INDIAN INSTITUTE OF TECHNOLOGY GANDHINAGAR

MA202 Spring Semester 2021

An enhanced cryptographic technique combined with steganography, to enhance Network Security System, using Numerical Methods

By:
Shreyshi Singh (19110032)
Xhitij Choudhary (19110044)
Ayush Anand (19110074))
Siddhi Surawar (19110170)
Somesh Pratap Singh (19110206)

Contents

1 Problem Statement

Cryptography is not a new concept. Rather, it has been in use for thousands of years now. It is the study of mathematical methods pertaining to aspects of security of information like data integrity, authentication, confidentiality, and data origin authentication. It does not only mean hiding information. Rather it is a set of techniques which can be implemented to code a data, send it to the intended receiver, and should be easily unpackable by them.

With the dawn of the Information Age, development of hacking techniques and sensitivity of data, security has become more important than ever. The most basic cryptography consists of a coded message and a key to decode it. Depending on whether the encryption key is same as the decryption key or not, cryptographic algorithms are divided into two parts - Secret Key Cryptography (or Symmetric Key Cryptography) and Public Key Cryptography (or Asymmetric Key Cryptography).

Steganography is also a method to encrypt data. Confused with cryptography many times, the difference between them is that steganography conceals data in a non susceptible way, so that it goes undetected. For instance, hiding the data in a normal looking image or file. The same is not true for cryptography, where encrypted data's existence is confirmed, but unknown. Due to the extra protection, steganography is a more preferred and implemented technique.

KEYWORDS
Diffie Hellman Key Exchange, Steganography, Cryptography, Bisection Method, Newton-Raphson method

2 Notations

Diffie-Hellman: Parameters p and g
p is a very large prime number, consisting of around 40 digits
g is an integer between 1...p
a: Random integer generated by Alice in 1...p-1
b: Random integer generated by Bob in 1...p-1

3 Assumptions

(i) We are assuming that we are only dealing with communications between two users at a moment. For multiple users, this algorithm has to be upgraded to a cyclic form and is still an open problem.

(ii) The attacker does not perform any active attacks, i.e, no hacker in the middle kind of attacks. All the person can do is read the encrypted data being transferred, but can't tamper with it.

(iii) Assumption made while applying the numerical methods: The error beyond 0.001 is ignored.

4 Pre-requisites

- **Secret Key Cryptography or Symmetric Key Cryptography (SKC)** : This is a conventionally used single-key encryption. In symmetric cryptography the same key is used for both encryption and decryption. This technique can be used to encrypt data, either by a single user to protect his/her files, or to be exchanged between users. If encrypted data is exchanged between two (or more) users, each must know the secret key to be used. Also, this key should be exchanged in a secure manner.

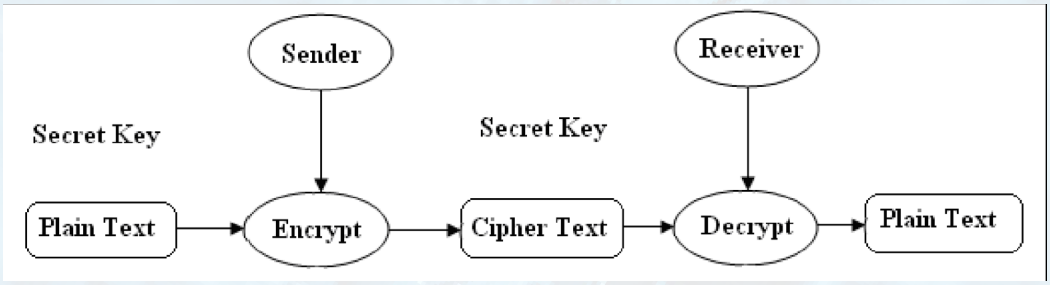

Figure 1 Symmetric Key Crytography , *Adapted from Ghosh A., Saha A., 2013, Regent Education Research Foundation*

- **Steganography** : Steganography is a way of hiding private information within something that appears to be nothing out of the usual. In the modern day sense, steganography refers to the information that has been concealed inside a digital image, video or audio file. Both steganography and encryption are used to maintain confidentiality of data. The main difference between them is that with encryption anybody can notice that two parties are communicating in secret. Steganography hides the very existence of a secret text. Adding encrypted copyright information to a file could be easy to remove but embedding it within the contents of the file itself can prevent it being easily identified and removed.
- **Digital Signature Standard (DSS) Technology** : A digital signature is an electronic analogue of a handwritten signature. It provides an assurance that the claimed signatory has signed the document or information. With the help of digital signature we can keep a check on whether or not the information was modified after it was signed. We can get these assurances on whether the data was received in a transmission or retrieved from storage. A well implemented digital signature algorithm meeting the requirements of this standard can provide these services.
- **Diffie-Hellman Key Exchange** : he Diffie-Hellman method is one of the most secure methods for the exchange of cryptographic keys over a public channel. For encrypted communication between two people, they first need to exchange secret keys by any physical means. This key exchange method enables two people who have no prior knowledge of each other to use their own secret keys and subsequently establish a communication jointly using a symmetric key cipher.

