Andrii GIZUN, Vladyslav HRIHA, Anna BLIDAR, Maria ROSHCHUK, Svitlana DERKACH

Supervizor: Andrii GIZUN[1]

IDENTYFIKACJA SYSTEMU ATAKÓW WEWNĘTRZNYCH

Streszczenie: W pracy przeanalizowaliśmy systemy identyfikacji cech wewnętrznych obiektów czy organizacji. Informacje te pozwoliły nam poprawić wyniki wykrywania poufnych informacji poprzez włączenie wskaźników głosowych, antropometrycznych, społeczno-psychologicznych i fizjologicznych do opracowanego dwustopniowego systemu identyfikacji ataków wewnętrznych. Na podstawie tych obliczeń matematycznych stworzono pierwszy dwustopniowy system identyfikacji cech osób mających dostęp do informacji poufnych, w tym wskaźniki społeczno-psychologiczne i fizjologiczne. Uzyskane wyniki można wykorzystać do integracji z systemem zarządzania organizacji w celu przeciwdziałania działaniom osób wewnętrznych, powodując szkody ekonomiczne i informacyjne.

Słowa kluczowe: inżynieria społeczna, insider, głos, typ psychologiczny, antropometria, system przedstawicielski

INSIDER ATTACKS SYSTEM IDENTIFICATION

Summary: In the work we have analyzed systems of identification of insider features. This information allowed us to improve the results of the detection of insider by integrating voice, anthropometric, socio-psychological and physiological indicators into the developed two-stage system of insider attacks identification. Based on these mathematical calculations, the first two-stage system of identification of insider traits was created, including socio-psychological and physiological indicators. The obtained results can be used for integration into the management system of organization to counteract insider activity, causing economic and information damage.

Keywords: social engineering, insider, voice, psychological type, anthropometry, representative system.

Introduction

80% of information leaks, resulting in serious information and financial losses, occur due to the human factor and social engineering. Insiders or other external attackers use them in order to psychologically affect employees of certain organizations.

[1] National Aviation University, Kyiv, Ukraine, gsmgrey1@gmail.com

Therefore, every organization needs reliable qualified workers who will not disclose the restricted-access information (RAI) of the organization to outsiders and will not fall under the influence of outsiders.
The degree of qualification and awareness of the applicant can be determined by examining his / her experience and documentation. The human sociological type influences the responsibility and competence of the worker. The sociological type is affected by different psychological parameters - these are the variables on which function, that is, the behavior of a particular person, depends. The presence of the human factor has the leading role in the organization of information protection theory. With its help, the theory of information security organization as a system of knowledge has a substantive hyperlink with the fundamental humanities - sociology, social psychology, social engineering.
One of the main tasks of modern diagnostics is the ability of technologies to determine the emotional state of a person with minimal impact on the human body. Such technologies are commonly referred to as non-invasive (do not affect the human body). Most of these technologies are based on different characteristics of the human body (analysis parameters). However, none of them provides a comprehensive assessment of the human condition, and their accuracy is not always sufficient (therefore, social and psychological indicators are taken into account in our system). There are many methods of personal identification alongside with these technologies, including various criteria related to human physiology. These methods analyze certain characteristics or signals that are generated by the human body. Given this analysis, they make the decision the system needs. Such technologies are used in automated recognition systems, access systems and so on. The human voice signal is used as the main deciding criterion very often. It can be analyzed with a wide variety of characteristics that are subsequently used for the analysis task. At the same time, it is known that any changes in the human body are reflected to some extent in the human voice. It can help determine the nature of transformations that occur in the human body as a whole. It is well known that the voice signal is individual for each person. Nevertheless, it is possible to build an automated two-stage system that identifies the person at the first level and determines the emotional state of the person at the second level. The development of such a system is very important.
For the first time, a two-stage system for identifying insider traits is proposed, which, based on the analysis of voice, anthropometric, physiological and socio-psychological parameters, recognizes the existing or potential insider.

