Data and Metadata. 2024; 4:365
doi: 10.56294/dm2024365

ORIGINAL

Securing biomedical audio data in IoT healthcare systems: an evaluation of encryption methods for enhanced privacy

Protección de los datos de audio biomédicos en los sistemas de salud de IoT: una evaluación de los métodos de cifrado para mejorar la privacidad

Mohammed Amraoui[1], Imane Lasri[2], Fouzia Omary[1], Mohamed Khalifa Boutahir[3], Yousef Farhaoui[3]

[1]Intelligent Processing and Security of System (IPSS Team), Department of Computer Science, Faculty of Sciences, Mohammed V University in Rabat. Morocco.
[2]Laboratory of Conception and Systems (Electronics, Signals and Informatics), Faculty of Sciences, Mohammed V University in Rabat. Morocco.
[3]Engineering science and technology laboratory, IDMS Team, Faculty of Sciences and Tech-niques, Moulay Ismail University of Meknes. Morocco.

Cite as: Amraoui M, Lasri I, Omary F, Boutahir MK, Farhaoui Y. Securing Biomedical Audio Data in IoT Healthcare Systems: An Evaluation of Encryption Methods for Enhanced Privacy. Data and Metadata. 2024; 4:365. https://doi.org/10.56294/dm2024365

Submitted: 02-02-2024 **Revised:** 07-04-2024 **Accepted:** 29-06-2024 **Published:** 30-06-2024

Editor: Adrián Alejandro Vitón Castillo

ABSTRACT

Communication technology have advanced quickly since the COVID-19 epidemic started, providing consumers with additional benefits and conveniences. Concerns over the privacy and confidentiality of this data have grown in importance as initiatives that promote the use of audio and video to enhance interpersonal interactions become more common. In the context of the Internet of Things (IoT), audio communications security is essential in the biomedical domain. Sensitive medical data may be compromised in these connections, which include exchanges between patients and doctors and broadcasts of vital signs. To protect patient privacy and reduce cybersecurity threats, strong security measures such as data encryption must be put in place. Our study attempts to address these issues in this environment. Comparative examination of the Chacha20, Salsa20, and Camellia encryption algorithms enabled us to ascertain that Chacha20 performs exceptionally well when it comes to audio file decryption and encryption speed. The results of our trials attest to this encryption method's astounding effectiveness and efficacy. We have also used the noise reduction technique, which is frequently used in audio security to enhance the quality of recordings and make it easier to identify significant information in audio signals. Then, Fourier transform technique, which is also used to analyze audio files and can be used to identify changes, extract hidden information, and authenticate audio files. By doing this, the audio files security and integrity are strengthened.

Keywords: Audio Security; Chacha20; Salsa20; Camellia; Noise Reduction; Fourier Transform.

RESUMEN

La tecnología de la comunicación ha avanzado rápidamente desde que comenzó la epidemia de COVID-19, brindando a los consumidores beneficios y comodidades adicionales. La preocupación por la privacidad y confidencialidad de estos datos ha aumentado en importancia a medida que las iniciativas que promueven el uso de audio y video para mejorar las interacciones interpersonales se vuelven más comunes. En el contexto del Internet de las Cosas (IoT), la seguridad de las comunicaciones de audio es esencial en el dominio biomédico. Los datos médicos confidenciales pueden verse comprometidos en estas conexiones, que incluyen intercambios entre pacientes y médicos y transmisiones de signos vitales. Para proteger la privacidad de los pacientes y reducir las amenazas a la ciberseguridad, se deben implementar medidas de seguridad sólidas, como el cifrado de datos. Nuestro estudio intenta abordar estos problemas en este entorno. El examen comparativo de los algoritmos de cifrado Chacha20, Salsa20 y Camellia nos permitió determinar que Chacha20 funciona excepcionalmente bien en lo que respecta al descifrado de archivos de audio y la velocidad de cifrado. Los resultados de nuestros ensayos dan fe de la asombrosa efectividad y eficacia de este método de cifrado. También hemos utilizado la técnica de reducción de ruido, que se utiliza con frecuencia en la seguridad del audio para mejorar la calidad de las grabaciones y facilitar la identificación de información significativa en las señales de audio. Luego, la técnica de transformada de Fourier, que también se usa para analizar archivos de audio y se puede usar para identificar cambios, extraer información oculta y autenticar archivos de audio. Al hacer esto, se fortalece la seguridad e integridad de los archivos de audio.

Palabras clave: Seguridad de Audio; Chacha20; Salsa20; Camelia; Reducción de Ruido; Transformada de Fourier.

INTRODUCTION

In the constantly evolving technological landscape, the integration of audio is becoming increasingly important, especially in the fields of biomedical and the Internet of Things (IoT). At the heart of this transition, voice communication is asserting itself as an immersive and natural alternative to text communication, thus improving the user experience. At the same time, advances in the field of IoT have paved the way for new applications, such as connected medical devices, which exploit voice communication to improve healthcare and remote medical monitoring. This convergence between biomedical and IoT raises unique challenges in terms of security and confidentiality of audio data, highlighting the need to adopt appropriate protection measures to prevent unauthorized access and guarantee the confidentiality of exchanges.[1]

The voice has the ability to express subtleties, dialects, and emotions that text alone cannot synchronous communication.[2] The use of phone calls, conference calls, online meetings, and voice mail systems are just a few examples where this is important.[3] Real-time audio communication speeds up decision-making, cooperation, and problem-solving. Speech makes communication easier for those who are physically or visually impaired.[4] The visually impaired and hard of hearing[5] can access information and spoken communication thanks to speech technology like screen readers, voice synthesizers, and automatic transcription systems.[6] However, the growing language use presents significant issues with regard to privacy and secrecy. To protect against illegal access and ensure the secrecy of talks, the security of voice data transferred is important. One of the frequent dangers of voice is that it might be recorded without the participants' knowledge, which violates their right to privacy. Unauthorized recordings may be used for surveillance, extortion, or the release of sensitive data.[7] Unauthorized third parties or malicious actors may be able to intercept audio data while it is in transit. Sensitive data may be revealed in this way.[8]

Automated analysis of audio data, including speech recognition, speaker identification, and sentiment analysis is now possible because to advances in artificial intelligence and machine learning. However, because these analyses have the potential to divulge private and secret information without the knowledge of the parties involved, confidentiality issues are raised.[9] To avoid secure storage and unwanted access, voice data must also be properly managed in addition to being stored and managed. Voice recordings may be lost or distorted as a result of a security breach or storage system malfunction.[10] It's important to implement the right security measures to reduce these dangers, such as employing secure communication protocols, limiting access and rights, encrypting voice data, and gaining authorization before collecting or analyzing voice data.[11] Strengthening privacy and data protection knowledge is also required on both the user's and the service provider's sides. The use of voice data by service providers is subject to strong privacy rules and procedures, security measures for the voice data they collect, store, and send, and valid user permission. Users should be made aware of your privacy policies and the security precautions in place to safeguard their voice data.

