MACSteg: Real-Time Voice Authentication and Deepfake Protection Using Device MAC Address Steganography

Sanaa Ahmed Kadhim[1,*], Zaid Ali Alsarray[1], Saad Abdual Azize Abdual Rahman[2], Massila Kamalrudin[3], Mustafa Musa[4]

[1]University of Information Technology and Communications, Biomedical Informatics College, 10066, Baghdad, Iraq

[2]Department of Computer Engineering Techniques, Almamoon University Collage, Baghdad, Iraq

[3]Universiti Teknikal Malaysia, Faculty of Information and Communication Technology, Melaka, Malaysia

[4]Universiti Teknikal Malaysia, Center of Research and Innovation Management, Melaka, Malaysia

Emails: dr.sanaa.ahmed@uoitc.edu.iq; zaid.ali-bmic@uoitc.edu.iq; Saad.a.azize@almamonuc.edu.iq; massila@utem.edu.my; mustafamusa@utem.edu.my

Abstract

The invention of deepfake applications make it possible to produce highly natural and real voice recordings which creates critical concerns about the credibility of audio telecommunications. The confirmation of the speakers' voices became essential especially for sensitive data such as financial, healthcare, and surveillance risk management services, authentication of speakers' voices became significantly crucial. To improve solutions to this issue, this paper presents MACSteg strategy which is a real-time, lightweight voice authentication technique by discreetly encapsulate device's MAC address within voice file using Quantization Index Modulation (QIM) stego-technique. Unlike many traditional strategies that degrade voice quality or produces noticed jitter and MACSteg technique preserve both clarity and efficiency. Implementations showed that the hidden MAC address stayed intact in spite of some typical voice processing such as compression, while interfered signals reformed by clatter or volume variations were consistently detected. The proposed system obtained a high signal-to-noise ratio (SNR) exceeding 70 dB. The illustrating that the alterations were inaudible, and maintained well in real-time submissions, giving only a processing delay of 0.01 milliseconds per each audio segment, this results indicate MACSteg's potential as a ascendable and effective approach for safeguarding voice authenticity, especially in circumstances where verification of speaker's voice is vital.

Keywords: Voice authentication; Steganography: QIM; MAC address hiding; Deepfake prevention; Secure audio communications

1. Introduction

Increasingly improvements of AI applications with deep learning make easy to generate human voices with high degree of reality and accuracy compared to the original speech. The techniques of "deepfake audios" can regenerate natural voice cadence, verbal attitude and even expressive gradation with remarkable precision. For these reasons, shielding substantial voice recordings from alterations became more convoluted, in this matter in essential sectors in which audio is basically used for authentication or for delicate communication like financial, governmental, security and remotely healthcare applications in which the misappropriation use of artificial audio can lead to identity burglary, data leakages, or misinformation spreading [1], [2].

Defense audio content. Many researchers worked on various systems over time, such as digital audio watermarking and crypto-systems with encryption, in spite of the protection obtained from these techniques they also have trade-

DOI: https://doi.org/10.54216/JCIM.170202
Received: March 30, 2025 Revised: June 20, 2025 Accepted: August 08, 2025

offs. for instance the degradation of audio quality caused by aggressive watermarks, and the delay caused by encryption or even decryption before playback which make it unsuitable for real-time applications [3], [4], [6]. These conciliations between audio integrity, time consuming, and security keeps the challenging for designers and developers of such systems [7].

And because of these challenges. This paper presents a real-time, audio authentication system "MACSteg" by embedding a device Mac address within the transmitted voice stream directly using Quantization Index Modulation (QIM) stego-technique, by declaiming MAC address as a inimitable device identifier, the proposed method empowers source confirmation without disturbing the clarity or sensitivity of the voice. QIM slightly modifies audio signals with significant data, which are inaudible to human ears although, retrival of embedded data is allowable. This approach consents for ongoing conversation verification and could indicate mistrustful modifications caused by altering or deepfake voice intrusion. For getting there, MACSteg introduces workable key answers for the trad-of between voice podcasting in real-time and speaker authenticity [8-11]. This study introduces many essential contributions:

