Novel Approaches to Text Steganography

Tanmay Bhagwat, Saurav Kanchan
Department of Computer Engineering, Ramrao Adik Institute of Technology, Navi Mumbai, Maharashtra, India-400706
Email address: tanmaybhagwat07@gmail.com, sauravnk30@gmail.com

***Abstract*— *Current techniques of steganography are not impervious to detection using thorough procedures. This paper proposes to ascertain security and precision in the process of Text Steganography by presenting a few unexplored techniques. It establishes reliable media to transfer any required message accompanied by hidden data. The process relies on RSA Algorithm to encrypt user data by generating subtle imperfections in the appearance or layout of the characters included in the memo. The message undergoes multilevel encryption and corresponding decryption using both the public and private keys, thus bearing invulnerability to cyber-attacks and security breaches. The receiver acquires the modified message in the sender's format, to fend off digital decryption from unauthorized users. As opposed to contemporary techniques of Text Steganography, the encrypted message has a relative constraint on the size as it depends on the size of the data to be hidden, bearing no definite restrictions on the amount of information to be conveyed.*

***Keywords*— *ASCII; cryptography; data hiding; decryption; encryption; RSA algorithm; steganography; text steganography.*

I. Introduction

Steganography is the practice of conveying a hidden purpose through a plain message in the most discreet way possible. In this context, the plain and the secret messages could be images, text, videos, documents or anything that can travel over digital media. Although the technological implementation of this concept is relatively new, the idea dates back to 440 BC when Herodotus suggested it in his renowned book of Histories [1]. The nature of data encryption has changed quite a while over some time, but the intent remains indistinguishable. While the channel of data transmission has changed from physical to digital over the better half of the previous century, its security and invulnerability have seen an exponential rise [2]. In times of tangible data hiding, there was always a threat of the real objective of the message to be exposed due to invasion by fickle third parties, which would not only fail your cause but also keep you from doing the same in the future. Initially, the digital version also did witness flaws through security breaches and manual trespassing. Thankfully, researchers involved thought of new solutions to overcome these mishaps [3]. The robustness of encryption algorithms in Steganography has seen a meteoric improvement along the decades, further evolving encryption techniques.

A. RSA Algorithm

One of the most widely known and practised algorithms in cryptography is the RSA Algorithm developed by Ron Rivest, Adi Shamir & Leonard Adleman at the Massachusetts Institute of Technology in 1977 [4]. It uses two keys, one for encryption and the other for decryption. Both the public and the private keys are generated by the receiver. Public Key can be viewed by any observer, hence this key is used by the sender to encrypt their data [5].

As the private key is kept a secret, it is only known to the receiver who can use it to decrypt the enciphered data. The digital signature of all the parties involved is also generated and authenticated for security purposes [6]. RSA is often used today in Image Steganography with the association of Least Significant Bit (LSB) Algorithm [7]. Here, the least significant pixels of the carrier image are replaced by the encrypted message to pass on the message without seeking unwanted attention from the viewer. Insertion of data depends on the resolution and number of bits present in the original image. As full proof as this method sounds, algorithms have been trained and implemented to identify the LSB pixels of an image and check for embedded data [8]. Hence, further work is necessary to keep it undetected. The equations for RSA are as follows.

$$p, q = primes \in (1, \infty) \quad n = p \times q$$

$$Totient\ function\ \Phi(n) = (p-1) \times (q-1)$$
$$1 < e < \Phi(n)\ such\ that\ e \div n \notin integer$$

$$d = (k * \Phi(n) + 1) / e\ where\ k \in integer$$
$$C = M^{e}\ mod\ n \quad M = C^{d}\ mod\ n$$