In this study, the topic of audio privacy and the value of putting in place suitable security measures are covered. To prevent unauthorized people from understanding the information, we provide a method based on encrypting audio data. The Chacha20,[12] Salsa20[13] and Camellia[14] audio encryption algorithms are introduced, and their effects on performance, security, and audio quality are assessed. The best solution for a certain need is chosen after a comparison of the available audio encryption techniques. It also discusses security concerns and assesses how well the algorithm withstands frequent assaults like brute force and dictionary attacks. We also used the Fourier transform,[15] which examine the distortion, noise, and fidelity metrics, as well as the factors affecting audio quality. The ultimate objective is to provide recommendations for choosing the best algorithm that is characterized by the speed of encryption and decryption of the audio, taking into account both the security of the data and the quality of the audio.

In recent years, the rapid advancement of audio processing technologies has highlighted the importance of securing and ensuring the integrity of audio data. Audio security involves techniques to protect audio content from unauthorized access, manipulation, and interception, crucial in fields like telecommunications,[16] surveillance,[17] forensics[18] and multimedia applications.[19] Researchers have explored innovative solutions to address these challenges.

Yazdanpanah et al.[1] used machine learning for voice steganalysis, presenting the Percent of Equal Adjacent Samples (PEAS) feature, achieving 99,82 % sensitivity. Shelke et al.[9] improved security with the Arnold transform and elliptic curve encryption for audio watermarking. Pleshkova et al.[17] used Public Key Infrastructure (PKI) to secure audio information delivery. Mcuba et al.[4] reviewed deep fake audio detection techniques, finding VGG-16 effective for MFCC features. Lin et al.[25] proposed a chaos-based cryptosystem for streaming audio and video, using dynamic key generators with AES CFB encryption. Lima et al.[27] introduced "event-based cryptography" for CPS automation networks, comparing ChaCha20 with RSA. Babiano et al.[18] tackled ransomware by recovering Salsa20 encryption keys from volatile memory.

Our research addresses the urgent need to secure audio transmissions amidst the rapid advancement of communication technologies, particularly during the COVID-19 pandemic. We evaluate the effectiveness of encryption algorithms like Chacha20,[12] Salsa20,[13] Camellia[14] and implement noise reduction and Fourier transform techniques[15] to enhance audio signal quality and security. Our methodology aims to provide insights to strengthen audio transmission security across various domains.

As part of our research into the security of audio transmissions, we have used a comprehensive methodology to evaluate the effectiveness of encryption algorithms, examine the Fourier transform technique for improving audio signal quality, and apply noise reduction techniques figure 1.

METHOD

Audio Data Preparation

We collected three standard audio files for our investigation:

- Audio 1: 3 seconds[29]
- Audio 2: 10 seconds[29]
- Audio 3: 120 seconds (biomedical audio file)[30]

Audio file encryption: to encrypt the audio files, we used the encryption methods ChaCha20, Salsa20, and Camellia. To ensure the data's security, each method has been utilized in conjunction with the relevant secret key.

Decryption of audio files: to decode the encrypted audio files, we employed the same encryption techniques. The original audio data was located by using the relevant secret keys.

Analysis of frequency characteristics: Chacha20 algorithm, which stands out for its performance, was used to apply the Fourier transform[15] to the original, encrypted, and decrypted audio files by examining the frequency elements of them. We were able to comprehend how encryption techniques affects the frequency distribution of audio recordings thanks to our investigation. The Fourier transform F of a signal x(t) is defined in equation 1:

$$X(f) = \int_{-\infty}^{\infty} x(t) \cdot e^{-j2\pi f t} dt \qquad (1)$$

Where X(f) represents the frequency domain representation of the signal x(t), and f is the frequency variable.

Noise Reduction: in order to enhance the original audio files' quality, we used a noise reduction approach. To get rid of or lessen undesirable aspects like background noise, interference, or distortion, our approach employs filters and signal processing algorithms.

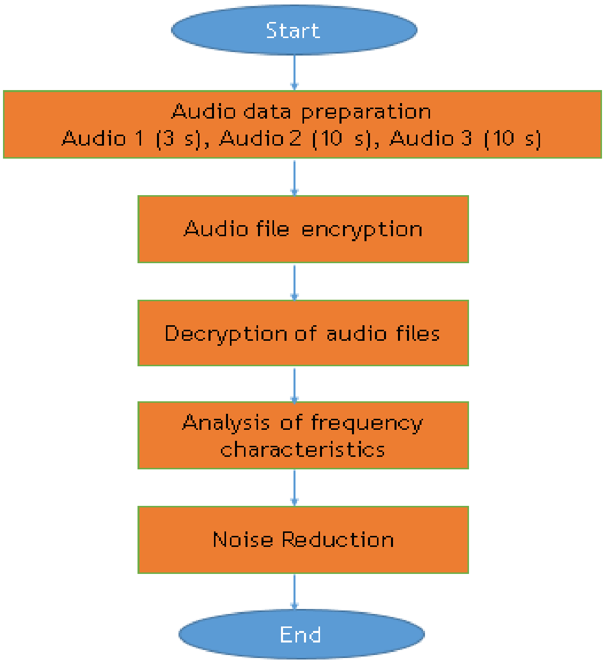

Start
Audio data preparation
Audio 1 (3 s), Audio 2 (10 s), Audio 3 (10 s)
Audio file encryption
Decryption of audio files
Analysis of frequency characteristics
Noise Reduction
End

