1. **INTRODUCTION**

****

The word steganography is derived from the Greek words “stegos” meaning “cover” and “grafia” meaning “writing” defining it as “covered writing”. Steganography is one such pro-security innovation in which secret data is embedded in a cover. The notion of data hiding or steganography was first introduced with the example of prisoners' secret message by Simmons in 1983.

Steganography and cryptography are closely related. Cryptography scrambles messages so they cannot be understood. Steganography on the other hand, will hide the message so there is no knowledge of the existence of the message in the first place. In some situations, sending an encrypted message will arouse suspicion while an ”invisible” message will not do so. Both sciences can be combined to produce better protection of the message. In this case, whenthe steganography fails and the message can be detected, it is still of no use as it is encrypted using cryptography techniques.

There exist two types of materials in steganography: message and carrier. Message is the secret data that should be hidden and carrier is the material that takes the message in it.

There are many types of steganography techniques. In this project, we are going to take a short look at text steganography technique-Plaintext Steganography.

* 1. **WHAT IS TEXT STEGANOGRAPHY?**

****

Figure 1: Steganography Meaning

Text steganography can involve anything from changing the formatting of an existing text, to changing words within a text, to generating random character sequences or using context-free grammars to generate readable texts. Textsteganography is believed to be the trickiest due to deficiency of redundant information which ispresent in image, audio or a video file. The structure of text documents is identical with what weobserve, while in other types of documents such as in picture, the structure of document isdifferent from what we observe. Therefore, in such documents, we can hide information byintroducing changes in the structure of the document without making a notable change in the concerned output.Unperceivable changes can be made to an image or an audio file, but, intext files, even an additional letter or punctuation can be marked by a casual reader. Storingtext file require less memory and its faster as well as easier communication makes it preferableto other types of steganographic methods.

* 1. **TYPES OF TEXT STEGANOGRAPHY METHODS?**

Some of the popular approaches of text steganography are:

* + 1. **Plaintext steganography**

In this technique the message is hidden within a plain text file using differentschemes like use of selected characters, extra white spaces of the cover text etc.

* **Use of selected characters of cover Text.**

Sender sends a series of integer number (Key) to the recipient with a prioragreement that the secret message is hidden within the respective position ofsubsequent words of the cover text. For example the series is ‘1, 1, 2, 3, 4, 2, 4,’and the cover text is **“A team of five men joined today”**. So the hiddenmessage is **“Atfvoa”**. A “0” in the number series will indicate a blank space in the recovered message. The word in the received cover text will be skipped if thenumber of characters in that word is less than the respective number in theseries (Key) which shall also be skipped during the process of message unhide.

* **Use of extra white space characters of cover text.**

A number of extra blank spaces are inserted between consecutive words of covertext. This numbers are mapped to a hidden message through an index of a lookup table. For example extra three spaces between adjacent words indicatethe number “3” which subsequently indicates a specific text of a look-up tablewhich is available to the both communicating parties as a prior agreement.

* + 1. **Line Shift**

In this method, secret message is hidden by vertically shifting the text lines to some degree. A line marked has two unmarked control lines one on either side of it for detecting thedirection of movement of the marked line. To hide bit 0, a line is shifted up and to hide bit 1, the line is shifted down. Determination of whether the line has been shifted up or downis done by measuring the distance of the centroid of marked line and its control lines. If thetext is retyped or if a character recognition program (OCR) is used, the hidden informationwould get destroyed. Also, the distances can be observed by using special instruments of distance assessment.

* + 1. **Word Shift**

In this method, secret message is hidden by shifting the words horizontally, i.e. left or right torepresent bit 0 or 1 respectively. Words shift are detected using correlation method thattreats a profile as a waveform and decides whether it originated from a waveform whose middleblock has been shifted left or right. This method can be identified less, because change ofdistance between words to fill a line is quite common. But if someone knows thealgorithm of distances, he can compare the stego text with the algorithm and obtain the hiddencontent by using the difference. Also, retyping or using OCR programs destroys the hidden information.

* + 1. **Syntactic Method**

This technique uses punctuation marks such as full stop (.), comma (,), etc. to hide bits 0 and 1.But problem with this method is that it requires identification of correct places to insert punctuation marks. Therefore, care should be taken in using this method as readers cannotice improper use of the punctuations.

