Method of speech signal scrambling based on matched wavelet filters

Oleksandr Lavrynenko[1,*,†]

[1] *National Aviation University, 1 Lubomyr Huzar ave., 03058 Kyiv, Ukraine*

Abstract

In this research study, a method of speech information protection using digital wavelet filter banks is proposed. An inverse scheme of single-level discrete wavelet transform is used to build the protection system. It includes digital synthesis and analysis filter banks. The filters used are synthesized using a key sequence. The key identifies the sender and receiver of the information and is used only in the filter synthesis stage. Also presented is a method for synthesizing matched wavelet filters satisfying the property of orthogonality of the wavelet basis, the presence of zero moments. The important requirements of the filters are the conditions of complete signal recovery and elimination of overlapping spectra. The results of the research show the effectiveness of synthesized wavelet filters in solving the problem of information protection. A speech protection algorithm is developed, which uses matched wavelet filters at the stage of building a bank of analysis-synthesis filters matched to the key. The algorithm has a simple implementation, and fast algorithms of digital signal processing (convolution, decimation, interpolation), allowing encrypting of the signal in real-time. The proposed algorithm is noise-resistant and can be used in channels with intensive interference. The algorithm is robust to time delays and hiccups, as well as distortions in the communication channel.

Keywords

speech signal, wavelet transform, packet wavelet transform, matched wavelet filter, speech scrambling, speech information protection, speech intelligibility, masking noise

1. Introduction

Information protection is an integral part of communication systems. Nowadays, more and more attention is paid to the protection of speech information, which is associated with the growth of speech communication in the modern information environment [1].

With the development of digital communication in radio engineering, gaming methods, and cryptographic algorithms have become widespread. Initially, the analog speech signal is converted into digital form [2]. An encryption algorithm is applied to the coefficients or signal parameters obtained after encoding [3]. Such systems have a high level of protection and require computational resources. Under interference conditions such algorithms do not work efficiently. A wide range of tasks requires algorithms that are applicable in the presence of sufficiently strong interference [4].

Along with mathematical methods of speech information protection, methods using digital signal processing (DSP) algorithms are widely demanded. Scrambling algorithms using fast linear orthogonal transforms (fast Fourier transform, fast wavelet transform) and discrete filter banks are of considerable interest [5]. As a rule, they are based on manipulations with spectral coefficients of linear transformation of signals. Such algorithms, when scrambling, cause a relatively small change in the signal bandwidth and very low residual intelligibility of the signal in the communication channel [6]. When using fast transformations increases the degree of information closure, but also increases the computational complexity of the processing algorithm, there is a delay in the signal [7, 8]. Orthogonal scramblers are not deprived of the common disadvantages of scramblers and introduce distortions in the recovered speech signal determined by the dispersion in the channel and synchronization error [9]. Thus, the problem of developing new fast algorithms for the protection of speech information operating under noise conditions is urgent.

2. Literature review and problem statement

Ukraine adheres to the following classification by the level of complexity of devices: maskers (simple), dynamic scramblers (medium complexity), and encryptors (high complexity) [10]. The maskers providing the tactical level of information protection include spectrum inverters, and static scramblers [11]. The proposed speech masking algorithm also belongs to this class.

Scramblers using filter banks are widely used among tactical closure systems [12]. In general, the traditional scheme contains M-channel analysis-synthesis filter banks, and forward and backward permutation blocks (Fig. 1) [13].

CPITS-II 2024: Workshop on Cybersecurity Providing in Information and Telecommunication Systems II, October 26, 2024, Kyiv, Ukraine

[*] Corresponding author.

[†] These authors contributed equally.

oleksandrlavrynenko@gmail.com (O. Lavrynenko)

0000-0002-7738-161X (O. Lavrynenko)

Mixing of signal segments according to a certain permutation rule takes place in the block P. Reverse permutation occurs at the input of the decoder in the block P^{-1}. The permutation rule is the key to the system. They have very low residual intelligibility of the scrambled signal in the communication channel, but introduce time delay and distortion in the reconstructed signal [14]. For its class, unlike other masking methods, the proposed method of speech signal scrambling based on matched wavelet filters has a high degree of information closure, high quality of reconstructed speech, and a sufficiently large number of keys.