Figure 1. Pipeline of our proposed methodology for audio security

Salsa20 algorithm

Salsa20,[13] a stream cipher, relies on a pseudorandom function built upon Add-Rotate-XOR (ARX) operations, integrating 32-bit addition, left rotation, and bitwise XOR. This cipher operates within a 16-word structure, with each word comprising 32 bits. The core structure can be depicted using the following equation:

$$s = \begin{pmatrix} s_0 s_4 \; s_8 \; s_{12} \\ s_1 s_5 \; s_9 \; s_{13} \\ s_2 s_6 s_{10} s_{14} \\ s_3 s_7 s_{11} s_{15} \end{pmatrix} = \begin{pmatrix} c_0\, k_3\, t_0\, k_5 \\ k_0\, c_1\, t_1\, k_6 \\ k_1\, v_0\, c_2\, k_7 \\ k_2\, v_1\, k_4\, c_3 \end{pmatrix} \qquad (2)$$

In this representation, c_0, c_1, c_2, c_3 denote constants, k_0, k_1,..., k_7 represent a 256-bit key, and IV= (t_0, t_1, v_0, v_1) represents a 64-bit counter (t_0, t_1) and a 64-bit nonce (v_0, v_1). The constants may vary depending on the key size, with a 256-bit key indicating 256-bit Salsa and a 128-bit key implying 128-bit Salsa.

The fundamental operation within Salsa20 [13] is the quarterround function, which is a non-linear operation consisting of four ARX rounds each round comprises addition (A), cyclic left rotation (R), and XOR (X). Addition involves two words with the result modulo 232, rotation shifts bits cyclically, and XOR combines the two words while suppressing carries.

Rounds in Salsa20 can be executed based on matrix columns and rows. A "rowround" applies four quarterrounds to each row, while a "columnround" applies four quarterrounds to each column of the initial state matrix. For example:

- Rowround: quarterround (s_0, s_1, s_2, s_3), quarterround (s_5, s_6, s_7, s_4), quarterround (s_{10}, s_{11}, s_8, s_9), quarterround (s_{15}, s_{12}, s_{13}, s_{14})
- Columnround: quarterround (s_0, s_4, s_8, s_{12}), quarterround (s_5, s_9, s_{13}, s_1), quarterround (s_{10}, s_{14}, s_2, s_6), quarterround (s_{15}, s_3, s_7, s_{11})

ChaCha20 algorithm

ChaCha20[12] a variant of Salsa20, is a high-throughput stream cipher designed for software platforms. The ChaCha family encompasses varying cipher instantiations tailored to balance "security versus performance", as specified in the original ChaCha family specification. This specification allows for adjustments in key length, nonce length, counter length, and the number of rounds. Let (p_i, c_i) stand for the plaintext and cipher text's corresponding 64-byte blocks. Using a 256-bit key of K = (k_0, k_1, k_2, k_3, k_4, k_5, k_6, k_7), ChaCha20 operates on words of 32 bits each time. The outputs block is a 512-bit key stream in length. The structure is represented using the following equation:

$$s = \begin{pmatrix} s_0 & s_1 & s_2 & s_3 \\ s_4 & s_5 & s_6 & s_7 \\ s_8 & s_9 & s_{10} & s_{11} \\ s_{12} & s_{13} & s_{14} & s_{15} \end{pmatrix} = \begin{pmatrix} c_0 & c_1 & c_2 & c_3 \\ k_0 & k_1 & k_2 & k_3 \\ k_4 & k_5 & k_6 & k_7 \\ t_0 & v_0 & v_1 & v_2 \end{pmatrix} \qquad (3)$$

The initial state is encapsulated within the first matrix, denoted by (s_0, s_1, s_2,..., s_{15}), representing the individual words at the outset. Meanwhile, the second matrix features predefined constants, c_0, c_1, c_2, c_3, and the key k_0, k_1,..., k_7, constituting a 256-bit key. Additionally, the initialization vector (IV), depicted as (t_0,t_1 ,v_0,v_1), delineates a 64-bit counter (t_0, t_1) and a 64-bit nonce (v_0, v_1). As in Salsa columnround and rowround are considered, but here in ChaCha Columnround and diagonalround are considered. It is considered as for odd Rounds first columnround is applied and for even rounds first diagonalround is applied. Columnround and diagonalrounds are as follows:

- Column round: quarterround (s_0, s_4, s_8, s_{12}), quarter round (s_1, s_5, s_9, s_{13}), quarterround (s_2, s_6, s_{10}, s_{14}), quarter round (s_3, s_7, s_{11}, s_{15})
- Diagonal round: quarterround (s_0, s_5, s_{10}, s_{15}), quarterround (s_1, s_6, s_{11}, s_{12}), quarterround (s_2, s_7, s_8, s_{13}), quarterround (s_3, s_4, s_9, s_{14})

Camellia algorithm

Mitsubishi Electric Corporation and the Nippon Telegraph and Telephone Corporation developed camellia algorithm [14] in concert. With the same degree of security as AES, this algorithm is a powerful rival. The Camellia algorithm, which runs on 128-bit data blocks, supports three different key sizes: 128 bits, 192 bits, and 256 bits. Figure 2 illustrates the architecture of Camellia Encryption Algorithm.

RESULTS

In the following subsections, our discussion will unfold in two primary domains. Initially, we'll delve into the encryption and decryption performance of audio files using various algorithms and analyze their efficiency across different file sizes. This analysis will include assessing the performance of ChaCha20, Salsa20, and Camellia encryption algorithms.

Subsequently, we'll shift our focus to enhancing audio security through alternative techniques beyond encryption. This will involve exploring the utilization of Fourier transform and noise reduction methods to bolster the security of audio transmissions in IoT healthcare systems.

Encryption/decryption performance for audio 1, audio 2, and audio 3

The three encryption algorithms (ChaCha20, Salsa20 and Camellia) were applied to the aforementioned audio (1 of 3 s, 2 of 10 s, 120 s). The original audio signals, the encrypted audio signals and the decrypted audio signals were visualized using a single figure 2.