* + 1. **White Steg**

This technique uses white spaces for hiding a secret message. There are three methods of hidingdata using white spaces. In Inter Sentence Spacing, we place single space to hide bit 0 and twospaces to hide bit 1 at the end of each terminating character [9]. In End of Line Spaces, fixednumber of spaces is inserted at the end of each line. For example, two spaces to encode one bitper line, four spaces to encode two bits and so on. In Inter Word Spacing technique, one spaceafter a word represents bit 0 and two spaces after a word represents bit 1. But, inconsistent useof white space is not transparent.

* + 1. **Spam Text**

HTML and XML files can also be used to hide bits. If there are different starting and closingtags, bit 0 is interpreted and if single tag is used for starting and closing it, then bit 1 is interpreted. In another technique, bit 0 is represented by a lack of space in a tag and bit 1 isrepresented by placing a space inside a tag.

* + 1. **SMS-Texting**

SMS-Texting language is a combination of abbreviated words used in SMS. We can hidebinary data by using full form of word or its abbreviated form. A codebook is made whichcontains words and their corresponding abbreviated forms. To hide bit 0, full form of the wordis used and to hide bit 1, abbreviated form of word is used.

* + 1. **Feature Coding**

In feature coding, secret message is hidden by altering one or more features of the text. A parserexamines a document and picks out all the features that it can use to hide the information.For example, point in letters i and j can be displaced, length of strike in letters f and t can bechanged, or by extending or shortening height of letters b, d, h, etc. . A flaw of thismethod is that if an OCR program is used or if re-typing is done, the hidden content would getdestroyed.

* + 1. **SSCE (Secret Steganographic Code for Embedding)**

This technique first encrypts a message using SSCE table and then embeds the cipher text in acover file by inserting articles a or an with the nonspecific nouns in English language using a certain mapping technique. The embedding positions are encrypted using the same SSCEtable and saved in another file which is transmitted to the receiver securely along with the stegofile.

* + 1. **Word Mapping**

This technique encrypts a secret message using genetic operator crossover and then embeds theresulting cipher text, taking two bits at a time, in a cover file by inserting blank spaces betweenwords of even or odd length using a certain mapping technique . The embedding positionsare saved in another file and transmitted to the receiver along with the stego object.

* + 1. **MS Word Document**

In this technique, text segments in a document are degenerated, mimicking to be the work of anauthor with inferior writing skills, with secret message being embedded in the choice ofdegenerations which are then revised with changes being tracked. Data embedding isdisguised such that the stego document appears to be the product of collaborative writing.

* + 1. **Cricket Match Scorecard**

In this method, data is hidden in a cricket match scorecard by pre-appending a meaningless zero

before a number to represent bit 1 and leaving the number as it is to represent bit 0.

**1.2.13. CSS (Cascading Style Sheet)**

This technique encrypts a message using RSA public key cryptosystem and cipher text is thenembedded in a Cascading Style Sheet (CSS) by using End of Line on each CSS style properties,exactly after a semicolon. A space after a semicolon embeds bit 0 and a tab after a semicolon embeds bit 1.

* 1. **PROJECT OVERVIEW**

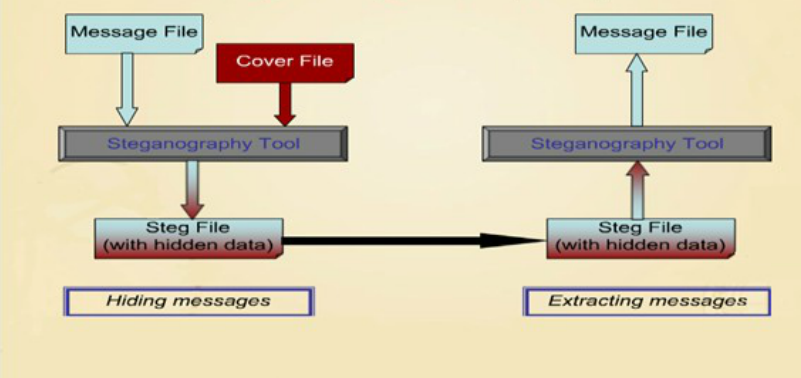
****

Figure 1.3: Process of Steganography

In this technique the message is hidden within a plain text file using differentschemes like use of selected characters, extra white spaces of the cover text etc.

