electronics

MDPI

Article

Securing IoT Data Using Steganography: A Practical Implementation Approach

Fatiha Djebbar

Department of Engineering Science, University West, 461 86 Trollhattan, Sweden; fatiha.djebbar@hv.se

Abstract: Adding network connectivity to any "thing" can certainly provide great value, but it also brings along potential cybersecurity risks. To fully benefit from the Internet of Things "IoT" system's capabilities, the validity and accuracy of transmitted data should be ensured. Due to the constrained environment of IoT devices, practical security implementation presents a great challenge. In this paper, we present a noise-resilient, low-overhead, lightweight steganography solution adequate for use in the IoT environment. The accuracy of hidden data is tested against corruption using multiple modulations and coding schemes (*MCSs*). Additive white Gaussian noise (*AWGN*) is added to the modulated data to simulate the noisy channel as well as several wireless technologies such as cellular, WiFi, and vehicular communications that are used between communicating IoT devices. The presented scheme is capable of hiding a high payload in audio signals (e.g., speech and music) with a low bit error rate (*BER*), high undetectability, low complexity, and low perceptibility. The proposed algorithm is evaluated using well-established performance evaluation techniques and has been demonstrated to be a practical candidate for the mass deployment of IoT devices.

Keywords: IoT cybersecurity; data protection; signal modulation; information hiding; digital audio; audio steganography

Citation: Djebbar, F. Securing IoT Data Using Steganography: A Practical Implementation Approach. *Electronics* **2021**, *10*, 2707. https://doi.org/10.3390/electronics10212707

Academic Editors: Carsten Maple, Matthew Bradbury and Munam Ali Shah

Received: 29 September 2021
Accepted: 1 November 2021
Published: 5 November 2021

1. Introduction

Over recent years, many definitions of the Internet of Things (IoT) have been presented. They generally refer to trends in integrating digital capabilities, including network connectivity, with physical devices and systems. The author of [1] views IoT as "a world of interconnected things that are capable of sensing, actuating, and communicating among themselves and with the environment (i.e., smart things or smart objects). In addition, IoT provides the ability to share information and autonomously respond to real/physical world events by triggering processes and creating services with or without direct human intervention". In contrast, Ref. [2] defines IoT as "a network that connects uniquely identifiable 'things' to the Internet. The 'things' have sensing/actuation and potential programmability capabilities. Through the exploitation of unique identification and sensing, information about the 'things' can be collected, and the state of the 'thing' can be changed from anywhere, any time, by anything". Similarly, in Ref. [3], IoT is defined as "the network of physical objects or 'things' embedded with electronics, software, sensors, and connectivity to enable objects to exchange data with the manufacturer, operator and/or other connected devices. The Internet of Things (IoT) refers to devices that are often constrained in communication and computation capabilities, now becoming more commonly connected to the Internet, and to various services that are built on top of the capabilities these devices jointly provide". IoT networks comprise billions of connected devices exchanging an exponentially growing global volume of data. As these devices handle sensitive data, ensuring their protection should be paramount. Steganography and cryptography techniques can be used to address cybersecurity concerns in IoT networks. In steganography, the interest is in concealing the existence of a message from a third party, while in cryptography, the purpose is to make a message unreadable to a third party. Additionally, the main objective

of steganographic systems is to provide a secure, undetectable, and imperceptible way to conceal a high rate of data in a digital medium. It is used under the assumption that it will not be detected if no one is attempting to uncover it. Steganography techniques manipulate the characteristics of digital media files and use them as carriers (covers) to hide secret information (payload). Covers can be images [4,5], audio [6–10], videos [11] and text [12,13]. Protocols [14,15] and storage devices [16] can also be used as carriers for hidden data. The payload is hidden in any type of digital cover using a key to secure the data and produce a stego file. Figure 1 illustrates a brief comparison between these two techniques.

	Cryptography	Steganography
Application	Secret communication using scrambled information	Secret communication using hidden information
Supported data	Text	Digital medium (e.g., Text, audio, image, video)
Secret key type	Single (private) Double keys (public)	Single (private)
Key size importance	Critical	Moderate
Processing time	Part of the roundtrip delay	Add processing time to the roundtrip delay
Usage	All communications types	Dependent on payload capacity
Human perception	Visible but unreadable	Invisible/Inaudible
Machine based attack	Cryptanalysis	Steganalysis
Attack result	Secret information recovered	Secret communication detected

Figure 1. Steganography versus cryptography.

Steganography applications are not limited to data protection. They have also been used maliciously by criminals, hackers, terrorists, and spies [17,18]. To defeat malicious interventions when communicating secretly, steganalysis techniques have been actively researched to counter steganography algorithms [19,20]. Steganalysis aims essentially to detect the existence of the payload and does not necessarily consider its successful extraction. Steganalysis techniques have a dual role: (1) they are regarded as attacks to break steganographic algorithms and (2) used to measure the strength of steganographic algorithms. Steganalysis attempts to detect or destroy the payload using audio/image processing and statistical analysis approaches. The work of a steganalyst can be very challenging, especially when the only information available is the stego file (blind steganalysis). Most of the steganalysis techniques presented lately are based on learning to differentiate between the cover- and the stego-audio signals. The learning process is performed by machine learning, for example, by using a support vector machine (SVM), [21] on a dataset fed with a set of features extracted from the cover and stego signals. A decision is then made on whether the tested signal is a cover or a stego file (Figure 2). Well-selected features will strengthen the discriminatory power between the cover- and stego-audio signals. The features are intended to capture the differences between the cover and stego signals due to data embedding. In blind steganalysis, the only available information is the received signal. In this case, a reference signal is created to provide an estimate of the cover signal.

The reference signal could be created by applying a de-noising method to the input signal or by applying second steganography using steganographic tools [22–25].

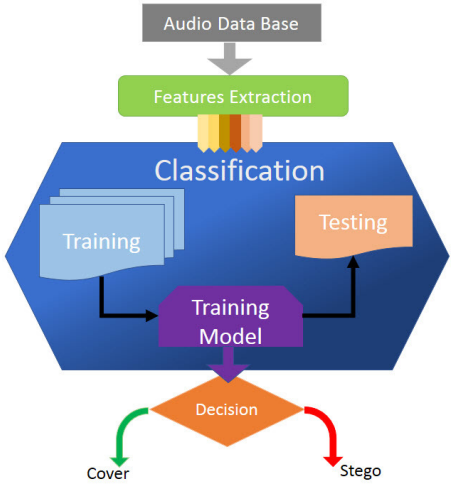


Figure 2. Cover and stego signal classification using SVM.

Existing data protection solutions such as steganography and cryptography are inapt for direct adaptation due to their computational complexity, application specificity, and inflexibility. Furthermore, the IoT environment is depicted as having: limited device capability, high data rate traffic, massive scalability and a large spectrum of heterogeneous devices [26]. IoT cybersecurity challenges, considering the aforementioned limitations, can be addressed using lightweight audio steganography algorithms. Image and video steganography requires a high data rate with additional energy for data transmission, increased processing time, energy consumption, and storage requirement. Audio signals (e.g., speech and music), in this case, are better candidates to facilitate the accommodation of IoT devices' constrained resources. In addition, the number of voice-enabled IoT devices is expanding across industries and has become a standard in connected devices such as: mobiles, tablets, sensors, wearable devices, and smart speakers. The shift from touch to voice is a need that is highlighted by the current pandemic to improve safety. In health care, there are hospitals in the US that allow parents to gain access to high-quality clinical information and specific treatment protocols on Alexa-enabled devices. Manufacturing plants, smart agriculture, construction sites and production lines also require hands-free mobility that IoT voice recognition systems can provide. These voice recordings are transmitted over untrusted public networks and then stored and processed on untrusted third-party cloud-based infrastructures. It is important to protect the information fed to or received by the voice-enabled IoT device such as authentication data, personal and private information.