- Development of MACSteg, a novel real-time audio authentication system that embeds the MAC address of the transmitting device into the voice stream using Quantization Index Modulation (QIM) steganography.
- Hardware-level source verification, enabling traceability and tamper detection based on unique MAC identifiers without requiring user-side encryption keys.
- High voice quality, getting signal-to-noise ratio (SNR) more than 70 dB and peak signal-to-noise ratio (PSNR) aver 80 dB. The results obtained from the system reveal the encapsulation of certain data does not interfere speech quality neither its clarity. The transmitted audio was kept comprehensible through many varied conditions.
- Robustness against manipulation, where common attacks such as noise addition, amplitude scaling, compression, and time stretching were either detected or successfully resisted by the system.
- Real-time performance validation, showing that MACSteg operates efficiently with negligible processing delay (0.01 ms per audio chunk), making it suitable for live audio applications.
- System's lightweight design makes it ideal for real-world deployment, including on low-resource platforms such as IoT devices and mobile systems and even the VoIP telecommunications with no need for any specific hardware.

The rest of the study organized as follows: Section 2 summarizes the previous researches of audio authentication and deep fake prevention using steganography. Reviewing the existent approaches showing their gains and drawbacks, also illustrating the gaps in their works that this study intends to solve. Section 3 introduces the system's methodology, showing the steps of embedding MAC address within the voice by QIM stego-technique, in addition to the technique of data retrieving and authentication. Section 4 illustrates the experimental implementation and the founding results corresponding to voice quality, forcefulness, and the performing of the proposed system. At last, Section 5 shows the conclusions came out from the main results; track the limitations, and presents suitable recommendations for future developments for the suggested MACSteg system.

2. Related Work

Safeguarding the secrecy of transmitted audio files has a highest priority for the susceptible areas like confidential, governmental, financial, and health care. For many years, many strategies were introduced to improve voice authenticity, such as digital audio watermarking and cryptography, and as newly used methods, biometrical verifications and methods depending on hardware factors [12-15]. Recently works concerned with conventional watermarking methods such as Least Significant Bit (LSB) or echo hiding, and phase coding, cryptographic techniques such as (RSA, and AES), and hybrid techniques for audio encryption [16- 21]. These techniques have the advantage of implementation simplicity, although, they have disadvantages in the robustness, especially if the voice exposed to compression or signal interfering. Lately, many researches preferred the use of cryptography methods (AES and RSA) to shield digital audio [3]. In spite of getting robust security from those methods, they are complicated and time consuming which make them inadequate for real-time applications, particularly for low resource abilities devices. Some studies suggested merged models balance the security and the performance, while few other researchers have worked on the challenges modeled by low-latency transformation systems.

One of the auspicious solutions that gives a good trad-of between the efficiency and simplicity is the use of Quantization Index Modulation (QIM) [5], [6]. It consents the embedding of some data within audio files causing a minimal preselected modification. Studies like [6] and [7], suggested that, the use of QIM steganography could reserve the voice quality while transforming audio signals. At the same time, authentication using hardware identifiers was implemented when using IoT devices and vehicles connections [10], [11]. These studies depend on some identifiers related to hardware devices to be used as a verification factor for endpoint communications.

DOI: https://doi.org/10.54216/JCIM.170202
Received: March 30, 2025 Revised: June 20, 2025 Accepted: August 08, 2025

Although, employing this perception in verifying transmitted audio still relatively unmapped. Using the idea of embedding hardware identifiers within digital voice streams beside steganography may release innovative potentials for securing and authorizing real-time voices. In parallel, the limitations and complications of preventing voice for deepfake strategies courage researchers for improving new systems and methods. System using biometric voice features, deep learning machines, and generative adversarial networks (GANs) were used to authorize speeches effectively [14–22]. However, many of these methods and systems are sensitive and implementation AL exhaustive. Some researchers suggested the embedding of specific data to be used for speech verification in real-time systems during communications.

