DESIGN AND IMPLEMENTATION OF A TOOL TO HELP COMPUTER FORENSICS INSTRUCTOR TO DEMONSTRATE COMMON DATA HIDING TECHNIQUES

GRADUATE PROJECT REPORT

Submitted to the Faculty of
The School of Engineering & Computing Sciences
Texas A&M University-Corpus Christi
Corpus Christi, TX

In Partial Fulfillment of the Requirements for the Degree of
Master of Science in Computer Science

by
Anudeep R. Kandi
Fall 2012

Committee Members

Dr. Mario Garcia ______________________________
Committee Chairperson

Dr. Dulal C. Kar ______________________________
Committee Member

Dr. David Thomas ______________________________
Committee Member
ABSTRACT

Data is one of the most important asset of any organization. It must be kept protected and secure from unauthorized access. Data hiding is one of the best practiced techniques in the industry for business e-transactions and maintaining information security. There are different data hiding techniques available to secure data from the intruder or hacker. In this research an application is developed with the intension to hide data that consists of the proposed hybrid model ENHARM and existing techniques like Steganography, different encryption methodologies and disk encryption techniques. All the methodologies that include the proposed application and existing techniques in a single web application that provides easy user access. The proposed data hiding technique provides Confidentiality, Integrity and Authentication. The approach used in the application is different from the existing techniques. The proposed model ENHARM is implemented by using the randomly generated 6X6 matrix and well known standard encryption algorithms like AES, blowfish, etc. to encrypt the plain text. The functionality of the ENHARM hybrid model is clearly described in the architecture presented later in this document. The data hiding application developed is user friendly and works efficiently. One of the main advantage of this project is that it can be used for faculty and students learning data hiding techniques.
# TABLE OF CONTENTS

Abstract ..................................................................................................................................... ii

Table of Contents .................................................................................................................. iii

List of figures .......................................................................................................................... v

1. Introduction ......................................................................................................................... 1
   1.1 Importance of Data ........................................................................................................ 1
   1.2 Data Hiding ................................................................................................................ 1
   1.3 Importance of Data Hiding ......................................................................................... 1
   1.4 Reasons for Data Hiding ............................................................................................. 2
   1.5 History ........................................................................................................................ 3
      1.5.1 Data Hiding ......................................................................................................... 3
      1.5.2 Types of Data Hiding Techniques ....................................................................... 4
         1.5.2.1 Steganography ............................................................................................ 4
            1.5.2.1.1 Pure Steganography ............................................................................. 5
            1.5.2.1.2 Secret Key Steganography ................................................................. 5
            1.5.2.1.3 Public Key Steganography ................................................................. 6
         1.5.2.2 Cryptography ............................................................................................... 7

2. Existing Systems ................................................................................................................. 8
   2.1 Data Hiding Technique in Images .............................................................................. 8
   2.2 Data Hiding Technique in Audio Frequencies ........................................................... 9
   2.3 Data Hiding Technique in Videos .............................................................................. 9
   2.4 Substitution Cipher .................................................................................................... 10
   2.5 Symmetric Cryptography ......................................................................................... 10
2.6 Asymmetric Cryptography .......................................................................................11
2.7 Hybrid Model for Data Hiding .............................................................................12
2.8 Internal concepts of Microsoft Windows version for file systems and disks ..........13
2.9 Aim and Objectives .............................................................................................14
3. System Design ........................................................................................................15
3.1 Framework ............................................................................................................15
3.2 System Architecture .............................................................................................15
4. Implementation of Tool ..........................................................................................17
4.1 Implementation of Common Data Hiding Techniques ...........................................17
  4.1.1 Steganography using Image ...........................................................................18
  4.1.2 Symmetric Encryption ..................................................................................21
  4.1.3 Changing File Extension .............................................................................22
  4.1.4 Caesar’s Cipher ............................................................................................22
  4.1.5 Vigenere Cipher .........................................................................................24
  4.1.6 One Time Padding ........................................................................................25
  4.1.7 File Compression ..........................................................................................26
  4.1.8 Lock the Personal Folder .............................................................................27
4.2 Implementation of Proposed Model ......................................................................29
  4.2.1 Functionalities of the Proposed Data Hiding Model (ENHARM) ...................30
    4.2.1.1 Encryption .............................................................................................31
      1. Input .............................................................................................................31
      2. Encryption Process ......................................................................................31
      3. Output .........................................................................................................33
4.2.1.2 Decryption..........................................................................................33

1. Input.............................................................................................................33

2. Decryption Process.....................................................................................34

3. Output..........................................................................................................36

4.2.2 Example for the Explanation of Proposed Model (ENHARM).................36

4.2.2.1 Encryption Module.............................................................................36

4.2.2.2 Decryption Module............................................................................38

5. Testing and evaluation ..................................................................................40

6. Conclusion ...................................................................................................55

7. Bibliography and References.......................................................................56
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Functionality of Pure Steganography</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Architecture Illustrating secret key Steganography</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>Architecture for the Integrated data hiding Tool</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>Home Page for the Implemented Tool</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>Existing data hiding techniques</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>Represents the steganography functionality</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>Steganography using Images</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>Symmetric Encryption Techniques</td>
<td>21</td>
</tr>
<tr>
<td>9</td>
<td>Changing File Extension Page</td>
<td>22</td>
</tr>
<tr>
<td>10</td>
<td>Implemented Caesar's Cipher</td>
<td>23</td>
</tr>
<tr>
<td>11</td>
<td>Vigenere Cipher Table</td>
<td>24</td>
</tr>
<tr>
<td>12</td>
<td>Vigenere Cipher Encryption Decryption Page</td>
<td>25</td>
</tr>
<tr>
<td>13</td>
<td>One time Padding foe Encryption and Decryption</td>
<td>26</td>
</tr>
<tr>
<td>14</td>
<td>File compression technique</td>
<td>27</td>
</tr>
<tr>
<td>15</td>
<td>Hiding the Personal Files into a Folder</td>
<td>28</td>
</tr>
<tr>
<td>16</td>
<td>Encryption Page for the Proposed Hybrid Model</td>
<td>29</td>
</tr>
<tr>
<td>17</td>
<td>Decryption Page for the proposed Hybrid Model</td>
<td>30</td>
</tr>
<tr>
<td>18</td>
<td>Proposed Architecture for Encryption Process Module</td>
<td>32</td>
</tr>
<tr>
<td>19</td>
<td>Proposed Architecture for Decryption Process Module</td>
<td>34</td>
</tr>
<tr>
<td>20</td>
<td>New User Registration Page</td>
<td>41</td>
</tr>
<tr>
<td>21</td>
<td>Login Page where the Registered user enters the Username and Password</td>
<td>42</td>
</tr>
<tr>
<td>22</td>
<td>Home page of the Implemented Tool</td>
<td>42</td>
</tr>
</tbody>
</table>
Figure 23: The Proposed Hybrid model for Encrypting and Decrypting the Text ........43