B. Text Steganography

In Text Steganography, common techniques are based on indentation manipulation and character formatting [9]. We thought of bringing these systems in a new light through minor upgrades. The modified methods possess the capability to process comparatively higher data size while also extending the threshold of data hiding. This paper plans to conceal required information in the transferred text through subtle but discernible imperfections which could be interpreted by the receiver. Imperfections are precise manipulations in data, which usually skip the grasp of a glance [10]. These changes can be spotted and decrypted only by a knowledgeable user. Our encryption method deals with five distinct types of imperfections, namely, shift up, shift down, shift left, shift right and highlight. Every single type of imperfection carries a certain meaning related to the hidden message. Numerous combinations of these subtle manipulations can be used together and classified as levels of imperfections [11]. Levels are allocated to a particular symbol depending on the number of imperfections it carries. If a character is both spaced towards the right and highlighted, it belongs to the second level of imperfections and so on. Using this technique, we can readily increase the scope of message encryption manifold.

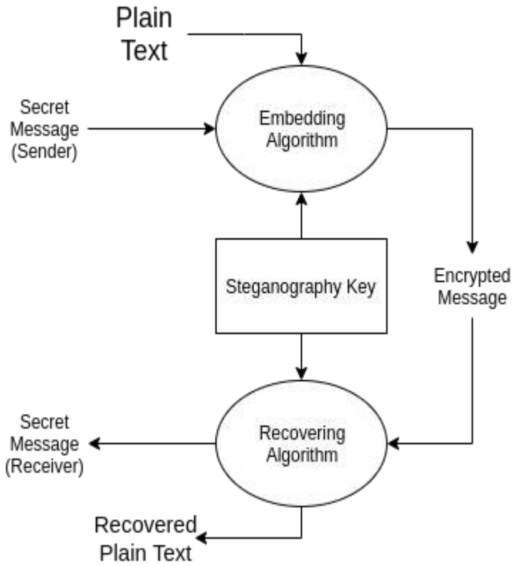

Fig. 1. Process of Text Steganography.

As the encryption using RSA is not completely impenetrable, we have devised a system where the secret ASCII message undergoes RSA Encryption [12]. The resultant data is converted to its binary form, which is further referred to for the sole purpose of appropriate allocation of imperfections. During decryption, these imperfections can be mapped with their binary equivalents to form a discernable string of binary data. This string is further retracted back to its decimal format. Viable for RSA Decryption, the resulting decimal number is deciphered to reveal the intended message.

II. Proposed Methodology

A. Encryption

ASCII stands for American Standard Code for Information Interchange, which is an international information exchange standard developed by the Internet Assigned Numbers Authority (IANA) [13]. The message goes through a series of steps for encryption, which help maintain a robust environment for communication. The process is initiated by converting the secret message into the globally accepted ASCII format. Thus, each character of the message is transformed into its corresponding ASCII code and the resultant string is concatenated.

Further, the RSA Algorithm is used to encrypt the produced ASCII Code [14]. The receiver generates the public and private keys from a set of random prime numbers. The public key is shared with all the coexisting users through an insecure channel. Any user willing to send a message uses the public key to encrypt the valid ASCII value. Once the encoded text is obtained, all the present characters are transformed to their binary equivalents. Due to this conversation, there is a significant increase in the length of the secret keyword. Binary messages tend to make the process of integrating the imperfection association much simpler and easy to comprehend. The aforementioned binary code is divided into groups of 5 digits, one each for every type of imperfection. If the keyword length is not a multiple of 5, it is appended with the number of zeros required to form a cipher having the length of the nearest greater multiple of 5. Each bit of the keyword, starting from left, is set or reset for shift up, shift down, shift left, shift right and highlight respectively. Consider the 5-bit keyword 10100. Here the plain text will be shifted slightly above the margin and towards left as well. (Refer to Fig. 2)