* **Use of selected characters of cover Text.**

Sender sends a series of integer number (Key) to the recipient with a prior agreement that the secret message is hidden within the respective position of subsequent words of the cover text. For example the series is ‘1, 1, 2, 3, 4, 2, 4,’ and the cover text is **“A team of five men joined today”**. So the hidden message is **“Atfvoa”**. A “0” in the number series will indicate a blank space in the recovered message. The word in the received cover text will be skipped if the number of characters in that word is less than the respective number in the series (Key) which shall also be skipped during the process of message unhide.

* **Use of extra white space characters of cover text.**

A number of extra blank spaces are inserted between consecutive words of cover text. This numbers are mapped to a hidden message through an index of a lookup table. For example extra three spaces between adjacent words indicate the number “3” which subsequently indicates a specific text of a look-up table which is available to the both communicating parties as a prior agreement.

* 1. **PURPOSE AND SCOPE**
     1. **Purpose :**

The growth of high speed computer networks and that of the Internet, in particular, has increased theease of Information Communication. Ironically, the cause for the development is also of the apprehension- use of digital formatted data. In comparison with Analog media, Digital media offers several distinctadvantages such as high quality, easy editing, high fidelity copying, compression etc. But this typeadvancement in the field of data communication in other sense has hiked the fear of getting the datasnooped at the time of sending it from the sender to the receiver. So, Information Security is becomingan inseparable part of Data Communication. In order to address this Information Security,Steganographyplays an important role. Steganography is the art and science of writing hidden messagesin such a way that no one apart from the sender and intended recipient even realizes there is a hiddenmessage.

* + 1. **Scope:**

In today’s world, we often listen a popular term “Hacking”. Hacking is nothing but an unauthorized access of data which can be collected at the time of data transmission. With respect to steganography this problem is often taken as Steganalysis. Steganalysis is a process in which a steganalyzer cracks the cover object to get the hidden data. So, whatever be the technique will be developed in future, degree of security related with that has to be kept in mind. It is hoped that Dual Steganography, Steganography along with Cryptography may be some of the future solution for this above mentioned problem.

* 1. **APPLICATIONS**
* Steganography can be used anytime you want to hide data. There are many reasons to hide data but they all boil down to the desire to prevent unauthorized persons from becoming aware of the existence of a message. With these new techniques, a hidden message is indistinguishable from white noise. Even if the message is suspected, there is no proof of its existence. In the business world steganography can be used to hide a secret chemical formula or plans for a new invention.
* Steganography can also be used for corporate espionage by sending out trade secrets without anyone at the company being any the wiser.
* Terrorists can also use steganography to keep their communications secret and to coordinate attacks. All of this sounds fairly nefarious, and in fact the obvious uses of steganography are for things like espionage. But there are a number of peaceful applications. The simplest and oldest are used in map making, where cartographers sometimes add a tiny fictional street to their maps, allowing them to prosecute copycats. A similar trick is to add fictional names to mailing lists as a check against unauthorized resellers.
* Most of the newer applications use steganography like a watermark, to protect a copyright on information. Photo collections, sold on CD, often have hidden messages in the photos which allow detection of unauthorized use. The same technique applied to DVDs is even more effective, since the industry builds DVD recorders to detect and disallow copying of protected DVDs.

1. **SOFTWARE AND HARDWARE REQUIREMENTS**
   1. **SOFTWARE REQUIREMENT:**

Operating Platform : Windows XP / 07 / 08

Database : Text document

Software : jdk1.7.0\_25

Front End Tool : Command Prompt

Language : Java

Documentation Tool : MS WORD 2007

* 1. **HARDWARE REQUIREMENT:**

Processor : Intel Pentium / i3 / i5 / i7

Speed : 2.30 GHz

Hard Disk : 6 MB

Random Access Memory : 128 MB

1. **PLATFORM USED**
   1. **JAVA**

Java is a set of several computer software products and specifications from Oracle Corporation that provides a system for developing application software and deploying it in a cross-platform computing environment. Java is used in a wide variety of computing platforms from embedded devices and mobile phones on the low end, to enterprise servers and supercomputers on the high end. While less common, Java applets are sometimes used to provide improved and secure functions while browsing the World Wide Web on desktop computers.