Figure 24: Encrypt Phase of the Proposed Hybrid Model ........................................44

Figure 25: Decrypt Phase of the Proposed Hybrid Model ........................................45

Figure 26: Browsing a File from the File System ...................................................46

Figure 27: Entering Text in the Text Area provided ............................................46

Figure 28: Entering the Text to Check the Md5 Functionality ..................................47

Figure 29: Generating a Symmetric Key ...............................................................48

Figure 30: A Specified Algorithm is Selected .......................................................49

Figure 31: After the Symmetric Encryption Algorithm is processed ......................50

Figure 32: Applying Random Matrix Functionality to the Secret Key ....................51

Figure 33: Displaying the Matrix after the Random Matrix functionality is Applied .....52

Figure 34: Entering the Encrypted text and Encrypted secret key in text boxes ..........53

Figure 35: Displays the Original Text after Decryption Process ............................54
1. INTRODUCTION

1.1 Importance of Data:

Data is an important asset for every organization. Data, plays a major role in decision making. It can be stored, processed and transferred between different entities. Information can be exchanged between any individuals or organizations. Data, being a valuable asset for organizations, must be securely maintained [3]. If the data is exchanged between departments or organizations, the sender and the receiver of the data should only have access to data. There shouldn’t be any involvement of intruder access to the data.

1.2 Data Hiding:

Data is a valuable resource for many organizations. It should not be accessible to unauthorized users. So, the data must be kept hidden from everybody except from the sender and receiver. The term data hiding means that confidential data is to be in encrypted form or the original message is to be hidden in some other media format[5]. There are different application software's available to encrypt data to hide information from unauthorized access. By using those applications, the original data content can be made available only to the sender and receiver. Sensitive information should be encrypted before it is transferred through a medium. The original information is retrieved by decrypting the encrypted information.

1.3 Importance of data hiding:

Now-a-days, it is not safe to exchange confidential data through any transmission medium without providing security to the message. Data hiding techniques overcome this problem by overriding the original message with encrypted information. There are different
data hiding techniques that are available like steganography, cryptography etc. These
techniques help in restricting access to important information to the third person and provide
access only to the authorized personnel [7].

The characteristics of data hiding are as follows:

Robustness: This is one of the major characteristics of data hiding. It states that the
information is not lost when the original message is encrypted. It shows that data hiding
techniques are very robust

Safety: It is very difficult for intruder to retrieve the secret information from the encrypted
text. The algorithms used are efficient and capable in providing security to the data.

Self-recovery ability: Data is not damaged or lost when receiver is trying to retrieve the
original information from encrypted text.

1.4 Reasons for Hiding Data:

The personal files and private information can be kept secret by using Data
Hiding Techniques. The private information includes e-commerce transactions of
organizations and company secrets, etc. It avoids misuse of data from the intruders. Data
hiding plays a major role when the confidential information is kept at service providers. It
also helps to prevent data from damage.
1.5 History:

1.5.1 Data hiding:

Now-a-days the information is exchanged between the individuals or organizations is through electronically. Senders usually choose internet as a medium to send confidential information to the receiver [4]. There might be a chance for hackers to retrieve the message while in transmission. To hide the secret message from intruder, the original message must be replaced with an encrypted message by using different standard algorithms and methods. To encrypt a message, there are various data hiding techniques available. If the message is encrypted, it is difficult for the intruder to retrieve the original message. The hacker must be aware of all the encryption techniques and also the approach to extract the correct information from hidden data. There are different kinds of hiding techniques present with various approaches among them; steganography and cryptography are the major methodologies. *Steganography is an art of hiding original messages in any digital media.* In Greek, Steganography refers to "concealed writing" which means hiding data [1]. Basically steganography hides the data in the form of images, audio and videos. It uses digital media for hiding data. It embeds original data into any digital media like image, audio, etc. and is then sent to the recipient. The receiver uses the same application for extracting hidden information.

Cryptography is another technique used for hiding data. In Greek, cryptology refers to “hidden secret”. New age cryptography in computer science is mainly based upon mathematical theories [2]. There are different algorithms for encryption of data for data hiding. Using cryptography, information is sent while ensuring Confidentiality, Integrity, and Authentication (CIA). Various algorithms such as AES(Advanced Encryption Standards),
and DES (Data Encryption Standards) etc. are used for encrypting original data. Besides these cryptographic and steganographic techniques, there are different encryption techniques like one time padding, Vigenere cipher, Caesar cipher (substitution methods) etc.. All these are different kinds of data hiding techniques that are implemented for hiding the message.

1.5.2 Types of Data Hiding Techniques:

1.5.2.1 Steganography:

Steganography is a method used for hiding data from unauthorized access. Only sender and receiver can know the existence of data. The data or file will be embedded into another format like an image or audio. The original file will look like an image or an audio file. This is a different form of hiding data from intruder. It makes use of digital media for embedding the original file into the image. The functionality of hiding data in steganography and cryptography are different. While steganography is applied, the data is embedded into another file; the secret message will disappear and can’t be viewed. In case of cryptography, the secret message is replaced by an encrypted text.