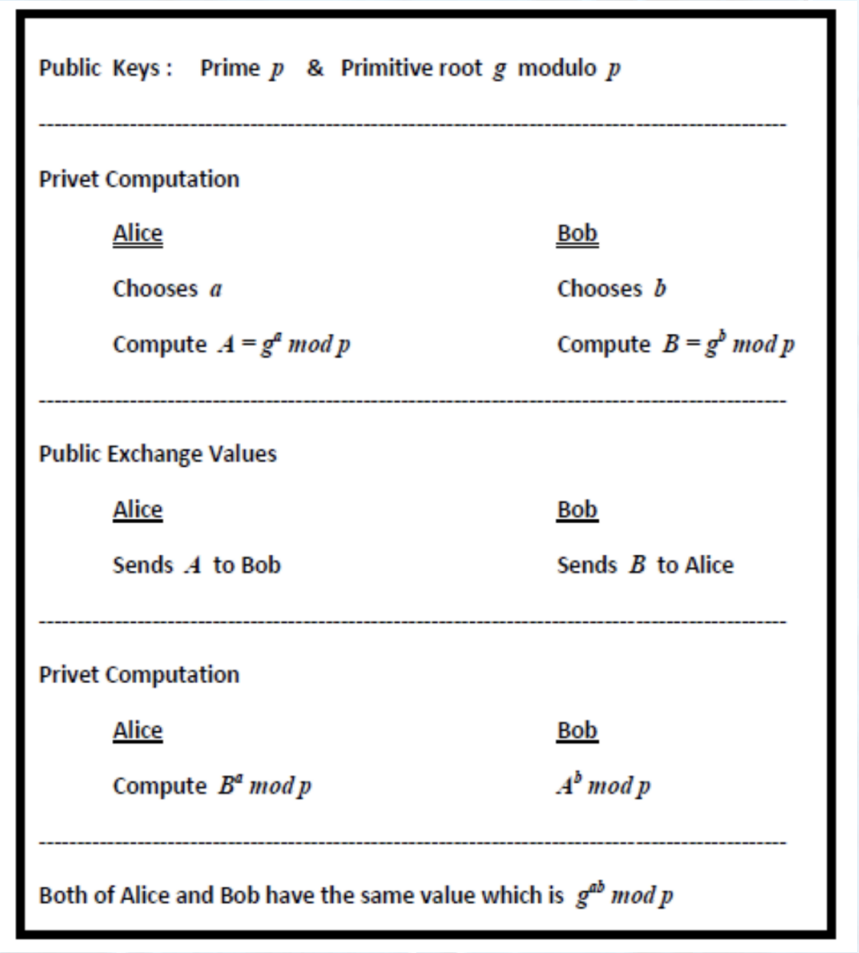

Figure 2 Diffie Hellman Key Exchange, *Adapted from Al-Siaq I.R., 2017, Research India Publications*

The attacker has the value of p, g, g^amodp and g^bmodp. The discrete logarithm problem is to compute a or b from gamodp or g^bmodp respectively.

5 Solution Methodology

5.1 Mathematical Concept

In the context of computer science, one-way functions are functions that are easy to compute (finding f(x) at a given x), but it is difficult to obtain the image of a random input. The words 'easy' and 'difficult' are in relation to a computing entity, 'easy' meaning cheap enough for a legitimate user, and 'difficult' meaning expensive for any malicious agent. So naturally, one-way functions provide a solid method to encode messages in cryptography.
Another important property is that, if f is a one-way function, then it is difficult to invert, but if we know the inputs and outputs, we can easily verify whether they are accurate or not.
We have two different numerical methods to approximate the root of the one-way function in both text and image cryptography. Further, we have compared the no. of iterations and CPU run time taken by each method.

5.2 A brief explanation of our solution approach

Many encryption algorithms already exist for Network Security. Our project aims to implement a symmetric encryption-decryption algorithm which is developed using steganography and the encryption and decryption is done using three different numerical methods. The encryption is done using the concept of one-way function and then finally steganography is used to encrypt it.

5.3 Algorithm for encryption and decryption

Algorithm for encryption

1. After reading the text file, we store it in a string.
2. We find the ASCII equivalent of each character (including spaces), and store it in array.
3. Using Diffie Hellman, we generate a secret key. The multiples of that secret key is the location of garbage values (if the secret key is 3, then garbage values are at indices 0,3,6,9,etc). Then, we add garbage values at specified locations.
4. Using one of the numerical methods, we find the solution for f(x)=ASCII value for each ASCII value, and store the solutions in another array.
5. This array is our ciphertext and is transmitted to the receiver.