While many research proposals have been presented in the steganography literature, audio steganography schemes addressing cybersecurity issues in IoT devices are still limited in number, with modest contributions. In this paper, we propose an audio steganography solution designed for IoT implementation through a scalable, noise-resilient, and lightweight algorithm that can accommodate mass deployment in an IoT environment. The remaining part of this paper is organized as follows. The related work is discussed in Section 2 and then our methodology for secure IoT communication is presented in Section 3. The experimental setup and performance evaluation results are given in Section 4 and we conclude the paper and state some future directions in Section 6.

2. Related Work

Several steganography techniques have been developed recently. Some of them adopt data hiding in the time domain [27,28], frequency domain (e.g., Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT)) [6,29], or

encoded domain (e.g., AMR, G.723.1, G729) [30,31]. Most of these techniques are based on least significant bit encoding (LSB). In its naive implementation, the LSB technique could be used as a simplified demonstration of the data hiding process using steganography. Figure 3 shows how we generated a stego signal by replacing the first LSB layer of the cover signal with the digital payload "2018". LSB allows the embedding of a high payload capacity with low perceptibility. However, since the embedding is applied in the LSB plane of the cover signal, the secret message could be easily removed by an unauthorized user. To improve the undetectability of secret data, Ref. [32] proposed a framework based on Generative Adversarial Networks (GANs) to implement optimal embedding for audio steganography in the temporal domain, whereas the authors of [33] proposed a generalized joint adaptive intra-frame and adaptive inter-frame steganography method (called AHCM) within compressed audio streaming and implemented an AdaMP3Stego algorithm in MP3 audio based on the psychoacoustic model. To ensure hidden data robustness against LSB removal and re-sampling attacks, Ref. [34] embedded secret audio into the cover audio by separating the processing of the amplitudes and signs of the secret audio. Ref. [35] combined scrambling and steganography to provide high security for speech hiding in speech. Similarly, Ref. [36] proposed hiding data under the hearing mask in high-frequency samples and [6] proposed a perturbation minimizing algorithm by finding an optimal embedding path and an optimal modification strategy. Discrete wavelet transform and sparse decomposition are used in [37]. There are also several audio steganography tools that are freely available over the Internet, i.e., Xiao Steganography [22], Steghide [23], S-Tools [24] and Camouflage [25].

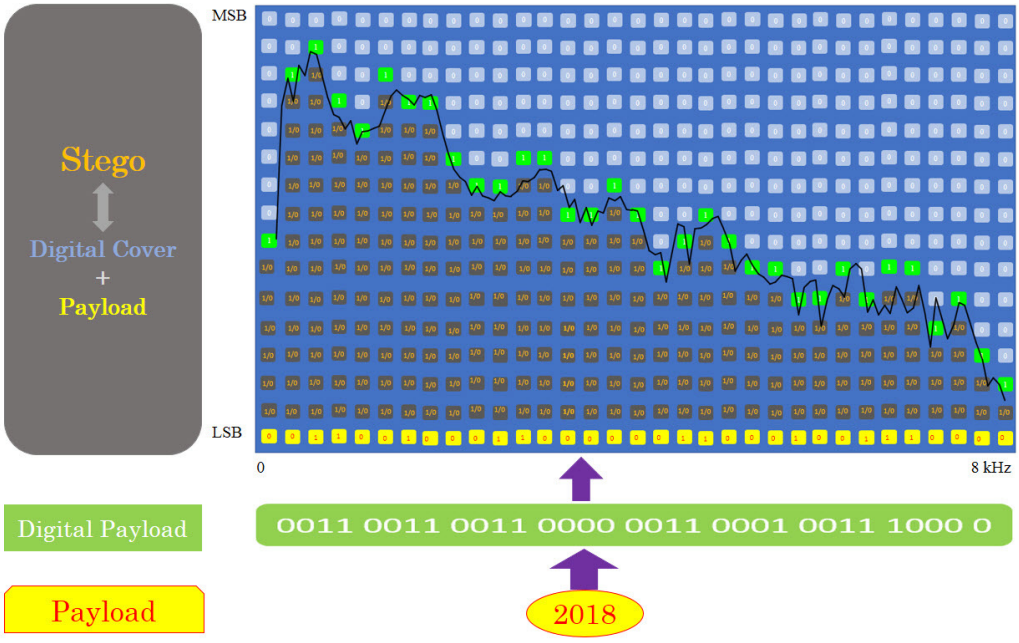


Figure 3. Data embedding in a signal spectrum using least significant bit technique.

In audio steganalysis, a second-order difference of pitch delay features was used in [38]. The authors of [39] modeled the deviation between the reference and the input signals by reversible Mel-cepstrum coefficients (R-MFCC), while the authors of [40] presented a steganalysis technique specifically designed for the steganographic method developed by [41]. A more generalized multi-layer architecture for the steganalysis of all mp3 encoders has been presented by [19]. The first layer of the architecture detects the encoder and the second layer performs the steganalysis. To distinguish between the cover and the stego signal, the authors of [20] used the entropy and the energy of the signal as the discriminator features. The authors claim that this combination enhances the performance of the classifier. A steganalysis method for quantization index modulation QIM steganography in a low-bit-rate encoded speech stream such as G.723.1 and G.729 is presented by [42]. The authors

used the correlation characteristics of split vector quantization VQ codewords of linear predictive coding filter coefficients, arguing that it leads to a stronger correlation network. They showed that this technique improves the steganalysis results for low-bit-rate encoded speech streams.

Although cyber-physical systems (CPSs)/IoT audio steganography techniques are non-existent in the state-of-the-art review, few CPSs/IoT image-based steganography attempts have been made. Ding et al. [43] used mobile edge computing to implement image steganography in IoT, assuming that mobile edge computing fulfills the high computing power, data storage capacity, and bandwidth requirements for the Internet of Things (IoT) through edge servers that process data close to data sources or users. To efficiently protect external product packing in IoT against anti-counterfeiting, Pu et al. [44] used fractional-order spatial steganography (FSS) and blind steganalysis. To promote secure data transfer in a smart IoT environment, a security scheme is advocated to employ a combined approach of lightweight cryptography and the variable least significant bit substitution steganography technique to conform to the intrinsic constraints of IoT devices [45]. Elhoseny et al. [46] proposed a hybrid security model for securing the diagnostic text data in medical images. Covington et al. [47] discuss the special characteristics, challenges, and peculiarities of ensuring security in IoT systems. Among these peculiarities are the distributed deployment infrastructure, the interoperability and heterogeneity of devices, and the large traffic volume of IoT elements. The authors state that these specific aspects of IoT play a major role in the increased probability of security attacks in IoT when compared to other systems in a controlled environment with well-defined security policies and tools. Owing to the intrinsic challenges imposed by the IoT characteristics, several proposals are presented in the current literature. Researchers, in [48,49], have proposed schemes not specifically designed for IoT, but rather for securing data in processing power- and memory size-constrained devices such as mobile phones and embedded devices using simple low-computational complexity cryptography and steganography algorithms. The work presented in [50] attempts to design a security framework for IoT. This framework consists of two algorithms: AES and image steganography. The authors addressed IoT security by proposing a two-tier security scheme. However, they appropriated a classical image steganography algorithm without any adaptation to meet the IoT-specific challenges. Similarly to [50], the authors of [46] proposed an integrated model between a steganography technique and a hybrid encryption scheme to ensure the security of medical data transmission. However, their solution provides for an algorithm to better accommodate IoT security at the cost of the heavy processing required by the presented scheme.