In steganography technique, data is securely transferred through a medium by embedding it into duplicate data which can be in the form of image, audio or video files [8]. There are different steganography types

1. Pure Steganography
2. Secret key Steganography
3. Public key Steganography
1.5.2.1.1 Pure Steganography:

This is the basic form of steganography, where information is embedded into another article without any key [8]. This method provides secrecy to the information. This method is not included with any kind of key technology such as embedding the message into an image or audio file in order to protect message from unauthorized access.

Figure 1 represents the functionality of the pure steganography, the original message is embedded into an image file using embedding functionality and a stego message is created. The stego message by using the De-embedding functionality retrieves the original message at the receivers side.

Figure 1: Functionality of Steganography [8]

1.5.2.1.2 Secret Key Steganography:

This method is another form of steganography technique which uses the same procedure of pure steganography by including the key functionality [8]. An individual key is provided at the time of embedding a message into an image or audio
files. While De-embedding, the same secret individual key is provided for extracting original message or data from an image or audio files.

Figure 2 represents the functionality of the secret key steganography. It is very similar to the pure steganography by including the key methodology.

![Architecture illustrating secret key steganography](image)

**Figure 2: Architecture illustrating secret key steganography [8]**

1.5.2.1.3 Public Key Steganography:

This method of steganography is very similar to the previous secret key steganography method. This technique makes use of two keys. One is private key or secret key and another one is public key [8]. All these two keys are mathematically linked. One key is used when embedding the original text and another key is used for de-embedding or decoding the plain text from embedded digital media. Among the two keys one is kept private and another key is public.
1.5.2.2 Cryptography

Cryptography or Cryptology is a Greek word which means “secret”. It is the study of various techniques confined to data hiding from the third parties. Cryptography concepts are related to various security terms such as confidentiality, integrity and authentication [2]. It makes use of computer technology to protect data from unauthorized or third parties and also in various applications like e-commerce, ATM cards etc. Different forms of cryptography techniques came into existence because of increase in usage of various computer communications. When data is transferred through an insecure medium, there is a need for cryptography technologies. In terms of cryptography terminologies, Encryption, Decryption is used.

**Encryption**: The process of transforming data from plain text into cipher text using an algorithm is called Encryption.

**Decryption**: Process of transforming information from cipher text to plain text is called Decryption

**Note**: Cipher text is also called encrypted text. This occurs after the encryption algorithm is processed.
2. EXISTING SYSTEMS

2.1 Data Hiding Technique in Images:

For hiding data in images, there exists a technique called Steganography. Steganography is a field of Computer Science for hiding data such that data can't be readable by anyone except the sender and the receiver. There are many ways to hide data; among them, data hiding in images is one of the well-known methods. In this method, the secret message is embedded into an image by using embedding functionality from the concepts of data structures [10]. The secret message bytes are embedded into a color image randomly. Retrieving the secret message will be difficult for the attacker as data is hidden in image using this technique. A Stego-Image will be created and sent to the receiver for retrieving the original secret message using the same methodology. Later, several approaches came into existence by enhancing the existing data hiding method. Author Kuo [9] proposed a technique called reversible data hiding based on the division of blocks in image for hiding the secret message in images. Data is divided into bytes and each byte is embedded into each and every pixel of the image for hiding. There exists a counter technique called steganalysis for detecting whether the steganography technique is used for data hiding. To overcome steganalysis, the authors are proposing new techniques that can't be detected by steganalysis. In this version, combination of steganography with cryptography is the new technique to hide confidential data. There are other models like bit plane spicing algorithm using the intensity of the pixel, halftone image algorithm helps in hiding large amount of data contents, fibonacci decomposition. Due to above enhancements in the models and increased efficiency in methodology, attackers face difficulty in retrieving the data.
2.2. Data Hiding Technique in Audio frequencies:

This method of hiding data in audio channels is based upon the concept of steganography. The working principle of hiding data in audio frequencies is similar to that of hiding data in images. There are different methods proposed for embedding data into audio channels. In the first model proposed [12], the data is hidden in Least Significant Bit (LSB) of audio channels instead of directly embedding data into audio frequencies. This method of hiding data using LSB will increase the difficulty level in retrieving data for intruders.

Another model of hiding data is an enhancement to the previous approach. XOR functionality is performed to the existing LSB method [13]. The LSB’s are XORed for hiding the sensitive data and making it difficult to be detected. Later, several approaches and new algorithms are proposed in enhancing the existing models to hide data using audio channels.

2.3. Data Hiding Technique in Videos:

This is another form of steganography technique used for hiding data through videos where a video is sent to receiver to retrieve the hidden data. In this case, it would be difficult for the intruder to get access to the hidden data in the video. Author Li, [14] proposed a model for hiding data by using H.264 encoded video sequences. The main advantage of this model is to increase the quality of the video when the data is encoded. The secret message is to be extracted from the encoded video correctly without any loss or damage. Usage of this model not only provides security to the data, but also makes it difficult for the intruder to detect the hidden data. A new model [15] is proposed by enhancing the
existing video compression techniques called context adaptive variable length coding. All these existing models help in increasing the efficiency of hiding data through a video format.

2.4. Substitution cipher:

The substitution cipher is one of the encryption techniques used to convert the sensitive data to a cipher text. It converts letter after letter or pair of letters or a mixture of letters to cipher text. The receiver will do reverse process to get the plain text back. There are different terms in the substitution cipher; like if the user encrypts the text letter by letter, it is called simple substitution cipher [17]. If the user encrypts a group of letters, it is called polygraphic substitution cipher. The usage of these many models will depend upon the user requirements. The examples of these kinds of encryption models are caesar cipher, vigenere cipher, etc.

2.5. Symmetric Cryptography:

Symmetric cryptography is also called secret key cryptography. It uses the same secret key for both encryption and decryption. If sender and receiver want to exchange the data, they should first decide which algorithm to use. The sender uses a key to encrypt the plain text to unreadable format and later the receiver also uses the same key for decrypting from cipher text to the original plain text. The difficulty in this model is sharing the key between sender and receiver [17].