1. Analysis of insider identification methods and systems

When studying a concept "insider identification system", we should first of all look at the interpretation of the basic term - "social engineering".
Social engineering – is manipulating a person or a group of people with a purpose of breaking into security systems and stealing sensitive information. Thus, in order to obtain authenticating information from bank customers, criminals use special methods of social engineering (SE), the prerequisite of which are cognitive prejudices that dictate human behavior in the social world. Therefore, the reliability of any computer system is not higher than the reliability of its computer operator. SE is a type of social programming.

Social programming is the use of the human psyche features in order to achieve any goal.
Reverse SE - creating the conditions under which a person will resort to a social engineer by herself.
Computer programs to check the veracity of statements have emerged as a means of combating terrorism; they had to be engaged in public airports, railway stations, supermarkets. However, now some companies have begun to use voice lie detectors when hiring employees, insurance companies with the help of the program "call bluff" communicate with customers and additionally check the justification of their claims for compensation.
In the commercial sphere of Ukraine, there are 3 companies involved in voice analysis for fraud detection: Nemesysco ltd, Nice System ltd, and Center of Speech Technologies. In order to develop an insider identification system, it is necessary to analyze existing systems, some of which are finished products of well-known companies.
K-Factor (HR1) is a personnel testing and evaluation system produced by Nemesysco. The system allows to automatically test employees of the company or applicants for work and on the basis of the analysis of truthfulness (falsehood) of the information presented by the respondents, to determine the presence of risk factors in their behavior and to predict their possible manifestation in the future.
The K-factor system is based on sense technology. The analysis is carried out in the following areas: loyalty to the organization; confidentiality and privacy; honesty; trustworthiness; predisposition to bribery and fraud; alcohol, drugs, gambling.
Sense. Patented SENSE technology based on multilevel voice analysis (L.V.A.). Unlike other voice analysis technologies, SENSE technology can analyze different levels of waves in the voice by conducting a deep analysis of the subject's emotions. SENSE technology can identify whether the interlocutor is agitated, confused, tense, willing to share information, focused. SENSE technology doesn't evaluate the context of the answers; it evaluates the psycho-emotional reactions during the test that are reflected in his or her speech. These reactions are measured and evaluated by Nemesysco-based algorithms.
Sense technology includes four main processes:
1. Analysis of the shape of the voice sound wave for measuring the existing features of high and low frequencies and their changes in comparison with the sample of voice.
2. The exact frequency spectrum of the data entered in comparison with the sample.
3. Calculation of the parameters set at the stage of preliminary preparation and used to create a basic sample, free from emotions.
4. Review and comparison of new parameters with the baseline status of the entity and report.
The entire set of input can then be analyzed using a statistical learning algorithm to predict the possibility of questionable or false answers. Another level of analysis can be used in certain cases to evaluate the conversation as a whole.
Microsoft Emotion API. The system recognizes emotions on the face of a person. You should upload a photo (max 4 MB in size) or link to a photo. The system recognizes the number of faces and determines the coefficient of 8 parameters of emotions: anger, fear, contempt, disgust, happiness, surprise, sadness and neutral expression. It supports formats: JPEG, PNG, GIF, BMP. A maximum of 64 faces can be detected on each image. Faces that have a frontal position are best recognized. The

experimental system and is freely available at: *https://azure.microsoft.com/ru-ru/services/cognitive-services/emotion/.*

HR1, Nemesysco. The HR1-credit system evaluates bank customers to identify the risk of a loan default. HR1 on the basis of SENSE technology operates with the help of L.V.A multilayered voice analysis. An algorithm for calculating the default risk was developed to analyze the possibility of granting a loan, taking into account the psychological aspect of the process of repayment of the loan. This system is intended to test individuals who want to buy goods on credit in the store or for bank customers who want to take out a loan. The test can be retaken at least in a month. If the results do not satisfy the worker, other technologies are recommended.

LX5000-S. The LX5000-S computer polygraph is the latest lie detector that analyzes the human emotional state and detects lies by several parameters of the LX5000-S analysis, combining the benefits of analog polygraph procedures, the latest computer technology, valid data processing algorithms and the world's most advanced polygraph software. The basic polygraph kit is capable of simultaneously logging nine channels of information retrieval.

