Ruslana ZIUBINA[1], Yuliia BOIKO[2]

Opiekun naukowy: Olexandr YUDIN[3]

OCENA WPŁYWU IMPULSOWEGO SZUMU NA METODY GŁOSOWEJ IDENTYFIKACJI SPIKERA

Streszczenie: Przeprowadzono ocenę efektywności opracowanych metod efektywnej szerokości spektrum oraz największej informacyjnej wagi głównego tonu w zadaniu identyfikacji i weryfikacji spikera przy istniejących przeszkodach różnego pochodzenia. Określono, że wskutek wpływu/obecności chaotycznych impulsowych szumów możliwość identyfikacji spikera w wielorakich zadaniach podjęcia decyzji gwałtownie obniża się przy zbiegnięciu się częstości głównego tonu głosowego sygnału i początkowej częstości impulsowej szumu. Metoda efektywnej szerokości spektrum zapewnia wysokie wskaźniki identyfikacji w warunkach wpływu takiego typu szumów dla zależnej od tekstu identyfikacji przy nie dużej liczbie możliwych hipotez.

Słowa kluczowe: częstość głównego tonu, metody identyfikacji, chaotyczne impulsowe przeszkody, prawdopodobieństwo identyfikacji

IMPACT ASSESSMENT OF IMPULSE NOISE ON THE METHODS OF SPEAKER RECOGNITION

Summary: Impact assessment of the developed methods of effective spectrum width and the largest information weight of the main tone in the problem of speaker identification and verification at the existing noise of different origin was held. It was determined, a result of the influence of chaotic impulse noise, the possibility of identifying a speaker in in multi-alternative detection tasks decreases sharply when the frequency of the main tone of the speech signal and the initial frequency of the impulse noise are coincided. The effective spectrum width method provides high identification value under the influence of this type of noise for text-dependent identification with a small quantity of possible hypotheses.

Keywords: the frequency of the main tone, methods of identification, chaotic impulse noise, probability of identification.

---

[1] Candidate of Sciences (Technical) Taras Shevchenko National University of Kyiv: Department of Cyber Security and Information Protection, email: ziubina@knu.ua
[2] Candidate of Sciences (Technical) Taras Shevchenko National University of Kyiv: Department of Applied Information Systems, email julia_boyko2010@ukr.net
[3] Professor, Doctor of Science (Technical), Taras Shevchenko National University of Kyiv: Department of Cyber Security and Information Protection, email yak333@ukr.net

1. Introduction

Recognition of a speaker or a person's voice is a biometric method based on an analysis of the individual characteristics of a person. The process of speaker recognizing includes vocal tract structure features and individual's behavior nature. In recent years, biometric protection using has got a development due to the growing danger of cyberattacks to information structures. In Ukraine, biometrics based on fingerprints acquired a special status with the introduction of biometric passports, and a number of attacks on state-owned enterprises prompt to reflect about the existing level of critical information resources protection. Using of authentication systems in the conditions of remote communication will enable access to classified information with a less of time for transportation to the right place. A voice-based biometric is characterized by the fact that it can have many passwords or do not depend on them at all, and data collection can take place dynamically, directly, when interacting with the system or for a specified time interval.

2. The main part

The developed methods of effective spectrum width [1-2] and the largest information weight of the main tone allow to identify the speaker in conditions of high noise level. To determine the qualitative indicators of the developed methods, the studies were carried out in the simulation of noise according to the law of Gaussian distribution [3]. The use of such noise is classical in speech signal processing systems, but despite the fact that the developed methods are based on the frequency characteristics of speech signal formation, it would be advisable to consider the possible presence of noise of a different nature. Another type of noise is chaotic impulse noise, which are formed considering the initial frequency. Using of chaotic impulse noise with a given frequency f_0 can affect the quality of the identification and verification methods in the case when it exactly coincides with the frequency of the main tone of the speaker's voice. Thus, we obtain the following problem of assessing the quality of the developed methods in the presence of chaotic impulse noise with a frequency that almost exactly coincides with the frequency of the main tone of the speaker's voice $f_0 \approx f_{ЧОТ}$.
The mathematical model of the chaotic impulse noise has the form:

$$HIN = \sum_{k=0}^{N_n-1} \sum_{n=0}^{N_u-1} f \begin{bmatrix} t_з + \frac{T_n}{N_u} \cdot n + P \cdot k \le t_i < t_з + \frac{T_n}{N_u} \cdot (n+1) + P \cdot k, \\ f \begin{Bmatrix} rnd(1) \ge 0.5, \\ Sm \cdot A\cos\left[2\pi f_0 t_j + \varphi_n \,|_{\varphi_n = X_n norm(N,M,D)}\right], \\ 0 \end{Bmatrix} \\ 0 \\ \end{bmatrix} \tag{1}$$

where $t_з$ – code structure delay interval, N_u – number of pulses, T_n – the duration of the code structure, N_n – number of structures, P – period of structures recurrence, φ_n – random phase shift [4].