The proposals above are not explicitly designed to resolve cybersecurity issues in IoT but are instead designed to accommodate devices with low processing power, memory size, or battery constraints in traditional networks. In particular, none of these techniques have addressed the issue of embedded data distortion during transmission due to the noisy transmission environment. Additionally, these techniques are image-based, leaving IoT audio steganography unexplored. Audio signals contain a high level of redundancy, allowing a high payload capacity with minimum disturbance. The number of voice-enabled IoT devices is immensely increasing as this feature has become a standard in connected devices such as: mobiles, tablets, sensors, wearable devices, and smart speakers. On the contrary to images and video, audio files require lower data rates and energy for transmission, resulting in a shorter processing time, better energy consumption, and fewer storage requirements. This facilitates the accommodation of constrained IoT devices. This work proposes an IoT audio steganography realizing the perceptible importance of audio in the current digital society and the application trends that depend heavily on cheap, low-power, and intelligent signal processing algorithms and systems such as autonomous vehicles, health care, Intelligent Transportation Systems (ITS), smart grid, environment, and smart cities. In addition, IoT cybersecurity includes new requirements and challenges that must be first addressed before the deployment of these devices:

1. Device capabilities: we adopted Fast Fourier Transform (FFT), which inherits its polynomial complexity $O(N \log N)$ [51], uses a small FFT size, and hides data through bits replacement, resulting in minimum overheads. In addition, data are hidden in the phase components of the audio signal, resulting in a better signal-to-noise ratio [52] and reduced re-transmission probability. Thereby, the proposed scheme addresses the processing limitations of IoT devices, suggesting it is a better candidate for the mass implementation of IoT devices.
2. Interoperability: There is a myriad number of nodes in IoT networks (laptops, sensors, RFID, etc.). The proposed security techniques should be capable of accommodating the full spectrum of heterogeneous devices without impeding their functionality.
3. Scalability: The presented solution must be scalable since IoT networks include a large number of devices.
4. Data traffic rate: The scheme is designed to achieve a high payload capacity. Even if some IoT devices require low data rates, the collective volume of data sent by the large number of these devices sums up to a massive amount.

We also extend our work [26] by investigating hidden data survival transmitted over noisy channels such as WiFi, Bluetooth, Radio Frequency Identification ($RFID$), and other IoT communication technologies. To cover several digital communication techniques used in IoT networks, we consider multiple modulations and coding schemes ($MCSs$) such as quadrature phase-shift keying ($QPSK$), binary phase-shift keying ($BPSK$), quadrature amplitude modulation ($16-QAM$) or ($64-QAM$), and the noisy environment is modeled as additive white Gaussian noise ($AWGN$), thus facilitating practical implementation in IoT networks. We further evaluate the proposed algorithm by measuring the bit error rate, throughput performance, time complexity, and statistical undetectability rate using state-of-the-the-art steganalysis methods.

3. Proposed Solution

3.1. Methodology

The proposed IoT steganographic scheme hides data (payload) in a cover signal (s_c) to generate a stego signal (s_s), which is then sent over the IoT network. For audio cover signals such as speech and music, low frequencies are typically much more potent than high-frequency components. They exhibit better signal-to-noise (SNR) ratios and have higher energy, making them more tolerable for data embedding. To further attenuate the noise induced by hidden data and enhance the stego signal quality, we use the signal spectrum properties whereby noise values 13 dB less than the original signal spectrum (frequency mask) are inaudible for all frequencies [53]. Hence, data embedding is prioritized at lower frequencies, higher energy bins, and at least 13 dB under the frequency mask. To satisfy steganography system criteria in terms of the payload (hiding capacity), statistical undetectability, stego signal quality, and security of embedded data against intentional removal attempts, we used a multi-key combination described in Figure 4. The detailed rationale is explained in the following:

– Payload: the hiding frequency band limit (FBL) is a key set to define where data will be hidden in each frequency frame. FBL is delimited by F_{HDmin} and F_{HDmax}, representing, respectively, the lower and higher frequency bin values used for data hiding. Similar to the sliding window concept, FBL could be increased or decreased to control the embedding location and the payload capacity. Additionally, we used variable least significant bit (SLB) as a second key to increase the payload capacity. SLB represents the lower limit of the embedding locations at each frequency frame.
– Stego signal quality: to increase the stego signal quality, we used only high-energy magnitude frequency bins for data hiding. For this, we defined a *threshold* value that acts as a third key and sets the minimum energy required for a frequency bin to be elected for data embedding. A high *threshold* value is expected to maintain good stego signal quality. Yet, there exists a trade-off between high *threshold* values and the resulting payload capacity and stego signal quality. To ensure a good quality of

the stego-audio signal, we also defined a distortion level (Δ) as an additional key to model the upper limit of the embedding areas. As an example, if we set the frequency mask (ρ) to 13 dB, the expected SNR value between the cover and the stego signals is:

$$\begin{aligned} SNR_{dB} &= 10log\left(\frac{|S_c(m,k)|^2}{\Delta(m,k)^2}\right) \\ &= 10log\left(\frac{|S_c(m,k)|^2}{\alpha * |S_c(m,k)|^2}\right) \\ &= 10log(\alpha)^{-2} = 2 * (\rho) = 26dB \end{aligned}$$

The noise added by the embedded data can be modeled as the difference between the stego (s_s) and cover signals (s_c) such as: $\Delta = s_s - s_c$, Δ can also be represented as a factorization of the cover signal by α, where $\Delta = \alpha * s_c$. Hence, the value of ρ could be increased to enhance the SNR value. The Δ value, on the other hand, has a dual effect on the stego signal: (1) preserving the statistical nature of the modified signal by shaping the noise created by the hidden data into an audio-like spectrum and (2) embedded data below Δ preserve the stego signal quality.

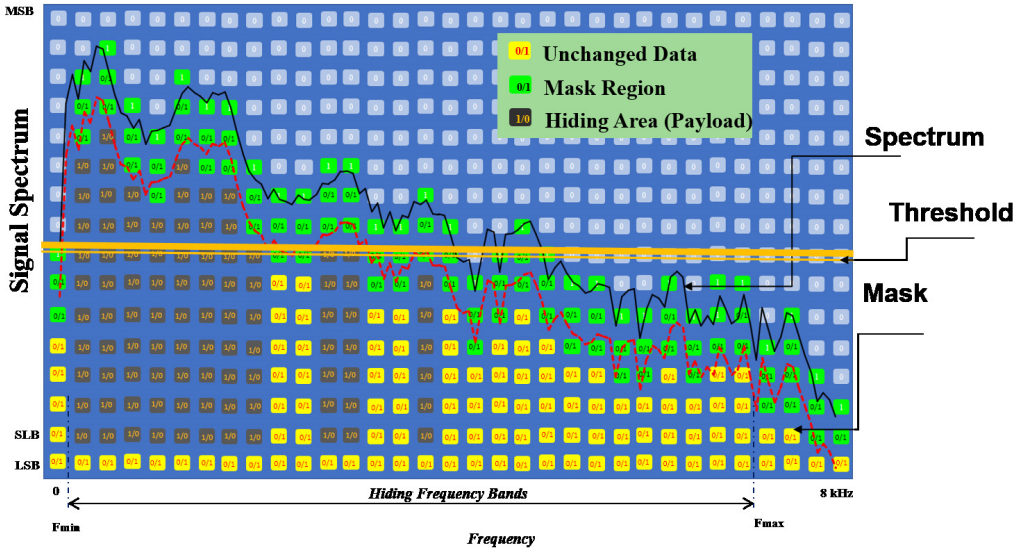


Figure 4. Three candidate embedding locations found in a frequency frame with given values of F_{HDmin}, F_{HDmax}, *Threshold*, ρ and *SLB*.

In addition to securing the embedded data locations, preserving the quality of the stego signal, increasing the payload and hidden data undetectability must be satisfied. Once the embedding locations are defined in the magnitude spectrum using the aforementioned keys, we created their replicas in the phase spectrum to hide our secret data. This is motivated by the following reasons:

- Only a few bits in each selected frequency component are modified, which results in a smooth transition while preserving the phase continuity. This means a minimum overhead with low imperceptibility.
- When phase coding is used, it gives a better signal-to-noise ratio [52], thereby reducing the probability of re-transmission, which in turn makes the proposed scheme a better candidate for the mass implementation of IoT devices.
- Phase coding is robust to common linear signal manipulation such as amplification, attenuation, filtering, and re-sampling [52], which provides better quality with no additional complexity cost.
- It allows opportunities to increase the hiding capacity to accommodate a large number of devices and higher IoT traffic.

