

A Quantum Approach for Text based Steganography

Darsana A.Naik¹, Pallavi N², Keerthi P³, Abhinav Kumar Rai⁴, Adarsh Kumar Jha⁵, Avaneesh Bindal⁶ and Shubhi Singh Chauhan⁷

^{1-2,4-7}M S Ramaiah Institute of Technology, Department of Computer Science and Engineering, Bangalore, India
Email: darshananaik@msrit.edu, Pallavi.n@msrit.edu

³Impact institute of Technology, Department of Computer Science and Engineering, Bangalore, India
Email: Keerthipakka85@gmail.com

Abstract— In our daily lives, maintaining the security of sensitive information has proven to be very difficult. Communications sent frequently through a communication channel like the Internet may catch the attention of crackers and hackers, leading to possible attempts to decrypt and reveal the hidden communications. With the development of digital communication, the Internet, and cloud-based storage, there is an increasing need for novel and effective methods for protecting the privacy and security of digital data. Steganography and cryptography are often used together as a solid privacy protection strategy. However, numerous studies on steganography have been established by various academics. Here, a text steganography method has been developed using quantum circuits and XOR-based symmetric cryptography. In this case, the quantum technique also includes boosting security using a quantum gate truth table to map the logic allows for the embedding of covert and hidden messages. Since text messages are generally regarded as being safe and not requiring further investigation, this technique offers an excellent data protection system. Steganography is an art form that can be used to transmit data securely through any form of open communication.

Index Terms— Text Steganography, Numeric Handwriting, Security and Quantum Steganography.

I. INTRODUCTION

In particular, quantum computers have demonstrated a significant advantage over traditional computers due to quantum mechanics, as evidenced by Grover's database searching technique [1] and Shor's prime factor decomposition algorithm [2].

Over the past few decades, steganography has gained popularity as a method of data concealment and protection against hackers. The term "steganographia" originates from Johannes Trithemius' (1462-1516) work on information concealment. Steganography is an ancient art form that involves hiding information using various covers, communication routes, or algorithms to create a stego without arousing suspicion or detection by eavesdroppers. Its roots can be traced back to ancient Rome, where messages were concealed in papaya leaves and artwork. During World War I, the Germans extensively utilized steganography in all its forms to transport sensitive data. Steganography evolved from cryptography, which focuses on concealing content through encryption, while steganography conceals communication within a channel using concealment techniques. Researchers in this field have explored various media such as text, images, videos, protocols, music, and sounds to develop the concept of steganography. Among these categories, text-based steganography is particularly challenging due to the lack of redundant information, making it both fascinating and demanding.

A. Quantum Steganography

The practice of concealing quantum information while ensuring the secrecy of any covert communication is referred to as quantum steganography. Quantum steganography involves hiding the communication within an apparently innocuous "cover text," which, in the context of quantum steganography, takes the form of a quantum state. The most advanced protocol for quantum steganography utilizes a harmless cover state that is encoded using a quantum error-correcting code and transmitted through a quantum channel. To disguise the message as channel noise, the sender (Alice) embeds a quantum state, consisting of a sequence of M quantum bits (qubits), into an N -qubit codeword[3]. The resulting state appears identical to an innocent state corrupted by typical channel errors. The recipient, Bob, is capable of decoding the hidden state. Prior to utilizing quantum steganography, Alice and Bob typically exchange a secret key, which can be a string of random bits or a series of entangled qubit pairs (ebits). By exploiting Eve's limited knowledge about the channel, which may arise from intentionally amplifying its noise level, Alice and Bob can achieve covert communication at a non-zero asymptotic rate. Even if Eve possesses specific knowledge about the channel, Alice and Bob's ability to communicate covertly is constrained by the number of times the channel has been utilized [4].

Operator	Gate(s)	Matrix
Pauli-X (X)	 	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Pauli-Y (Y)		$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$
Pauli-Z (Z)		$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Hadamard (H)		$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$
Phase (S, P)		$\begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$
$\pi/8$ (T)		$\begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix}$
Controlled Not (CNOT, CX)		$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$
Controlled Z (CZ)	 	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$
SWAP	 	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
Toffoli (CCNOT, CCX, TOFF)		$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

Figure 1: All predefined gates and its matrix.