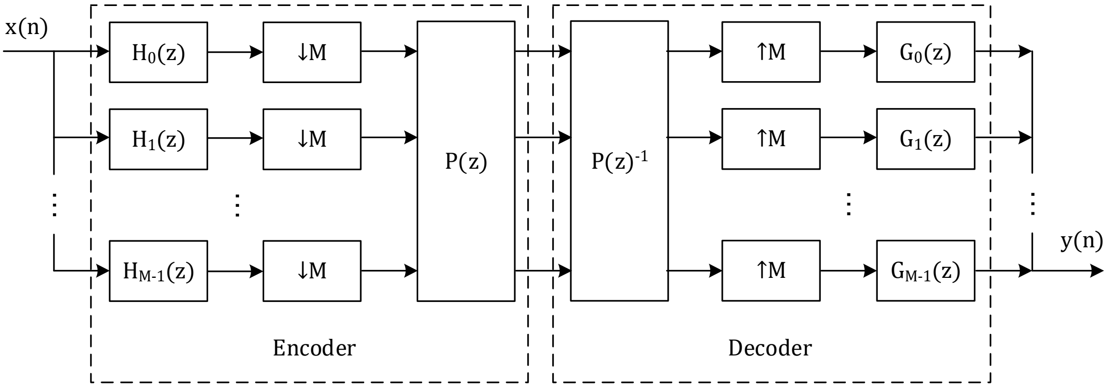

Figure 1: Scrambler circuit as an M-channel filter bank

This allows to use of its advantages for modernization of the existing fleet of radio stations. For example, the use of such a task on more complex devices with guaranteed information closure encryptors [15], is unjustified and, as a rule, requires radical technical solutions affecting the design of devices. When modernizing radios, the reliability of the protection system is important its simple design, and the insignificance of material costs.

3. Proposed method

In this section, a speech information protection algorithm using digital wavelet filter banks is proposed [16]. To build the protection system, an inverse scheme of single-level discrete wavelet transform is used. It includes digital synthesis and analysis filter banks (Fig. 2). The filters used are synthesized using a key sequence. The key identifies the sender and receiver of information and is used only at the filter synthesis stage [17].

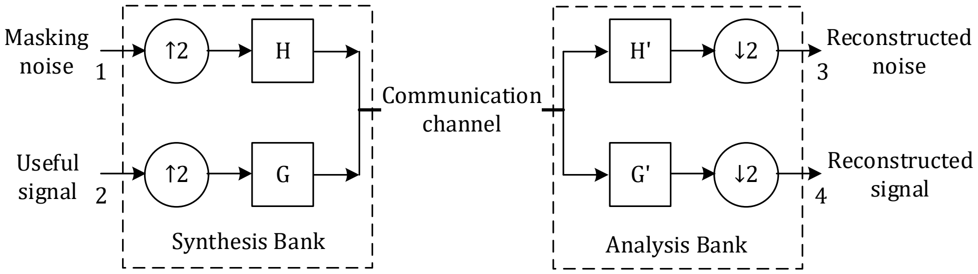

Figure 2: Block diagram of the information protection system

The useful signal is input to input 2 of the synthesis bank, then to the sampling frequency expander and filter-interpolator G. This shifts the frequency range of speech to the high-frequency region, which is positive for its closure. For other types of data, the useful signal can be fed to any input of the synthesis bank. A masking additive white Gaussian noise (AWGN) of higher power is fed to input 1, then to the sampling rate expander and filter interpolator H. The signal images transformed by the filters are mixed to form a noise-like mixture [18]. The extraction of a useful signal from such a mixture occurs after it passes through the filter-decimator $\bar{G}$ And the sampling frequency compressor in the analysis bank. The transform over the signal changes the bandwidth of the signal as a result of interpolation. The frequency of the signal in the channel is increased by a factor of 2 compared to the input signal. The system has the property of exact recovery [19].

The additive mixture of interference and useful signal images can be easily separated based on the orthogonality of the approximating and wavelet functions. The choice of the key initially determines the shape of these functions and the frequency properties of the corresponding wavelet filters. To ensure reliable information protection, the key is chosen to be noise-like and can be generated using a pseudo-random number generator [20].

Such an information protection system can be realized only due to the orthogonality properties of matched wavelet filters (MWFs) [21]. Unknown basis functions must be used for mixing with signal samples. A similar protection system can be built on the well-known Daubechies, and Haar wavelet filters, but it will be easy to crack.

To build a robust system, quadrature-mirror wavelet filters (QWFs) with unique frequency responses are required [22]. Such filters can be matched with wavelet filters. They must satisfy the conditions imposed on wavelet filters: orthogonality of the wavelet basis, and presence of zero moments.

The zero moments of the frequency response of the approximating filter H can be introduced a priori to solve the synthesis problem. If the filter has zero moments, the expression for $H(\omega)$ can be written in the form: $H(\omega) = \left(1 + e^{-j\omega}\right)Q(\omega)$, where $Q(\omega)$ is some function [23].