3.2. Hiding Algorithm

The proposed IoT steganographic scheme generates a stego signal which is sent over the IoT network. To accommodate several wireless technologies used in IoT networks (e.g., cellular, WiFi and vehicular communications), the stego signal is modulated using orthogonal frequency division multiplexing ($OFDM$) and one of the multiple modulations and coding schemes ($MCSs$) such as $QPSK$, $BPSK$, $16-QAM$ or $64-QAM$ is chosen in a way to obtain the maximum bandwidth while considering the capability of the transmitting IoT device. The stego signal is then sent over an additive white Gaussian noise ($AWGN$) channel to modulate the noisy environment. At the receiver, the signal is demodulated and the embedded data are extracted.

To generate the stego signal, we divided the cover signal into M frames of 4 ms and N samples each, $s_c(m,n)$, $1 \leq m \leq M$ and $1 \leq n \leq N$. FFT was applied to each frame in the frequency domain such as $S_c(m,k) = FFT(s_c(m,n))$ and the magnitude spectrum $|S_c(m,k)|$ was isolated. The hiding band range is specified using $F_{HDmin} \leq k \leq F_{HDmax}$, where F_{HDmin} and F_{HDmax} are the minima and the maxima of FBL (i.e., for a sampling frequency of 16 kHz, F_{HDmin} and F_{HDmax} are 1 and 28, respectively). The *threshold* value was set as the minimum energy required for a magnitude frequency component to be selected for data hiding. The distortion level of the magnitude spectrum $\Delta(m,k)$ was set at ρ_{dB} below the magnitude spectrum, where ρ_{dB} is the amount of attenuation from the original spectrum and is approximately 13 dB.

The embedding process in the phase and the generation of the stego channel using $QPSK/OFDM$ modulation is described in Algorithm 1.

Algorithm 1: Compute $|\phi_s(m,k)|$ and modulate $s_s(m,n)$

Input: $s_c(m,N)$
Output: $s_s(m,n)_{Modulated}$
1 $m = 1, n = 1$
2 **while** $m \leq M$ **do**
3 **while** $n \leq N$ **do**
4 $|S_s(m,n)| \leftarrow |S_c(m,n)|$
5 **while** $F_{HDmin} \leq k \leq F_{HDmax}$ **do**
6 **if** $10 * log_{10}(|S_c(m,k)|) \geq threshold_{dB}$ **then**
7 **if** $\Delta(m,k)_{dB} \geq SLB_{dB}$ **then**
8 $\phi_s(m,k) \leftarrow \phi_c(m,k) + \delta(m,k)$
9 $s_s(m,n) = iFFT(|S_c(m,k)|e^{j\phi_s(m,k)})$
10 $s_s(m,n)_{Modulated} \leftarrow QPSK/OFDM(s_s(m,n)) + AWGN$
11 **return** $s_s(m,n)_{Modulated}$

$\delta(m,k)$ represents the modification in the phase value induced by the embedded bits in a given phase component such as: $\phi_s(m,k) \leftarrow \phi_c(m,k) + \delta(m,k)$. The embedding in a given phase component is defined as follow: $|\phi_c(m,k)| = (a_n 2^n + a_{n-1}2^{n-1} + a_{n-2}2^{n-2} + ...a_2 2^2 + a_1 2^1 + a_0 2^0)$, where $a_n = \{0,1\}$ represents the bit value of the cover phase at a given LSB position, $n = \{8,16\}$ depending on the quantization value of the audio signal and $\delta(m,k) = (d_i 2^i + d_{i-1}2^{i-1} + ... + d_0 2^0)$, where $d_i = \{0,1\}$ is the binary representation of the payload that will be injected into a given phase component. The value of stego phase is recalculated as: $|\phi_s(m,k)| = (a_n 2^n + a_{n-1}2^{n-2} + d_i 2^i + ...d_1 2^1 + d_0 2^0 + a_1 2^1 + a_0 2^0)$. Finally, the new phase is multiplied with its magnitude to produce the stego spectrum therefore $S_s(m,k) = |S_c(m,k)|e^{j\phi_s(m,k)}$. The inverse $iFFT$ transformation is applied on the segment to get the new stego signal segment $s_s(m,n)$. The principal parts of our audio hiding technique are represented in Figure 5, where the embedding of area blocks illustrate steps 3 to 6 in Algorithm 1.

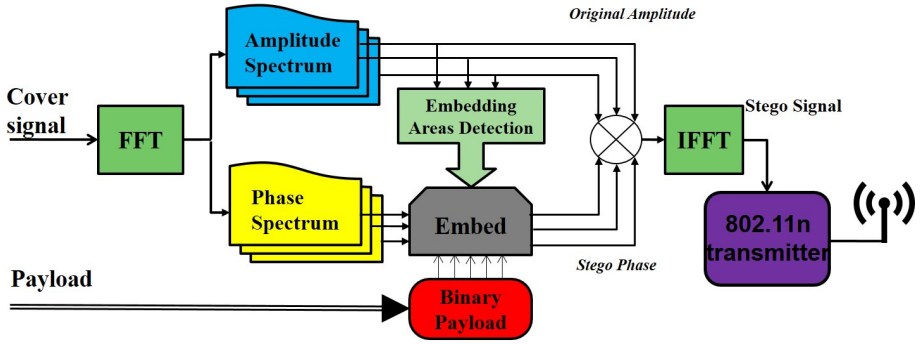


Figure 5. Block diagram for the embedding process.

3.3. Data Extraction Algorithm

At the receiver side, we first demodulated the stego channel and then we extracted the embedded data from the phase components of the stego signal. To achieve this, we first found out the embedding locations in the magnitude spectrum $|S_s(m,k)|$ using the embedding keys: $Threshold$, $\Delta(m,k)$, SLB and FBL. Then, we mapped the embedding locations to the phase spectrum. Secret data segments were extracted and then reassembled as shown in Algorithm 2:

Algorithm 2: Demodulate $s_s(m,n)$ and Extract $\delta(m,k)$

Input: $s_s(m,n)_{Modulated}$
Output: $\delta(m,k)$
$m = 1, n = 1$
while $m \leq M$ **do**
  while $n \leq N$ **do**
    $s_s(m,n)_{Demodulated} \leftarrow OFDM/QPSK(s_s(m,n))$
    $FFT(s_s(m,n))$
    $|\phi_s(m,n)| \leftarrow |\phi_c(m,n)|$
  while $F_{HDmin} \leq k \leq F_{HDmax}$ **do**
    if $10 * log_{10}(|S_c(m,k)|) \geq threshold_{dB}$ **then**
      if $\Delta(m,k)_{dB} \geq SLB_{dB}$ **then**
        Extract $\delta(m,k)$ from $|\phi_s(m,k)|$
return $\delta(m,k)$

Data extraction from the phase spectrum is shown by the block diagram presented in Figure 6.

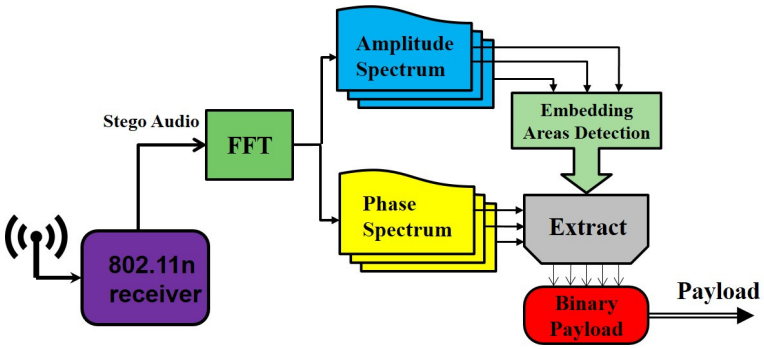


Figure 6. Block diagram for data extraction.