In addition, the x- and y-axes of the figure are labeled ("Time" for the x-axis and "Amplitude" for the y-axis). This operation concerns the ChaCha20 algorithm, which represents the same thing for the salsa20 and camellia algorithms, giving similar images for the audio studied. The encryption and decryption results are as follows:

Audio 1 (3 s)

- Amplitude of the original audio signal: Min = -8 587, Max = 9 002
- Amplitude of the encrypted audio signal: Min = -32 767, Max = 32 767
- Amplitude of the decrypted audio signal: Min = -8 587, Max = 9 002

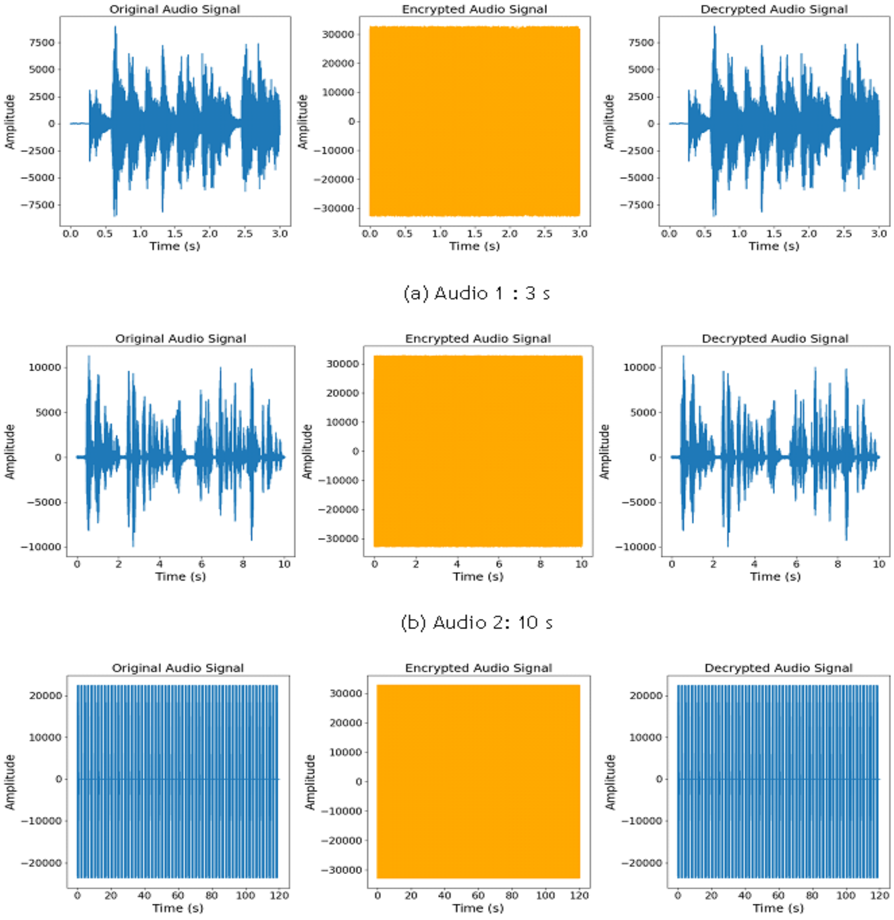
Audio 2 (10 s)

- Amplitude of the original audio signal: Min = -9 968, Max = 11 312
- Amplitude of the encrypted audio signal: Min = -32 768, Max = 32 767
- Amplitude of the decrypted audio signal: Min = -9 968, Max = 11 312

Audio 3 (120 s)

- Amplitude of the original audio signal: Min = -23 702, Max = 22 502
- Amplitude of the encrypted audio signal: Min = -32768, Max = 32 767
- Amplitude of the decrypted audio signal: Min = -23 702, Max = 22 502

Figure 3 presents a comparative analysis of the encryption and decryption times for audio samples of different durations: audio 1 lasting 3 seconds, audio 2 lasting 10 seconds and biomedical audio 3 lasting 120 seconds, respectively. In figure 3, when applied to audio 1, ChaCha20 appears as the most efficient algorithm, with encryption and decryption times of 0,001038 and 0,000971 seconds, respectively. Salsa20 follows with slightly slower times at 0,001329 seconds for encryption and 0,001104 seconds for decryption, while Camellia shows the slowest performance, with encryption and decryption times of 0,003607 and 0,00401 seconds, respectively. For Audio 2, ChaCha20 maintains its lead with encryption and decryption times of 0,00280 and 0,002818 seconds. Salsa20 follows with encryption and decryption times of 0,006191 and 0,00672 seconds, respectively, and Camellia shows the slowest performance, with encryption and decryption times of 0,009982 and 0,015021 seconds. the 120 s audio gives an overview of the encryption and decryption times. Salsa20 demonstrates an encryption time of 0,086522 seconds and a decryption time of 0,074040 seconds. Camellia has an encryption time of 0,157103 seconds and a decryption time of 0,160258 seconds. As for ChaCha20, it displays an encryption time of 0,063818 seconds and a decryption time of 0,059758 seconds. Despite variable audio durations, ChaCha20 consistently demonstrates superior performance in terms of encryption and decryption times over three digits, highlighting its effectiveness as a cryptographic solution for different sizes of audio. In terms of security, the three algorithms - Salsa20, ChaCha20 and Camellia - offer robust encryption capabilities, Camellia being particularly renowned for its strength close to the Advanced Encryption Standard (AES). However, the analysis highlights ChaCha20 as the fastest option, emphasizing the importance of selecting an algorithm that meets both speed and security requirements.



(a) Audio 1 : 3 s

(b) Audio 2: 10 s

(c) Audio 3: 120 s

Figure 2. Encryption/decryption of audio 1 for 3 s, audio 2 for 10 s, audio 3 for 120 s

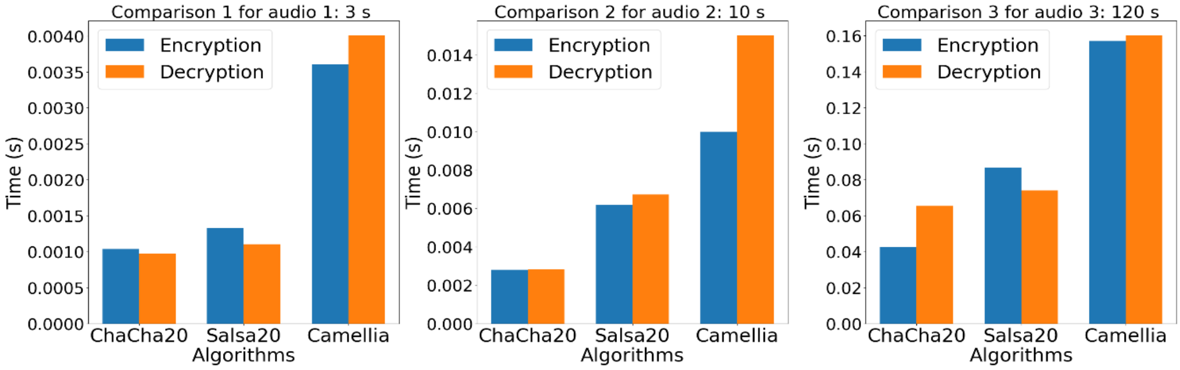


Figure 3. Encryption and Decryption results for different algorithms

Enhancing audio security through Fourier transform and noise reduction techniques