B. Quantum Gates

A quantum gate, also known as a quantum logic gate, is a simple quantum circuit that operates on a few qubits. These gates are the quantum counterparts of classical logic gates used in traditional digital computers. Unlike many classical logic gates, quantum logic gates are reversible. In the context of conventional digital circuits, classical logic gates are utilized, while qubits serve as the fundamental components of quantum circuits[7]. Although reversible gates have been applied to classical computing, matrix symbols are commonly used to represent quantum gates.

A quantum gate can be represented by a $2^k \times 2^k$ unitary matrix, where " k " represents the number of qubits. The input and output of a gate both consist of the same number of qubits. Researchers have explored various types of quantum gates, including phase shift gates, Hadamard gates, swap gates, Pauli-Z gates, Pauli-Y gates, Pauli-X gates, controlled gates, Toffoli gates, Fredkin gates, and more[8]. In this context, Hadamard, swap, and controlled gates have been extensively employed.

C. Cover Data

In this instance, sender (A) wants to send recipient (B) a standard message. Since a third party might be able to read this message, any intruder or listener could steal its contents. In this scenario, we present a protocol that utilizes the techniques described below. To transmit a classical message (secret data), the sender creates a quantum state that corresponds to the message. The length of the cover data, denoted by $N = 2n$, is determined by the random choice of " n " made by the sender. By applying established steganography techniques, the sender encodes the secret message into the cover data and transmits it to the receiver. The receiver decodes the message using the

same algorithm, which can include methods such as Bernstein-Vazirani, the quantum Fourier Transform, and others[5].

D. Text Steganography

Text has been a traditional and challenging form of media used in literature. In earlier times, secret messages were often concealed within letters and telegrams. Plain text serves as the cover data in quantum-based text steganography, appearing as innocuous information. The steganography protocol is characterized by three key aspects. Firstly, it can utilize any plain text that has not been involved in any previous secret communication between the parties. The content of the plain text does not need to be altered in any way to create the stego data. Secondly, embedded messages are incorporated as components of entangled states rather than as openly accessible information (messages that appear innocent)[6]. Finally, quantum keys are shared between the sender and receiver, which are used by the receiver to extract the secret messages from the stego data.

Text steganography can be broadly categorized into three types:

1. *Format-Based Methods*: These methods involve physically modifying the format of the text to conceal information.
2. *Random and Statistical Generation*: Steganographers provide techniques for users to create their own cover texts. Data hiding can be achieved, for instance, by arranging characters in a random order or using statistical features to determine the frequency of a specific letter in a line of text.
3. *Language Steganography*: This method hides messages using language-specific techniques and may also employ the grammar of the language.

II. PROPOSED SYSTEM

The framework presented provides guidance on embedding and extracting data in a proposed steganography scheme that conceals a secret text within a cover text message. To differentiate text steganography from the more specialized field of "linguistic steganography," the definition of text steganography remains deliberately ambiguous. In this context, we introduce a steganography protocol based on quantum algorithms[9].

The system takes a plaintext message as an argument via a command line interface. A quantum circuit is utilized to encode and convert the received input message, treating it as a bit stream for subsequent transformations. By overlaying the bit stream with a cover object, such as a carrier text, a stego image of the bit stream is created. Prior to transforming it into a stego picture, a customized invertible XOR operation is applied to symmetrically encrypt the bit stream. The length of the stego text matches the length of the cover text thanks to the replacement technique. To enhance security, privacy, secrecy, integrity, and to prevent message detection during transmission, the suggested technique combines steganography and cryptography[10]. Inspiration for this approach is drawn from two well-known quantum algorithms, Deutsch-Jozsa and Bernstein-Vazirani.

A. Methodology

The proposed system involves two stakeholders: SENDER and RECEIVER. The user should possess sufficient knowledge of steganography systems to interact with the system. To begin, the user inputs the plaintext that requires encryption. Next, the user selects a text in which they wish to hide the secret message. The secret message is then encoded into the stego text using the proposed embedding technique on quantum circuits. At the receiver's end, the receiver receives the mapping technique and the encoded stego text corresponding to the secret message. Both of these objects are necessary for extracting the secret message. The mathematical function acting as a symmetric XOR operation is employed for decrypting the information..