Algorithm for decryption

1. Read the encrypted file.
2. Receiver obtains the ASCII values by simply calculating f(x), at non-garbage positions.
3. Put the values in the polynomial function and get the functional values (ASCII values).
4. Converting ASCII values back to characters, we have successfully sent the text.

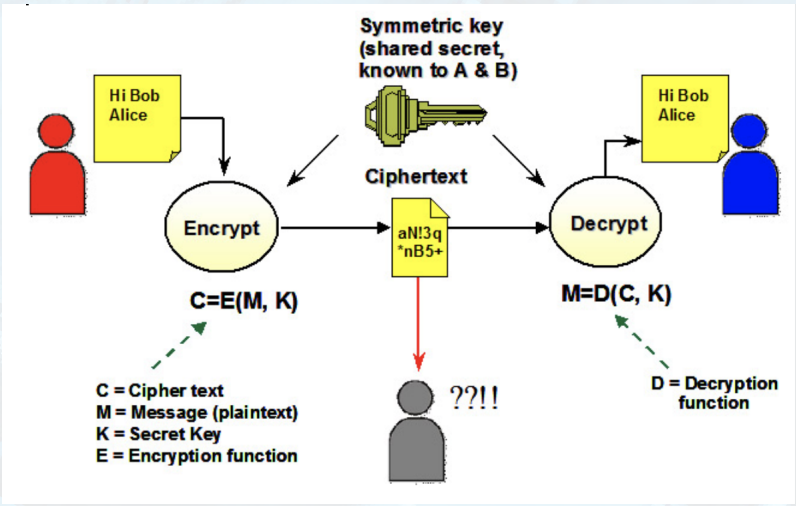

Figure 3 How cryptography works?

5.4 Solution methodology for TEXT cryptography

Following are the steps for encrypting and decrypting a text file:

1. Read the text file and convert into string
2. Deduce the ASCII value for each character of the string
3. Obtain a secret key using Diffie Hellman Algorithm. This key will be known to ONLY Alice and Bob
4. Add garbage at indices which are multiples of the secret key
5. For each ASCII value in the array of the ASCII values, find solution to f(x)=ASCII value and store it in an array
6. Send this array to Bob
7. Bob Gets the ASCII values by evaluating f(x), where x is each element of the received array
8. By converting back only the non-garbage values (through the secret key, which is known only to Alice and Bob), Bob has received the message.

5.5 Solution methodology for IMAGE cryptography

Following are the steps for encrypting and decrypting an image file

1. Select the image to be sent.
2. Image is first converted into grayscale
3. By calculating the average brightness of a single pixel, the grayscale image is converted into many symbols (such as'!','?', etc) . These symbols form our text file.
4. Now, we have a text file. From this point, the solution methodology for TEXT cryptography is followed to encrypt and send.

Example: If we input the first image shown below , we get its ASCII art equivalent as the second image