Improved audio security thanks to Fourier transform and noise reduction techniques.The use of the Fourier transform makes it possible to examine the frequency components of the original audio signals, encrypted and decrypted over different audio transmission durations (from 3 seconds; from 10 seconds and 120 seconds). This mathematical method makes it possible to identify the frequencies of the signal, thus improving the understanding of the audio content through visuals representation. Before encryption, the original audio files undergo noise reduction using the "noisereduce" package, aimed at eliminating unwanted noise and improving the quality of the audio signal for later analysis. These strategies are essential to ensure the security of audio information.[31] Noise reduction not only improves listening and analysis by eliminating disturbances, but also helps to preserve the confidentiality of the information contained in the audio files, thus reducing the risk of compromise of sensitive data. In addition, the effects of the Fourier transform include the analysis of frequency components, the detection of anomalies, the evaluation of security and the validation of integrity. This improves our understanding and assessment of the security of encrypted audio transfers. Figures 4 and 5 illustrate our proposed idea to improve audio security by using Fourier transform and noise reduction techniques. Each phase must be applied to audio 1 of 3 s, audio 2 of 10 s and audio 3 of 120 s. In our case, we applied these techniques only for audio 1 of 3 s. First the encryption when sending(sender): when transmitting an audio file, the encryption process is important. The sender encrypts the audio file using an encryption algorithm such as ChaCha20 before sending it to the recipient. For example, a 3-second audio file, a 10-second audio file, and a 120-second biomedical audio file would be encrypted to ensure their security during transit.In figure 4, the Fourier transforms of the three audio signals (original, encrypted and decrypted) have been obtained.

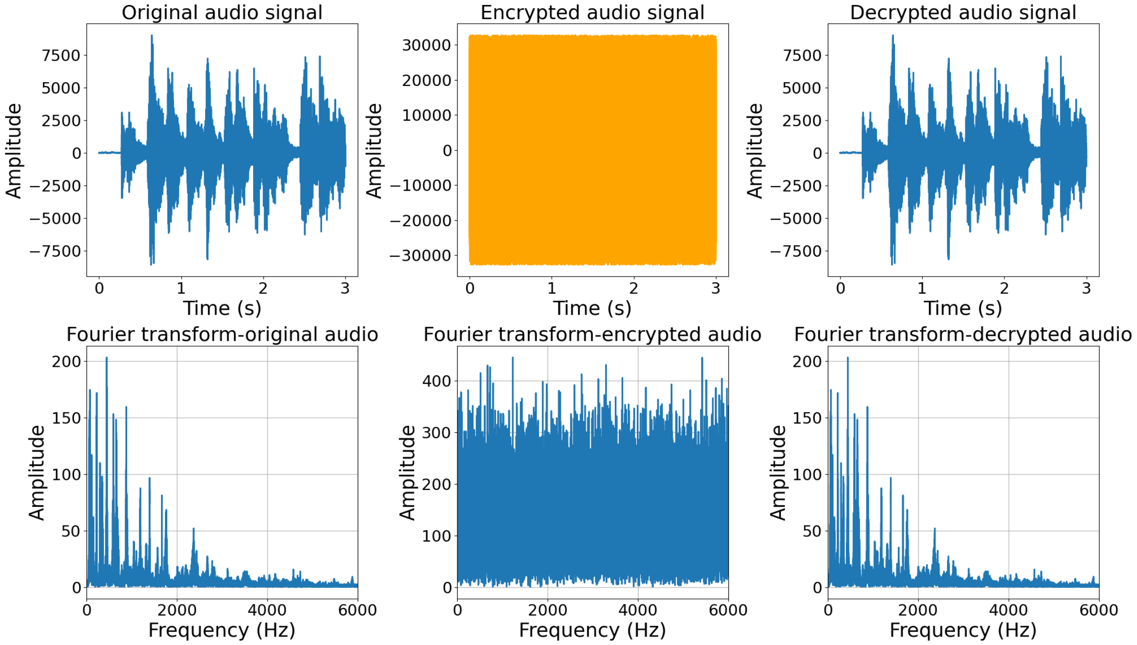


Figure 4. Fourier transform of audio 1 (original signal, encrypted signal, decrypted signal)