4. Performance Evaluation

4.1. Simulation Setup

The proposed IoT steganography technique was implemented with MatLab. We used a set of 50 wav signals with a confidence ratio of 95%. The signals were sampled at 8, 16, and 44.1 kHz. Each signal was quantized at 16 bits with a duration of 3 to 10 s and a data rate varying between 128 Kbps, 256 Kbps, and 705 Kbps. Stego signals were generated by embedding data at the frequency bands [0.25–4] kHz, [0.25–7] kHz, and [0.5–22] kHz. In each experiment, only one source was used and the results were averaged over all sources.

4.2. Scenarios

Four scenarios were designed to analyze the adequacy of the proposed algorithm for the IoT environment and to evaluate its performance by conducting a comparative study between the stego channel and the cover signal.

1. Perceptual undetectability against payload capacity: In this scenario, the perceptual similarity between the original cover and the stego channel is evaluated against the payload capacity using the PESQ test and $SegSNR$. We aimed to maximize the payload capacity (Kbps) while maintaining a good stego signal quality to determine the scheme's capability to send a high payload capacity. The PESQ is used to measure the quality of the stego signal. The output of the test is a number $\in$ [1 4.5]. An output value of 4.5 indicates that the stego signal is the same as the cover signal. A value of 1 indicates the severest degradation. Achieving a high payload capacity while maintaining perceptual undetectability confirms that the algorithm is capable of accommodating a high data traffic rate summed up from the mass of IoT device transmissions.
2. Resilience to noise channel: In this scenario, we measured the bit error rate (BER) in the stego channel due to noise-induced by the injected payload and the channel. A lower BER corresponds to low retransmission probability and therefore the scheme capability to send high traffic load.
3. Scheme application dependencies: We propose a new performance measure to evaluate the average embedding ratio (AER) against different types of applications (e.g., speech, music, and video). An AER value analysis will allow us to prove that our algorithm is not application-specific and can accommodate heterogeneous and high traffic.
4. Statistical undetectability and time complexity: In this scenario, we measure the performance of our algorithm against steganalysis attacks to determine how well the system can distinguish between stego and cover signals. The accuracy of our predictions is measured by F-measure and the receiver operating characteristic (ROC). In this scenario, the time complexity of our scheme is also measured.

4.3. Performance Metrics

The following performance evaluation metrics were adopted to evaluate the proposed algorithm:

- Perceptual evaluation of the signal quality ($PESQ$) measure, defined in ITU-T P862.2 [54].
- Segmental signal-to-noise ratio ($SegSNR$) in dB (Equation (5)).
- Bit error rate (BER): calculated as the percentage number of retrieved binary bits after demodulating the stego channel and have been altered due to noise, divided by the total number of bits in the transmitted payload.
- Average embedding ratio (AER): AER is defined as:

$$Embedding\ rate\ (\%) = \frac{total\ embedded\ bits}{total\ bits\ of\ the\ signal} \cdot 100 \quad (1)$$

AER is used to calculate the ratio of the embedded data to the cover signal size.

- F-measure and the area under the ROC curve are two of the most popular computational methods to find a balance between false positives and false negatives. The F-measure is calculated using precision (PP) and recall (R) such as:

$$F - measure = 2 \cdot \frac{PP \cdot R}{(PP + R)} \tag{2}$$

$$PP = \frac{TP}{(TP + FP)} \tag{3}$$

$$R = \frac{TP}{(TP + FN)} \tag{4}$$

The ROC is the fraction of true positives (TPR = true positive rate) versus the fraction of false positives (FPR = false positive rate). In this experiment, TP, FP, and FN are defined as follows:

– True positive (TP): stego-audio signal classified as stego-audio signal;
– True negatives (TN): cover-audio signal classified as cover-audio signal;
– False negatives (FN): stego-audio signal classified as cover-audio signal;
– False positives (FP): cover-audio signal classified as stego-audio signal.

5. Results and Discussion

5.1. Scenario-1: Perceptual Undetectability and Payload Capacity

In this scenario, the effectiveness of the proposed algorithm is evaluated on signal frames sampled at 64. We set the algorithm keys' values to maximize the hiding capacity while maintaining the speech quality, i.e., threshold = -20 dB, ρ = 15 dB, F_{HDmin} and F_{HDmax} were set to 1 and 28 for 4 ms frame length. The distortion between the stego and cover signals is calculated and averaged over several frames. $SegSNR$ value for one modified speech frame of 4 ms is given by the following equation:

$$SegSNR_{dB} = 10 \log_{10} \left(\frac{\sum_{k=1}^{28} |S_c(m,k)|^2}{\sum_{k=1}^{28} |S_c(m,k) - S_s(m,k)|^2} \right) \tag{5}$$

The summation was performed over the signal on a frame basis. To evaluate the results, the following criteria were used. First, the capability of embedding a larger quantity of secret data (Kbps) was sought while the naturalness of the stego signal was retained. Second, the hidden data were fully recovered from the stego channel after being sent over the wireless channel.

Figure 7 shows that for $SegSNR$, the values range from 42 to 48 dB and from 4.38 to 4.41 for the $PESQ$ while registering a payload up to 24 Kbps. These results guarantee a high payload with very good quality of the stego channel regardless of the LSB layer depth.

Additional analysis to Figure 7 indicates that the quality of the stego channels is maintained even when the embedding rate is maximized. Hence, to achieve a higher secret traffic rate required in IoT applications and maintain the stego channel quality, hiding at the first LSB layer is to be adopted if the payload survives channel noise as we discuss in the next section.

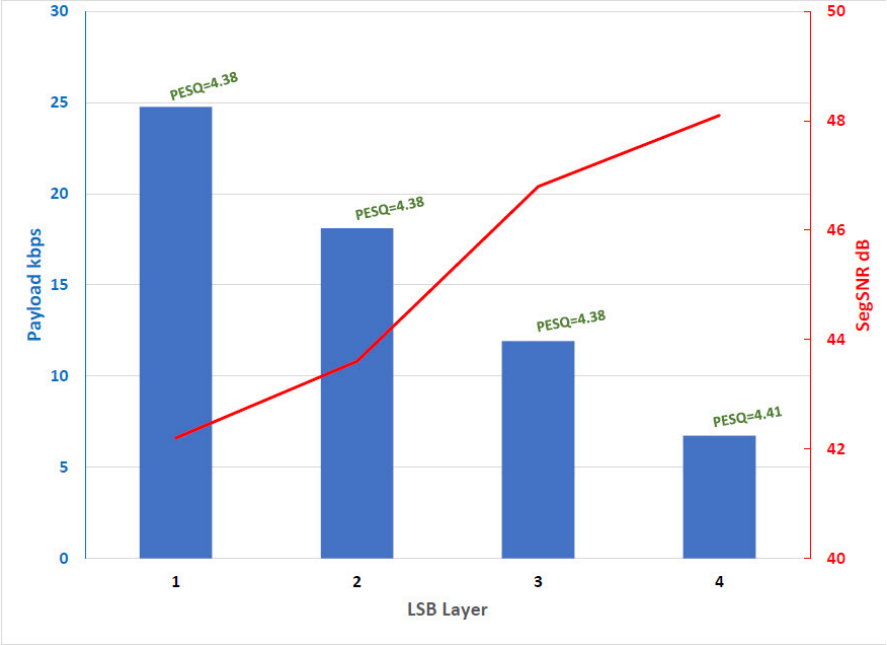
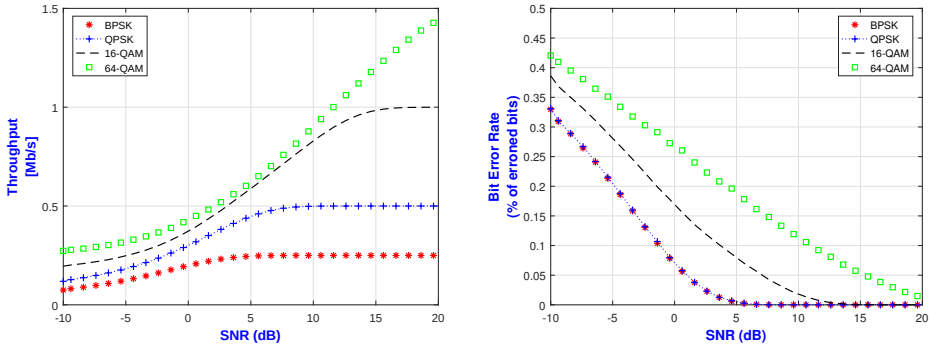


Figure 7. SNR, $PESQ$ and payload (Kbps) test results using frames of 4 ms length and varying LSB key varies from 1st to 4th LSB layer.