Figure 4 Real Image

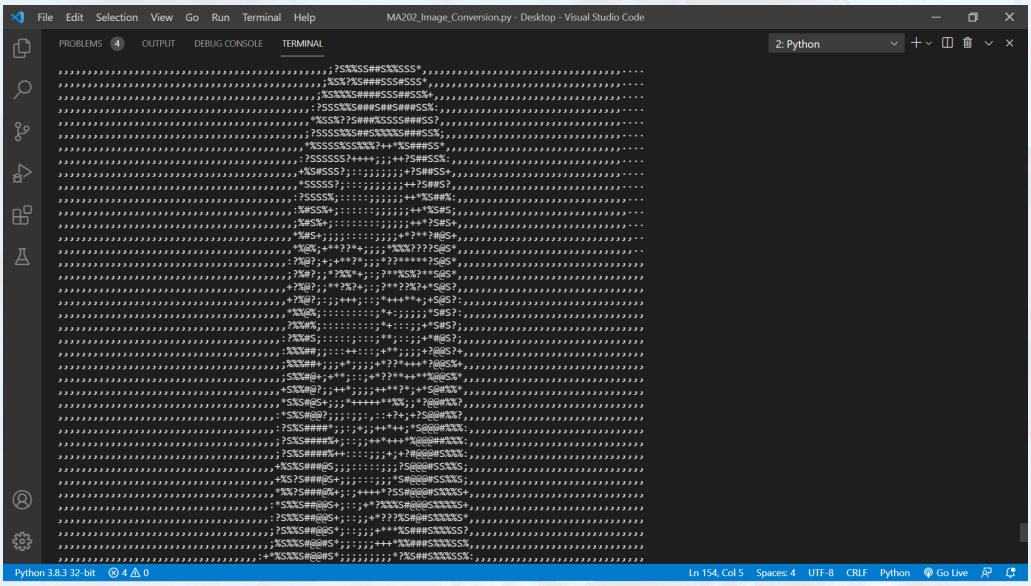

Figure 5 ASCII Equivalent

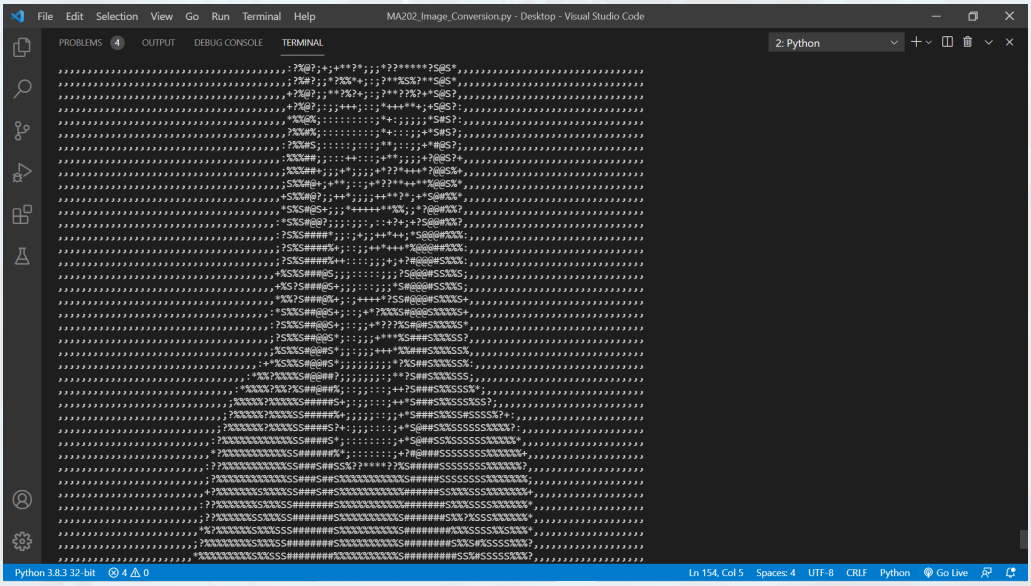

Figure 6 ASCII Equivalent

6 Numerical Solutions to solve for root of one-way function

6.1 Newton Raphson Method

Newton Raphson method is one of the most efficient and widely used methods for finding the root of a function. If one aassumes the initial root to be x_i , then the tangent from the point x_i , $f(x_i)$ denotes the improved estimate of the root. The figure given below best represents the statement given above.

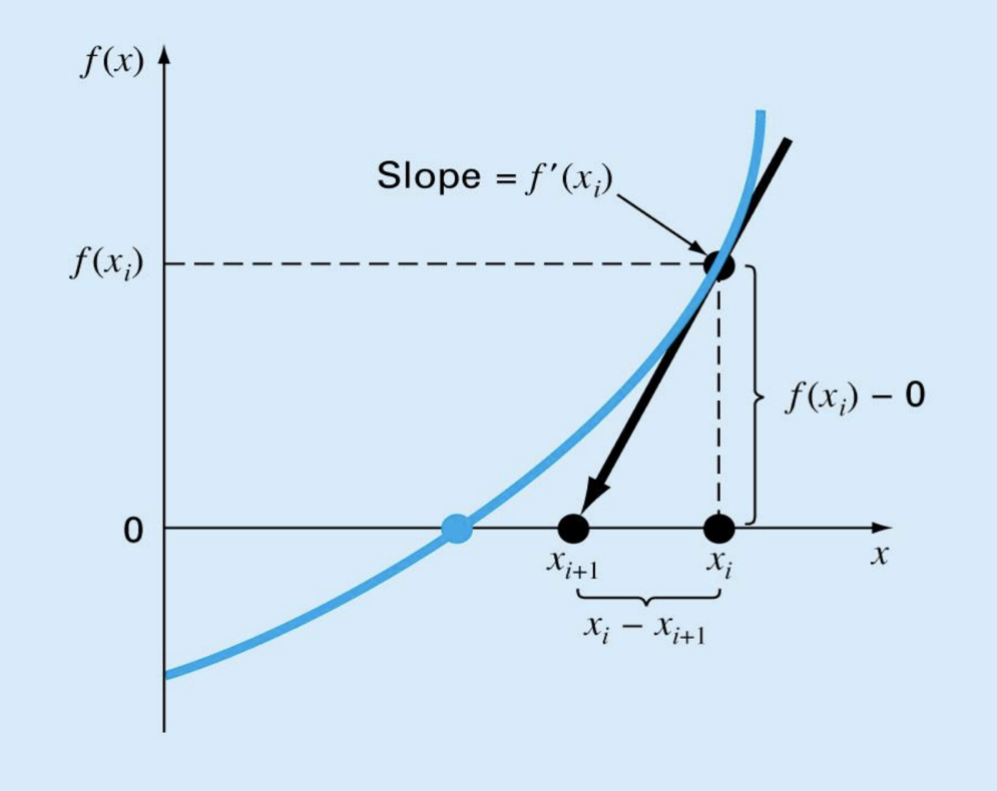

Figure 7 Graphical representation of Newton Raphson Method , *Adapted from Chapra Steven C., 7th edition*

$$f'(x_i) = \frac{f(x_i) - 0}{x_i - x_{i+1}} \tag{1}$$

$$x_{i+1} = x_i - \frac{f(x_i)}{f'x_i} \tag{2}$$

It is interesting to note that the Newton Raphson Method can be derived from the Taylor series expansion for $f(x_{i+1})$.