3. Methodology

MACSteg is developed and implemented especially for audio transmitted in real-time, hopping to gain a low latency and solid originality authentication. The main idea of the proposed system is to embed the unique address, Media Access Control (MAC), as a hardware signature of the origin device used in transmitting speech immediately within audio segments. This can be done with the use of Quantization Index Modulation (QIM) as an embedding, stego-technique that works properly keeping the balance between robustness and voice quality.

The proposed method works through Three-stage workflow: getting MAC address to be considered as a device identifier, embedding the identifier within audio signal (sender side), then, verify this identifier by the authorization party (receiver side). Firstly the MAC address of the sender device is captured and changed from 6-digit Hexa to a binary format of 48-bit. This stream of bits will be spread at multiple segments of audio QIM embedding technique. For verification and the receiver will extract the stream of spread bits, regenerate the hidden address, then compare the regenerated identifier with the original of the audio source.

The key aspect of this method converges the ability of unceasingly authenticate audio source keeping the original sound clarity and quality, it consumes only a trifling computational resource; make it as considerable method appropriate for time-sensetive systems. MACSteg is particularly well suited for secure voice-over-IP (VoIP) calls, encrypted conferencing systems and defense-related communications contexts where verifying the speaker's identity is essential. The entire process is summarized in Figure 1.

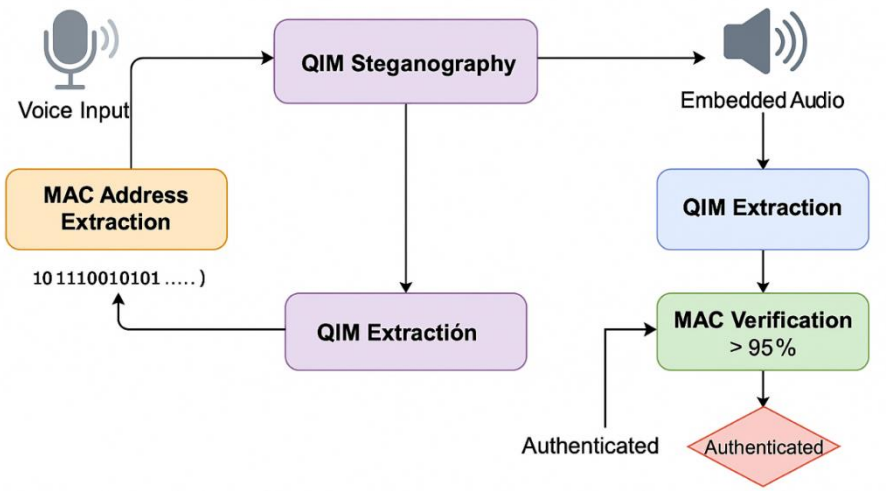

Figure 1. Proposed MACSteg system

3.1. MAC Address Embedding

The system first captures the MAC address of the broadcasting device and converts it to a 48-bit binary sequence. Audio is segmented into frames, and QIM is used to embed each bit at defined intervals.

QIM uses two quantizes:

- $Q0(x) = \Delta \cdot \text{round}(x / \Delta)$
- $Q1(x) = \Delta \cdot \text{round}((x - \Delta/2) / \Delta) + \Delta/2$

Where Δ is the quantization step size. The choice of quantizer depends on the bit to embed.

DOI: https://doi.org/10.54216/JCIM.170202
Received: March 30, 2025 Revised: June 20, 2025 Accepted: August 08, 2025

3.2. Extraction and Verification

The receiver extracts bits from the predefined positions using inverse QIM and reconstructs the MAC address. A match ratio (>95%) against the original is considered authentic.

3.3. Real-Time Simulation

The system processes audio in 100ms chunks. Embedding is distributed across chunks to maintain flow and support latency-sensitive scenarios like VoIP.

4. Experimental Results

4.1. Setup

Experiments were conducted in Google Colab using synthetic speech-like signals. Evaluation metrics included SNR, PSNR, MSE, and bit-match accuracy.

Libraries and tools used were successfully installed:

- librosa, soundfile, numpy, matplotlib, scipy, uuid, getmac, and dependencies.