Let us present the developed method of synthesizing the MWF. The theory of MWF is developed from the following problem. It is required to construct for $f(n)$ a set of orthogonal quadrature-mirror wavelet filters in such a way that at its wavelet decomposition, the output of the detailing filter is zero, i.e. all detailing coefficients of the wavelet domain should be equal to zero (Fig. 3).

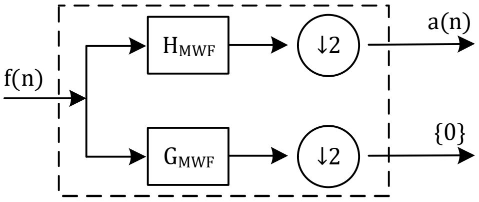

Figure 3: MWF synthesis problem

The procedure of wavelet transform of the signal $f(n)$ in the frequency domain can be written in the following form:

$$\begin{cases} H(\omega)F(\omega) + H(\omega+\pi)F(\omega+\pi) = A(\omega) \\ G(\omega)F(\omega) + G(\omega+\pi)F(\omega+\pi) = D(\omega), \end{cases} \quad (1)$$

where $F(\omega)$, $A(\omega)$, $D(\omega)$ are the Fourier images of the sequence $f(n)$, interpolated approximating and detailing wavelet transform coefficients, respectively, and $H(\omega)$ and $G(\omega)$ are the frequency response (FR) of the decomposition filters [24].

In order to prevent elision, we can assume that the relationship between the H and G filters is set to be fair for the QWF:

$$G(\omega) = e^{-j\omega} \cdot H^{*}(\omega+\pi). \quad (2)$$

This will immediately satisfy one of the constraints placed on filters:

$$G(\omega+\pi) \cdot \overline{G(\omega)} + H(\omega+\pi) \cdot \overline{H(\omega)} = 0,$$

where $\overline{G(\omega)}$, $\overline{H(\omega)}$ are the FRs of the corresponding recovery filters, with $\overline{G(\omega)} = G^{*}(\omega)$ and $\overline{H(\omega)} = H^{*}(\omega)$.

Another important property, the orthogonality property of the wavelet basis for filters in the frequency domain is written in the form:

$$|H(\omega)|^2 + |H(\omega+\pi)|^2 = 2. \quad (3)$$

Solving the system (1) under the assumption that$D(\omega) = 0$, and using the relation (2), we get

$$H(\omega) = \frac{A(\omega) \cdot F^{*}(\omega)}{|F(\omega)|^2 + |F(\omega+\pi)|^2}.$$

The filter $H(\omega)$ is almost built, it remains to find the condition on $A(\omega)$. This can be done using (3), given that $A(\omega) = A(\omega+\pi)$. As a result, we obtain

$$|A(\omega)|^2 = 2 \cdot \left(|F(\omega)|^2 + |F(\omega+\pi)|^2\right).$$

Let $A(\omega)$ be a real analytic function, then

$$A(\omega) = \sqrt{2} \cdot \sqrt{|F(\omega)|^2 + |F(\omega+\pi)|^2}.$$

The final result is represented as:

$$H(\omega) = \frac{\sqrt{2} \cdot F^{*}(\omega)}{\sqrt{|F(\omega)|^2 + |F(\omega+\pi)|^2}}. \quad (4)$$

As a result of solving the problem, digital wavelet filters matched to the input sequence have been found. Such filters are called matched wavelet filters [25] since their impulse response is formed taking into account the properties of the processed signal. In our case, it is a key sequence. Thus, the information about the key is embedded in the filters themselves.

Thus, we can present a speech protection algorithm that relies on the dual use of masking noise. The difference between the proposed system and its first variant is the addition of masking noise to the signal at the input (before the transmultiplexer) and the inverse transformation at the output. For this purpose, in addition to the transmultiplexer, including expanders and compressors of the sampling frequency, analysis, and synthesis filters, adders are introduced into the system. Fig. 4 shows a detailed block diagram of the second variant of the scheme.

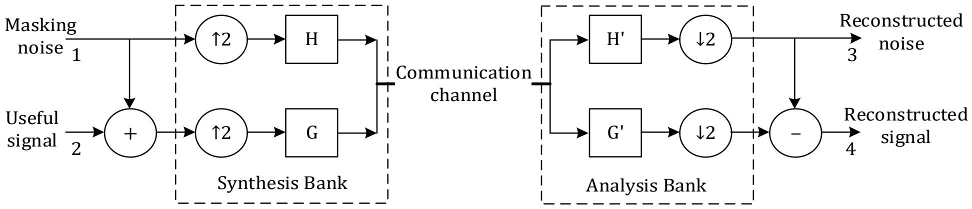

Figure 4: Block diagram of a speech protection system with double masking

The masking noise [26] is fed to input 1, then to the sampling rate expander and filter-interpolator H. At the same time, it is also fed to the adder. The useful signal is fed to input 2, mixed in the adder with the higher power masking noise. The mixture then goes to the sampling rate expander and filter interpolator G. The signal images transformed by the filters are mixed to form a noise-like mixture. The reconstruction of the signal in the analysis bank is done in reverse order. The system has the property of accurate recovery. The key point is the uniqueness of the analysis and synthesis filter banks. Wavelet filters are synthesized in the previously described way.