B. Algorithm

a. Algorithm for Message Encryption/Decryption

- Making registers with n qubits, where ' n ' can be up to 23, after which a memory fault is raised.
- Produce a random number between [0, 65536] using the qubits that are available.
- Using a quantum key generator, get the key using the random number. Let's call the key "K"
- Select the binary form of the key.
- In the circuit, use an X-gate wherever the key's binary representation of '1' appears.
- The sender should be chosen at random.
- To establish a random basis, generate random numbers between [0, 1], and then add the "H" and "S" flags to the basic vector in accordance with the following rule: Let ' n ' be the random number generated between [0, 1].
- Place the barrier after determining the sender's random-bases vector.

- Create a random bases vector for the receiver using the same rule as described before for the sender.
- Use the Quasm Simulator to simulate the quantum hardware at the backend and count the number of hits.
- Use the basis vectors of the transmitter and receiver to perform key shift.
- Key shifting is performed according to the below mentioned rule:
- Let the receiver basis vector be VRB and sender's be VSB.
- Let the shifted key be K^-

$$K^- = \begin{cases} k_i, & \text{where } V_{Sbi} = V_{Rbi} \\ \phi & ; \text{Otherwise} \end{cases} \quad (1)$$

Denotes we don't include any bit. Here, i [0, length VSB]

b. Algorithm for Message Embedding

Encode the secret message into the Quantum Circuit using the following steps:

- Convert the secret message to a binary vector first.
- Trim the binary string to fit on the device with 'n' qubits. Now, for each bit in the secret message's bit string, complete the following steps:
- Build a circuit containing 'n' qubits and one ancilla qubit, as well as n classical bits to write the output to.
- Put ancilla bits in state 1 $\rightarrow$ using H-gate and Z-gate.
- Before querying the Oracle, use the Hadamard gate (H-gate).
- Use the barrier to deliver instructions to the transpiler.
- On the reversed secret message bit, use the inner product oracle.
- Use the barrier once again to divide the ckt.

Apply Hadamard gates after querying the oracle. We will see the following transformation

$$a \xrightarrow{H^{\otimes n}} \frac{1}{\sqrt{2^n}} \sum_{x \in \{0,1\}^n} (-1)^{a \cdot x} |x\rangle \quad (2)$$

Now, qubits are measured which collapse them to either 0 or 1 classical bit. Collect all the classical bits in the circuit array.

C. Algorithm for Message Extracting

Encrypt the message and key using steganographic encoding by first computing the XOR. Encode the message bit string's leftmost bits into the carrier message using the following rule:

- If the bit in the bit string is one, change the letter to upper case.
- Using the reverse of the preceding method, decode the secret message from the new carrier message.

III. RESULT ANALYSIS

The performance of the model is enhanced as the number of qubits is increased, but the time taken for the whole process is increased. The overall increment in stego image size for text-based steganography is less compared to image / audio-based steganography for the same length of secret information.