4.2. Quality and Robustness

To assess the system's robustness against manipulation, various signal-level attacks were simulated. While compression maintained a perfect 100% match of the embedded MAC address, other manipulations such as noise injection, amplitude scaling, and time stretching were effectively detected and flagged as tampered. The small difference values, including a maximum difference of approximately 0.0049 and a mean difference close to zero, further confirm that the embedded data introduces only minor, statistically negligible changes to the audio waveform, as shown in Table 1.

Table 1: Experimental Results

Metric	**QIM Δ = 0.005**	**QIM Δ = 0.01**	**Interpretation**
SNR (dB)	87.13	66.19	High values = minimal distortion
PSNR (dB)	>90	80.77	Preserved peak signal quality
MSE	≈0	≈0	Low error between original & modified
Max Difference	0.004942	0.004942	Indicates minor sample-level changes
Mean Difference	0.000002	0.000002	Extremely small signal shift

4.3. Real-Time Capability

The real-time processing capability of the MACSteg system was evaluated by measuring the execution time for audio chunks of 100 milliseconds each. Results indicate that the average and maximum processing times were both maintained at just 0.01 milliseconds per chunk. This remarkably low computational demand demonstrates that the system can handle streaming audio without introducing any perceivable delay, making it well-suited for real-time applications such as secure VoIP communications. The system achieved a real-time factor exceeding 19,000×, meaning it operates over 19,000 times faster than the duration of the audio being processed. Such performance confirms the system's ability to function on low-power or edge devices while preserving responsiveness. These findings validate MACSteg as a viable solution for latency-sensitive environments where speed and security must coexist without compromise as shown in Table 2.

Table 2: Real-Time Performance Summary

Parameter	Value
Chunk Duration	100 ms
Average Processing Time	0.01 ms
Max Processing Time	0.01 ms
Real-Time Factor	19021.8x
Real-Time Capable	Yes

DOI: https://doi.org/10.54216/JCIM.170202
Received: March 30, 2025 Revised: June 20, 2025 Accepted: August 08, 2025

4.4. Visual Results and Discussion

To further validate the transparency and robustness of the proposed MACSteg approach, a series of visual analyses were conducted. These plots demonstrate the comparative integrity between the original and modified audio signals. Each subfigure presents a different aspect of the signal transformation, offering a comprehensive view of the steganographic impact.

4.4.1 MACSteg: Visual Analysis of Audio Authentication, Figure 2-7 below consolidates these findings.

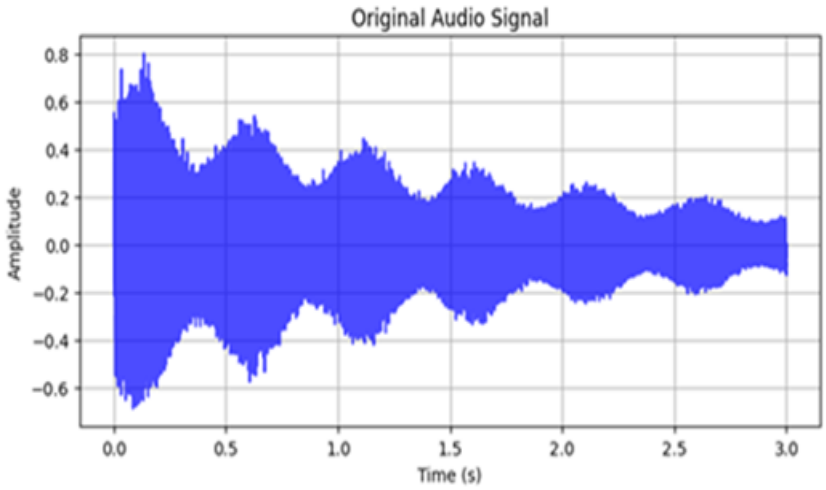

Figure 2. Original audio signal

This plot shows the waveform of the original audio without any modifications. The amplitude variation over time gives insight into how the natural voice signal fluctuates, preserving clarity and original characteristics. This baseline is crucial for comparison after embedding.