A generalized speech information protection scheme can also be proposed to close the conversations of several users. It is known [27] that wavelet decomposition over both subbands yields a complete balanced tree (Fig. 5). If the initial block of wavelet filters is orthogonal, then the scheme corresponding to any level of the full tree decomposition is orthogonal. Such a scheme as a whole, as well as its separate block, has the property of accurate signal recovery. An inverse scheme consisting of separate blocks can also be constructed for the full wavelet tree [28].

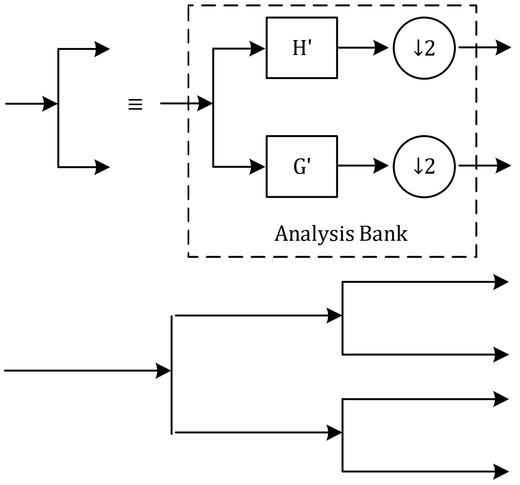

Figure 5: Two-level decomposition of the frequency-time plane using wavelet packets

In this paper, an inverse scheme for two levels of decomposition is considered. It consists of 3 pairs of analysis and synthesis banks and has 4 inputs (Fig. 6). The circuit can be used to protect speech information in several ways: (1) masking noise is fed to one of the inputs, the other inputs have useful signal; (2) useful signal is fed to one of the inputs, the other inputs have masking noise. In the first case, the number of users using a common channel for secure transmission of speech information increases to 3. In the second case, due to the imposition of several noises increases the security of the system. Along with this two-level discrete wavelet transform scheme [29] complicates the speech protection system. Only the first case is considered in this paper.

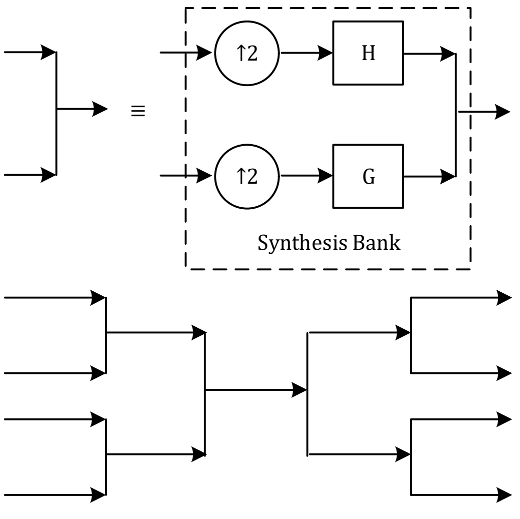

Figure 6: Frequency-time plane partitioning using wavelet packets in a multi-user scheme

The circuit uses MWFs synthesized by key sequence. For synthesizing each pair of filter banks a common key for all pairs is used. Masking noise is fed to one of the inputs of the system, the other inputs are approximately equal in power useful signals. This principle is used to close the conversations of three users at once. As a masking noise is used AWGN [30] of higher power.

4. Results and discussion

The research on the system of speech information protection using MWF was carried out on speech signals. The system was analyzed for non-recursive (FIR filters) and recursive (IIR filters) systems. The case of operation of the protection algorithm in the conditions of application of the ITU-T G.711 standard for coding signals in the channel with 8, 16, and 32 bits is considered [31]. The operating parameter of the system is the masking noise power. To analyze the influence of masking noise, the parameter M is the ratio of signal and masking noise power in dB is introduced. To study the noise immunity of the system, the parameter N is the ratio of signal and external noise power in the communication channel in dB was introduced:

$$M = 10 \cdot lg\frac{P_{\text{signal}}}{P_{\text{masking noise}}};\ N = 10 \cdot lg\frac{P_{\text{channel signal}}}{P_{\text{channel noise}}}.$$

Estimates of speech parameters used in the paper are discussed.