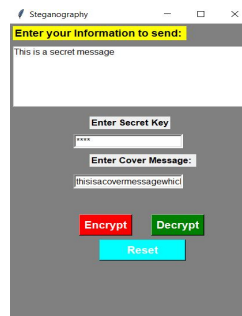


Figure 2: Proposed Stenography System

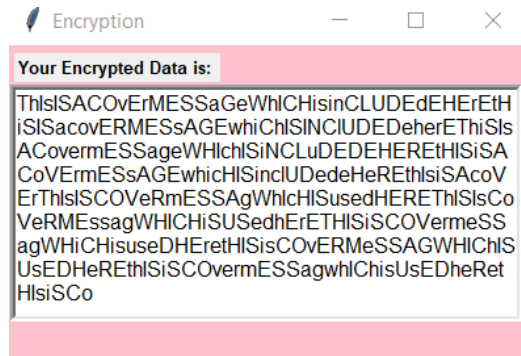


Figure 3: Encrypted Data

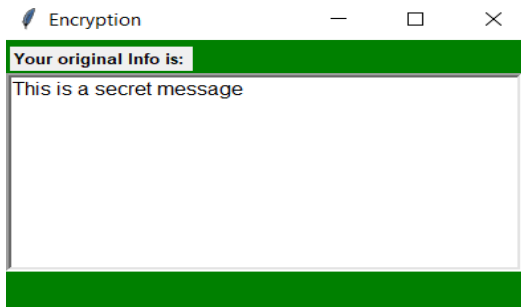


Figure 4: The original Message

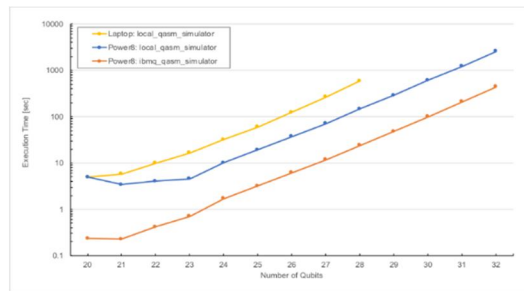


Figure 5: Number of Qubits vs. Execution Time

IV. CONCLUSION

Quantum steganography, which is based on a stunning blend of notions from quantum physics and information theory, is a remarkable demonstration of the dialogue between classical and quantum informatics. Its security premise is based on deep classical information theory theorems as well as a thorough understanding of Heisenberg's uncertainty principle and entanglement theory. Data concealing in multimedia can aid in giving confirmation of content's origin and delivery. Image and multimedia carriers are a widespread trend that sets off security concerns. Text to Text Steganography is a lesser-known and very simple yet efficient form of steganography. It has been proved that almost any type of text information can be used to disguise plain text conversations in our work. We also established that the file size increase is in bytes, as opposed to image steganography, where the resultant size is in Kilobytes. Combining colour and font steganography, as well as introducing particular standards such as font styling, can improve the process.

The purpose of this paper is to examine the security of several information-hiding strategies that take advantage of the flexibility of certain property features of quantum systems. The scenario considered involves two people engaged in a covert exchange of messages in the presence of the wicked invader's passive observation. The limitations of such steganographic techniques in giving security cover to covert communication are demonstrated, followed by a proposal for their security augmentation justified by mathematical annotation. The study's implicit finding elucidates that in today's High-Tech immersive computing environment, steganographic techniques, along with quantum bits evolved, must consider the security component of information whether it is at rest (permanent

storage) or in transit. With all of these developments, it can be used in the future and expanded with more secure algorithms and intricate cover pictures.

ACKNOWLEDGMENT

I thank the organization and who supported us in completion of the work.

REFERENCES

- [1] Banerjee, I., Bhattacharyya, S., & Sanyal, G. (2012, August). Text Steganography through Quantum Approach. In International Conference on Information Processing (pp. 632-643). Springer, Berlin, Heidelberg.
- [2] Luo, Y., Huang, Y., Li, F., & Chang, C. (2016). Text steganography based on ci-poetry generation using Markov chain model. *KSII Transactions on Internet and Information Systems (TIIS)*, 10(9), 4568-4584.
- [3] Singh, P., Chaudhary, R., & Agarwal, A. (2012). A Novel Approach of Text Steganography based on null spaces. *IOSR Journal of Computer Engineering*, 3(4), 11-17.
- [4] Xiang, L., Yang, S., Liu, Y., Li, Q., & Zhu, C. (2020). Novel linguistic steganography based on character-level text generation. *Mathematics*, 8(9), 1558.
- [5] Banerjee, I., Bhattacharyya, S., & Sanyal, G. (2012). A Procedure of Text Steganography Using Indian Regional Language. *International Journal of Computer Network & Information Security*, 4(8).
- [6] Shaw, B. A. (2010). Quantum steganography and quantum error-correction. University of Southern California.
- [7] Banerjee, I., Bhattacharyya, S., & Sanyal, G. (2013). Study and analysis of text steganography tools. *International Journal of Computer Network and Information Security*, 5(12), 45.
- [8] Xiang, L., Yang, S., Liu, Y., Li, Q., & Zhu, C. (2020). Novel linguistic steganography based on character-level text generation. *Mathematics*, 8(9), 1558.