5.2. Scenario 2: Resilience to Noise

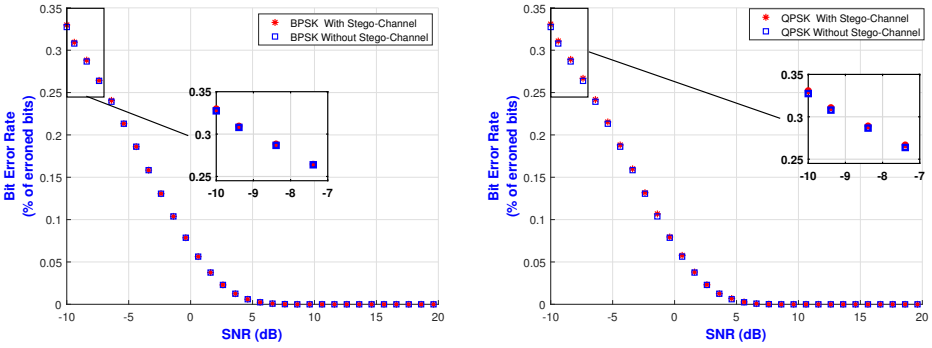
In this scenario, we designed two experiments to assess the robustness of the proposed algorithm for implementation in IoT. The first experiment computes the bit error rate (BER) on our stego channel in the $AWGN$ channel and the second experiment measures the dependency of the expected throughput of the stego channel on the signal-to-noise ratio (SNR). In each experiment, all modulations and coding schemes such as $BPSK$, $QPSK$, $16-QAM$, and $64-QAM$ are tested. Our results show that the proposed algorithm while using $64-QAM$ and $16-QAM$ modulations achieves a better average throughput performance than other modulation schemes (Figure 8a), i.e., at $SNR = 10$ dB, we registered almost 1 Mbps using $64-QAM$ modulation against 0.5 Mbps in $QPSK$ modulation. In all modulations schemes, throughput values increase as SNR increases. A further analysis of the behavior of the proposed algorithm using all modulation and coding schemes in a noisy environment (Figure 8b) shows that the BER level proportionally increases when the noise level alleviates in all tested modulation schemes. In addition, Figure 9 demonstrates that the algorithm is resilient to noise even at low SNR and is almost equivalent to the performance of the channel without engaging the stego channel. There is, however, an inherited percentage of uncontrollable erroneous bits due to channel noise, which is not related to our algorithm. Figure 9a–d show that the proposed algorithm did not induce additional BER, since the BER rate is almost the same with and without engaging our algorithm. The BER shown in the above figures is due to the channel condition which is not the focus of this paper. In $BPSK$ modulation, Figure 9a, for instance, at $SNR = -10$ dB, we registered 0.3274% of BER while using the channel without engaging data hiding versus 0.3298% in the presence of the stego channel. The BER values overlap, in the channel with and without the proposed steganographic algorithm, starting from $SNR = -5.4$ dB. This behavior was observed for all modulation schemes. There will be, however, a successful payload retrieval trade-off between the maximum throughput and the channel noise level. This trade-off does not impact the quality of the retrieved hidden data if the latter are a signal or image. The visual and auditory human system tolerates a certain level of degradation. However, the payload of nature text or number, while using $16-QAM$ or $64-QAM$ with $SNR = -10$ dB and throughput around 0.25 Mbps (Figure 8a), will survive at a relatively higher $AWGN$ power i.e., 6 dB and up. Overall, our algorithm shows resilience to noise for SNR values, from low to high dB values. Therefore, the BER

analysis demonstrates the scheme's capability in terms of reducing the IoT retransmission probability and therefore increasing the traffic load.

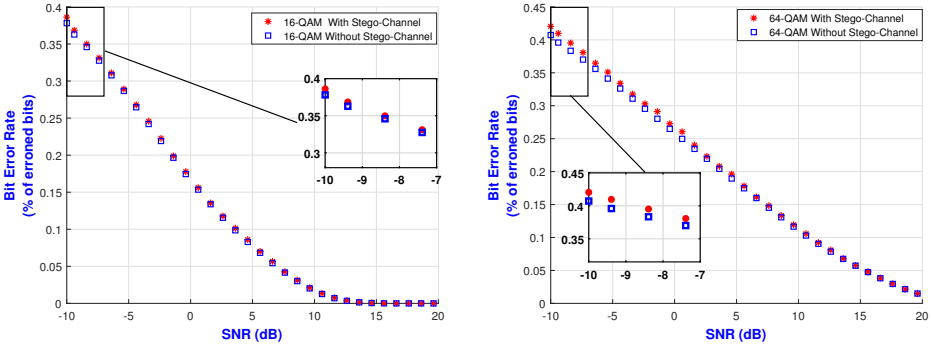


(**a**) (**b**)

Figure 8. Hidden channel throughput as a function of signal to noise ratio (**a**) and BER response to AWGN addition of the proposed algorithm using different channel modulations schemes (**b**).



(**a**) (**b**)



(**c**) (**d**)

Figure 9. Hidden channel *BER* as a function of signal to noise ratio using BPSK (**a**), QPSK (**b**), 16QAM (**c**) and 64QAM (**d**) modulation schemes.

5.3. Scenario 3: Average Embedding Ratio (AER)

We propose another performance measure to evaluate the average embedding performance of the algorithm against different types of applications (speech, music, and video) and its adaptability for high traffic load. The results of the *AER* measure presented in Figure 10 are almost the same as those for all tested signals. We registered embedding ratios of 31.96, 35, 36.2% in speech, music, and video, respectively. These results indicate that our algorithm is not application-specific and can accommodate heterogeneous traffic, which makes it suitable for IoT. In addition, the increased *AER* values show clearly that

the proposed algorithm is a good candidate for IoT by accommodating high traffic and enabling a practical IoT implementation.

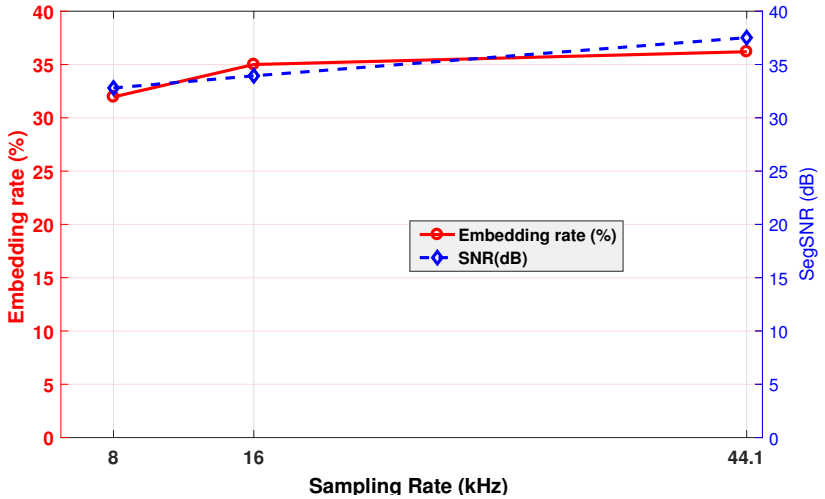


Figure 10. Payload capacity and SNR responses to different sampling frequencies.