As a result, we have a time representation of generated chaotic impulses, which will act as a noise (Fig.1).

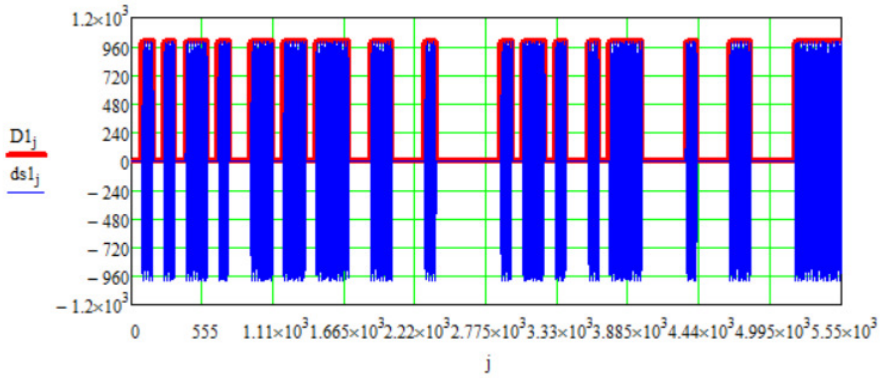

Figure 1. Time representation of generated chaotic impulses

The result of the addition of the information signal $Sh02$ and the created noise sequence was an additive mixture $Sh01$ (Fig. 2).

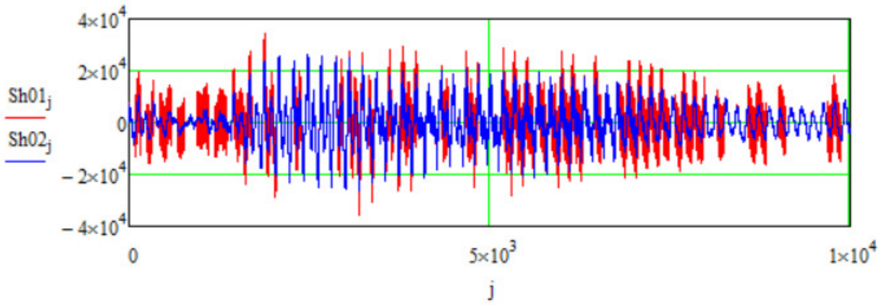

Figure 2. Hourly representation of the additive mixture $Sh01$ and the information signal $Sh02$

To assess the quality of the developed methods of effective spectrum width and the largest information weight of the main tone, the *SNR* ratio indicators that were calculated to evaluate the effectiveness of the developed methods in the use of white noise were chosen.

Table 1. Determination of the probability of correct speaker identification in multi-alternative decision support systems with $H = 8$

SNR, dB	Method of effective spectrum width		Method the largest information weight of the main tone	
	Text-dependent identification	Text-independent identification	Text-dependent identification	Text-independent identification
– 2,9	1	0,22	0,34	0,079
– 5,23	1	0,275	0,335	0,127
– 6,67	0,99	0,25	0,3	0,05
– 7,88	0,99	0,22	0,29	0,14
– 8,9	0,99	0,216	0,258	0,15
– 10,18	0,96	0,109	0,25	0,14
– 11,52	0,94	0,22	0,202	0,08
– 12,65	0,79	0,16	0,205	0,12
– 14,6	0,63	0,12	0,185	0,12
– 16,15	0,57	0,094	0,18	0,12

In the case of 8 hypotheses participating in the experiment, the correct identification indicators show high results with the use of the method of effective spectrum width (MESW), as the hypothesis is compared by analyzing and enumeration of all the informative constituents of the spectrum, and only for text-dependent identification. In the case of the method of the largest information weight of the main tone (MLIWMT) with such a number of hypotheses, the method can not cope with the task. The main reason for these results is the dependence on the frequency of the maine tone $f_{ЧОТ}$, which in this case is approximately equal f_0 . Thus, the use of the method of the largest information weight of the main tone is not effective for the identification of the speaker in the presence of chaotic pulse noise in the information system of providing IT services.