$$f(x_{i+1}) = f(x_i) + f'(x_i)(x_{i+1} - x_i) + \frac{f''(\xi)}{2}(x_{i+1} - x_i)^2 \tag{3}$$

When the function intersects the x-axis $f(x_{i+1}) = 0$

Equation (2) can be obtained by truncating the higher power terms in equation (3)

The Newton Raphson method is quadratically convergent i.e. the error is roughly proportional to the square of the previous error.

$$E_{t,i+1} \approx \frac{-f''(x_r)}{2f'(x_r)} E_{t,i}^2 \tag{4}$$

Where $E_{t,i}$ denotes the difference $x_r - x_{i+1}$, x_r is the true value

As mentioned above, Newton Raphson method is one of the most efficient methods however it also has some pitfalls.

The following cases show that extending the tangent can lead you farther away from the root even though you lie sufficiently close to it. The problem is amplified tremendously in the naughty case of zero slope.

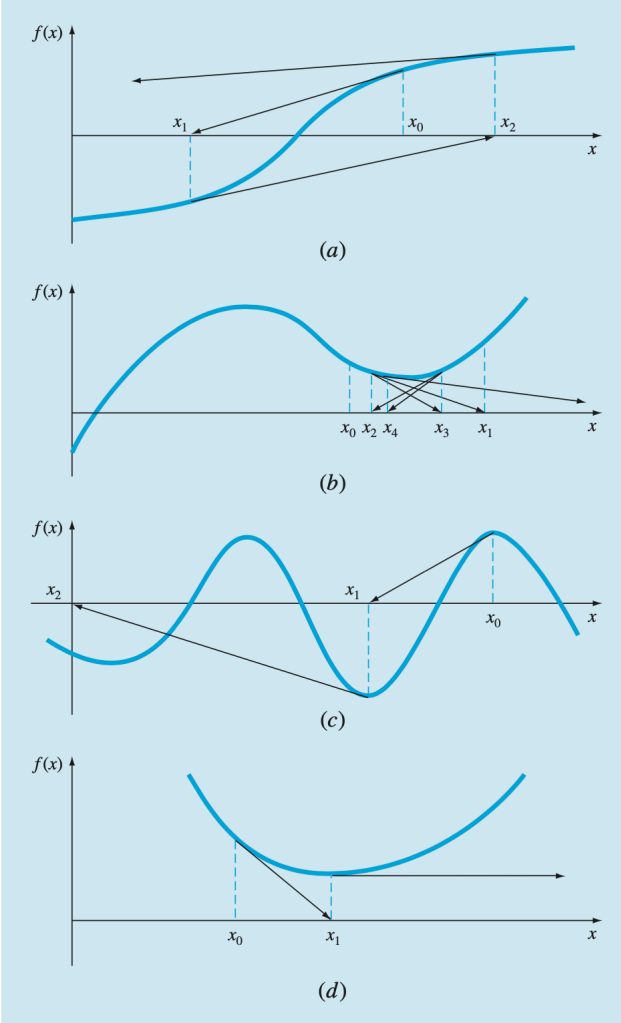

Figure 8 Cases where Newton Raphson Method shows poor convergence, *Adapted from Chapra Steven C., 7th edition*

6.2 Bisection Method

If f(x) is real and continuous in an interval from x_l tox_u and $f(x_l)$ has a sign opposite to that of $f(x_u)$, then there exists at least one root of $f(x)$ in the interval x_l to x_u .
The condition can be expressed mathematically as $f(x_l)f(x_u) < 0$ for all x belongs to $[x_l, x_u]$, then there exists at least one root of $f(x)$ in the interval.
The bisection method is a type of incremental search method where the interval is divided into half. If the function changes sign over an interval then we estimate the root of the function to be the midpoint of the interval and repeat the process until we get appropriately

precise results.

The method to calculate root is as shown below:
Step 1
Choose x_l and x_u such that $f(x)$ changes sign in the interval $[x_l, x_u]$.
Step 2
Estimate the root as

$x_r = \frac{x_l + x_u}{2}$

Step 3
If $f(x_l)f(x_u) < 0$, then the root lies in the lower subinterval. Set $x_u = x_r$ and return to step2.
If f$f(x_l)f(x_u) > 0$, then the root lies in the upper subinterval. Set $x_l = x_r$ and return to step2.
If $f(x_l)f(x_u) = 0$, then x_r is the root. Terminate the computation.