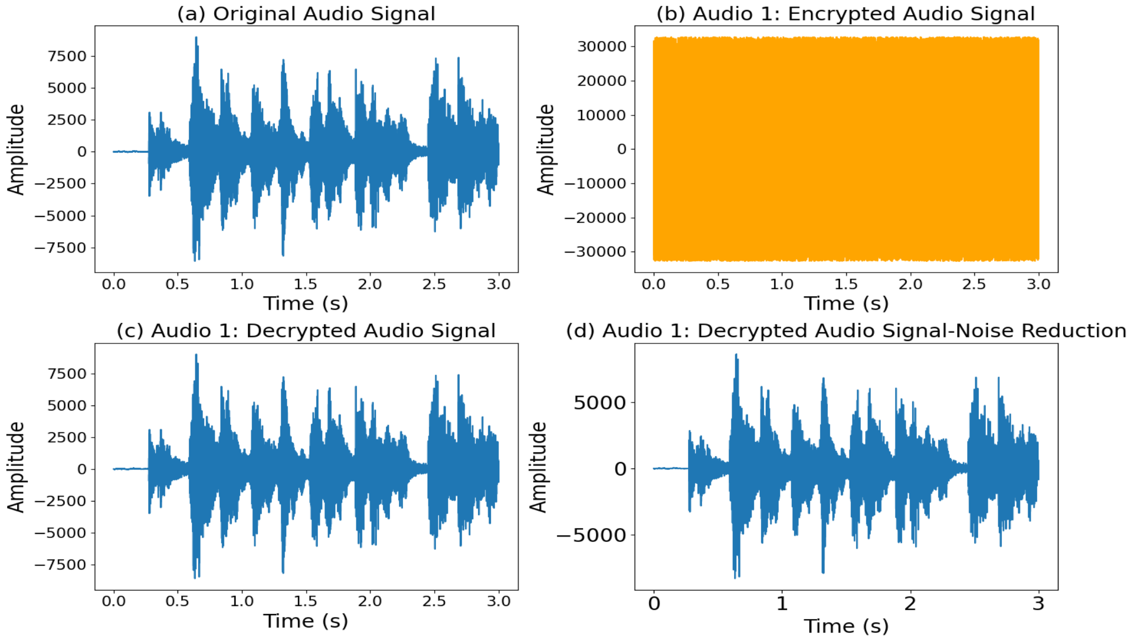


Figure 5. Noise reduction technology for better audio quality

In figure 5, the technique for reducing the noise of the decrypted audio signal has been applied. We have analyzed three algorithms previously in terms of processing time to evaluate the effectiveness and speed of various audio encryption and decryption methods in the context of the Internet of Things (IoT). We have found that ChaCha20 is the fastest algorithm while maintaining sufficient security. the images (a), (b), (c) and (d) of figure 5 which are respectively the original audio signal, the encrypted signal, the decrypted adio signal and the decrypted audio signal with noise reduction. We used the Fourier transform approach to analyze and transmit the audio data in the frequency domain after encrypting the audios using ChaCha20. This allowed for efficient compression and better bandwidth management, which is essential for IoT devices with limited resources.

The noise reduction approach was used throughout the decryption process to improve the quality of the audio signal by removing unwanted noise and artifacts. By combining these techniques, we have been able to improve the quality of the transmitted audio, guarantee safe and excellent communication between the sender and the recipient, and maximize the speed and efficiency of audio encryption and decryption. These methods are crucial for Internet of Things applications such as voice assistants, real-time communication systems and voice messaging devices integrated into connected elements, where speed, security and audio quality are important considerations. Thanks to the integration of these methodologies, we have been able to optimize the performance and reliability of communications between networks of connected devices, as well as guarantee safe and efficient audio transmission in an IoT (internet of things) context.

Prior to encryption, the audio stream was examined using the Fourier transform to determine each frequency component, enabling us to distinguish and isolate unwanted frequencies like background noise. [33,37] The audio quality was then enhanced before to encryption by using noise reduction techniques to stifle these undesirable frequencies. Once the audio stream was decrypted, we utilized the Fourier transform once more to confirm and modify its frequency components, so restoring its original quality. Lastly, to ensure that the final audio is clear and crisp for transmission or listening, we conducted noise reduction to remove any remaining noise that could have been created during the encryption and decryption operations.

DISCUSSION

Table 1 presents a range of studies that highlight ongoing efforts to address security challenges in audio data processing and transmission. Each study offers unique insights and methodologies tailored to specific aspects of audio security, spanning from steganalysis to cryptographic ransomware mitigation. One notable finding is the effectiveness of elliptic curve cryptography (ECC) compared to RSA in audio encryption, as demonstrated by Shelke et al.[9] ECC's ability to provide a smaller key size and larger key space while maintaining security underscores its suitability for securing audio data in various applications. Additionally, Mcuba et al.[4] emphasize the importance of robust classification architectures in distinguishing between original and cloned audio features. The high accuracy achieved with the VGG-16 architecture highlights the significance of advanced machine learning techniques in audio security. Yazdanpanah et al.[1] focus on steganalysis in speech signals, employing Support Vector Machine (SVM) with a Gaussian membership function (GMF) for detection. Their findings demonstrate high specificity (81,2 %) and sensitivity values for different embedding ratios. Furthermore, Pleshkova et al.[17] explore the use of Public Key Infrastructure (PKI) for secure audio information transmission. Although their study does not specify a dataset, they evaluate the decrypted audio quality using objective (SNRseg) and subjective measures, emphasizing the importance of audio fidelity in secure transmission protocols. Moreover, Babiano et al.[18] highlight the evolving nature of ransomware threats and advocate for innovative mitigation strategies. Their approach, which successfully recovers Salsa20 keys from volatile memory, offers a promising avenue for countering ransomware attacks and mitigating their impact on organizations and individuals. Our approach combines noise reduction and Fourier transform with the ChaCha20 encryption algorithm for audio processing. We applied this technique to three different audio samples with varying lengths and dimensions:

- Audio 1 (3 seconds, size: 1058,75 kilobits) Encryption Time: 0,001038 seconds Decryption Time: 0,000971 seconds
- Audio 2 (10 seconds, size: 3529,44 kilobits) Encryption Time: 0,002800 seconds Decryption Time: 0,002818 seconds
- Audio 3 (120 seconds, size: 3529,44 kilobits) Encryption Time: 0,063818 seconds Decryption Time: 0,059758 seconds

ChaCha20 effectively encrypts audio data and decrypts it even more quickly. Naturally, encryption takes slightly longer for larger data sizes. By leveraging Fourier transform and noise reduction techniques, we enhance the efficiency and clarity of transmitted audio. This strategy is particularly advantageous in the Internet of Things (IoT) domain, where rapid and high-quality data transfer is critical.

Our study focuses on assessing the security of biomedical and non-biomedical audio data in IoT healthcare systems using the ChaCha20, Salsa20, and Camellia algorithms.

With encryption and decryption times of 0,001038 and 0,00097 seconds for Audio 1 and 0,002800 and 0,002818 seconds for Audio 2, ChaCha20 showed higher efficiency. The longer biological Audio 3 (120 seconds, 3529,44 kilobits) took 0,059758 seconds to decrypt and 0,063818 seconds for ChaCha20 to encrypt. These outcomes demonstrate how well ChaCha20 protects sensitive medical data. The length and volume of audio files affect Fourier and noise reduction processing, which affects overall application efficiency in healthcare.