The visual result of the developed methods for $H = 8$ is shown in Fig. 3.

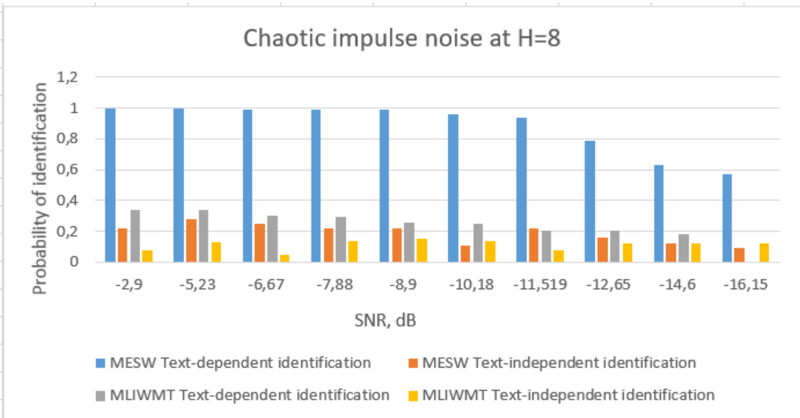

Figure 3. Indicators of the probability of correct speaker identification using chaotic impulse noise by the developed methods at $H = 8$

Table 2. Determination of the probability of correct speaker identification in multi-alternative decision support systems at $H = 6$

SNR, dB	Method of effective spectrum width		Method the largest information weight of the main tone	
	Text-dependent identification	Text-independent identification	Text-dependent identification	Text-independent identification
– 2,9	1	0,139	0,18	0,15
– 5,23	1	0,25	0,23	0,25
– 6,67	0,99	0,35	0,24	0,26
– 7,88	0,99	0,33	0,35	0,18
– 8,9	0,99	0,304	0,325	0,21
– 10,18	0,98	0,27	0,36	0,315
– 11,519	0,84	0,114	0,32	0,17
– 12,65	0,84	0,19	0,23	0,17
– 14,6	0,47	0,07	0,197	0,114
– 16,15	0,37	0,1	0,226	0,18

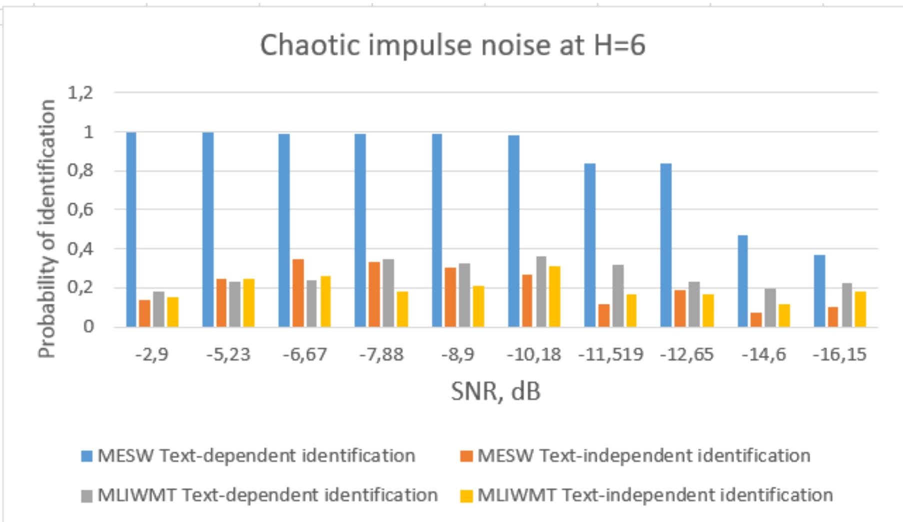

Figure 4. Indicators of the probability of correct speaker identification using chaotic impulse noise by the developed methods at $H = 6$

Using developed methods for a multi-alternative task of identifying a speaker in conditions of available chaotic impulse noise at $H = 6$ (Figure 4.) showed that, as in the previous experiment, high performance gives the method of effective spectrum width for text-dependent identification, and in the case of the method of the largest information weight of the main tone, the identification remains unlikely, due to the significant effect of the noise on the frequency of the main tone.