For the bisection method, the approximate error is found as follows:

$$\epsilon_a = \frac{|x_r^{new} - x_r^{old}|}{|x_r^{new}|} 100\% \tag{5}$$

From the following figure we can infer the curves for approximate error and true error. The curves do not coincide but it can be easily observed that the curve for approximate error does cover the downward trend of the true error curve. As evident from the curve that ϵ_a is always greater than ϵ_t, we can safely terminate the computation when ϵ_a falls below the specified error.

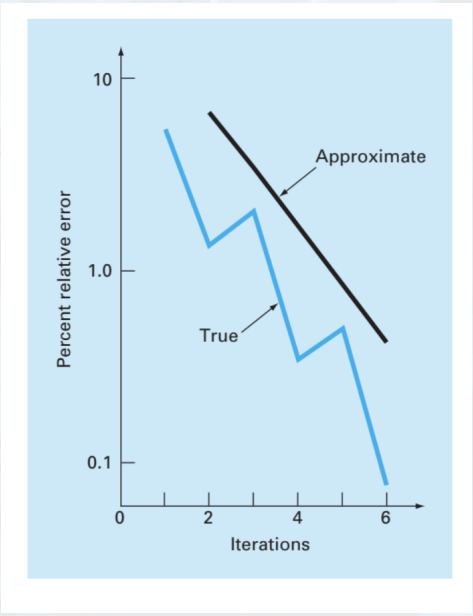

Figure 9 Errors for the Bisection method, *Adapted from Chapra Steven C., 7th edition*

The Bisection Method has several advantages in the form that it is easy to implement, always ends up in finding the root given the initial conditions are satisfied. We also have an idea

about the number of iterations in the algorithm if we are given the initial absolute error. There are however associated several limitations with it. Bisection method is generally slower than the other methods. Another major problem is that we should know about the interval where the function changes sign. This is sometimes not easy to find and takes a large amount of time. The algorithm for the bisection method takes no use of the information of the given values of $f(x_l)$ and $f(x_u)$. Say if the value of $f(x_l)$ is close to zero then the root will be closer to x_l. If we could have used this information then we could have been able to decrease the implementation time.

7 Results and Discussion

7.1 Observations

- For standard encryption/decryption functions, the algorithm for symmetric key works much faster than their asymmetric counterparts. It is designed for efficient processing of large datasets. It enhances speed and performance. Asymmetric cryptography is relatively expensive computationally.
- We find that Newton-Raphson method is a better numerical method as compared to the bisection method for approximate root finding.
  From our observations, while solving for the root of one way function, we found out that the Newton-Raphson method takes less number of iterations than the Bisection method for encrypting every 10 characters. This difference further magnifies as the number of characters increase. The same is shown below in the graph between the number of iterations and the number of characters for both the methods.
  However, the observations are bound to some uncertainty if the initial guess value in the Newton method is not taken smartly.