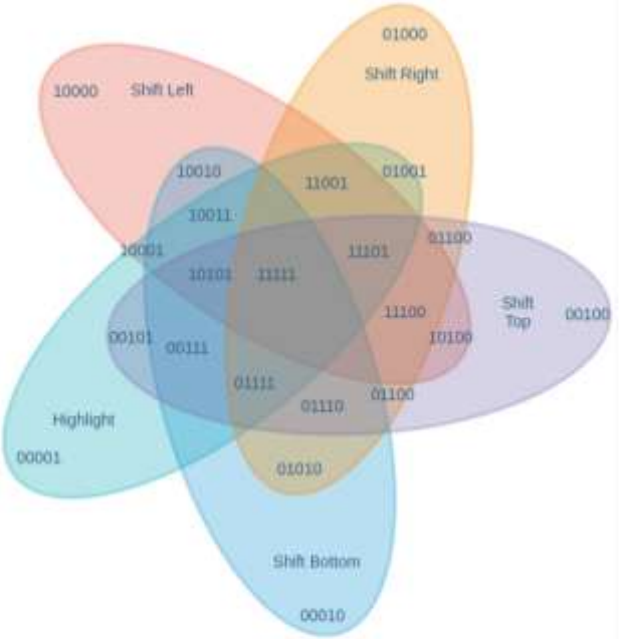

Fig. 2. Imperfections and their combinations

B. Decryption

The deciphering methodology works exactly in a reverse manner of the encryption model. Authorized decryption can only be performed by the intended receiver as he/she is the sole possessor of the private key, which is a salient factor for decryption. The digitally generated imperfections in the text are accurately detected by our decryption algorithm. These imperfections are converted to their corresponding 5-bit binary code groups. All of these ‘n’ groups are combined to make an array of ‘5n’ binary elements. This resultant string is then transformed to its corresponding decimal value. The RSA Decryption process using the private key comes into the picture at this point of time. Performing modulo operations on the private key and the key elements consistently generated during encryption results into the original ASCII value [15]. The ASCII table is referred to convert these values to their pertinent characters. The receiver attains this outcome as a result of the decryption process, which typically matches the message hidden by the sender.

The steganographic process discussed above succeeds to provide an extremely secure channel for private message exchange. Due to the utilization of the robust dual encryption and decryption standards i.e. ASCII as well as RSA, the system is proven invulnerable to third party attacks. The process manages to arrange a reliable channel for confidential communication while keeping the user text rather inconspicuous or difficult to grasp for the naked eye [16]. A simulation of this methodology is seen in Fig. 3.