Encrypting the data using symmetric cryptography is based on two methods like stream cipher and block ciphers. Stream cipher functionality is to encrypt data one bit at a time. In block cipher, encryption of data is done by considering a block of bits at a time and a
secret key is provided. Decryption also undergoes same functionality. There are different algorithms proposed for this model like Data Encryption Standards (DES), Advanced Encryption Standards (AES), 3DES for encryption and decryption.

This model will share same secret key for both encryption and decryption and is called shared key cryptography or shared secret cryptography. Mathematically, it is represented as Dsk(Esk(M))=M.

Using this model the authentication can be provided using the same key and can be encrypted quickly. If the key is revealed, it will become very easy for the intruder to decrypt the message. Keys must be changed regularly to avoid damage to data.

2.6. Asymmetric Cryptography:

Asymmetric Cryptography is also called public key cryptography. It uses two different keys for encrypting and decrypting data. One kind of key is called public key which is distributed to everyone and another key is called private key which is evaluated mathematically for creating. Whereas in symmetric cryptography, there will be only one key shared among the sender and receiver. The secret key should be kept private and must be known only to the authorized user.

There should be two keys determined which are mathematically related and one key is not sufficient to decrypt the original text. One key is used to encrypt the plain text to cipher text and the other key is used to decode the cipher to plain text. The standard algorithms for asymmetric cryptography are RSA, Diffie-Hellman, Elgamal, etc. All these algorithms are based on the mathematical logics to generate the keys from which only a few public key algorithms are secure and practical.
There are some disadvantages when using asymmetric cryptography. Evaluation of a key by using mathematical logics is very complicated to understand. So, the time consumption for encrypting the text is high in this method.

2.7. Hybrid model for data hiding:

To increase the efficiency of encryption, when sensitive data is shared among the organizations, PGP (Pretty good Privacy) proposed a hybrid model by combining the advantages of symmetric and asymmetric cryptography. The combination of cryptography and steganography techniques also comes under a hybrid model [18]. PGP first compresses the plain text, which helps in saving time in transmission and maintaining disk space and also strengthens the security of the data while cracking the cipher text. A one-time randomly generated session key also called shared key is provided to encrypt the plain text to cipher text using standardized algorithms. The session key along with compression methodology increases the resistance and security to the cipher text. Then the session key is also encrypted using the public key cryptography concepts. The encrypted session key along with the cipher text is sent to the receiver. Initially receiver will decrypt the session key and later cipher text is converted to plain text by using the decrypted session key. This is the working principle of the PGP or hybrid model of cryptography.

Limitations of existing hybrid models:

1. The existing hybrid models are very complex to process and difficult to understand.
2. It requires a lot of man power for making evaluations.
3. More time is consumed when encrypting the text using hybrid models.
2.8. Internal concepts of Microsoft Windows version for file systems and disks:

These are some of the data hiding methods that are available in the Microsoft windows operating system to hide data in local files systems. Alternating Data Stream, cipher, changing file extension, etc. are some of the most common techniques for hiding in windows OS [7]. These techniques are used only for hiding files in the local file system or hard drives. These files can't be sent or opened in any another PC. In steganography and cryptography, hidden or encrypted files can be sent to the receiver and can be opened using similar methodology that is used by sender. These applications are executed in the command prompt of windows operating system.
2.9 Aim and Objectives of the project

Research Objective:

The objective of this project is to develop a web based tool by integrating the most common data hiding techniques into a single application and proposing a hybrid model ENHARM for hiding data.

Objectives:

- Developing a web based tool by integrating different data hiding techniques into a single application.
- Providing an application for the students to learn different data hiding techniques
- Developing a hybrid model ENHARM based on random matrix for hiding data, that consumes less time for encryption and involves less complexity.
- Testing the functionality of the project.
3. SYSTEM DESIGN

3.1 Framework:

The tool is developed by considering some common data hiding techniques that are in use. The functionality and features of these techniques are different from each other. It is provided with an user friendly interface. The developed tool is a web application used for hiding confidential data. The tool also consists of a proposed hybrid model ENHARM. The security is provided to the information in multi layers by using MD5, symmetric encryption algorithms like AES, DES and randomly generated matrix. By implementing all these major techniques into a single application, the user can access them at a same place.

3.2 System Architecture:

Figure 3 represents the architecture of the data hiding tool that is implemented. The user should enter the correct login details for accessing the application and was able to see the home page. The "common data hiding techniques" denotes the existing techniques that are used for hiding the data like steganography using image, cryptography standard algorithm AES, DES some encryption methods like caesar cipher, vigenere cipher, one time padding, etc. The proposed model ENHARM is a hybrid data hiding model to hide the confidential data.
Figure 3: Architecture for the Integrated Data hiding Tool
4. IMPLEMENTATION

The Implementation of Data hiding tool consists of two modules.

1. Implementation of the common data hiding techniques.
2. Implementation of the proposed model (ENHARM)

4.1 Implementation of Common Data Hiding Techniques:

Figure 4 represents the homepage of the tool consisting of implemented proposed model and some of the existing commonly used data hiding techniques. The left panel of the page displays the different methods like the proposed ENHARM model and existing hiding techniques.

Figure 4: Home page for the Implemented Tool

The Implemented tool consists of Proposed Hybrid model and also some of the most common data hiding techniques. This common techniques includes the steganography
using image, some encryption techniques like caesar cipher, Vigenre cipher, etc and changing file extensions.

Figure 5 represents the implemented user interface for existing data hiding techniques including steganography using image, Symmetric Encryption, etc. The detail explanation of each and every method is explained

![Figure 5: Existing Data Hiding Techniques](image)

### 4.1.1 Steganography using Image:

It is a method of hiding secret information into a digital media. Hiding information can be in various forms like images, audio, video, text. In the developed application, hiding the secret messages is through images. By embedding the secret message
into a cover medium, it is very difficult for the third person to detect the existence of message in an image, that stego-image is sent to the receiver. Functionality of hiding data into images is shown in Figure 6. cover image is the file where the user wants to hide a secret message using Embedding functionality, and a stego image is created. The secret message is retrieved from the stego image using the De-embedding functionality at receivers side.

