Approaches to the protection of audio files in BULGARIAN LABLING CORPUS

Dimitar Popov[1], Velka Popova[1], Krasimir Kordov[1] and Stanimir Zhelezov[1]

[1]*Konstantin Preslavsky University of Shumen, 115 Universitetska str., 9700 Shumen, Bulgaria*

Abstract

This article discusses the problems of protection of multimodal corpora with human speech, which are being developed in the Laboratory of Applied Linguistics (LabLing) at the University of Shumen. Ensuring the protection of audio files in the BULGARIAN LABLING CORPUS, which are provided for free access to users, is the main goal of this study. To achieve this goal, two approaches have been chosen - cryptographic and steganographic. Cryptographic and steganographic methods for protection of BULGARIAN LABLING CORPUS audio files have been proposed and verified. It has been proven that the algorithms for the implementation of the proposed methods have a high level of reliability and security, which makes them extremely suitable for the purpose of this study.

Keywords

spontaneous speech corpus, labling corpus, sound files protection, cryptography, steganography

1. Introduction

This article discusses the problems of protection of multimodal corpora with human speech, which are being developed in the Laboratory of Applied Linguistics (LabLing) at the University of Shumen. LabLing is part of the consortium of the Bulgarian national research infrastructure for resources and technologies for language, cultural and historical heritage, integrated within CLARIN and DARIAH (CLaDA-BG – https://clada-bg.eu/en).

One priority of the researchers from LabLing is to monitor the individual speech development of a group of Bulgarian children by conducting longitudinal observations. The methodology of Brian MacWhinney [1] was used for optimal multifaceted visualization of speech through the interactive multimodal system of the CHILDES platform, where the integrated presentation of empirical data through transcripts of audio and video recordings is possible, which are simultaneously linked to several modes of communication [2]. As a significant result of the work of the LabLing team the publication of the pilot version of the first Bulgarian CHILDES - corpus can be highlighted – BULGARIAN LABLING CORPUS (https://childes.talkbank.org/access/Slavic/Bulgarian/LabLing.html).

These data will definitely be of great importance for the formation and creation of a national interdisciplinary electronic infrastructure in the process of integration and development of electronic resources in the Bulgarian language. Therefore, the construction of LabLing CORPUS is a priority task of the CLaDA-BG consortium. In relation to that, there is an immediate need to ensure the protection of audio files in BULGARIAN LABLING CORPUS, which are provided for free access to users. This paper seeks a solution to this problem in the paradigm of cryptography and steganography.

2. Cryptographic protection of audio files of the Laboratory of Applied Linguistics

2.1. Why cryptography?

The use of cryptography for this purpose is one of the most common methods of information protection in the transmission of data on computer networks and in the exchange of information in communication channels between remote objects.

Cryptographic means of protection are special methods and means for transforming information, as a result of which its content is masked. Cryptographic transformations change the components of the messages (letters, words, numbers) in an implicit form through special algorithms, code keys or hardware solutions.

2.2. Which cryptographic methods are used?

Cryptography uses two types of cryptographic methods depending on the secret keys that are used for encryption

Language Technologies and Digital Humanities in Bulgaria (LTDH-BG) CLaDA-BG 2021 Conference, First International CLaDA-BG Conference, 6 - 7 September 2021, Varna, Bulgaria

labling@shu.bg (D. Popov); v.popova@shu.bg (V. Popova); krasimir.kordov@shu.bg (K. Kordov); s.zhelezov@shu.bg (S. Zhelezov)

0000-0001-5998-6263 (D. Popov); 0000-0001-6222-2416 (V. Popova); 0000-0002-4397-8303 (K. Kordov); 0000-0001-8761-2592 (S. Zhelezov)

and decryption - symmetric and asymmetric.

The concept of asymmetric cryptographic methods was introduced by W. Diffie and M. Hellman of Stanford University nearly 40 years ago. They propose the idea of creating cryptographic systems using a public key, thus giving a new direction in the development and research of cryptographic methods [3]. In their article, they justify and define the use of public key systems called Public Key Systems (PKS). All asymmetric cryptographic algorithms are characterized by the use of key pairs, and the recognition of one of the keys cannot be used to calculate the other. The encryption process is performed with the public key, which is non-secret, and the decryption process is performed with a secret (private) key known only to the recipient of the message. In symmetric encryption, the same secret key is used to encrypt and decrypt messages.

The communication process proceeds in the following steps:

1. The sender of the message uses an encryption method, transforming the incoming message with the secret key;
2. As a result, an encrypted message is received, which is sent to the recipient;
3. The encrypted message reaches the recipient, who uses the secret key to recover the incoming message.