Omilia, NovaIT. This technology is used by AlfaBank, OTPBank for automatic recognition of the voice of customers calling to the bank and for the performance of the necessary operations that meet the needs of the banking customer by providing banking services. The system transmits call details in SIP, XML, SOAP, database access. The system understands the voice in: Russian, Ukrainian + Surzhik, Belarusian + Surzhik, Kazakh + Surzhik, Polish, Spanish, Greek, English (USA, Canada, England).

БД Cohn-Kanade. The International Database Extended Cohn-Kanade Facial Expression Dataset, University of Pittsburgh, USA (327 dynamic images) contains framesets for each dynamic image (video sequence). An existing approach to neural network recognition of human facial expressions on static images that also uses this base described by T. McLaughlin lies within the use of deep neural networks. The highest accuracy achieved was 67%.

The results of the emotional state detection systems analysis are presented in table 1.

Table 1. The results of analysis of the emotional state detection systems

System	**Criteria for comparison**				
	Analysis parameter	Technology	Platform/ Method of implementation	Requirements	Scope
K-factor	Voice	Sense	Windows XP/ Program	200 МБ on HD	Personnel security
Microsoft Emotion API	Face expression	БД Conh-Kanade	Internet/ Program	Size: max 4 MB	Detection of terrorists
LX5000-S	Breathing	LXSOFTWARE	Windows/ Combined	ПЗ LXSOFTWARE	Personnel security
Omilia	Voice	NLU	Linux Java, DB – MySQL/ Combined	Duration - 6 months	Bank security

HR1	Voice	Sense	Windows XP/ Program	200 МБ on HD	Bank security
Voice pathology detection system	Voice	Static processing of loud sounds	Windows/ Program	Microsoft Sound Recorder	Medicine
Driver status monitoring system	Voice	Parameterization	No data	No data	Driving

System	**Criteria for comparison**					
	Speed	Mobility	Cost	Anthropometry	Psychotype	Result
K-factor	15-20 min	-	1920$	-	-	The forecast of human behavior
Microsoft Emotion API	5s	+	For free	-	-	Human emotions in points
LX5000-S	360 samples per second	-	Closed data	-	-	Recognition of lies
Omilia	Instantly	+	200 000$	-	-	Automatic recognition of voice
HR1	5-7 min	-	725$	-	-	Identification of the risk of a loan default
Voice pathology detection system	No data	-	No data	-	-	Detection of speech pathologies of speakers
Driver status monitoring system	No data	+	No data	-	-	Detection of the driver's condition

2. Development of a two-stage insider identification system

Insider identification is carried out in stages and includes physical voice assessment, non-invasive voice evaluation (including acoustic evaluation of voice recordings by specialized computer systems), invasive tests, and photographing. The first two steps are a surface survey, however, the acoustic evaluation of the records makes it possible to assess the quality of the voice and to detect irregularities invisible during the mirror examination.

At the first stage of the TSIIS (two-stage insider identification system), the interviewee has to be identified. For the identification we use the method of

anthropometric research - anthropometry. Anthropometric indicators include somatometric and somatoscopic features. Somatometric determine the length, body weight and circumference of the chest, stomatoscopic - the state of subcutaneous fat, development (excessive, moderate, insufficient), distribution (uniform, uneven), what places have the biggest deposits of fat, the development of the muscular system spine, chest, leg shape, posture.

Among the main parameters of the voice signal we chose *voice frequency* (F_0) and *amplitude distribution of the input signal over time space* (Y_i)

To calculate these parameters, we propose to use a three-factor mathematical model, which is as follows:

$$F_0 = c_F \times A^{a_F} \times H^{h_F} \times W^{w_F}, \quad (1)$$

where F_0 – fundamental frequency, Hz;
A – age of a person;
H – height of the person, sm;
W – weight of the person, kg;
$c_{F,}\, a_F, h_F, w_F$ – factor frequency coefficients determined by the nonlinear multiple regression method.