PESQ (Perceptual Evaluation of Speech Quality) is used to automatically evaluate the quality of speech transmitted in telecommunication environments. To obtain the evaluation, the source signal and the signal at the system output are compared. The evaluation is graded on the MOS scale (mean opinion score, ITU-T recommendation P.800), which covers the range from 1 (poor) to 5 (excellent). The acceptable quality of the reconstructed signal corresponds to a PESQ score greater than 2.5 points.

Expert evaluation of speech intelligibility Q is introduced to determine speech intelligibility in the channel and at the system output. Based on the results of listening, the experts evaluate the intelligibility of the signal. The traditional 5-point scale was used, where the best sound quality corresponds to the highest score. At one point of the scale, the useful signal is completely unintelligible. Acceptable quality of the recovered signal corresponds to the assessment $Q > 3$ [32]. The Q score agrees well with the PESQ speech quality score.

It is considered what requirements the protection system should meet.

The main purpose of the system is reliable closure of speech information with the possibility of its full recovery. Based on the purpose, for the proposed speech protection system the recovered signal should have good intelligibility. The signal in the channel on the contrary should be completely unintelligible. Such conditions are fulfilled for a certain interval of values of the parameter M. The lower limit of the parameter M is determined from the conditions when distortions of the reconstructed signal become unacceptable for perception. It is estimated by the PESQ criterion and the expert evaluation of speech intelligibility Q. The upper limit is determined from the conditions when the useful signal in the channel becomes completely unintelligible and is based on the evaluation of Q. The masking noise power satisfying these conditions is selected from the interval formed by the intersection of the shaded regions in Figs. 7 and 8. The value of the PESQ score, which is chosen to be greater than 2.5, is also considered.

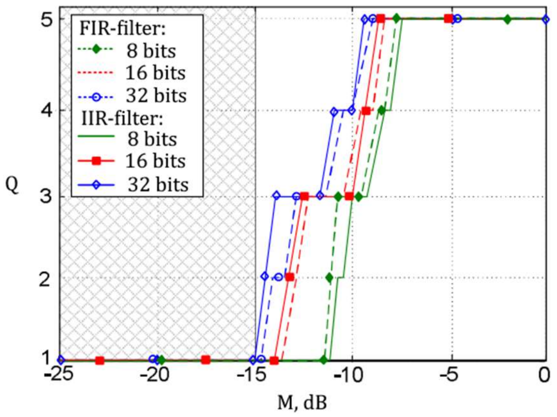

Figure 7: Dependence of speech intelligibility estimation in the communication channel on M

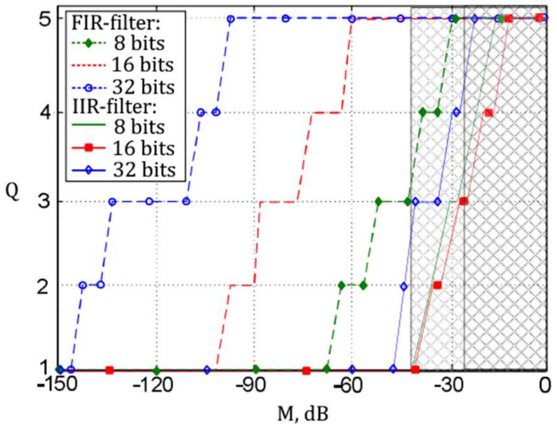

Figure 8: Dependence of the system output speech intelligibility estimate on M

Thus, the selection interval for the parameter M is:

(a) $-34 < M, ∂Б < -15$ for FIR filters.

(b) $-20 < M, ∂Б < -15$ for IIR filters.

The noise immunity of the system is considered as the performance of the system under information distortion in the presence of noise. Generalized, as an external noise is used AWGN, modeling data distortion. Based on the values of $PESQ > 2.5$ points (Fig. 9), it follows that acceptable quality of the transmitted signal is achievable at $N > 25$ dB.

It is established that the algorithm is robust to external noise. The dependence of the signal-to-noise ratio (SNR) at the system output on the SNR in the channel is linear.

The resilience of the system, i.e., the extent to which it is secure against the tampering of the contents of the negotiation, is the most important and challenging issue.

The technical unrealizability of the cracking system is confirmed by the results of a direct search of key combinations, which did not yield any results for a limited time interval (a week). The search was conducted on a test computer (Windows 11 operating system; 11th Gen Intel (R) Core (TM) i7-11370H 3.3 GHz processor; 16 GB memory) and the algorithm was simplified. As in most modern defense systems, the system persistence is determined by the amount of key information. In this work, we used keys with dimensionality b from 150 to 250 bits (20 samples). Accordingly, there are 2^b variants of the key sequence.