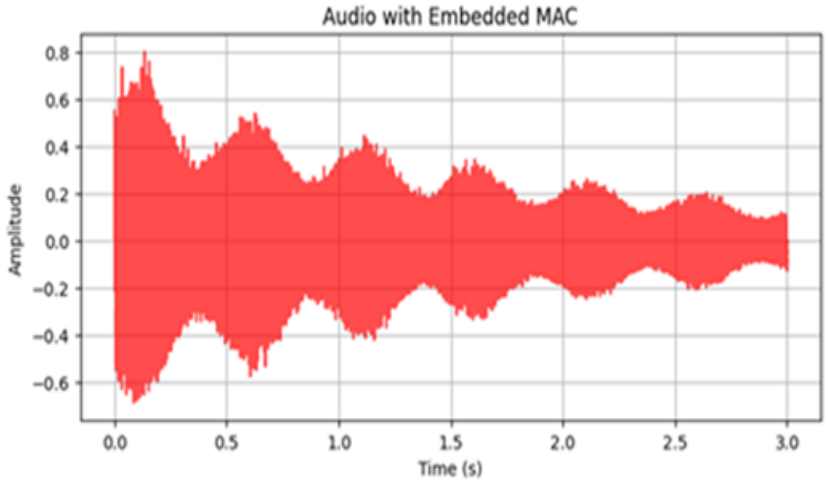

Figure 3. Audio with Embedded MAC

Here, we see the same audio signal after the MAC address has been embedded using QIM steganography. Visually, it looks almost identical to the original, which indicates that the process introduces minimal perceptual changes. This supports the goal of imperceptibility in voice security.

DOI: https://doi.org/10.54216/JCIM.170202
Received: March 30, 2025 Revised: June 20, 2025 Accepted: August 08, 2025

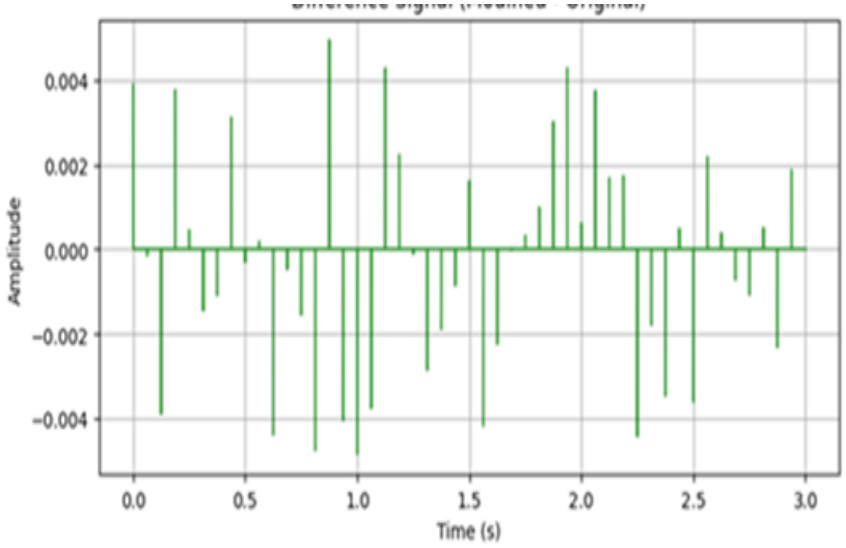

Figure 4. Difference Signal (Modified - Original)

This plot highlights the tiny changes introduced into the audio by subtracting the original signal from the modified one. The very small amplitude spikes demonstrate that the embedding process made nearly inaudible alterations, confirming the transparency of the method.