Figure 6: Representing the Steganography functionality

Figure 7 represents the customized Steganography method using images. This method will hide and unhide the data using the embedding and de-embedding functionality. The secret message can be entered in text area provided or another option to upload a file. The file or text will be hidden into an image that is selected and is sent to the receiver. By doing the reverse process, the receiver can open the original file or text. For hiding a file the
user should provide the secret message and cover image file, for retrieving the secret message the user must upload a stego image file in the option provided.

Figure 7: Steganography using Images
4.1.2 Symmetric Encryption:

The Symmetric encryption is a part of cryptography that encrypts and decrypts the secret message. There are different standard algorithms exists for encrypting the message like AES (Advanced Encryption Standards), DES, Blow fish etc. The sender and receiver shares the same secret key for encrypting and decrypting the message.

Figure 8 represents the Symmetric Encryption Technique using the AES algorithm. The sender and receiver will share the same secret key for encryption and decryption. This methods is implemented by using the standard AES algorithm in a customized way using the PHP scripting language. The functionality of this method is mainly divided into two main modules, they are encryption and decryption. At the time of encryption the user should provide the text or a file and the password, while decryption the user should enter the same secret key to extract the original content.
4.1.3 Changing the File Extension:

This is a data hiding technique by changing the properties of a file, like file extension, date etc. It is difficult to open a file unless the correct file format is known.

Figure 9 represents the changing file extension page, where the original file in the .doc format extension is changed to .PDF or .JPG, or .MP3 format. The users only knows the correct file format. The file doesn't open if it is not in the correct format extension. This is one form of hiding methodology in the local file systems.

![Figure 9: Changing File Extension Page](image)

4.1.4 Caesar's cipher:

Caesar cipher is also called as shift cipher. It is an ancient form of cryptography. The letters in the plain text are replaced by another fixed letter for producing a cipher text. This is a substitution form of encryption technique.
Figure 10 represents the Caesar's cipher or the shift cipher. It is a type of substitution cipher in which each letter is replaced by another letter of the plain text for conversion into cipher text. The user must enter the text in the text area provided for encrypting or decrypting the message.

For example

It works like A is replaced by C, B is replaced by D, the above example is for shift 2, that represents the character is replaced by another character that comes after 2 places. This undergoes till the plain text ends.

Figure 10: Implemented Caesar's cipher
4.1.5 Vigenere Cipher:

Figure 12 represents the Vigenere cipher for encrypting the sensitive information. This is an old form of encryption technique for hiding the data. It is very easy to understand and implement. Based upon the figure 11, the plain text is compared to the key provided and a cipher text is generated. The reverse process is to decrypt the cipher text to plain text using the same keys. This method of encryption and decryption is called as Vigenere cipher.

<table>
<thead>
<tr>
<th>A B C D E F G H I J K L M N O P Q R S T U V W X Y Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>A B C D E F G H I J K L M N O P Q R S T U V W X Y Z</td>
</tr>
<tr>
<td>B C D E F G H I J K L M N O P Q R S T U V W X Y Z A</td>
</tr>
<tr>
<td>C D E F G H I J K L M N O P Q R S T U V W X Y Z A B</td>
</tr>
<tr>
<td>D E F G H I J K L M N O P Q R S T U V W X Y Z A B C</td>
</tr>
<tr>
<td>E F G H I J K L M N O P Q R S T U V W X Y Z A B C D</td>
</tr>
<tr>
<td>F G H I J K L M N O P Q R S T U V W X Y Z A B C D E</td>
</tr>
<tr>
<td>G H I J K L M N O P Q R S T U V W X Y Z A B C D E F</td>
</tr>
<tr>
<td>H I J K L M N O P Q R S T U V W X Y Z A B C D E F G</td>
</tr>
<tr>
<td>I J K L M N O P Q R S T U V W X Y Z A B C D E F G H</td>
</tr>
<tr>
<td>J K L M N O P Q R S T U V W X Y Z A B C D E F G H I</td>
</tr>
<tr>
<td>K L M N O P Q R S T U V W X Y Z A B C D E F G H I</td>
</tr>
<tr>
<td>L M N O P Q R S T U V W X Y Z A B C D E F G H I J</td>
</tr>
<tr>
<td>M N O P Q R S T U V W X Y Z A B C D E F G H I J K</td>
</tr>
<tr>
<td>N O P Q R S T U V W X Y Z A B C D E F G H I J K L</td>
</tr>
<tr>
<td>O P Q R S T U V W X Y Z A B C D E F G H I J K L M</td>
</tr>
<tr>
<td>P Q R S T U V W X Y Z A B C D E F G H I J K L M N</td>
</tr>
<tr>
<td>Q R S T U V W X Y Z A B C D E F G H I J K L M N O</td>
</tr>
<tr>
<td>R S T U V W X Y Z A B C D E F G H I J K L M N O P</td>
</tr>
<tr>
<td>S T U V W X Y Z A B C D E F G H I J K L M N O P Q</td>
</tr>
<tr>
<td>T U V W X Y Z A B C D E F G H I J K L M N O P Q R</td>
</tr>
<tr>
<td>U V W X Y Z A B C D E F G H I J K L M N O P Q R S</td>
</tr>
<tr>
<td>U V W X Y Z A B C D E F G H I J K L M N O P Q R S T</td>
</tr>
<tr>
<td>V W X Y Z A B C D E F G H I J K L M N O P Q R S T U</td>
</tr>
<tr>
<td>W X Y Z A B C D E F G H I J K L M N O P Q R S T U V</td>
</tr>
<tr>
<td>X Y Z A B C D E F G H I J K L M N O P Q R S T U V W</td>
</tr>
<tr>
<td>Z A B C D E F G H I J K L M N O P Q R S T U V W X Y</td>
</tr>
</tbody>
</table>

Figure 11: Vigenere cipher table [19]

The demonstration for vigenere cipher technique is clearly described in the example below.