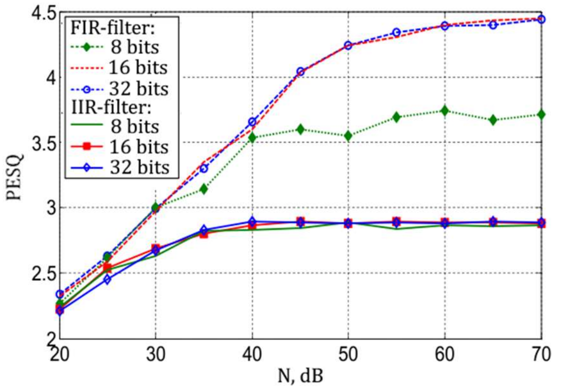

Figure 9: Dependence of $PESQ$ on N

There is a dependence on the key length in the system (Fig. 10). With its increase the quality of the reconstructed signal deteriorates from excellent to good (at 150 samples and further). This behavior of the system can be explained by the accumulation of errors as a result of convolution with a filter with a long impulse response. It is acceptable to use a key with a dimensionality of 20–30 samples.

The physical unrealizability of the hacking system is based on the fact that to date there are no systems that allow to distinguish the signal against the background of noise at the values of the parameter M used in this work. The application of known methods of noise suppression did not give a positive result. It is technically very difficult to isolate a useful signal.

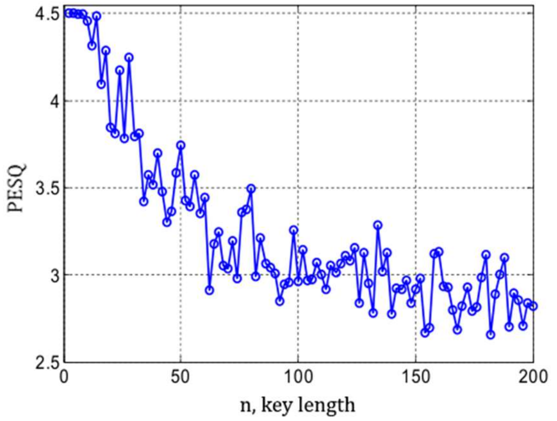

Figure 10: Dependence of PESQ on the key dimension

To ensure reliable protection of information, the key is chosen to be noise-like and can be generated using a pseudo-random number generator. When the key length increases, the frequency response of filters becomes more complicated, and the degree of information closure increases. It is found that it is inefficient to use very long keys because the algorithm performance decreases as a result of long convolutions. The method of key distribution is a separate non-trivial task. Most often participants agree on the key to be used beforehand.

5. Conclusions

Based on the research conducted on the speech information protection system, the following results are obtained in this paper:

(1) A method for synthesizing matched wavelet filters satisfying the property of orthogonality of the wavelet basis, the presence of zero moments is presented. The important requirements on the filters are the conditions of complete signal recovery and elimination of overlapping spectra. The results of the studies show the effectiveness of synthesized wavelet filters in solving the problem of information protection.

(2) A speech protection algorithm is developed that uses MWF at the stage of building a bank of analysis-synthesis filters consistent with the key. The algorithm has a simple implementation, and fast algorithms of digital signal processing (convolution, decimation, interpolation), allowing encrypting of the signal in real-time. The proposed algorithm is noise-resistant and can be used in channels with intense interference, with the value of the parameter $N > 25$ dB. The algorithm is robust to time delays and hiccups, distortions in the communication channel.

(3) Operating parameters of the protection system are obtained, satisfying the high degree of system closure and acceptable quality of the recovered signal. Operating parameters are calculated based on PESQ estimation, and expert estimation Q. Such estimations allow us to qualitatively describe the processes occurring in the speech protection system. The parameter setting interval M: $-34 < M,\text{дБ} < -15$ for FIR filters and $-20 < M,\text{дБ} < -15$ for IIR filters is obtained.

(4) The system is analyzed for FIR and IIR filters. The quality of the decoded signal is lower when using FIR filters, because of the error amplification in recursive systems.

(5) The case of operation of the protection algorithm in conditions of application of ITU-T G.711 standard for signal coding is considered. The results are obtained taking into account the quantization of the signal in channels 8, 16, and 32 bits. It is found that 8 and 16-bit quantization of encrypted signal in the channel provides necessary conditions for correct protected transmission of speech information. There is no need to increase the number of quantization levels (for example, up to 32 bits).