The functional diagram of the system of identification of insider features is presented in fig. 1.

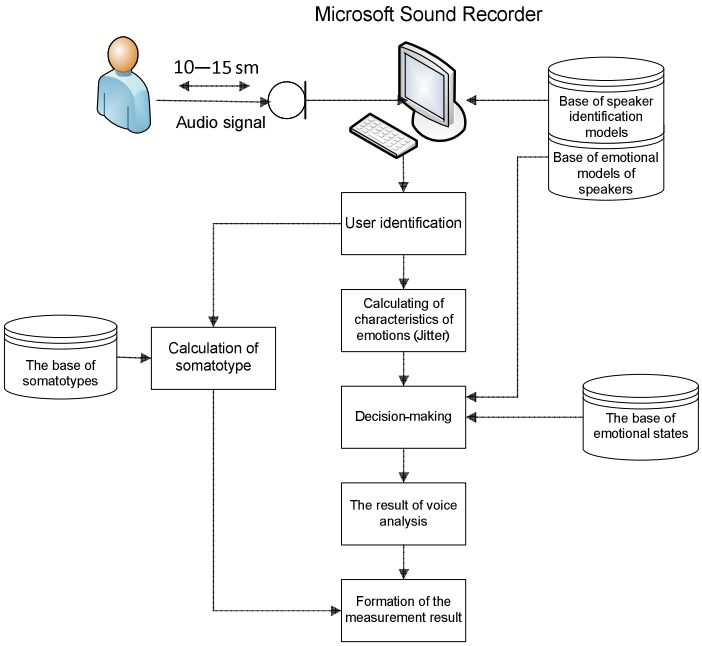

Figure 1. Functional diagram of the system of identification of insider features

The amplitude distribution of the input signal in time space cannot generally be described by a mathematical formula because of the large amount of data. However, we believe that the amplitude of each calculation can be described by a model similar to the fundamental frequency:

$$Y_i = c_{y,i} \times A^{a_{y,i}} \times H^{y,i} \times W^{w_{y,i}}, \quad (2)$$

where Y_i – the amplitude of the i-th oscillation of the human voice signal;
A – age of a person;
H – height of the person, sm;
W – weight of the person, kg;
$c_{y,i,}\, a_{y,i}, h_{y,i}, w_{y,i}$ – factor amplitude coefficients determined by the multiple regression method.

A multiple regression method for determining frequency and amplitude ratios is used despite the fact that the human voice is a dynamic signal that changes over time. Changes in a person's voice signal can occur both in single voice mode between frames and between voice recordings at different times. We consider these changes primarily due to changes in human anthropometry.

In addition to identifying the speaker, voice parameterization is used to determine a person's emotional state. Among these parameters are Jitter, Shimmer, harmonic to noise ratio, level of excitement, standard deviation.

The second stage is the process of calculating the characteristics of the expression of emotions. **Jitter** - unwanted phase and / or frequency random distortions during signal transmission. They arise due to the instability of the clock generator, changes in the parameters of the transmission line in time and different speed of propagation of frequency components of the same signal.

In digital systems there is a manifestation in the form of random rapid (with a frequency of 10 Hz or more) changes in the location of the digital signal edges in time, which leads to asynchronous behavior and, as a consequence, distortion of the transmitted information. For example, if the front has a low slope or "lag" in time, then the digital signal seems to be delayed, shifts relative to a significant point in time - the point in time at which the signal is evaluated.

Jitter is one of the major problems in the design of digital electronics devices, in particular digital interfaces. Inaccurate calculation of jitter can lead to its accumulation when passing a digital signal on the path and, ultimately, to the inability of the device.

The following parameters were calculated for each group of entries in the Matlab environment: 1. The average variation of the Jitta pitch wavelength (μs). 2. Relative variation of Jitt pitch (%).

The point of intersection of zero of the real signal front for different pulses varies in time.

The jitter component can be drawn from a clock or self-timed digital signal and analyzed separately. Among the most useful ways of studying the effect of jitter is to study the frequency response and identify the main frequency components of jitter.