The symmetric approach in cryptography is also referred to as conventional cryptography. Symmetric cryptographic algorithms are divided into block and stream. In streaming encryption, each character is transformed independently of the others using the secret key, and in the block approach, the message is divided into blocks, and the transformation of the characters in the block is highly dependent.

To choose a cryptographic method, it is important to compare the advantages and disadvantages of the two types of cryptographic methods [3]. In terms of speed, symmetrical methods are preferred to asymmetrical ones.

2.3. Use of pseudo-random sequences for encryption

Pseudo-random number generators are a class of cryptographic primitives that are a major building component of any symmetric cryptographic system that performs streaming encryption. A true random sequence is that sequence of bits (0 and 1) for which the knowledge of the arbitrary subset of its elements does not give any information about the other bits. Examples of such sources are: the decay of the nucleus of a radioactive element, the thermal noise of a diode or resistor, the sound from a microphone or video input from a camera, the instability of the oscillator frequency, and others. The use of such sources is associated with a number of technical difficulties [4, 5], so the so-called pseudo-random series (PRS), and the generators of such PRS are called pseudo-random generators (PRG). Traditionally, linear-feedback shift registers (LFSR) and feedback with carry shift registers (FCSR) are used as approaches in the construction of PRG.

In recent years, chaotic maps have been used in the construction of PRG. This is due to their chaotic behavior and better cryptographic protection performance [6, 7]. This paper describes the implementation of pseudo-random generator based on two maps of this type - duffing map and circle map. The method itself is described in detail in [8]

The main steps of the algorithm that implements the method are the following:

1. The generator is initialized;
2. The samples are converted into binary form;
3. The samples are encrypted with pseudo-random sequence;
4. The samples are combined into an encrypted file.

2.4. Verification of the method

The main purpose of the cryptographic analysis is restoring the plain message from the encrypted message. In this section, in order to prove the audio encryption efficiency, we performed various empirical tests to compare plain files and their corresponding encrypted files.

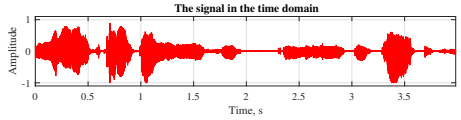
Waveform Plotting

One of the most common approaches, concerning audio signal analysis is waveform plotting to display the audio signal amplitude distributed in time. To compare the plain audio files with the encrypted ones we present the visualization of one of the tested files. Figure 1(a) represents the waveform of a normal file before encryption, Figure 1(b) represents the changes in the file after encryption and Figure 1(c) demonstrates the restored file after decryption.

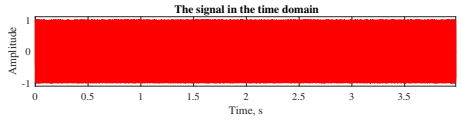
The difference between the plain file plot and the encrypted file plot is an indication of successful encryption. Furthermore, the strong difference also means the original file cannot be restored even partially.

Spectrogram Plotting

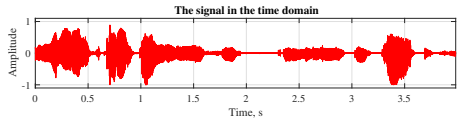
The spectrogram plotting is another important approach for analyzing audio signals. In this case the main focus is the frequency of the sound against time domain. Comparing plain files with encrypted files allows us to see the difference between the files and to evaluate the proposed audio encryption algorithm. Figure 2(a) shows the spectrogram of a plain file, Figure 2(b) represents the changes in the file after encryption and Figure 2(c)



(a) Plain audio file.

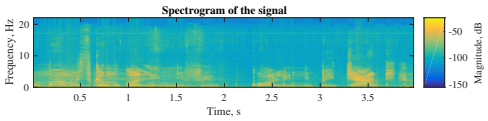


(b) Encrypted audio file.

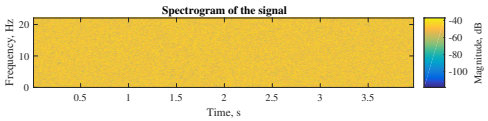


(c) Decrypted audio file.

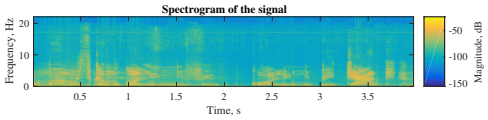
Figure 1: Waveform Plotting.



(a) Plain audio file.



(b) Encrypted audio file.



(c) Decrypted audio file.

Figure 2: Spectrogram Plotting.

demonstrates the restored file after decryption. The spectrogram plot of the encrypted file means the frequency of the original signal in the plain file is completely destroyed. This test is another indicator of the high encryption properties of the proposed audio encryption algorithm.