Table 1. Comparison of our approach with state-of-the-art methods in audio steganalysis and cryptography security techniques

Authors	Type	Best Method	Dataset	Results
Yazdanpanah et al.[1]	Speech	Percentage of Equal Adjacent Samples (PEAS) speech steganalysis with Gaussian member-ship function (GMF)	18000 noisy and noise-free speech instances	Specificity : 81,2 %, Sensitivity : 78,36 % for 12,5 %, 81,4 % for 25 %, 93,74 % for 37,5 %, and 99,82 % for 50 %.
Shelke et al.[9]	Audio	Modified elliptic cryptography and the Arnold transform	Audio signals 4148 samples	When compared to RSA, elliptic curve cryptography is discovered to be more fitting for encryption
Pleshkova et al.[17]	Audio	Public Key Infrastructure (PKI) and the modular exponentiation algorithm		(SNRseg) and subjective measures of sound similarity between original and decrypted audio
Mcuba et al.[4]	Speech	MFCC with VGG-16	Baidu Silicon Valley AI Lab dataset	Test accuracy of 86,906 %
Babiano et al.[18]	Audio	Salsa20		The article's methods successfully recover Salsa20 keys from volatile memory
Our approach	Audio	ChaCha20 with Fourier transform and noise reduction techniques	Audio 1 of 3s and size 1058,75 kilobits Audio 2 of 10s and size 3529,44 kilobits Audio 3 of 120s and size 3529,44 kilobits	Encryption Time (s): 0,001038 Decryption Time (s): 0,000971 Encryption Time (s): 0,002800 Decryption Time (s): 0,002818 Encryption Time (s): 0,063818 Decryption Time (s): 0,059758

CONCLUSIONS

In conclusion, ensuring the security of shared data is still important as communication technologies advance. In order to increase the security of biomedical audio communications in Internet of Things healthcare systems, our study examined a number of strategies. After comparing many encryption algorithms, including Camellia, Salsa20, and Chacha20, we were able to determine that Chacha20 is the most effective since it can encrypt and decode audio data quickly while also protecting privacy. Furthermore, we have investigated complementary techniques like noise reduction, which are important to enhancing the recording quality and obtaining important information from the audio signals. We thoroughly examined the audio recordings based on the Fourier transform analysis, which enabled us to spot changes, identify noise attacks, and unearth buried information. Our research emphasizes how important it will be to implement cutting-edge technology in order to adequately safeguard audio data. The Chacha20 method is a viable option for encrypting audio recordings. We may confidently traverse the changing digital world by incorporating these strategies into our communication habits and knowing that our audio data is safe and secure in IoT healthcare systems. Future research will focus on advancing security and privacy in IoT healthcare audio communications through the exploration of novel encryption methods and innovative techniques. This involves refining existing encryption techniques to ensure robust data security without compromising audio quality.

BIBLIOGRAPHIC REFERENCES

1. Yazdanpanah S, Chaeikar SS, Jolfaei A. Monitoring the security of audio biomedical signals communications in wearable IoT healthcare. Digital Communications and Networks. 2022; S2352864822002437. doi: 10.1016/j.dcan.2022.11.002.

2. Lin C-H, Hu G-H, Chen J-S, Yan J-J, Tang K-H. Novel design of cryptosystems for video/audio streaming via dynamic synchronized chaos-based random keys. Multimedia Systems. 2022;28(5):1793-1808. doi: 10.1007/s00530-022-00950-6.

3. Chen Y, et al. SoK: A Modularized Approach to Study the Security of Automatic Speech Recognition Systems. arXiv. 2021. doi.org/10.48550/arXiv.2103.10651

4. Mcuba M, Singh A, Ikuesan RA, Venter H. The Effect of Deep Learning Methods on Deepfake Audio Detection for Digital Investigation. Procedia Computer Science. 2023;219:211-219. doi: 10.1016/j.procs.2023.01.283.

5. Lasri I, Riadsolh A, Elbelkacemi M. Facial emotion recognition of deaf and hard-of-hearing students for engagement detection using deep learning. Education and Information Technologies. 2023;28(4):4069-4092. doi.org/10.1007/s10639-022-11370-4

6. Mawalim CO, Titalim BA, Okada S, Unoki M. Non-intrusive speech intelligibility prediction using an auditory periphery model with hearing loss. Applied Acoustics. 2023;214:109663. doi: 10.1016/j.apacoust.2023.109663.

7. Hazzaa F, Shabut AM, Ali NHM, Cirstea M. Security Scheme Enhancement for Voice over Wireless Networks. Journal of Information Security and Applications. 2021;58:102798. doi: 10.1016/j.jisa.2021.102798.

8. Castillo VS. Analysis of the scientific production on the implementation of artificial intelligence in precision agriculture. LatIA 2023;1:1-1. https://doi.org/10.62486/latia20231.

9. Alsabhany AA, Ridzuan F, Azni AH. The Progressive Multilevel Embedding Method for Audio Steganography. J. Phys.: Conf. Ser. 2020;1551(1):012011. doi: 10.1088/1742-6596/1551/1/012011.

10. Shelke R, Nemade M. Audio Encryption Algorithm Using Modified Elliptical Curve Cryptography and Arnold Transform for Audio Watermarking. In: 2018 3rd International Conference for Convergence in Technology (I2CT). IEEE; 2018. p. 1-4. doi: 10.1109/I2CT.2018.8529329.

11. Gu Z, Liu Y. Scalable Group Audio-Based Authentication Scheme for IoT Devices. In: 2016 12th International Conference on Computational Intelligence and Security (CIS). IEEE; 2016. p. 277-281. doi: 10.1109/CIS.2016.0070.

12. Singh DHP, Mettu DL, Kuchipudi VB. Analytical Approaches for Voice Recognition: Security Oriented Techniques in Cyber Attacks. JOURNAL OF CRITICAL REVIEWS. 2020;7(19).

13. Bernstein DJ. ChaCha, a variant of Salsa20. 2008. Disponible sur: https://cr.yp.to/chacha/chacha-20080128.pdf (Consulté le 16 décembre 2023).

14. Bernstein DJ. The Salsa20 family of stream ciphers. 2007. Disponible sur: http://cr.yp.to/papers.html#salsafamily (Consulté le 16 décembre 2023). doi: 10.1007/978-3-540-68351-3_8

15. Aoki K, Ichikawa T, Kanda M, Matsui M, Moriai S. Camellia: A 128-Bit Block Cipher Suitable for Multiple Platforms — Design and Analysis. Selected Areas in Cryptography 2000. pp. 39-56. doi: 10.1007/3-540-44983-3_4

16. Bracewell RN. The Fourier transform and its applications. New York: McGraw-Hill; 1986. doi:10.2307/2314845.

17. Abdallah HA, Meshoul S. A Multilayered Audio Signal Encryption Approach for Secure Voice Communication. Electronics. 2022;12(1):2. doi: 10.3390/electronics12010002.

18. Gamboa AJP, Díaz-Guerra DD. Artificial Intelligence for the development of qualitative studies. LatIA 2023;1:4-4. https://doi.org/10.62486/latia20234.