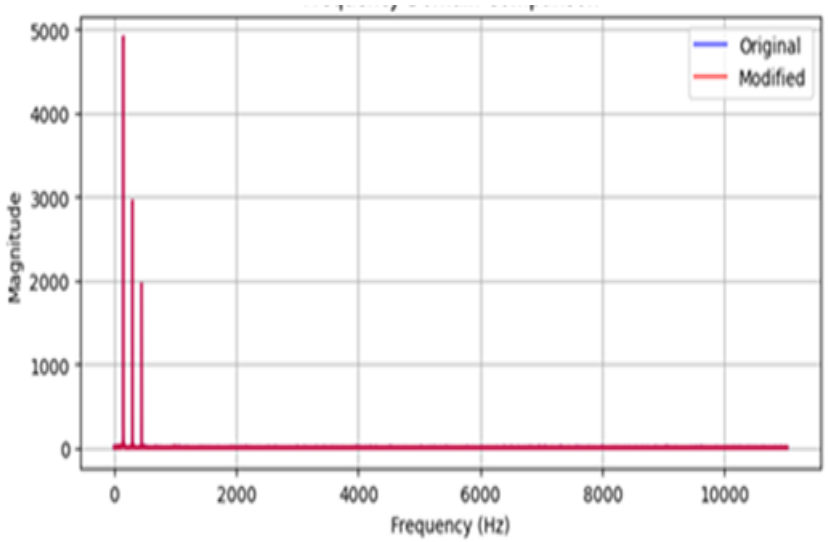

Figure 5. Frequency Domain Comparison

This graph compares how the original and modified signals look in the frequency domain. Both lines overlap significantly, meaning the spectral properties of the audio are preserved. This ensures that the voice still sounds natural even after embedding the MAC address.

DOI: https://doi.org/10.54216/JCIM.170202
Received: March 30, 2025 Revised: June 20, 2025 Accepted: August 08, 2025

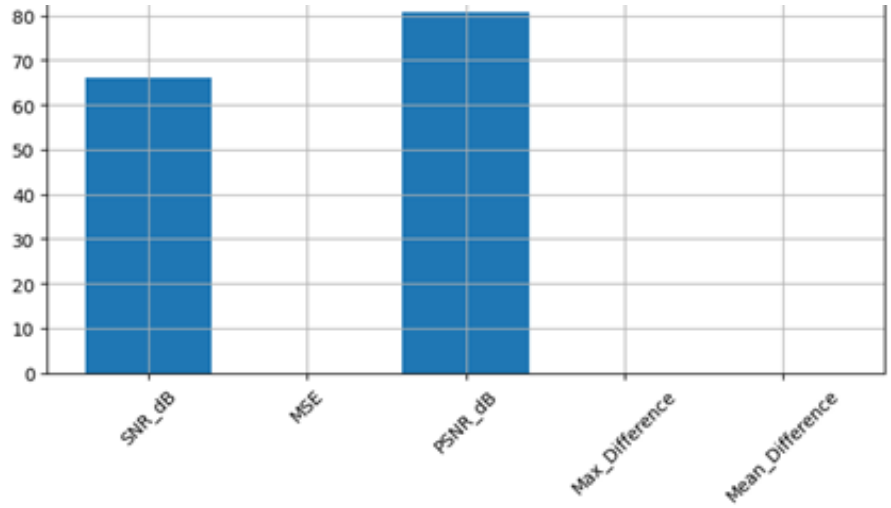

Figure 6. Audio Quality Metrics

Figure 6 illustrates the essential factors that indicates digital voice quality after embedding MAC identifier: MSE (mean square error) is almost zero while SNR (Signal-to-Noise Ratio) and PSNR (Peak SNR) showed high values (>70, > 80 respectively), meaning the audio did not lose its clarity or suffer from distortion. Max Difference and Mean Difference had values near zero indicating high similarity between the original and the stego-audio.