Java is a computer programming language that is concurrent, class-based, object-oriented, and specifically designed to have as few implementation dependencies as possible. It is intended to let application developers "write once, run anywhere" (WORA), meaning that code that runs on one platform does not need to be recompiled to run on another. Java applications are typically compiled to bytecode (class file) that can run on any Java virtual machine (JVM) regardless of computer architecture. Java is, as of 2014, one of the most popular programming languages in use, particularly for client-server web applications, with a reported 9 million developers. Java was originally developed by James Gosling at Sun Microsystems (which has since merged into Oracle Corporation) and released in 1995 as a core component of Sun Microsystems' Java platform. The language derives much of its syntax from C and C++, but it has fewer low-level facilities than either of them.

The original and reference implementation Java compilers, virtual machines, and class libraries were developed by Sun from 1991 and first released in 1995. As of May 2007, in compliance with the specifications of the Java Community Process, Sun relicensed most of its Java technologies under the GNU General Public License. Others have also developed alternative implementations of these Sun technologies, such as the GNU Compiler for Java (bytecode compiler), GNU Classpath (standard libraries), and IcedTea-Web (browser plugin for applets).

1. **CONCLUSION**

We believe that steganography when combined withencryption provides a secure means of secret communication between two parties. Our application, with its text analysis and ranking capability is a significant improvement on current steganography tools.

1. **REFRENCE**

http://security.stackexchange.com/questions/20414/steganography-to-hide-text-within-text

[1] S. Changder, D. Ghosh, and N. C. Debnath, “Linguistic approach for text steganography throughIndian text,” 2010 2nd *Int. Conf. on Computer Technology and Development*, 2010, pp. 318-322.

[2] Ming, Chen, Z. Ru, N. Xinxin, and Y. Yixian, “Analysis of CurrentSteganography Tools: Classifications & Features”, Information Security Centre, Beijing University of Posts & Telecommunication, Beijing, December 2006.

[3] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, “Techniques for data hiding,” *IBM Systems*

*Journal*, vol.35, pp. 313-336, 1996.

1. **APPENDIX**

**Appendix A – Source Code**

CODE FOR COVER TEXT:

import java.io.\* ;

public class writefile

{

String name ;

String textline ;

public writefile(String name1 , String data)

{

name = name1 ;

textline = data ;

}

public void writeTofile() throws IOException

{

FileWriter write = new FileWriter(name);

PrintWriter print\_line = new PrintWriter(write);

print\_line.printf("%s" +"%n" , textline);

print\_line.close();

}

}

CODE FOR ENCRYPTION:

import javax.swing.JOptionPane;

import java.io.\*;

import java.io.Console;

class sender{

public static void main(String ... ar) throws Exception

{

Console c=System.console();

String key,ctext;

System.out.println("\t\t\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*");

System.out.println("\t\t\t\tTEXT STEGANOGRAPHY");

System.out.println("\t\t\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*");

System.out.println("ENCRYPTION DETAILS:\n\n");

ctext=c.readLine("ENTER COVER TEXT:");

key=c.readLine("\n\n\n\n\nENTER KEY:");

writefile wf=new writefile("cover.txt",ctext);

wf.writeTofile();

System.out.println("\n\n\t\t\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*");

System.out.println("\n\nFILE HAS BEEN GENERATED!!!");

System.out.println("\nKEY IS :"+key);

}

}

CODE FOR DECRYPTION:

import java.io.\* ;

import java.io.Console ;

public class Decrypter

{

public static void main(String [] ar) throws Exception

{

Console input = System.console();

System.out.println("\n\nDECRYPTION DETAILS:\n\n");

String key = input.readLine("\nPLEASE ENTER THE KEY IN ORDER TO GET THE CODE:\n\n\n ");

receiver obj = new receiver("cover.txt" , key);

obj.decode();

}

}

class receiver

{

String name ;

String key ;

/\* ------------------------ Constructor definition -------------------------------- \*/

public receiver(String name1 , String key1)

{

name = name1 ;

key = key1 ;