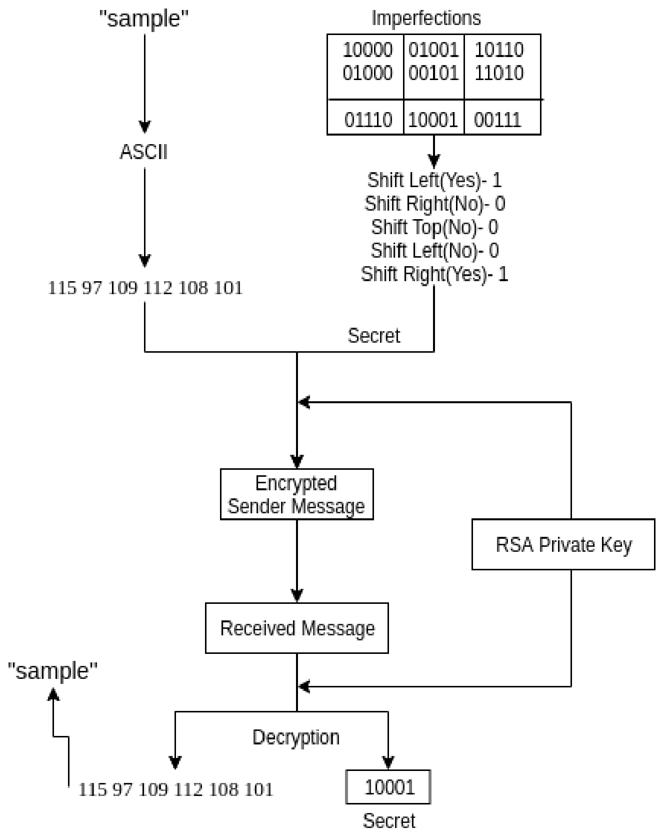

Fig. 3. Encryption-Decryption using an example

III. Conclusion

Encryption protects the message by converting it into an unintelligible format while steganography is the method of hiding the message from the attacker. The paper presents a novel technique of information hiding which includes both encryption and steganography. RSA is an asymmetric key-based algorithm is used to convert the plain-text to an unintelligible format and steganography is used to hide the binary equivalent of the encrypted ciphertext into an image by exploiting the imperfection of characters. Combination of both encryption and steganography makes the proposed model resistant to the standard cryptanalytic attacks and also increases the confidentiality of the message being sent. The method can be improved further by introducing levels of each imperfection which will increase its embedding rate.

References

[1] F. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information Hiding: A survey" (PDF)," 1999, Proceedings of the IEEE. [Online]. Available: 1062-78.CiteSeerX10.1.1.333.9397.doi:10.1109/ 5.771065.Retrieved2008-09-02

[2] M. Dobsicek, "Modern Steganography," *8th International Student Conference on Electrical Engineering FEE CTU*, 2004.

[3] C. P. Sumathi, T. Santanam, and G. Umamaheswari, "A study of various steganographic techniques used for information hiding," 2014, arXiv preprint.

[4] R. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems" (PDF)," *Communications of the ACM*, vol. 21, no. 2, 1978. [Online]. Available: 120-126.

[5] B. Alese and E. Falaki, "Comparative Analysis of Public-Key Encryption Schemes," *International Journal of Engineering and Technology*, vol. 2, 2012.

[6] R. C. Merkle, "A certified digital signature," in *Conference on the Theory and Application of Cryptology*. Springer, 1989, pp. 218–238.

[7] R. Halder, S. Sengupta, S. Ghosh, and D. Kundu, "A Secure Image Steganography Based on RSA Algorithm and Hash-LSB Technique," *IOSR Journal of Computer Engineering (IOSR-JCE*, vol. 18, no. 1.

[8] A. Sarkar and S. Karforma, "A new pixel selection Technique of LSB based steganography for data hiding," pp. 12–125.

[9] S. Bhavana and K. L. Sudha, "Text Steganography using LSB insertion method along with Chaos Theory," *International Journal of Computer Science, Engineering and Applications*, vol. 2, no. 2, pp. 145–145, 2012.

[10] R. Mustafa and C. Bach, "Information Hiding in Images Using Steganography Techniques," 2013. [Online]. Available: 10.13140/RG.2. 1.1350.9360.

[11] T. Morkel, J. H. Eloff, and M. S. Olivier, "An overview of image steganography," *ISSA*, pp. 1–11, 2005.

[12] S. Sahu, J. Singh, and J. Ashraf, "Encryption and decryption of text data with rsa cryptography using matlab."

[13] N. Freed and J. Postel, "Iana charset registration procedures," *RFC 2278*, vol. 19, 1998.

[14] Steef, Ahmad, et al. "RSA Algorithm With a New Approach Encryption and Decryption Message Text by ASCII." *International Journal on Cryptography and Information Security*, vol. 5, no. 3/4, 2015, pp. 23–32., DOI:10.5121/ijcis.2015.5403.

[15] Kessler, Gary C. "An overview of cryptography." *the Handbook on Local Area Networks, Auerbach* (1998).

[16] Hopper, Nicholas J., John Langford, and Luis Von Ahn. "Provably secure steganography." In Annual International Cryptology Conference, pp. 77-92. Springer, Berlin, Heidelberg, 2002.

Tanmay Bhagwat- Core Committee Member and Member of Research Wing for Computer Society of India

Saurav Kanchan- Member of Research Wing for Computer Society of India