Lets assume the plain text is "graduateproject" and key provided by the user is "masters". Then the cipher text generated based on the vigenere table is as below.

Plain text: graduateproject

Key: masters
evaluation:

\[
\text{graduate project}
\]

\[
\text{masters mastersm}
\]

The cipher text is \( srawylqpjhnuf \)

Figure 12 represents the Vigenere cipher where the user enters the text in the provided text area. The encryption and decryption process workflow is mainly based on figure 11.

![Vigenere cipher Encryption Decryption page](image)

**Figure 12: Vigenere cipher Encryption Decryption page**

4.1.6 One Time Padding:

One time padding is another kind of encryption technique in cryptography. A secret key is randomly generated. Length of the key is same as that of the plain text. Each character of the plain text is replaced with another character by doing the modular addition.
Figure 13 represents one time padding encryption technique for encrypting the plain text and decrypting the cipher text. In the figure 13, user will enter the text and by checking on the encrypt and decrypt options the process will run. The working principle of one time pad is as below Let us consider the 'M' as Message, 'C' as cipher text, 'K' as secret key and

\[ C = (M+K) \mod 26 \]
\[ M = (C-K) \mod 26 \]

By following this methodology the plain text is converted to cipher text and can maintain message confidential. The Key once used for a message can't be used again. The message can encrypted or decrypted using this key only.

**Figure 13: One time padding for Encryption and Decryption**

4.1.7 Compress the file:

This technique will compresses the file and changes the file into an unknown format like .cmp. It is very difficult to open the file without using the same application. By doing the decompression in the same application the original file can be retrieved.
Figure 14 represents the compress the file technique, it reduces the file size and changes the extension of the file with .cmp. It is very difficult to retrieve the correct information from the compressed file. Initially it chooses the file and compress it with filename.cmp. For retrieving the original content the receiver must use the similar function and upload the file with .cmp extension. Figure 14 represents that there undergoes two options encrypt and decrypt. While encrypt option is checked the user should choose a file for hiding. The chosen file will be compressed and changes the file extension with .cmp. For retrieving the original file, the decrypt option is checked and a file is chosen with a .cmp extension.

![File compression technique](image)

**Figure 14**: File compression technique

### 4.1.8 Lock the personal folder:

By using this technique, the files can be hidden in the personal computers or hard drives. This model is provided with a password to open the hidden files. By entering the
folder name and password a batch file is available. The batch file will creates a new folder. Drag all the files that wants to be hidden into the same folder and lock the file using the batch file. To open the hidden files, enter the password in the command prompt that is available after clicking on the batch file.

Figure 15 represents that personal folders can be hidden in the file systems or disk drives by using this technique. A folder name and a password should be entered. After entering the credentials a batch file is created and available for download. The batch file will open that folder with the name that is entered in the application. That folder will hides all the confidential information.

![Figure 15: Hiding the personal Files into a Folder](image-url)
4.2 Implementation of Proposed model ENHARM:

The Implementation of a proposed hybrid model for data hiding includes the combination of Md5 Hashing functionality, Encryption Algorithm and newly proposed Random Matrix Algorithm. All these algorithms provide multi layers of security to the confidential data. The screen shots for the proposed implemented tool are as follows

Figure 16 represents the encryption page for the proposed hybrid model where the confidential data is transformed into the cipher text which is sent to the receiver. The process of encrypting the text is different from the existing techniques of data hiding.. The user must enter the text and chooses an algorithm either AES or blow fish for encryption. The secret key is again encrypted by using the random matrix functionality. The receiver must have the valid matrix and also aware of the approach to decrypt the encrypted text.

Figure 16: Encryption Page for the Proposed Hybrid Model
Figure 17 represents the decryption page for the proposed model. The receiver must have the cipher text and encrypted key for decrypting the original message. Receiver should follow the similar approach of the encryption to retrieve the original message from the provided encrypted text. Figure 17 screen shot displays that the user must enter the encrypted text and encrypted text secret key for decryption.

4.2.1 Functionalities of the Proposed Model (ENHARM):

The proposed system is divided into two main modules

1. Encryption:
   Process of transforming plain text to encrypted text.
2. Decryption:

Reverse process of Encryption or transformation of cipher text into plain text.

4.2.1.1 Encryption:

1. Input:

In the encryption module, the plain text or confidential data is given as input to the proposed system. The data consists of lines of text that can be entered in text area provided and there also exists an option to choose a file for encryption.

2. Encryption process:

In the proposed system, security to the confidential data is provided in multi-layers. The approach and the substitution methodology used in the proposed system are different from existing hybrid models. In the process of encryption, transformation from plain text into cipher text is clearly described in figure 18.

Figure 18 specifies the Encryption architecture for the proposed model ENHARM
The proposed system is a hybrid encryption process with a new approach emerged from the hashing algorithm, symmetric algorithm and randomly generated matrix. The encryption process mainly undergoes three phases.

Phase 1:

A cipher text is generated after passing the plain text through MD5 algorithm. The converted message is used as secret key in the next step of encryption in symmetric encryption algorithm.
Phase 2:

The symmetric key that is generated by MD5 is used in symmetric encryption algorithm. Plain text is again converted into an unreadable format by using the symmetric encryption algorithms Advanced Encryption Standard (AES) and Blow Fish. The cipher text is then sent to the intended recipient.

Phase 3:

The message digest is again encrypted using randomly generated matrix functionality. After symmetric encryption algorithm is applied, the symmetric key is converted to another cipher text.

3. Output:

After the Encryption process is done, the results obtained are cipher text of the plain text by using Encryption algorithm and cipher text of the symmetric key by using random matrix functionality.

4.2.1.2 Decryption:

1. Input:

In the Decryption module, the encrypted text of symmetric key and plain text are given as an input to the proposed system. The cipher text of symmetric key is first decrypted in order to open original plain text.
2. Decryption Process:

Technical definition for the term decryption is reverse process of encryption. Transformation of cipher text to plain text is called Decryption. To provide more security, encryption is done in the multi-layer. The same process is to be done in a reverse manner for obtaining original plain text. The proposed system also provides privacy and protection to the data by indulging confidentiality, integrity and availability. The decryption process is clearly described in the following architecture for obtaining original data without any loss or damage.