}

/\* ------------------------- Decoding Code ---------------------------- \*/

public void decode() throws IOException

{

File f = new File(name);

FileInputStream fis = new FileInputStream(f);

int data ;

int words = 0 ;

int k = 0 ;

char []ch = new char[100] ;

char []keyArr = key.toCharArray();

/\* ------------------- Reading File Contents ------------------------ \*/

while(true)

{

data = fis.read();

ch[k] = (char)data;

k++ ;

if(data == -1)

break;

}

/\* ------------------- End of file reading ------------------------ \*/

char []ch2 = new char[k] ;

for(int i = 0 ; i<k-1 ; i++)

ch2[i] = ch[i] ;

String fullmessage = new String(ch2) ;

fullmessage = fullmessage.trim() ;

/\* ------------- Calculation of number of words in the cover text----------- \*/

for(int i = 0 ; i<k ; i++)

{

if((int)ch[i] == 32)

words++ ;

else

continue;

}

if(words+1 < keyArr.length)

{

System.out.println("Key length is greater than the cover text words");

System.exit(0);

}

int l = 0 ;

int p = 0;

int m ;

String []hiddenMsg = new String[words+1] ;

/\* ------------------ Extracting subtrings from the main cover text string -------------------- \*/

for(m =0 ; m<fullmessage.length() ; m++)

{

if(fullmessage.charAt(m) == ' ')

{

hiddenMsg[p] = fullmessage.substring(l , m);

l = m+1 ;

p++;

}

else

continue;

}

hiddenMsg[p] = fullmessage.substring(l , m) ;

char []code = new char[keyArr.length] ;

System.out.println("/\*\*\*\*\*\*\*\*\*\*\*\*\*\* YOUR CODE \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/\n");

System.out.print(" ");

for(int i = 0 ; i <keyArr.length ; i++)

{

int z = (int)keyArr[i] - 49 ;

if(hiddenMsg[i].length() < z+1)

continue;

else{

code[i] = hiddenMsg[i].charAt(z) ;

System.out.print(code[i]);

}

}

System.out.println();

}

}

**Appendix B – Screen Shots:**

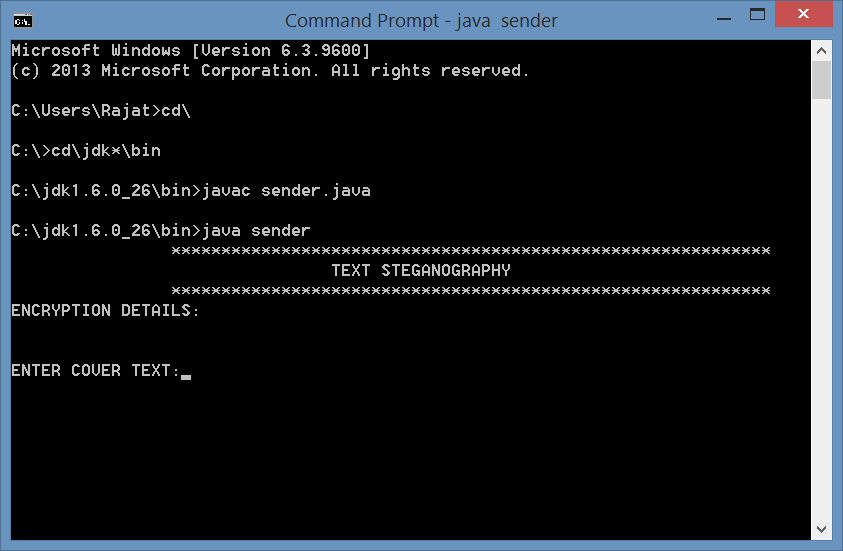
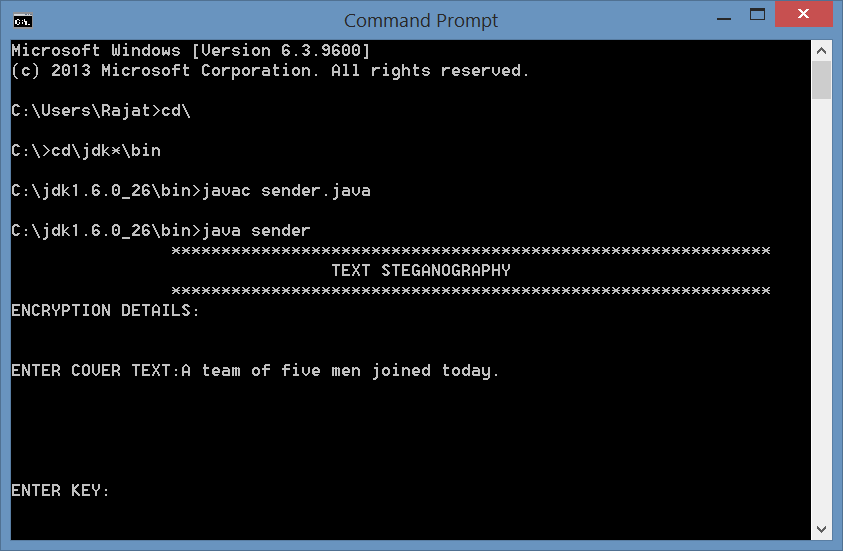
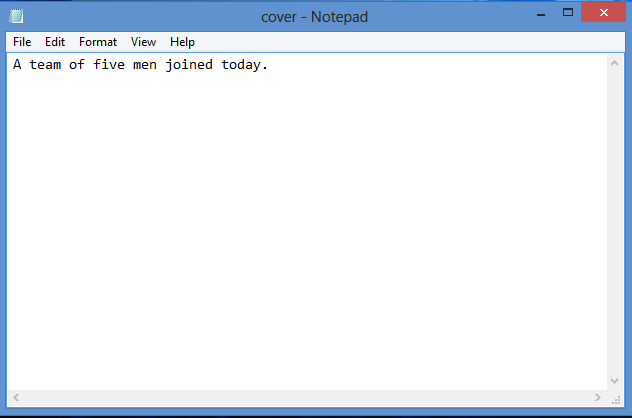
****