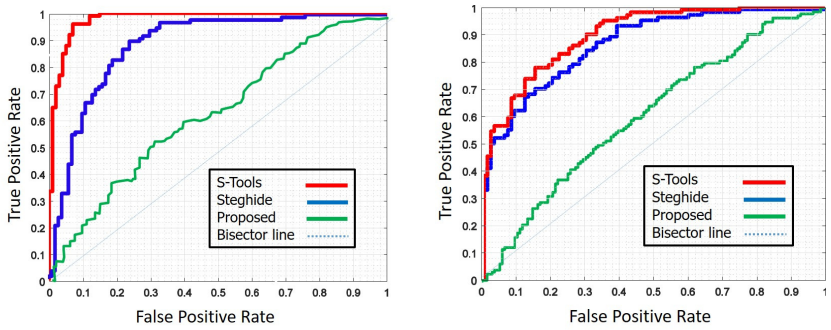
5.4. Scenario 4: Time Complexity and Statistical Undetectability Rate

In this experiment, we assessed the proposed IoT audio steganography based on the time complexity and undetectability rate. Our datasets consist of 500 training and 500 testing audio samples (speech and music) with a frequency of 44.1 kHz and quantized at 16 bits. The duration of each signal is 10 s. All covers were embedded with random messages using the proposed IoT steganography algorithm. We employed the SVM library tool [21] with radial basis function (RBF) kernels [55]. We used two successful audio steganalysis methods: the first is 2D-Mel (second-order derivative-based Mel-cepstrum) [56], which is an efficient method, and the second is EE-AS (Energy-Entropy Audio Steganalysis) [20] which proved to increase the detection rate of steganography work. The accuracy of our predictions was measured by F-measure and the receiver operating characteristic (ROC).

In the absence of a practical implementation of the IoT audio steganography technique, we compared our method with the well-known audio steganography tools Steghide and S-tools in terms of undetectability. In Table 1, we record the overall accuracy, where higher score values are interpreted as a high detection rate. We registered 59.3% compared to 73.2% in Steghide and 81.7% in S-tools (average score between EE-AS and 2D-Mel). These records and ROC curves in Figure 11 demonstrate that we were able to achieve a higher level of undetectability scores.

Table 1. F-measure and ROC result.

Hiding Methods	Steganalysis Method	F-Measure	ROC
S-Tools	2D-Mel	0.706	0.725
	EE-AS	0.909	0.91
Steghide	2D-Mel	0.63	0.67
	EE-AS	0.791	795
Proposed	2D-Mel	0.564	0.569
	EE-AS	0.591	0.593



(**a**) (**b**)

Figure 11. ROC curves for EE-AS (**a**) and 2D-Mel (**b**) tested on S-Tools, Steghide and the proposed algorithm.

Our scheme has also low time complexity since it is based on FFT. For N samples, FFT takes a maximum time complexity of $O(N \log N)$ [51]. In addition to the complexity time, the computational time needed is also small considering we only use audio frames of 4 ms, which require a small FFT size.

6. Conclusions

In this paper, we designed an audio steganography scheme to address the security issues of IoT, which takes into account the collective characteristics of IoT devices. The algorithm is lightweight, noise-resilient, and provides a high payload, which makes it suitable for the deployment of IoT systems. The proposed scheme is based on locating areas to hide data in high frequencies of the audio phase, leading to an increased payload capacity, reduced cover signal disturbance and preserved naturalness of the stego file. The scheme is resilient to noise, where the payload capacity is independent of the signal type, and of low computational complexity, making it appropriate for capability-constrained IoT devices. We utilized audio signals as covers to expand the IoT steganography application range, currently limited to images, and to allow data protection within the increased number of voice-enabled devices. The performance results show that we achieved a low detectability rate of 59.3% and a high payload capacity of 24 kbps at 32 dB SNR. By varying the LSB depths, the $SegSNR$ and $PESQ$ scores gradually increased from 42 to 48 dB and from 4.38 to 4.41, respectively. The simulation and implementation results also indicate that our algorithm (in the presence of the stego channel) is resilient to noise and its performance is very close to the performance of the channel with $BPSK$, $QPSK$, $16-QAM$ and $64-QAM$ modulation schemes in the absence of the stego channel. In all, the proposed scheme can meet IoT steganography requirements and challenges by being able to provide for data protection, survive IoT communication channel noise, and accommodate a large number of devices and the IoT's large traffic volume.

Funding: This research received no external funding.

Data Availability Statement: No new data were created in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Hassan, Q.F. *Internet of Things A to Z: Technologies and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2018.
2. Minerva, R.; Biru, A.; Rotondi, D. Towards a Definition of the Internet of Things (IoT). *IEEE Internet Initiat.* **2015**, *1*, 1–86. Available online: https://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Issue1_14MAY15.pdf (accessed on 13 April 2021).
3. IETF. Internet of Things. Available online: https://www.ietf.org/topics/iot/ (accessed on 12 April 2021).
4. Evsutin, O.; Dzhanashia, K. Algorithm of Information Embedding into Digital Images Based on the Chinese Remainder Theorem for Data Security. *Cryptography* **2020**, *4*, 35. [CrossRef]

5. Zhou, W.; Zhang, W.; Yu, N. A New Rule for Cost Reassignment in Adaptive Steganography. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 2654–2667.
6. Ing, X.; Huang, W.; Zhang, M.; Zhao, I. A topography structure used in audio steganography. In Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 20–25 March 2016; pp. 2134–2138.
7. Guerchi, D.; Djebbar, F. Narrowband Speech Hiding using Vector Quantization. *Int. J. Inf. Commun. Eng.* **2009**, *5*, 5–8.
8. Balgurgi, P.; Jagtap, S. Audio Steganography Used for Secure Data Transmission. In *International Conference on Advances in Computing*; Springer: New Delhi, India, 2016; Volume 174, pp. 699–706.
9. Djebbar, F.; Ayad, B. Audio Steganograpgy by Phase Modification. In Proceedings of the Eighth International Conference on Emerging Security Information, Systems and Technologies, Lisbon, Portugal, 16–20 November 2014.
10. Djebbar, F.; Ayad, B.; Abed-Meraim, K.; Habib, H. Unified phase and magnitude speech spectra data hidingalgorithm. *J. Secur. Commun. Netw.* **2012**, *6*, 961–971.
11. Balaji, R.; Naveen, G. Secure data transmission using video Steganography. In Proceedings of the IEEE International Conference on Electro/Information Technology (EIT), Mankato, MN, USA, 15–17 May 2011; pp. 1–5.
12. Shirali-Shahreza, M.; Shirali-Shahreza, S. Persian/Arabic Unicode Text Steganography. In Proceedings of the 2008 the Fourth International Conference on Information Assurance and Security, Naples, Italy, 8–10 September 2008; pp. 62–66.
13. Djebbar, F.; Guerchi, D.; Abed-Maraim, K.; Hamam, H. Text-in speech spectrum steganography. In Proceedings of the 2010 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010), Kuala Lumpur, Malaysia, 10–13 May 2010.
14. Seo, O.J.; Manoharan, S.; Mahanti, A. Network Steganography and Steganalysis—A Concise Review. In Proceedings of the 2nd International Conference on Applied Theoretical Computing and Communication Technology, Bengaluru, India, 21–23 July 2016.
15. Jiang, Y.; Tang, S.; Zhang, L.; Xiong, M.; Yip, Y.J. Covert Voice over Internet Protocol Communications with Packet Loss Based on Fractal Interpolation. In Proceedings of the 2016 ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), New York, NY, USA, 22–26 August 2016; Volume 12, p. 20.
16. Neuner, S.; Voyiatzis, A.G.; Schmiedecker, M.; Weippl, E.R. Timestamp Hiccups: Detecting manipulated filesystem timestamps on NTFS. In Proceedings of the ACM Proceedings of the 12th International Conference on Availability, Reliability and Security, Reggio Calabria, Italy, 29 August–1 September 2017; pp. 1–6.
17. Steganography in the News. Available online: https://www.technologyreview.com/s/419833/russian-spies-use-of-steganography-is-just-the-beginning/ (accessed on 13 September 2021).
18. Federal News Radio. Available online: http://www.zdnet.com/article/terrorists-and-steganography/ (accessed on 12 September 2021).
19. Ghasemzadeh, H. Multi-layer architecture for efficient steganalysis of UnderMp3Cover in multiencoder scenario. *IEEE Trans. Inf. Forensics Secur.* **2018**, *14*, 186–195. [CrossRef]
20. Djebbar, F.; Ayad, B. Energy and Entropy Based Features for WAV Audio Steganalysis. *J. Inf. Hiding Multimed. Signal Process.* **2017**, *8*, 168–181.
21. SVM. Available online: https://www.csie.ntu.edu.tw/~cjlin/libsvm/ (accessed on 12 September 2021).
22. Xiao-Steganography. Available online: http://download.cnet.com/Xiao-Steganography/3000-2092_4-10541494.html (accessed on 13 September 2021).
23. Steghide. Available online: http://steghide.sourceforge.net/ (accessed on 13 September 2021).
24. S-Tools Version 4.0. Available online: http://info.umuc.edu/its/online_lab/ifsm459/s-tools4/ (accessed on 13 September 2021).
25. Camouflage. Available online: http://camouflage.unfiction.com/Download.html (accessed on 13 September 2021).
26. Djebbar, F.; Abu-Ali, N. Lightweight Noise Resilient Steganography Scheme for Internet of Things. In Proceedings of the 2017 IEEE Global Communications Conference, Singapore, 4–8 December 2017; pp. 1–6.
27. Banerjee, S.; Roy, S.; Chakraborty, M.S.; Das, S. Avariable higher bit approach to audio steganography. In Proceedings of the International Conference on Recent Trends in nformation Technology (ICRTIT), Chennai, India, 25–27 July 2013; pp. 46–49.
28. Shirali-Shahreza, S.; Shirali-Shahreza, M. Steganography in Silence Intervals of Speech. In Proceedings of the Fourth IEEE International Conference on Intelligent Information Hiding and Multimedia Signal, Harbin, China, 15–17 August 2008; pp. 605–607.
29. Qi, Q.; Sharp, A.; Peng, D.; Yang, Y.; Sharif, H. An active audio steganography attacking method using discrete spring transform. In Proceedings of the IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), London, UK, 8–11 September 2013; pp. 3456–3460.
30. Huang, Y.; Liu, C.; Tang, S.; Bai, S. Steganography Integration Into a Low-Bit Rate Speech Codec. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 1865–1875. [CrossRef]
31. Huang, Y.F.; Tang, S.; Yuan, J. Steganography in Inactive Frames of VoIP Streams Encoded by Source Codec. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 296–306.
32. Abdulrazzaq, S.T.; Siddeq, M.M.; Rodrigues, M.A. A Novel Steganography Approach for Audio Files. *SN Comput. Sci.* **2020**, *1*, 97. [CrossRef]
33. Yi, X.; Yang, K.; Zhao, X.; Wang, Y.; Yu, H. AHCM: Adaptive huffman code mapping for audio steganography based on psychoacoustic model. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 2217–2231.