With a small amount of jitter, the front of the meander shifts backwards or forwards by a small amount in time. With increasing jitter, the displacements reach large values. The jitter amplitude is called the time offset value and is measured in units of time: either in fractions of a second (nanoseconds, picoseconds), or in interval units (UI, unit interval). For those who are experiencing jitter measurements for the first time, lettering along the axes of the graph can be confusing - often both the vertical and horizontal axes delay time.

Jitter frequency is called the frequency at which the phase shift occurs. Just as in the case of noise or interference, the signal delivered by the jitter can be a pure sine wave, a complex oscillation, or a completely random process.

Another important process of the system is the detection of employee somatotype. The psychological type of a person is influenced by his physiological structure, which William Sheldon associated with types of temperaments. According to his theory, people are divided by physiological constitution into 3 different types (somatotypes): endomorphs, mesomorphs, and ectomorphs.

Endomorphs have the best developed digestive organs. Physiological features of this type are rounded body shapes; large internal organs; subcutaneous fat is present on the abdomen and hips; soft body; relaxed arms and legs; muscles are not developed.

Mesomorphs have developed muscles, bones and cartilage. These are athletic people with large chest, straight posture, broad shoulders, muscular arms and legs, and low in fat.

Ectomorphs have a rather lean body, the muscles are poorly developed, with long arms and legs; the thorax is narrow, and there is almost no fat layer. They have a well-developed nervous system.

Insider propensity ratios on the psychological type scale are presented in Table. 2.

The responsible employee is most suitable for ectomorphs that correspond to the cerebrotic psychological type by 80%. Somatotype of this type corresponds to a scale of 1-1-7, which denotes the cerebrotonic organism in its pure form. But in nature, this type is rare. Therefore, a worker who has the overwhelming amount of cerebrotonia in the scale of his somatotype will be suitable for working with IOD. We have given each somatotype a factor of 0 to 6. People with a factor of 0 are pure-looking cerebrotonics and most responsible workers. Workers with a factor of 6 are the strongest insiders. If on the scale of psychological type the indicator of viscerothenia or somatotonia exceeds cerebrotonia, then such employee does not fit in a responsible position. Thus, people with psycho-type ratios of 0 to 3 are allowed to access restricted information.

Table 2. Insider propensity ratios on a psychological type scale

Psychological type scale (points)													
1-1-7	**0**	1-3-5	**2**	1-4-4	**3**	5-1-3	**4**	5-2-2	**5**	3-4-2	**5**	6-2-1	**6**
1-2-6	**1**	3-1-5	**2**	4-1-4	**3**	1-5-3	**4**	1-6-2	**5**	7-1-1	**6**	4-4-1	**6**
2-1-6	**1**	2-3-4	**3**	4-2-3	**4**	3-3-3	**4**	6-1-2	**5**	1-7-1	**6**	4-4-1	**6**
2-2-5	**2**	3-2-4	**3**	2-4-3	**4**	2-5-2	**5**	4-3-2	**5**	2-6-1	**6**	5-3-1	**6**

3. Experimental study

To perform the insider identification experiment, a database of records of long vowels / a /, / o /, / e /, / and / Ukrainian male and female speakers was created. The speakers were asked questions about their name, age, anthropometric indicators (height,

weight) and gender. This helped us to identify the respondent by determining formulas 1 and 2 and observed the value of the shake parameter during the talk of the speakers. As the passport names, height and weight of the respondents were already known, none of them tried to provide false information by voice. Therefore, the Jitter parameter was measured for each loud sound at rest of the subjects [8, 9, 10].

At the second level, the human psychological type on the basis of physiological data was determined. This parameter is K_P. Recording conditions: the microphone is located 10-15 cm from the mouth. The recording software - Microsoft Sound Recorder (Wave PCM, 32 kHz sampling rate, mono, 16 bits) [11, 12, 13, 14, 15].