Figure 6.1: Enter Cover Text



F

Figure 6.2: Creating Cover Text

**** Figure 6.3 : Notepad file of Cover Text

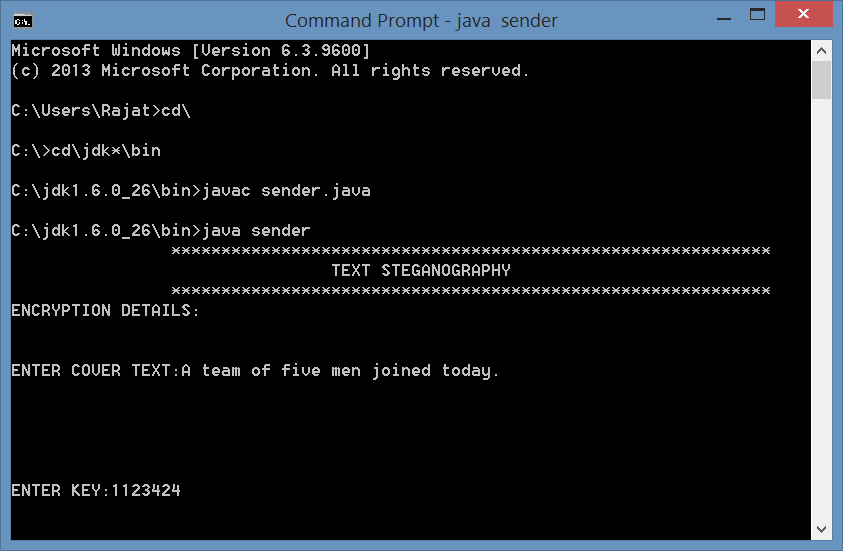


Figure 6.4 : Key for Cypher Text

