility W, % | Signal-to-noise ratio $q_i$ in octave bands | Signal-to-noise ratio in the frequency band 180 ... 5600 Hz |
|--------------------------------------------|----------------------------|---------------------------------------------|----------------------------------------------------------|
|                                            |                            | 250, 500, 1000, 2000, 4000                 |                                                          |
| «White» noise                              | 20                         | +0.8, -2.2, -10.7, -18.2, -24.7             | -10.0                                                   |
|                                             | 30                         | +3.1, +0.1, -8.4, -15.9, -22.4              | -7.7                                                    |
|                                             | 40                         | +5.1, +2.1, -6.4, -13.9, -20.4              | -5.7                                                    |
| «Pink» noise                               | 20                         | -5.9, -5.9, -11.4, -15.9, -19.4             | -8.8                                                    |
|                                             | 30                         | -3.7, -3.7, -9.2, -13.7, -17.2              | -6.7                                                    |
| Noise with the decline in the spectral density of 6 dB per octave | 20                         | -14.1, -11.1, -7.4, -11.9, -15.6             | -13.0                                                   |
|                                             | 30                         | -12.0, -9.0, -11.5, -13.0, -13.5             | -10.8                                                   |
| Noise speech-like interference             | 20                         | -3.9, -7.9, -12.9, -15.9, -16.9             | -9.0                                                    |
|                                             | 30                         | -1.7, -5.7, -10.7, -13.7, -14.7             | -6.8                                                    |
|                                             | 40                         | +0.1, -3.9, -8.9, -11.9, -12.9             | -5.0                                                    |

Acoustic and vibroacoustic masking systems use interference of both “white” and “pink” noises, as well as speech-like interference. In protection complexes, interference is used to mask speech, which is similar in structure to masked speech. This can be interference from an external source or interference from a speech-like noise synthesizer by a phonemic cloner. The interference generated by such a synthesizer is not just speech-like, the phonemic cloner ensures the formation of such interference that maximally matches the sounds of the speech of a particular person or group of people whose conversations are protected from eavesdropping.

3. Model description
A model of a speech-like noise generator based on the Matlab mathematical package is proposed. The model is built on the basis of standard objects of the Simulink visual design environment. In the presented model, it is proposed to multiply the transition areas by linear functions varying from 1 to 0 and from 0 to 1. However, due to the fact that hearing sensitivity is nonlinear, it is more efficient to use spline functions of a higher order (up to the third degree).
Figure 2. Model of a speech-like interference generator with 6 shaper stages.

Studies of the model have shown that an increase in the number of allophone cascades leads to a change in both the spectrum of the generated speech-like interference and its amplitude. These changes are clearly visible in the transition from a three-stage scheme to a four-stage one.

Figure 3. Simulation results of a speech-like interference generator with a different number of shaper stages.

Figure 4 shows the spectra of the speech message of the interlocutors and the speech-like interference synthesized by the generator. The simulation result shows a sufficient similarity of these spectra.
4. Conclusion

Thus, a model of a speech-like interference generator based on the Matlab mathematical package has been synthesized. The model is built on the basis of standard objects of the Simulink visual design environment. The simulation result shows a sufficient similarity of the spectra of the speech message of the interlocutors and the speech-like interference synthesized by the generator.

References

[1] Tsaregorodtsev A V and Taraskin M M 2017 Methods and means of information protection in public administration (Moscow: Prospect)
[2] Lynkov L M, Golikov V F and Borbotko T V 2013 Fundamentals of information protection and intellectual property management (Minsk: BSUIR)
[3] Crow V A, Vorona V A and Kostenko V O 2016 Methods and means of protecting information from leakage through technical channels, Computational nanotechnology
[4] Buzov G A, Kalinin S V and Kondratiev A V 2005 Protection against information leakage through technical channels (Moscow: Hotline-Telecom)
[5] Zheleznyak V K, Makarov Yu K and Khorev A A 2000 Some methodological approaches to assessing the effectiveness of protection of speech information (Moscow: Special equipment)
[6] Khorev A A 2013 Methods of protecting the allocated premises from the leakage of speech (acoustic) information through technical channels: vibroacoustic protection systems (Moscow: Special equipment)