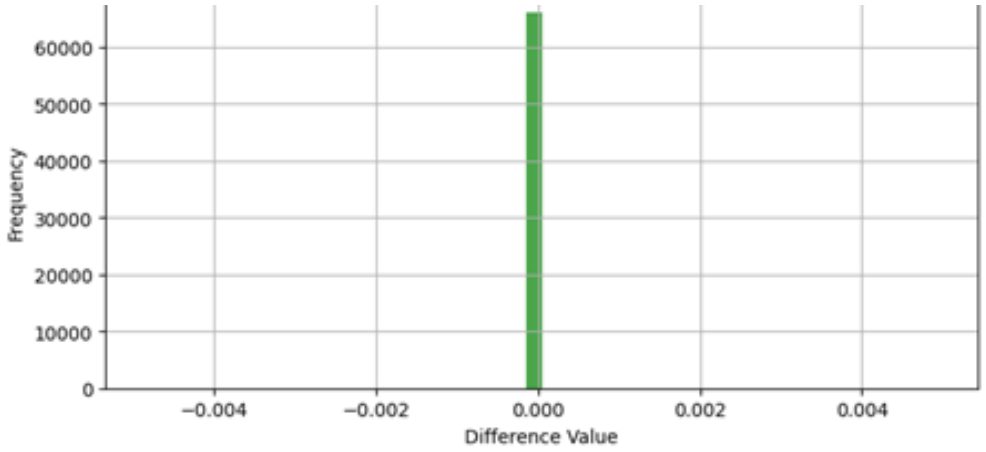

Figure 7. Distribution of Sample Differences

The histogram in Figure 7 showed that the difference between the stego-audio and the original one was tiny such that all the values of modified bits were near zero, indicating that, the modification occurred on the original audio while embedding MAC identifier was not destroying or corrupting process. The clustering values were evenness and constricted which further approve that the embedded process cause no distortion to the voice. By all previously discussed metrics, MACSteg preserves audio reliability at the same time ensures signal verification and integration. The process of embedding MAC address for verification makes imperceptible changes and because no distortion while modifications. Statistical metrics indicates minor and insignificant modification values. These results show evidence of MACSteg system to be proper for real-time or time-sensitive systems, in which low latency and quality are critical. In addition, the results are along with previous researches; determine that an efficient steganography technique, which is used for real-time systems, should have a balance between robustness and quality of the modified audio signal [6–11].

DOI: https://doi.org/10.54216/JCIM.170202
Received: March 30, 2025 Revised: June 20, 2025 Accepted: August 08, 2025

5. Conclusion

In this artical introduces MACSteg to be an effective and implementable system as in a solution for deep fake voice copying and authentication, which are essential in real-time applications by insertion of the device's MAC address continuously during transmission within audio signals via QIM-based steganographic technique, the proposed system empowers consistent source confirmation keeping speech quality and clarity unchanged. The system of maintains the robustness even with some audio interfering situations like noise, compression, and modifications occurred by time modifications which are usually used in deepfake circumstances. The results illustrate MACSteg's opportunities to be solid groundwork for audio communication security within an environment in which individual's identity confirmation is crucial. A low-latency, and minimum computations required for system performance make it proper and suitable for incorporation into developable platforms, such as VoIP communication systems, mobile software, and online conferencing applications.

References

[1] S. Deepikaa and S. Ramakrishnan, "VoIP steganography methods, a survey," *Cybern. Inf. Technol.*, vol. 19, no. 1, pp. 73-87, Mar. 2019, doi: 10.2478/cait-2019-0004.

[2] R. Aloufi, H. Haddadi, and D. Boyle, "On-device voice authentication with paralinguistic privacy," *Comput. Soc.*, 2022, doi: 10.48550/arXiv.2205.14026.

[3] Agarwal, P. R. Singh, and S. Katiyar, "Secured audio encryption using AES algorithm," *Int. J. Comput. Appl.*, vol. 178, no. 22, 2019.

[4] S. Roy et al., "Audio steganography using LSB encoding technique with increased capacity and bit error rate optimization," in *Proc. 2nd Int. Conf. Comput. Sci., Eng. Inf. Technol.*, 2012.

[5] P. Li et al., "IDEAW: Robust neural audio watermarking with invertible dual-embedding," in *Proc. Conf. Empir. Methods Natural Lang. Process.*, Nov. 2024, pp. 4500-4511.

[6] Y. Tian et al., "Watermarking algorithm based on quantization index modulation and singular value decomposition," *Adv. Mater. Res.*, vol. 271-273, pp. 536-540, 2011, doi: 10.4028/www.scientific.net/AMR.271-273.536.

[7] O. Oloyede et al., "Review and analysis on audio steganography techniques," *Int. J. Eng. Comput. Sci.*, vol. 6, no. 1, pp. 22-29, 2024, doi: 10.33545/26633582.2024.v6.i1a.106.

[8] D. Tristan et al., "Steganographic model to conceal the secret data in audio files utilizing a fourfold paradigm: Interpolation, multi-layering, optimized sample space, and smoothing," *J. Saf. Sci. Resilience*, vol. 6, no. 2, 2025.