(6) The estimation of the degree of speech information closure, determined mainly by the ratio of masking noise and useful signal levels, is carried out. From the analysis of signal spectrograms in the channel, it follows that it is very difficult to distinguish a useful signal from the mixture. The amount of key information, which determines the number of tests required for key selection, has been estimated. In this work, we used keys of dimension b from 150 to 250 bits, and the number of key combinations is 2^b. Direct search of key combinations for the allotted time did not give results. The information can be decrypted only by knowing the key on the receiving side with which it was encrypted.

Future research requires analyzing the system behavior depending on the type of masking noise, i.e., investigating noise, structural, and combined interference. To find out which noise interference of white noise type, a mixture of white and pink noise, and structural interference of “speech chorus” type are the most suitable for the proposed information protection system.

References

[1] S. B. Sadkhan, A. Salah, The Trade-off between Security and Quality using Permutation and Substitution Techniques in Speech Scrambling System, in: International Conference of Computer and Applied Sciences (CAS) (2019) 244–249. doi: 10.1109/CAS47993.2019.9075489.

[2] L. Kriuchkova, et al., Experimental Research of the Parameters of Danger and Protective Signals Attached to High-Frequency Imposition, in: Cybersecurity Providing in Information and Telecommunication Systems II, vol. 3550 (2023) 261-268.

[3] Y. Huang, Y. Wang, Multi-Format Speech Perception Hashing based on Time-Frequency Parameter Fusion of Energy Zero Ratio and Frequency Band Variance, in: 3rd International Conference on Electronic Information Technology and Computer Engineering (2019) 243–251. doi: 10.1109/EITCE47263.2019.9094822.

[4] D. Abdulrida, N. A. Abbas, Speech Descrambling Based on Chaotic Parameter Estimation, in: 1st Babylon International Conference on Information Technology and Science (BICITS) (2021) 33–38. doi: 10.1109/BICITS51482.2021.9509926.

[5] A. M. Raheema, S. B. Sadkhan, S. M. Abdul Sattar, Performance Evaluation of Voice Encryption Techniques Based on Modified Chaotic Systems, 6th International Engineering Conference Sustainable Technology and Development (IEC) (2020) 135–140. doi: 10.1109/IEC49899.2020.9122933.

[6] N. Hayati, et al., End-to-End Voice Encryption Based on Multiple Circular Chaotic Permutation, in: 2nd International Conference on Communication Engineering and Technology (ICCET) (2019) 101–106. doi: 10.1109/ICCET.2019.8726890.

[7] O. Romanovskyi, et al., Prototyping Methodology of End-to-End Speech Analytics Software, in: 4th International Workshop on Modern Machine Learning Technologies and Data Science, vol. 3312 (2022) 76–86.

[8] I. Iosifov, et al., Transferability Evaluation of Speech Emotion Recognition Between Different Languages, Advances in Computer Science for Engineering and Education 134 (2022) 413–426. doi: 10.1007/978-3-031-04812-8_35.

[9] G. Konakhovych, et al., Method of Reliability Increasing Based on Spare Parts Optimization for Telecommunication Equipment, Lecture Notes in Networks and Systems 992 (2024) 296–309. doi: 10.1007/978-3-031-60196-5_22.

[10] H. Ye, et al., A Voice Encryption Method Based on Complex Bao Chaos System, International Conference on Computing, Communication, Perception and

Quantum Technology (CCPQT) (2022) 268–273. doi: 10.1109/CCPQT56151.2022.00053.

[11] V. Kuzmin, et al., Method for Correcting the Mathematical Model in Case of Empirical Data Asymmetry, Lecture Notes in Networks and Systems 657 (2023) 249–260. doi: 10.1007/978-3-031-36201-9_21.

[12] K. P. Pushpavathi, B. Kanmani, FIR Filter Design using Wavelet Coefficients, International Conference on Wireless Communications Signal Processing and Networking (WiSPNET) (2019) 410–415. doi: 10.1109/WiSPNET45539.2019.9032718.

[13] S. C. Venkateswarlu, N. U. Kumar, A. Karthik, Speech Enhancement using Recursive Least Square based on Real-Time Adaptive Filtering Algorithm, in: 6th International Conference for Convergence in Technology (I2CT) (2021) 1–4. doi: 10.1109/I2CT51068.2021.9417929.

[14] O. Holubnychyi, O. Lavrynenko, D. Bakhtiiarov, Well-Adapted to Bounded Norms Predictive Model for Aviation Sensor Systems, Lecture Notes in Networks and Systems 736 (2023) 179–193. doi: 10.1007/978-3-031-38082-2_14.