Figure 19 represents the decryption architecture for the proposed model ENHARM
The decryption process is followed by three main phases for retrieving the original data. The step by step process of decryption process is explained below.

Phase 1:

Initially the Cipher text of the symmetric key is to be decrypted using randomly generated matrix of order 6*6. The transformed text is used as a symmetric key for the encryption algorithm for conversion of cipher text to plain text.

Phase 2:

The secret key that is used for encrypting plain text to cipher text is the same key that is used for conversion of cipher text to plain text. The key is extracted from the randomly generated matrix. There could be an option for choosing same encryption algorithm that is used in encryption in order to decrypt the original text from cipher text. The original text is retrieved by applying the same reverse process.

Phase 3:

The original data is passed through the MD5 hash function to obtain a message digest. The message digest is also called secret key or symmetric key.

Comparison:

The secret key that is retrieved after random matrix functionality and another secret key that is generated after the hashing is applied to plain text must be same, to display the original data to the receiver.
3. Output:

By applying the decryption process, the text obtained is the plain text. The decryption process performs the transformation of cipher text to plain text. The plain text is displayed in text area provided. To retrieve the correct information, secret key is to be generated using the random matrix, with respective to encryption algorithm used in encryption.

4.2.2. Example for Proposed System ENHARM:

4.2.2.1 Encryption Module:

Input:

In this module, plain text is given as input to the proposed system.

Ex: Plain Text : "graduate project final report"

Encryption Process:

Phase 1:

The plain text is encrypted using MD5. The produced encrypted text is used as secret key for the encryption algorithm. The secret key generated is a 32 bit string.

Secret key: 6499d57cf5bcaff49b561ce37995bd40

Phase 2:

The Encryption algorithm is chosen and secret key is generated after the MD5 is applied which is used for locking the encrypted text. The encryption algorithms like AES, DES, etc are used to produce cipher text.
Phase 3:

The secret key is again encrypted using the randomly generated matrix. The working procedure is explained below.

The secret key is "6499d57cf5bcaff49b561ce37995bd40".

The matrix of order 6*6 (row*column) is generated randomly which consists of 26 alphabets and numbers from 0 till 9.

Let's assume the randomly generated matrix is

```
  m  w  h  2  8  n
 y  z  p  d  i  6
 7  v  3  u  s  f
 q  c  5  j  1  k
A  9  x  g  0  e
L  o  4  t  b  r
```

The position of the character is to be determined using the row*column of a matrix.

For example the position of character "a" is expressed as 5*1. By using the concept of substitution methodology, the character "a" is replaced by another character. The position of the character "a" is to be reversed by column*row, and that original character is to be interchanged by another character of position column*row

Actual position for a = 5*1

Replacement position for a = 1*5 = "m"
Applying the above concept for encrypting the secret key into another cipher text,

Generated secret key is: "6499d57cf5bcaff49b561ce37995bd40"

Encrypted secret key is: "ofiicuhd4ued844fieuogdb3hiuecf0"

Output:

The cipher text of the plain text and secret key are the outputs generated by undergoing the encryption process and randomly generated matrix functionality.

4.2.2.2 Decryption Module:

Input:

The cipher text of the secret key and plain text are inputs to the proposed solution.

Decryption process:

Decryption is the reverse process of the encryption. To retrieve the original text, the receiver must be aware of the approach and algorithms used in the encryption. By reversing the process, original text is displayed.

Phase 1:

The secret key is decrypted by using the random matrix functionality in the reverse process.

The matrix used in the encryption is

m w h 2 8 n
y z p d i 6
7 v 3 u s f
q c 5 j 1 k
Encrypted secret key is: "ofiicuhd4ued844fieuogdb3hiuecf0"

Generated secret key is: "6499d57cf5bcaff49b561ce37995bd40"

Phase 2:

There will be an option to chose the encryption algorithm used and a secret key is provided. The original text is displayed. The decryption process in the proposed system is very easy and provides security to data.

Plain text: "graduate project final report"

Output:

The output generated is the original plain text or file after the decryption process is done.
5. TESTING AND EVALUATIONS

Testing is the process of finding errors during the program execution. By testing a program, bugs can be identified. Different kinds of testing cases are performed to identify the errors and solve them in order to make the program function properly. Testing software also helps to check whether the application is reliable and works correctly or not.

The project is tested to identify the errors while executing the program. Implementation of a project is divided into modules. Testing is to be done for every module to make sure all the modules are working correctly.

These test cases are used to check the working of the project:

Test case 1:

To verify whether the user can successfully register,

There are some things to be checked and verified

- To access a tool, the user must register by opening a registration page
- The user must enter all the information that is required
- New user registration is done and can access the tool

Figure 20 represents the interface for new user registration in the application. New user enters the details such as username, E-mails, password to register. After entering the details, by clicking on the submit button the user is successfully registered.
Test Case 2:

To check if the existing user can successfully login or not, the following things need to be verified.

- The registered user must be able to open the login page
- The user must be able to enter login credentials to open an application
- In case of wrong username or password, the user is able to see that the entered username and password are incorrect
- The registered user should be able to successfully login and view the page

Figure 21 represents the login page for the registered user to enter the username and password. The username and the password credentials are entered in this page to open the application.
Figure 21: Login page where the registered user enters the username and password.

Figure 22 represents the home page that consists of most common data hiding techniques and the proposed hybrid model.

Figure 22 : Home page for the Implemented tool.
Test case 3:

Checks whether the user is able to view the encryption or decryption phase when the user clicks on Encrypt or Decrypt check boxes in the proposed hybrid model. The following things need to be observed.

- The user can view only the Encrypt phase when user check in the Encrypt option
- The user can view only the Decrypt phase when user check in the Decrypt option

Figure 23 represents the proposed hybrid model, where the user can encrypt or decrypt the confidential information.