34. Bharti, S.S.; Gupta, M.; Agarwal, S. A novel approach for audio steganography by processing of amplitudes and signs of secret audio separately. *Multimed. Tools Appl.* **2019**, *78*, 23179–23201.
35. Ballesteros, D.M.; Renza, D. Secure speech content based on scrambling and adaptive hiding. *Symmetry* **2018**, *10*, 694. [CrossRef]
36. Djebbar, F.; Abed-Maraim, K.; Guerchi, D.; Hamam, H. Dynamic energy based text-in-speech spectrum hiding using speech masking properties. In Proceedings of the 2nd International Conference on Industrial Mechatronics and Automation (ICIMA), Wuhan, China, 30–31 May 2010; Volume 2, pp. 422–426.
37. Ahani, S.; Ghaemmaghami, S.; Wang, Z.J. A Sparse Representation-Based Wavelet Domain Speech Steganography Method. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2015**, *23*, 80–91. [CrossRef]
38. Ren, Y.; Yang, J.; Wang, J.; Wang, L. AMR Steganalysis Based on Second-Order Difference of Pitch Delay. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 1345–1357. [CrossRef]
39. Hamzeh, G.; Mehdi, T.; Khassb, M.; Khalil, A. Audio steganalysis based on reversed psychoacoustic model of human hearing. *Digit. Signal Process.* **2016**, *51*, 133–141.
40. Yan, D.; Wang, R.; Yu, X.; Zhu, J. Steganalysis for MP3Stego using differential statistics of quantization step. *Digit. Signal Process.* **2013**, *23*, 1181–1185. [CrossRef]
41. Petitcolas, F.A.P. MP3Stego. 2002. Available online: http://www.cl.cam.ac.uk/fapp2/steganography/mp3stego/index.html (accessed on 13 September 2021).
42. Li, S.; Jia, Y.; Kuo, C.C.J. Steganalysis of QIM Steganography in Low-Bit-Rate Speech Signals. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2017**, *25*, 1011–1022.
43. Ding, X.; Xie, Y.; Li, P.; Cui, M.; Chen, J. Image Steganography Based on Artificial Immune in Mobile Edge Computing With Internet of Things. *IEEE Access* **2020**, *8*, 136186–136197. [CrossRef]
44. Pu, Y.; Zhang, N.; Wang, H. Fractional-Order Spatial Steganography and Blind Steganalysis for Printed Matter: Anti-Counterfeiting for Product External Packing in Internet-of-Things. *IEEE Internet Things J.* **2019**, *6*, 6368–6383. [CrossRef]
45. Das, R.; Chatterjee, P. Securing data transfer in IoT employing an integrated approach of cryptography & steganography. In Proceedings of the Proceedings of the International Conference on High Performance Compilation, Computing and Communications, Kuala Lumpur, Malaysia, 22–24 March 2017; pp. 17–22.
46. Elhoseny, M.; Ramírez-González, G.; Abu-Elnasr, O.M.; Shawkat, S.A.; Arunkumar, N.; Farouk, A. Secure medical data transmission model for IoT based healthcare systems. *IEEE Access* **2018**, *6*, 20596–20608.
47. Covington, M.J.; Carskadden, R. Threat implications of the Internet of Things. In Proceedings of the 2013 5th International Conference on Cyber Conflict (CYCON 2013), Tallinn, Estonia, 4–7 June 2013; pp. 1–12.
48. Katagi, M.; Moriai, S. Lightweight cryptography for the internet of things. *Sony Corp.* **2008**, *2008*, 7–10.
49. Stanescu, D.; Stangaciu, V.; Ghergulescu, I.; Stratulat, M. Steganography on embedded devices. In Proceedings of the 2009 5th International Symposium on Applied Computational Intelligence and Informatics, Timisoara, Romania, 28–29 May 2009.
50. Srivastava, A.K.; Agarwal, A.; Mathur, A. Internet of Things and its enhanced data security. *Int. J. Eng. Appl. Sci. (IJEAS)* **2015**, *2*, 257986.
51. Shukla, S.K.; Prasad, M.V. *Lossy Image Compression: Domain Decomposition-Based Algorithms*; Springer Science and Business Media: Berlin/Heidelberg, Germany, 2011; ISBN 1447122186.
52. Bender, W.; Gruhl, D.; Morimoto, N.; Lu, A. Techniques for Data Hiding. *IBM Syst. J.* **1996**, *35*, 313–336.
53. Ayad, B. Noise Suppressor. U.S. Patent 7,889,874 B1, 15 February 2011.
54. ITU-T Recommendation P.862: Perceptual Evaluation of Speech Quality (PESQ): An Objective Method for End-to-End Speech Quality Assessment of Narrow-Band Telephone Networks and Speech Codecs. Available online: http://www.itu.int/rec/T-REC-P.862/en (accessed on 20 September 2021).
55. Vapnik, V. *Statistical Learning Theory*; Wiley: Hoboken, NJ, USA, 1998.
56. Liu, Q.; Sung, A.H.; Qiao, M. Temporal derivative-based spectrum and mel-cepstrum audio steganalysis. *IEEE Trans. Inf. Forensics Secur.* **2009**, *4*, 359–368.