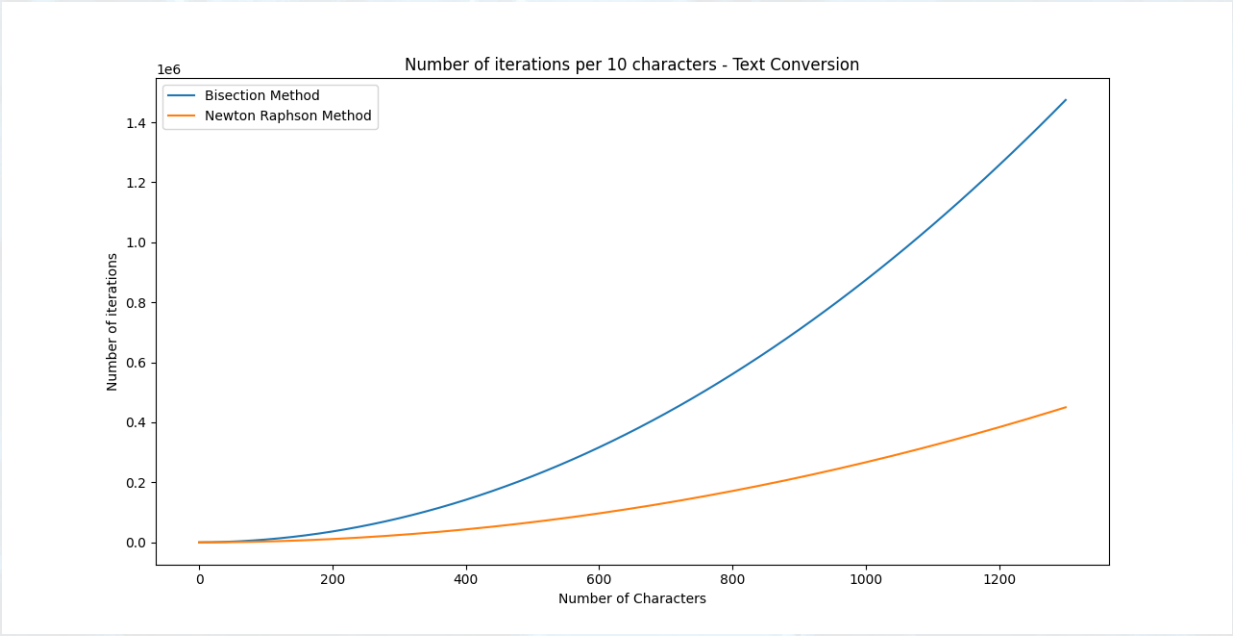

Figure 10 Graph between the number of iterations and number of characters, comparing the Newton-Raphson Method and Bisection Method ,for Text encryption

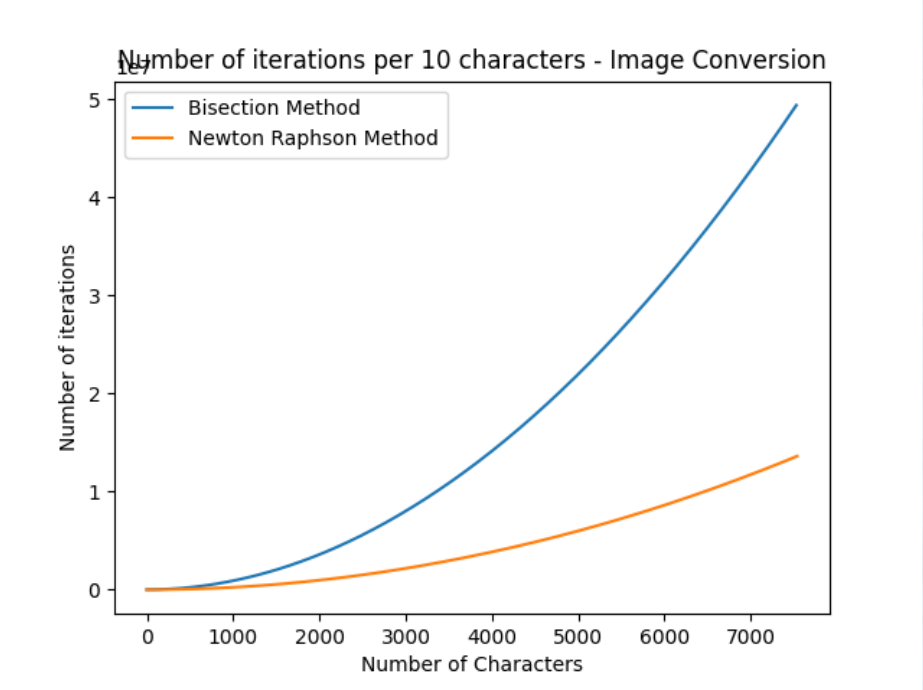

Figure 11 Graph between the number of iterations and number of characters, comparing the Newton-Raphson Method and Bisection Method, for Image encryption

- One of the most interesting observations we drew, was from the graph between the computation time and number of characters for both the methods.
  The graph comes out to be step-like in shape. For a smaller number of characters, the computation time is less for the bisection method might turn out to be less than Newton-Raphson Method. This complexity is inherent to the Newton-Raphson method because of the derivatives involved in its application. However, in the long run Newton-Raphson method is more efficient than the bisection method, as evident from the graph.

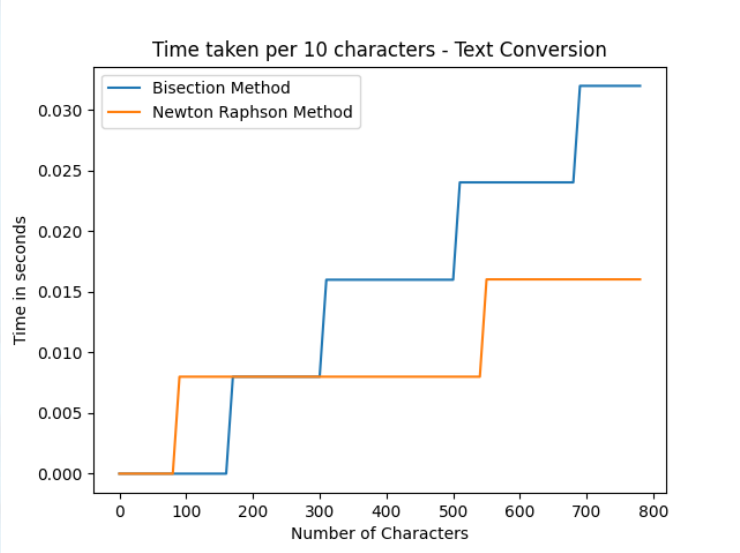

Figure 12 Graph between the Computation time and number of characters, comparing the Newton-Raphson Method and Bisection Method, for Text encryption