Figure 23: The Proposed Hybrid Model for Encrypting and Decrypting the text.
Figure 24 represents whether the user can view the Encrypt phase when the user check the Encrypt option. By clicking this option the user can encrypt the confidential information. The user is able to enter text and the encryption algorithm is processed in this page.

![Image of the Encrypt phase of the Hybrid model](image)

**Figure 24: The Encrypt phase of the Hybrid model.**

Figure 25 represents whether the user is able to view the decrypt phase when the user checks the decrypt option. By opening the decrypt phase, the user can decrypt the encrypted information. Decrypting the original text, user needs to provide the secret key and encrypted text.
Test case 4:

Testing the file options

In this case, it tests whether the options "Upload File" or "Enter Text" are working properly or not. The following things need to be observed.

- Considering Upload File option, checks whether the correct file is chosen or not
- In Enter Text option, it tests whether the typed text is correctly taken from keyboard or not

Figure 26 represents that when user wants to encrypt a file from the file system, the user needs to browse a file. It tests whether the user is able to browse a file or not.
Figure 26: Browsing a file from the file system.

Figure 27 represents that the user is able to enter a text in the text area provided. It tests whether it is taking the correct input that is typed from the keyboard or not.

Figure 27: Entering text in the text area provided.
Test case 5:

Testing the MD5 Functionality

- Text should be entered in the text area provided
- The entered text can be edited or deleted
- By taking an action and applying the MD5, the encrypted text is displayed

Figure 28 represents, that the confidential information needs to be entered in the text area provided to convert it into encrypted text.

![Figure 28: Entering the Text to check the MD5 functionality](image)

Figure 29 represents that after applying the MD5 functionality, the encrypted text or secret key is displayed for the text that is entered.
Figure 29: Generating a Symmetric key

Test case 6:

Testing the Symmetric Encryption functionality

To test this phase, the following things need to be checked and verified.

- Whether the user can chose the specified algorithm from the list
- The algorithm is working properly or not
- The user must provide a secret key that is extracted from the Md5 encryption technique
- The user is able to see the encrypted text

Figure 30 describes that user must select a symmetric encryption algorithm from the list. In this test case the user can successfully select an algorithm.
Figure 30: A Specified Algorithm is Selected

Figure 31 represents that the user is able to enter the secret key for the algorithm by selecting from the list. By clicking on the “apply encryption button”, the user must view the encrypted text in the box.
Test case 7:

Testing the Random matrix functionality

The secret key is encrypted using the random matrix functionality. The following things need to be tested and validated.

- Checks whether the user is able to enter secret key used in the text area provided
- Checks whether the random matrix functionality is working properly or not
- Tests whether the user is able to see the encrypted text of the secret key or not
- By clicking a button "Apply Random Matrix", it checks whether the matrix is displayed or not
Figure 32 represents that a Random Matrix Functionality is applied for the secret key provided in the box. The secret key represents the encrypted text of the Md5.

![Image of data hiding tool](image)

**Figure 32: Random Matrix Functionality is applied for the Secret key.**

Figure 33 clearly describes that after applying the functionality, the random matrix is displayed and the encrypted text of the secret key is also displayed. The matrix helps in converting the secret key into encrypted text.
Test case 8:

Testing whether the required information can be entered into the text boxes or not.

The following things need to be checked and verified.

There are some things to be checked and verified

- The user must be able to provide the encrypted text and secret key.

- Checks whether the user is able to enter the encrypted text and secret key in the text are provided.

  Figure 34 represents that the user is able to enter the required information in the text boxes provided.
Test case 9:

Testing the Decryption process functionality

It checks whether the original content is displayed or not after the decryption process is successfully completed.

There are some things to be checked and verified

- Checks whether the original text is being displayed or not after the decryption process is applied
- Need to check whether the correct algorithm, matrix is being displayed
- Compare and check whether the secret key after decryption and secret key after Md5 are same or not

Figure 34: Entering the encrypted text and secret key in the text boxes.
Figure 35 clearly describes that by clicking on the Decrypt button, the original text, algorithm used and matrix are displayed.

Figure 35: Displays the original text after the Decryption process.
6. CONCLUSION

In this project, a new hybrid data hiding tool is proposed to hide the confidential data from the intruders by providing multiple layers of security. Since cyber crime is increasing day by day, it is necessary to provide more security to the data from unauthorized access. The procedure and approach used in this technique are completely different from the existing mechanisms. The tool is based upon the combination of hashing, standard encryption algorithms and random matrix functionality. This tool is developed to provide more security to the sensitive data by making use of matrix functionality. In this project, the confidential data is hidden by providing multiple levels with the combination of encryption algorithms and random matrix algorithm.

The tool is provided with a user friendly interface to encrypt and decrypt the data. The approach used in this technique is very difficult for the hacker to decode confidential information. Using the functionality of the random matrix, more security is provided to the data and it is also easy for the receiver to decode the original content with in a short span of time and also with less complexity.

The Implementation of a tool consists of the proposed hybrid model and most common data hiding techniques. The complete application is password protected for availability only to the intended users. Integration of all these techniques into a single application helps the user in accessing different hiding techniques from the same place. It can also be used for teaching purpose as a virtual lab for the students can be created where they learn and practice all these techniques.
7. BIBILOGRAPHY


[6] David Bernstein, James Madison University, "An Introduction to Information Hiding 


[8] Zaidoon Kh. AL-Ani, A.A.Zaidan, B.B.Zaidan and Hamdan.O.Alanazi, "Overview: Main 

Based on Block Division”, Congress on Image and Signal Processing, Vol. 1, 27-30 May 
2008, pp. 365-369

using linked list concepts and LSB technique”, Journal of American Science, Vol. 7, No. 6, 
2011, pp. 97-100.

“A Sub sampling and Interpolation Technique for Reversible Histogram Shift Data Hiding”, 
Image and Signal Processing, Lecture Notes in Computer Science, Vol. 6134, 2010, 
Publisher: Springer Berlin/Heidelberg, pp. 384-393.