The proposed cryptographic algorithm relies on permutation-substitution architecture realized by using chaotic circle map and modified rotation equations. Extended cryptographic analysis is performed for testing the proposed method for security. The waveform plots and the spectrograms of the tested audio files demonstrate the changes in encrypted files compared to plain files.

3. Steganographic protection of audio files of the laboratory of applied linguistics.

Part of the activity of the laboratory of applied linguistics is related to the creation and processing of audio files with human speech. Recordings are subject to copyright of the laboratory and steganography methods may be used for their protection.

3.1. Why steganography?

Steganography is a scientific field of application, a set of technical skills and art for ways to hide the fact of transmission (availability) of information [9]. High-tech steganography is a term used by some authors to summarize the directions for hiding messages using communication and computer technology, nanotechnology and modern advances in biology. Steganography encompasses methods using redundancy in the binary representation of multimedia information.

Steganographic methods (stegomethods) allow hiding data in different containers: text documents (electronic articles, books, letters), in graphic files (drawings, banners, photographs), in video files (videos, movies, animation), in audio files (music works, speech, natural sounds), in the code of HTML pages, in movie subtitles, in messages transmitted via SMS, MMS, chat, blogs, etc. Text messages can be hidden in unused areas of Flash memory, hard drives, and optical drives. Given that each type of container has different formats, and various methods can be used to hide the information, it is clear how multidimensional the steganographic tasks are.

Each of the multimedia containers has its own characteristics and each of them requires the use of specific methods for embedding and retrieving hidden information. Multimedia steganography is one of the most studied areas of computer steganography. It covers methods using the excess in the binary representation of visual and sound information. Digital images, digital music and digital video are represented by matrices of numbers that encode the intensity of colors or sound signals in space and/ or time. The lower digits of digital readings contain a very small payload for the current parameters of sounds and images. Filling them with other data does not significantly affect the quality of perception of the image or sound by people.

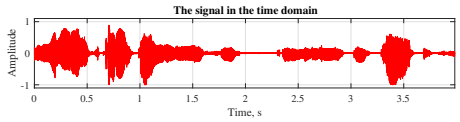
3.2. Which steganographic embedding methods and algorithms are used?

Digital images, digital music and digital video are represented by matrices of numbers that encode the intensity of colors or sound signals in space and/ or time. The lower grades of digital readings contain a very small payload for the current parameters of sounds and images. Filling them with other data does not significantly affect the quality of perception of the image or sound by people. When used with image containers, for example, about 100 KB of information can be embedded in an 800 KB image file without significantly altering the container image. In an audio container - quantized sound with a duration of 1 sec., with a frequency of 44 KHz and an accuracy of 8 bits, stereo mode, the popular LSB method allows you to hide a message of about 10 KB. This leads to about 1% change in the value of the amplitude of the sound signal, which is practically impossible to detect by most people when listening to the audio file.

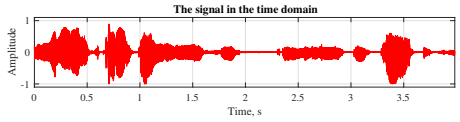
In recent years, many scientific developments have examined the possibilities for steganographic exchange of sensitive information provided by different types of multimedia containers - graphic files, audio files and video files. Most of them rely on a combination of different methods of steganography, and sometimes cryptographic methods are implemented in order to increase security. A typical example of this is the use of pseudo-random number generators for scattered steganographic embedding [10].

3.3. Use of pseudo-random sequences in LSB embedding

Numerous scientific publications in recent years have paid serious attention to spread spectrum steganography methods. Methods using the selection of pseudo-random positions to embed hidden information have become especially popular. These methods use different approaches to generate these pseudo-random sequences. Some of them are based on feedback shift registers [11], Others are based on different types of chaotic maps. In [8] is proposed a method of steganographic embedding in images using a pseudo-random generator based on duffing map and circle map. From these studies it is clear that this method has a high level of security and reliability, and is also fast. This makes it very suitable for use in steganographic protection of audio files. The realization of the method in steganographic protection of audio files with human speech is presented in the current development. For this purpose, empirical material from the database of the Laboratory of Applied Linguistics is used. The algorithm for its implementation contains the following steps:

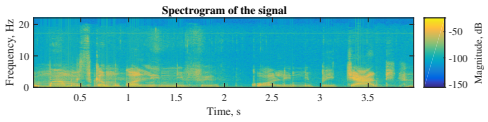


(a) Plain audio file.