The range for finding the pitch of the voice from the signal is selected from 80 Hz (lower pitch for bass) to 600 Hz (upper pitch for tenors). Once the dominant pitch frequency is detected, further analysis is conducted in the range (F0 ± F0 / 2), which covers all the frequencies of the vocal oscillations. Because normal voice communications during phonation can change the oscillation frequency by an octave lower or higher than the dominant pitch frequency [3]. Peaks and pits were detected in the received signal by calculating the first derivative. The distance between two adjacent peaks is the wavelength of the pitch. The duration values of all waves form the sequence P0(n), where n is the number of pitch waves in the signal. The following parameters were calculated for each entry group in the Matlab environment:

1. Average variation of Jitta pitch wave duration (μs).
2. Relative variation of Jitt pitch (%).

The distance between two adjacent peaks is the wavelength of the pitch. The duration values of all waves form the sequence P0(n), where n is the number of pitch waves in the signal. The following Jitta and Jitt parameters were calculated for each entry group in the Matlab environment:

According to the monitoring system for suspicious transaction types, the Very Suspicious response module requires one of the authentication methods. The authentication type information for the operations used in the experiment is shown in Table. 3.

Table 3. The results of the detection of worry by the system of identification of insider features

Sound	***Jitta, мкс***		***Jitt, %***	
	No worry	Worry	No worry	Worry
/a/	70	88	0,94	1,7
/o/	50,8	112,25	0,78	2,025
/e/	63,8	121	0,94	2,075
/i/	53,6	95	0,8	1,675

Subsequently, or in parallel with the first stage of the system, the main features of the human psychological type were identified and a method for finding the coefficient of psychological type by the physiological structure was proposed [16]. The results of the system of identification of insider features are presented in table. 4.

Table 4. The results of the system of identification of insider features

Speaker	K_J	K_A	K_P	K_R	Result
1	0	0	3	0	0
2	0	1	1	0	1
3	**1**	**0**	**4**	**4**	**4**
4	0	1	5	2	3
5	1	1	0	4	2
6	0	0	4	2	2
7	0	0	6	2	2
8	1	1	5	0	2
9	0	1	2	0	0
10	1	0	0	0	1
11	0	0	5	2	2

where K_j – the Jitter parameter, K_A – voice parameter, K_p – somatotype parameter, K_R – the leading modality parameter.

Thus, the method of calculating the indicator of the worries of the person using the Jitter variable was presented, as well as a method of identifying and evaluating the representative system of the person. On the basis of these data, a set of parameters has been developed that is integrated into the system, which allows to determine the level of propensity to insider activity: "no", "low", "medium", "high", "critical" [17]. The general results of the experimental study are presented in Fig. 2.

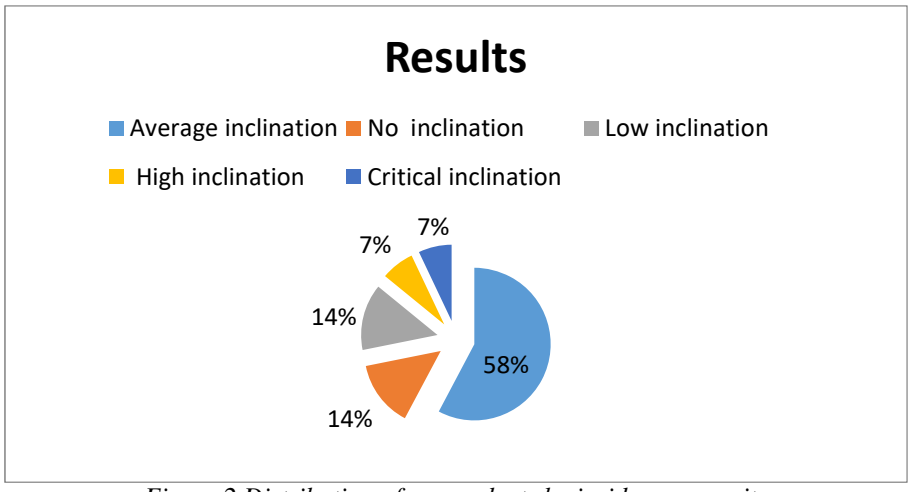

Figure 2 Distribution of respondents by insider propensity

4. Conclusions

The basic concepts and technologies of insider detection, which became the basis for creating voice identification and detecting the worries of the interviewee, were analyzed.