Table 3. Determination of the probability of correct speaker identification in multi-alternative decision support systems at $H = 3$

SNR, dB	Method of effective spectrum width		Method the largest information weight of the main tone	
	Text-dependent identification	Text-independent identification	Text-dependent identification	Text-independent identification
– 2,9	1	1	0,917	0,9
– 5,23	1	0,98	0,85	0,7
– 6,67	1	0,96	0,755	0,604
– 7,88	1	0,97	0,66	0,35
– 8,9	1	0,93	0,502	0,321
– 10,18	0,99	0,75	0,57	0,32
– 11,51	0,98	0,59	0,55	0,32
– 12,65	0,95	0,37	0,53	0,3
– 14,6	0,83	0,33	0,48	0,27
– 16,15	0,53	0,19	0,3	0,22

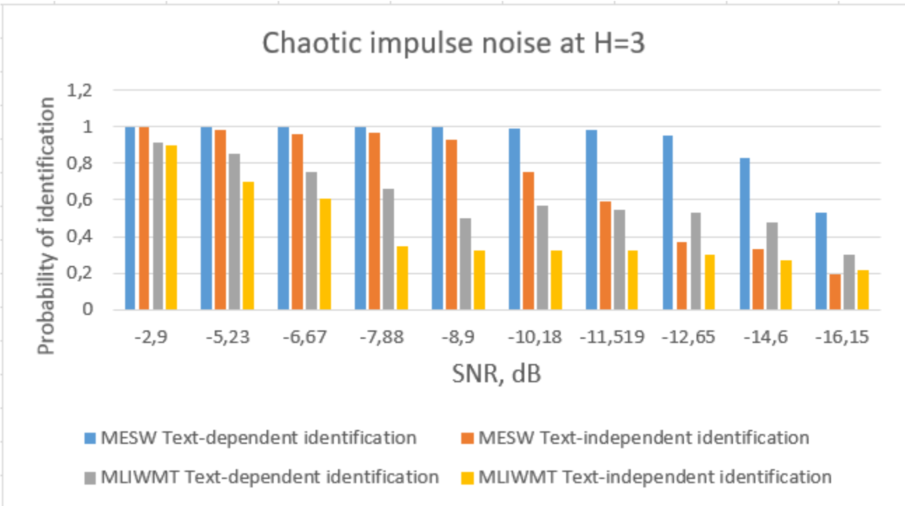

Figure 5. Indicators of the probability of correct speaker identification using chaotic impulse noise by the developed methods at $H = 3$

With regard to the determination of the probability of correct identification of the speaker in the multi-alternative decision support systems at $H = 3$ then in this case, the methods give an opportunity to carry out precise identification in conditions of $SNR = -10{,}18$ dB for text-dependent and at $SNR = -2{,}9$ dB for the text-independent method of effective spectrum width. For the method of the largest information weight of the main tone, the $SNR = -2{,}9$ dB value makes it possible to identify an accuracy of 90% for the text-independent and with an accuracy of 91% for text-dependent identification (Pic. 5).

3. Conclusions

The evaluation of the effectiveness of the developed methods makes it possible to conclude that high rates of audio signals identification in multi-alternative problems of decision-making in information systems are provided by the features of the selected feature space.

The presence of Gaussian noise in the communication channel enables text-dependent identification with a probability of 98% at $H = 8$ and $H = 6$ even in the case when the noise exceeds the signal by 1.67 times, and for text-independent-by 1.1 times. Method of largest information weight of the main tone gives these indicate for $H = 8$ and $H = 6$ even in situations when noise exceeds the signal at 1.43 times for text-independent identify, and for text-dependent a 96% probability of correct identification is ensured when the signal exceeds the noise in 1,1 times.

Certainly, the peculiarity of the chaotic impulse noise formation significantly affects the indicators of speaker identification by using the developed methods. However, despite this effect, the methods have demonstrated high efficiency and made it possible to identify the voice signal with a probability of more than 90% for both text-dependent and text-independent case when the noise exceeds the information signal by 0.83 times.

REFERENCE

1. ЮДІН О. К., ЗЮБІНА Р. В.: Метод ефективної ширини спектру //Наукоємні технології, 37(2018)1, 55– 60.
2. ЮДІН О. К., ЗЮБІНА Р. В.: Класифікація методів ідентифікації частоти основного тону //Наукоємні технології, 33(2017)1, 13-21.
3. ЮДІН О. К., ЗЮБІНА Р. В.: Оцінка ефективності методів ефективної ширини спектру та найбільшої інформаційної ваги основного тону в задачах ідентифікації та автентифікації аудіо сигналів //Наукоємні технології, 35(2017)3, 209-214.
4. АНТИПЕНСКИЙ Р.: Разработка моделей преднамеренных помех сигналам с дискретной модуляцией //Компоненты и технологии, 78(2007), 138– 143.