[15] M. Joorabchi, S. Ghorshi, Y. Naderahmadian, Speech Denoising Based on Wavelet Transform and Wiener Filtering, in: 8th International Conference on Frontiers of Signal Processing (ICFSP) (2023) 43–46. doi: 10.1109/ICFSP59764.2023.10372899.

[16] A. Saleh, S. B. Sadhkan, A Proposed Speech Scrambling based on Haar Transform and Permutation, 2nd International Conference on Engineering Technology and its Applications (2019) 31–36. doi: 10.1109/IICETA47481.2019.9013013.

[17] O. Lavrynenko, et al., A Method for Extracting the Semantic Features of Speech Signal Recognition Based on Empirical Wavelet Transform, Radioelectronic and Computer Systems 3(107) (2023) 101–124. doi: 10.32620/reks.2023.3.09.

[18] I. K. Alak, S. Ozaydin, Speech Denoising with Maximal Overlap Discrete Wavelet Transform, International Conference on Electrical and Computing Technologies and Applications (ICECTA) (2022) 27–30. doi: 10.1109/ICECTA57148.2022.9990250.

[19] M. Chandni, D. Govind, Effectiveness of Wavelet Synchrosqueezed Transform for Improved Epoch Estimation from Telephonic Speech Signals Using Zero Frequency Filtering, in: 18th India Council International Conference (INDICON) (2021) 1–5. doi: 10.1109/INDICON52576.2021.9691652.

[20] O. Lavrynenko, et al., A Wavelet-Based Steganographic Method for Text Hiding in an Audio Signal, Sensors 22(15) (2022) 5832. doi: 10.3390/s22155832.

[21] G. Yang, Y. Song, J. Du, Speech Signal Denoising Algorithm and Simulation Based on Wavelet Threshold, in: 4th International Conference on Natural Language Processing (ICNLP) (2022) 304–309. doi: 10.1109/ICNLP55136.2022.00055.

[22] S. B. Sadkhab, A. M. Raheema, S. M. Abdul Sattar, Design and Implementation Voice Scrambling Model Based on Hybrid Chaotic Signals, in: International Conference of Computer and Applied Sciences (CAS) (2019) 193–198. doi: 10.1109/CAS47993.2019.9075626.

[23] O. Yu. Lavrynenko, et al., Application of Daubechies Wavelet Analysis in Problems of Acoustic Detection of UAVs, in: 6th Workshop for Young Scientists in Computer Science & Software Engineering, vol. 3662 (2024) 125–143.

[24] G. Singh, et al., Novel Architecture for Lifting Discrete Wavelet Packet Transform with Arbitrary Tree Structure, Transactions on Very Large Scale Integration Systems 29(7) (2021) 1490–1494. doi: 10.1109/TVLSI.2021.3079989.

[25] C. Zhang, et al., Research on Extracting Algorithm of Speech Eigenvalue based on Wavelet Packet Transform and Gammatone Filter, in: 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) (2019) 165–169. doi: 10.1109/ITNEC.2019.8729292.

[26] J. Guo, et al., Modeling and Simulation of Power Grid Voltage Harmonic Detection Method based on the Improved Wavelet Packet Transform, Chinese Control Conference (CCC) (2021) 6716–6721. doi: 10.23919/CCC52363.2021.9549731.

[27] D. Bakhtiiarov, et al., Method of Binary Detection of Small Unmanned Aerial Vehicles, in: Cybersecurity Providing in Information and Telecommunication Systems, vol. 3654 (2024) 312–321.

[28] C. Zhang, et al., Research on Extracting Algorithm of Speech Eigenvalue Based on Wavelet Packet Transform and Gammatone Filter, in: 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) (2019) 165–169. doi: 10.1109/ITNEC.2019.8729292.

[29] O. Lavrynenko, et al., Method of Remote Biometric Identification of a Person by Voice based on Wavelet Packet Transform, in: Cybersecurity Providing in Information and Telecommunication Systems, vol. 3654 (2024) 150–162.

[30] D. Sun, et al., Damage Degree Assessment Based on Lamb Wave and Wavelet Packet Transform, Chinese Control and Decision Conference (CCDC) (2019) 3179–3184. doi: 10.1109/CCDC.2019.8833396.

[31] A. Dutt, P. Gader, Wavelet Multiresolution Analysis Based Speech Emotion Recognition System Using 1D CNN LSTM Networks, IEEE/ACM Transactions on Audio, Speech, and Language Processing 31 (2023) 2043–2054. doi: 10.1109/TASLP.2023.3277291.

[32] O. Lavrynenko, et al., Remote Voice User Verification System for Access to IoT Services Based on 5G Technologies, in: 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) (2023) 1042–1048. doi: 10.1109/IDAACS58523.2023.10348955.