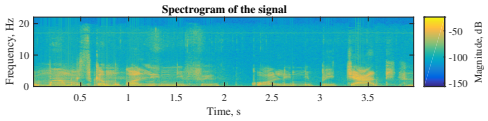


(b) Stego audio file.

Figure 3: Waveform Plotting.



(a) Plain audio file.



(b) Stego audio file.

Figure 4: Spectrogram Plotting.

1. The text information is transformed to vector V of binary sequence using ASCII table values of the characters;
2. Initializing the generator;
3. Choosing random samples (in chaotic order) from an audio, using the constructed pseudo-random generator for embedding information;
4. Using traditional LSB samples modification for hiding the values of the vector V, leaving no traces of steganography.

3.4. Verification of the method

Waveform Plotting

To compare the plain audio files with the stego ones we present the visualization of one of the tested files. Figure 3(a) represents the waveform of a normal file before and Figure 3(b) represents a stego file after embedding.

The lack of differences in the two files indicates that the selected method leaves no traces. Therefore, the main task of steganography, the hidden steganographic information to be invisible, is fulfilled.

Spectrogram Plotting

Comparing plain files with stego files shows that there is no visible difference between the files and allows to evaluate the proposed audio stego algorithm. Figure 4(a) shows the spectrogram of a plain file and Figure 4(b) represents the changes in the file after embedding. This test is another indicator of the high stego level of the proposed audio steganographic algorithm.

4. Conclusion.

Ensuring the protection of audio files in the BULGARIAN LABLING CORPUS, which are provided for free access to users, was the main goal of this study. To achieve this goal, two approaches were chosen - cryptographic and steganographic.

From the research described in this article, it is clear that the proposed cryptographic algorithm meets all modern requirements for cryptographic protection of information. This algorithm has a proven high level of reliability and security, which makes it extremely suitable for achieving the goal in the present study.

The same conclusion could be drawn for the proposed steganographic method. During the verification of the proposed method, it was demonstrated that the information that marks the audio files in the BULGARIAN LABLING CORPUS as part of the database of LabLing laboratory is completely invisible. This proves the applicability of the chosen method and the validity of the implemented stego algorithm in the protection of LabLing audio files.

Acknowledgments

This research was partially funded by the Bulgarian National Interdisciplinary Research e-Infrastructure for Resources and Technologies in favor of the Bulgarian Language and Cultural Heritage, part of the EU infrastructures CLARIN and DARIAH – CLaDA-BG, Grant number DO1-272/16.12.2019.

References

[1] MacWhinney, B., Wagner, J.: Transcribing, searching and data sharing: The CLAN software and the TalkBank data repository. In: Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion, Ausgabe 11. pp.154–173 (http://www.gespraechsforschung-ozs.de) (2010)

[2] Popova, V., Popov, D.: Multimodal Presentation of Bulgarian Child Language. In: Speech and Computer (SPECOM 2015). 17th International Conference, SPECOM 2015, Athens, Greece, September 20-24, 2015, Proceedings (A. Ronzhin, R. Potapova, N. Fakotakis, eds.). Springer International Publishing Switzerland, pp. 293–300 (2015).

[3] Diffie, W., Hellman, M.: New directions in cryptography. IEEE transactions on Information Theory, **22**(6), 644-654 (1976).

[4] Tamimi, A. A., Abdalla, A. M.: An audio shuffle-encryption algorithm. In Proceedings of the World Congress on Engineering and Computer Science, San Francisco, CA, USA, 22—24 October 2014.

[5] Sathiyamurthi, P., Ramakrishnan, S.: Speech encryption using chaotic shift keying for secured speech communication. EURASIP Journal on Audio, Speech, and Music Processing, **2017**(1), pp.1-11 (2017).

[6] Kordov, K. M.: Modified Chebyshev map based pseudo-random bit generator. In AIP Conference Proceedings, **1629**(1), pp. 432–436. American Institute of Physics (2014).

[7] Kordov, K.: Signature attractor based pseudorandom generation algorithm. Advanced Studies in Theoretical Physics, 9(6), pp. 287–293 (2015).

[8] Kordov, K., Zhelezov, S.: Steganography in color images with random order of pixel selection and encrypted text message embedding. PeerJ Computer Science, **7**, e380 (2021).

[9] Stanev, S., Szczypiorski, K.: Steganography Training: a Case Study from University of Shumen in Bulgaria. International Journal of Electronics and Telecommunications, **62** (2016).

[10] Kordov, K.: A novel audio encryption algorithm with permutation-substitution architecture. Electronics, **8**(5), 530 (2019).

[11] Kordov, K.: Modified pseudo-random bit generation scheme based on two circle maps and XOR function. Applied Mathematical Sciences, **9**(3), pp. 129-135 (2015).