19. Pleshkova S, Kinanev D, Bekiarski A. Secure Audio Information Transmission with Encryption Algorithms in PKI. In: 2018 International Conference on High Technology for Sustainable Development (HiTech). IEEE; 2018. p. 1-4. doi: 10.1109/HiTech.2018.8566659.

20. Fernandez De Loaysa Babiano L, Macfarlane R, Davies SR. Evaluation of live forensic techniques, towards Salsa20-Based cryptographic ransomware mitigation. Forensic Science International: Digital Investigation. 2023;46:301572. doi: 10.1016/j.fsidi.2023.301572.

21. Medeiros VN, Silvestre B, Borges VCM. Multi-objective routing aware of mixed IoT traffic for low-cost wireless Backhauls. J Internet Serv Appl. 2019;10(1):9. doi: 10.1186/s13174-019-0108-9.

22. Kubilay MY, Kiraz MS, Mantar HA. KORGAN: An Efficient PKI Architecture Based on PBFT Through Dynamic Threshold Signatures. The Computer Journal. 2021;64(4):564-574. doi: 10.1093/comjnl/bxaa081.

23. Shaaban OA, Yildirim R, Alguttar AA. Audio Deepfake Approaches. IEEE Access. 2023;11:132652-132682. doi: 10.1109/ACCESS.2023.3333866.

24. Chavez-Cano AM. Artificial Intelligence Applied to Telemedicine: opportunities for healthcare delivery in rural areas. LatIA 2023;1:3-3. https://doi.org/10.62486/latia20233.

25. Bansal V, Pahwa G, Kannan N. Cough Classification for COVID-19 based on audio mfcc features using Convolutional Neural Networks. In: 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON). IEEE; 2020. pp. 604-608. doi: 10.1109/GUCON48875.2020.9231094.

26. Seo S, Kim C, Kim J-H. Convolutional Neural Networks Using Log Mel-Spectrogram Separation for Audio Event Classification with Unknown Devices. JWE. 2022. doi: 10.13052/jwe1540-9589.21216.

27. Sudha V, Ganeshbabu TR. A Convolutional Neural Network Classifier VGG-19 Architecture for Lesion Detection and Grading in Diabetic Retinopathy Based on Deep Learning. Computers, Materials & Continua. 2020;66(1):827-842. doi: 10.32604/cmc.2020.012008.

28. Lin Y-J, et al. Artificial Intelligence of Things Wearable System for Cardiac Disease Detection. In: 2019 IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS). IEEE; 2019. pp. 67-70. doi: 10.1109/AICAS.2019.8771630.

29. Hameed ME, Ibrahim MM, Manap NA, Mohammed AA. A lossless compression and encryption mechanism for remote monitoring of ECG data using Huffman coding and CBC-AES. Future Generation Computer Systems. 2020;111:829-840. doi: 10.1016/j.future.2019.10.010.

30. Lima PM, Da Silva CKP, De Farias CM, Carvalho LK, Moreira MV. Event-based cryptography for automation networks of cyber-physical systems using the stream cipher ChaCha20. IFAC-PapersOnLine. 2022;55(28):58-65. doi: 10.1016/j.ifacol.2022.10.324.

31. Rathod C, Advani N, Gonsai A. Comparative Analysis of Encryption and Decryption Algorithms for Audio. 2018. doi: 10.9756/IAJCS/V10I1/1810098.

32. Cano CAG, Troya ALC. Artificial Intelligence applied to teaching and learning processes. LatIA 2023;1:2-2. https://doi.org/10.62486/latia20232.

33. CS 101- Sample Sound Files. Disponible sur: https://www2.cs.uic.edu/~i101/SoundFiles/ (Consulté le 8 mars 2024). https://www.soundjay.com/human/sounds/heartbeat-01a.mp3

34. Shen M, Tang Z. Audio Signal and Troubleshooting System Based on Wireless Sensor. International Journal of Online and Biomedical Engineering (iJOE). 2018;14(06):113-125. doi: 10.3991/ijoe.v14i06.8702.

35. Telegram N, Sahu PC, Panda S, Kandasamy N. USRP Based Digital Audio Broadcasting Using OFDM in Virtual and Remote Laboratory. International Journal of Online and Biomedical Engineering (iJOE). 2019;15(13):77-85. doi: 10.3991/ijoe.v15i13.8761.

36. Willems C, Meinel C. Tele-Lab IT-Security: an Architecture for an online virtual IT Security Lab. International Journal of Online and Biomedical Engineering (iJOE). 2008;4(2):31-37. doi: 10.3991/ijoe.v4i2.497.

37. Panyapanuwat P, Kamonsantiroj S, Pipanmaekaporn L. Similarity-preserving hash for content-based audio retrieval using unsupervised deep neural networks. International Journal of Electrical and Computer Engineering. 2021;11(1):879. DOI: 10.11591/ijece.v11i1.pp879-891.

38. Maradithaya S, Katti A. Sentimental analysis of audio based customer reviews without textual conversion. International Journal of Electrical and Computer Engineering (IJECE). 2024;14(1):653-661. DOI: http://doi.org/10.11591/ijece.v14i1.pp653-66.1

39. Khalifa Boutahir M, Hessane A, Lasri I, Benchikh S, Farhaoui Y, Azrour M. Dynamic Threshold Fine-Tuning in Anomaly Severity Classification for Enhanced Solar Power Optimization. Data and Metadata. 2023; 2:94-94. DOI: https://doi.org/10.56294/dm202394.

40. Lasri I, Riadsolh A, El Belkacemi M. Toward an effective analysis of COVID-19 Moroccan business survey data using machine learning techniques. In: the 13th International Conference on Machine Learning and Computing. 2021. pp. 50-58. DOI: 10.1145/3457682.3457690.

FINANCING

The authors did not receive financing for the development of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Mohammed Amraoui, Imane Lasri, Fouzia Omary, Mohamed Khalifa Boutahir, Yousef Farhaoui.
Research: Mohammed Amraoui, Imane Lasri, Fouzia Omary, Mohamed Khalifa Boutahir, Yousef Farhaoui.
Methodology: Mohammed Amraoui, Imane Lasri.
Drafting - original draft: Mohammed Amraoui, Imane Lasri.
Writing - proofreading and editing: Mohammed Amraoui, Imane Lasri.

https://doi.org/10.56294/dm2024365