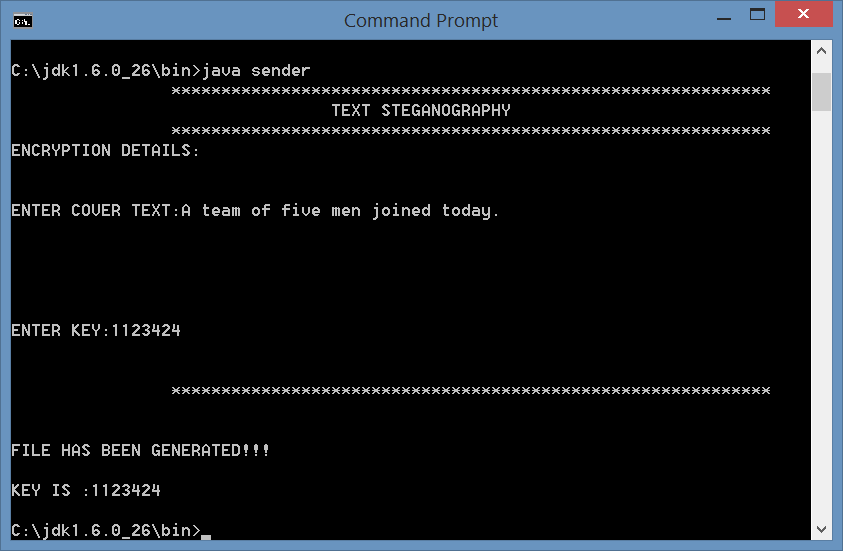
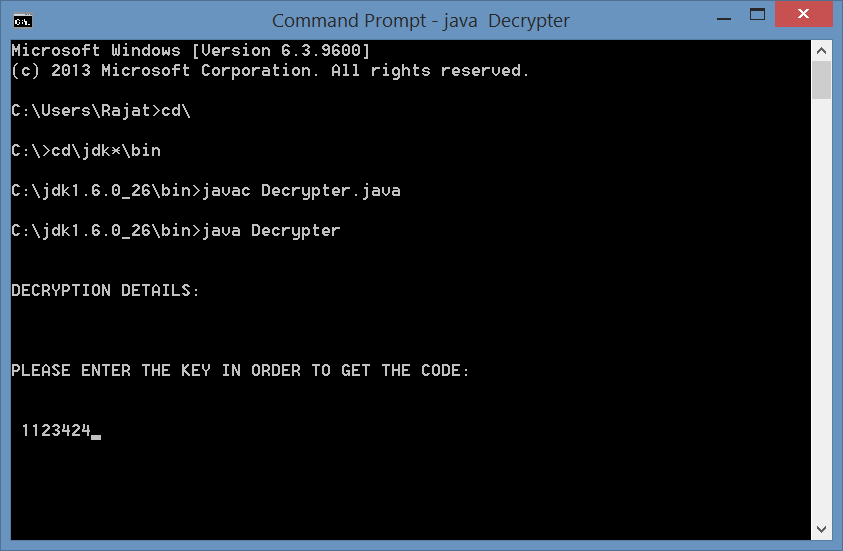
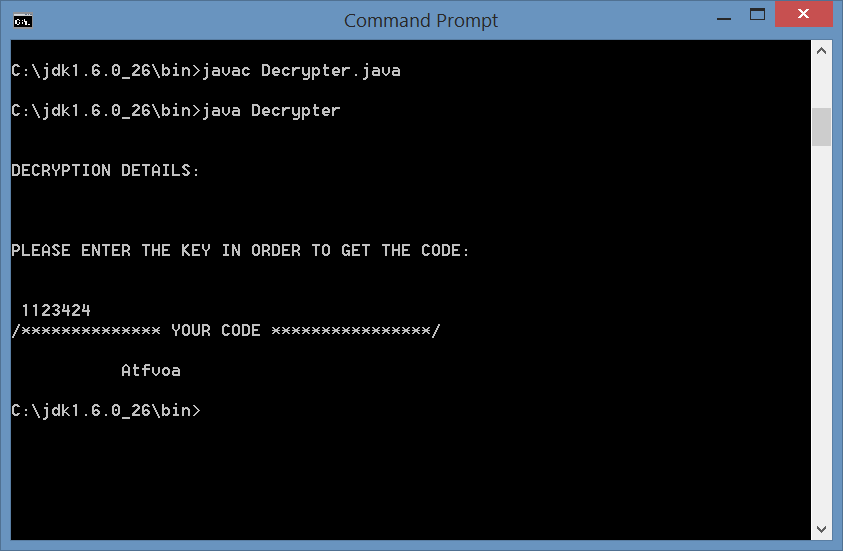


Figure 6.5 : Sender Side

Figure 6.6 : Enter key at Receiver sideFigure 6.7 : Decrypted Hidden Message