[9] E. Brooks, L. Jacob, and L. Lewis, "Leveraging GenAI for biometric voice print authentication," *SMU Data Sci. Rev.*, vol. 9, no. 1, 2025. [Online]. Available: https://scholar.smu.edu/datasciencereview/vol9/iss1/3

[10] Choudhury et al., "A novel steganalysis method based on histogram analysis," in *Lect. Notes Elect. Eng.*, vol. 315, 2015, pp. 779-789, doi: 10.1007/978-3-319-07674-4_73.

[11] M. Zadpe, "Lightweight voice authentication for IoT devices using MFCC and CNN on edge hardware," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 13, no. 5, pp. 6144-6151, May 2025, doi: 10.22214/ijraset.2025.71643.

[12] Phipps et al., "Securing voice communications using audio steganography," *Int. J. Comput. Netw. Inf. Secur.*, vol. 14, no. 3, Mar. 2022, doi: 10.5815/ijcnis.2022.03.01.

[13] L. H. Palivela, V. Dharmalingam, and P. Elangovan, "Voice authentication system," in *2023 Int. Conf. Data Sci., Agents Artif. Intell. (ICDSAAI)*, Chennai, India, 2023, pp. 1-6, doi: 10.1109/ICDSAAI59313.2023.10452482.

[14] J. M. Garcia, "Exploring deepfakes and effective prevention strategies: A critical review," *Psychol. Educ. A Multidiscipl. J.*, vol. 33, no. 1, pp. 93-96, Mar. 2025, doi: 10.70838/pemj.330107.

[15] M. Mustak, J. Salminen, M. Mäntymäki, A. Rahman, and Y. K. Dwivedi, "Deepfakes: Deceptions, mitigations, and opportunities," *J. Bus. Res.*, vol. 154, 2023, Art. no. 113368, doi: 10.1016/j.jbusres.2022.113368.

[16] Noreen, M. S. Muneer, and S. Gillani, "Deepfake attack prevention using steganography GANs," *PeerJ Comput. Sci.*, vol. 8, Oct. 2022, doi: 10.7717/peerj-cs.1125.

DOI: https://doi.org/10.54216/JCIM.170202
Received: March 30, 2025 Revised: June 20, 2025 Accepted: August 08, 2025

[17] S. Alanazi and S. Asif, "Exploring deepfake technology: creation, consequences and countermeasures," *Hum.-Intell. Syst. Integr.*, vol. 6, pp. 49-60, 2024, doi: 10.1007/s42454-024-00054-8.

[18] O. Vaidya et al., "Deep fake detection for preventing audio and video frauds using advanced deep learning techniques," in *2024 IEEE Recent Adv. Intell. Comput. Syst. (RAICS)*, Kothamangalam, India, 2024, pp. 1-6, doi: 10.1109/RAICS61201.2024.10689785.

[19] G. Vecchietti, G. Liyanaarachchi, and G. Viglia, "Managing deepfakes with artificial intelligence: Introducing the business privacy calculus," *J. Bus. Res.*, vol. 186, 2025, Art. no. 115010, doi: 10.1016/j.jbusres.2024.115010.

[20] S. A. A. and S. A. A., "Preventing unauthorized access to special applications using signed audio," *Int. J. Civ. Eng. Technol.*, vol. 10, no. 1, pp. 2733-2738, Jan. 2019.

[21] S. A. Kadhim et al., "Developing a new encryption algorithm for images transmitted through WSN systems," *Eastern-Eur. J. Enterprise Technol.*, vol. 124, no. 9, 2023.

[22] H. S. Aljohani, A. A. Alzahrani, and M. A. A. Hossain, "A comprehensive survey on security challenges in Internet of Things: Current trends and future directions," *J. Netw. Comput. Appl.*, vol. 210, p. 103440, 2024, doi: 10.1016/j.jnca.2024.103440.

DOI: https://doi.org/10.54216/JCIM.170202
Received: March 30, 2025 Revised: June 20, 2025 Accepted: August 08, 2025