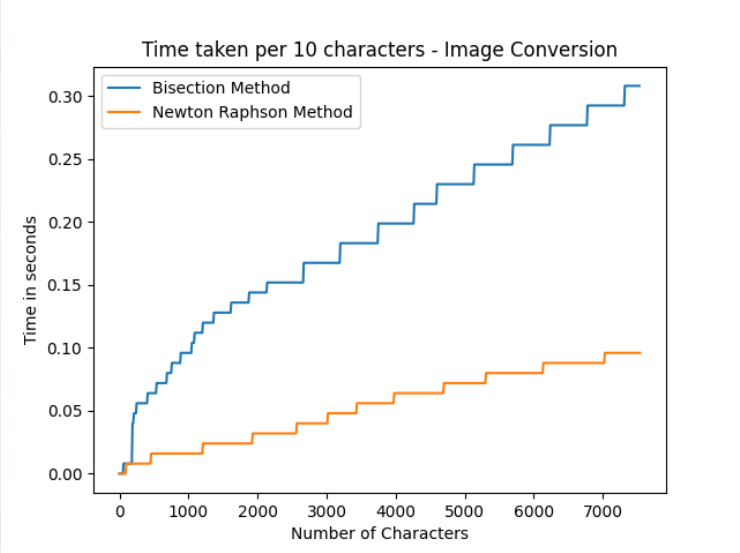

Figure 13 Graph between the number of iterations and number of characters, comparing the Newton-Raphson Method and Bisection Method ,for Text encryption

7.2 Innovations

- We have implemented a new encryption-decryption algorithm using the concept of Diffie Hellman key exchange and steganography together.
- Diffie Hellman allows 2 users to publicly exchange their chosen numbers and end up with a secret key, known only to the two of them.
- We used this secret key to put garbage values in our encrypted array of data and thus, implementing steganography. Garbage values (random integers) are inserted into the ASCII array at indices which are integral multiples of the secret key (if the key is 3, then garbage indices will be 0,3,6,9,12,etc).
- Since this key is known only to Alice and Bob, Bob knows the indexes at which garbage values are present so as to insulate the original message from them.
- To any hacker, this would seem like a simple text. But, it has garbage values to protect the real data. Hence, data has been concealed smartly among garbage values. This is a simple implementation of steganography.

7.3 Shortcomings

- We send the image by converting it to ASCII Art, encrypting and sending it, and then decrypted by the receiver. The receiver can best receive the grayscale image, but not the original, colored image. We did not put much thought into it as this would leave us venturing into an entirely new software and problem statement.
- We used a sample one way function for the implementation. Again, obtaining or generating one way functions is a tedious task and out of the scope of our problem.

8 Conclusions and Future scope

- We have proposed a new approach to network security by combining cryptography and steganography. The algorithms are simple and secure. Secret key conception is introduced using the idea of one-way function along with numerical methods. Finally a numerical method based encryption-decryption algorithm is developed using steganography to enhance the Network Security System.

- The comparative study suggests that Newton-Raphson method is more efficient than bisection method for generating an array of encrypted data for text or an image cryptography.

- Further, we can try implementing other numerical methods (such as, Fixed-point method, Muller's method, Brent's method, Dekker's method, etc.) and do a similar comparative study for them.

- The conversion of the ASCII image back to the original image is tedious and not feasible and may result in distortions of the image. At the most, the ASCII image can be mapped back to grayscale.

9 Bibliography

1. AL-Siaq, I. R. (2017). Public Key Cryptosystem Based on Numerical Methods. Global Journal of Pure and Applied Mathematics, 13(7), 3105-3112.
2. Anton, T., 2021. The need to manage both symmetric and asymmetric keys. [online] Cryptomathic.com. Available at: <https://www.cryptomathic.com/news-events/blog/the-need-to-manage-both-symmetric-and-asymmetric-keys> [Accessed 7 May 2021].
3. Ghosh, A., Saha, A. (2013). ANumerical METHOD BASED ENCRYPTION ALGORITHM WITH STEGANOGRAPHY. ACER, 2013, 149157.
4. In.mathworks.com. 2021. Polynomial roots - MATLAB roots- MathWorks India. [online] Available at: <https://in.mathworks.com/help/matlab/ref/roots.html> [Accessed 7 May 2021].
5. Kenekayoro, P. T. (2011). One way functions and public key cryptography. African Journal of Mathematics and Computer Science Research, 3(6), 213-216.
6. PyPI. 2021. image-to-Ascii. [online] Available at: <https://pypi.org/project/image-to-Ascii/> [Accessed 7 May 2021].

Image References

https://securitycerts.org/images/symmetric-alice-bob.jpg
https://www.shutterstock.com/photos