A two-stage insider identification system that takes into account voice, age, anthropometric and physiological data for identifying a potential or existing insider has been developed.
In the course of the experimental study of the system, one explicit insider of 11 people was identified. Comparing the results of the developed system with the results of the existing one, it became clear that the system adequately evaluates the speakers.

REFERENCE

1. ARDAKANI M. S., SOLEYMANI Z., TORABINEJAD F., KHODDAMI S. M., HEYDARI M. A.: Fundamental frequency, jitter, and shimmer of adult stutters and nonstutters voice. Audiology, Vol.16, No.2, 2008.
2. SCALASSARA P.R., MACIEL C.D., PERREIRA J.C.: Predictability analysis of voice signals. IEEE Engineering in Medical and Biology magazine 2009.
3. JIANGLIN WANG, CHEOLWOO JO: Vocal folds Disorder Detection using Pattern Recognition Methods. Proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale, Lyon, France August 23-26, 2007.
4. BALWANT A. SONKAMBLE D., DOYE D.: Speech Recognition Using Vector Quantization through Modified K-means LBG Algorithm. Computer Engineering and Intelligent Systems Vol 3, No.7, 2012.
5. Аналіз, розпізнавання і інтерпретація мовних сигналів. Т.К. Вінцюк. 2010.
6. PICONE J., POPOV R.: Методи моделювання сигналу в розпізнаванні мови. Кемерово. 2000.
7. Чисельні методи Н.С. Бахвалов. Головна редакція фізико- математичної літератури вид-ва «Наука», М., 1975.
8. Голос як один з каналів невербальної комунікації І. Б. Єсін. Юр. психологія. 2008.
9. Автоматичне розпізнавання говорить по голосу Г. С. Рамішвілі. Радио и связь, 1981. - 244 с. - Бібліогр. : С. 58 - 207.
10. РЕВТОВ Ю.І.: Розпізнання людини по голосу і особливостям мови в процесі розслідування: дис. . канд. юр. наук. - Мінськ, 1998. - 45 с.
11. ТЕПЛОВ Б.М.: Розум полководця: Проблеми індивідуальних відмінностей / Б. М. Теплов.-М .: АПН РРФСР, 1961. Чистовіч, Л. А. Фізіологія мови. Сприйняття мови людиною / Л. А. Чис-товіч і ін .. Л.: Наука, 1976. - 388 с.
12. .NET API 2015 [Електронний ресурс]. – 2018. – Режим доступу: https://uk.wikipedia.org/wiki/.NET APIс. – Дата доступу: квітень 2018.
13. Розробка через тестування [Електронний ресурс]. – 2018. – Режим доступу: https://en.wikipedia.org/wiki/Test-driven_development. – Дата доступу: квітень 2018.
14. Model-View-Controller [Електронний ресурс]. – 2018. – Режим доступу: https://ru.wikipedia.org/wiki/Model-View-Controller. – Дата доступу: квітень 2018.
15. ПОТАПОВА Р.К.: О возможности перцептивно-слухового распознавания состояния «агрессия» по устной речи / ПотаповаР. К., Комалова Л.Р. // Вестник МГЛУ. – 2014. – №13 (699) . – С.202-214.

16. BROSE A.: Affective states contribute to trait reports of affective well-being / Brose A., Lindenberg U.,Schmiedek F. // Emition .- 2013(5) .- p. 940-948.
17. KÄCHELE M.: Prosodic, Spectral and Voice Quality Feature Selection Using a Long-Term Stopping Criterion for Audio-Based Emotion Recognition / M. Kächele, D. Zharkov, S. Meudt and F. Schwenker // Pattern Recognition (ICPR), 2014 22nd International Conference on,Stockholm . – 2014. – pp. 803-808.